The EMBO Journal Vol.18 No.13 pp.3820–3833, 1999

DDP1, a single-stranded nucleic acid-binding protein of *Drosophila*, associates with pericentric heterochromatin and is functionally homologous to the yeast Scp160p, which is involved in the control of cell ploidy

Alfred Cortés, Dori Huertas, Laura Fanti[1,2], Sergio Pimpinelli[2], Francesc X.Marsellach, Benjamí Piña and Fernando Azorín[3]

Departament de Biologia Molecular i Cel.lular, Centre d'Investigació i Desenvolupament—CSIC, Jordi Girona Salgado 18–26, 08034 Barcelona, Spain, [1]Istituto di Genetica, Università di Bari, 70126 Bari and [2]Istituto Pasteur, Fondazione Cenci Bolognetti and Dipartimento di Genetica e Biologia Molecolare, Università 'La Sapienza', 00185 Roma, Italy

[3]Corresponding author
e-mail: fambmc@cid.csic.es

The centromeric dodeca-satellite of *Drosophila* forms altered DNA structures *in vitro* in which its purine-rich strand (G-strand) forms stable fold-back structures, while the complementary C-strand remains unstructured. In this paper, the purification and characterization of DDP1, a single-stranded DNA-binding protein of high molecular mass (160 kDa) that specifically binds the unstructured dodeca-satellite C-strand, is presented. In polytene chromosomes, DDP1 is found located at the chromocentre associated with the pericentric heterochromatin but its distribution is not constrained to the dodeca-satellite sequences. DDP1 also localizes to heterochromatin in interphase nuclei of larval neuroblasts. During embryo development, DDP1 becomes nuclear after cellularization, when heterochromatin is fully organized, being also associated with the condensed mitotic chromosomes. In addition to its localization at the chromocentre, in polytene chromosomes, DDP1 is also detected at several sites in the euchromatic arms co-localizing with the heterochromatin protein HP1. DDP1 is a multi-KH domain protein homologous to the yeast Scp160 protein that is involved in the control of cell ploidy. Expression of DDP1 complements a *Δscp160* deletion in yeast. These results are discussed in view of the possible contribution of DNA structure to the structural organization of pericentric heterochromatin.

Keywords: centromere/*Drosophila* dodeca-satellite/heterochromatin/ssDNA/vigilins

Introduction

The centromere is a specialized chromosomal structure which is essential for the accurate segregation of chromosomes during mitosis and meiosis (for a review see Choo, 1997). In all the species studied to date, with the exception of the budding yeast *Saccharomyces cerevisiae*, centromeric DNAs are characteristically enriched in highly repetitive satellite DNA sequences that, as was recently shown, are essential for proper centromere functioning both in humans (Harrington *et al.*, 1997; Ikeno *et al.*, 1998) and in *Drosophila* (Murphy and Karpen, 1995; Sun *et al.*, 1997). Remarkably, centromeric satellite DNAs are poorly conserved through evolution and, even within the same species, different chromosomes contain different arrays of satellite DNA sequences at the centromere (reviewed in Choo, 1997). Furthermore, centromeric satellites can also be found at other chromosomal locations showing no detectable centromere function (Lohe *et al.*, 1993; Sullivan and Schwartz, 1995), and vice versa, DNA sequences located outside of the centromere might acquire centromere activity acting as neo-centromeres (Voullaire *et al.*, 1993; Sacchi *et al.*, 1996; du Sart *et al.*, 1997; Williams *et al.*, 1998). These observations suggest that the primary nucleotide sequence might not be the only determinant of centromere formation and function.

Centromeres of higher eukaryotes are embedded within large blocks of heterochromatin and the formation of a centromere-specific high order chromatin structure, of largely unknown molecular characteristics, appears to be essential for centromere function (Zinkowski *et al.*, 1991; Sunkel and Coelho, 1995; Sun *et al.*, 1997). Centromeric satellites are likely to play a crucial role in the formation of this high order chromatin structure. Interestingly, although showing no significant homology at the level of their nucleotide sequences, many centromeric DNAs share common structural properties. For instance, several AT-rich satellites that are frequently found at the centromere of higher eukaryotes (for reviews see Rattner, 1991; Kalitsis and Choo, 1997), were shown to be intrinsically curved reflecting common bendability properties (Martínez-Balbás *et al.*, 1990). Repeated DNAs showing a marked pu/py strand asymmetry ('telomere-like' satellites) have also been found at the centromere of many vertebrates, insects and plants (Fry and Salser, 1977; Taparowsky and Gerbi, 1982; Novak, 1984; Meyne *et al.*, 1990; Richards *et al.*, 1991; Abad *et al.*, 1992; Grady *et al.*, 1992; Alfenito and Birchler, 1993; Carmena *et al.*, 1993). Several of these telomere-like centromeric satellites were shown to form altered DNA structures in which the G-rich strand (G-strand) forms very stable intramolecular hairpins (Grady *et al.*, 1992; Catasti *et al.*, 1994; Chou *et al.*, 1994; Ferrer *et al.*, 1995; Ortiz-Lombardía *et al.*, 1998).

The *Drosophila* dodeca-satellite (GTACGGGACCGA) is a member of the telomere-like centromeric satellites family. The dodeca-satellite is found at the centromere of chromosome 3 in *Drosophila melanogaster* and of several other chromosomes of different *Drosophila* species (Abad *et al.*, 1992; Carmena *et al.*, 1993). It was shown earlier that the dodeca-satellite G-strand forms very stable fold-back structures, stabilized by the formation of tandem G·A pairs, while the complementary C-strand remains

basically unstructured (Ferrer *et al.*, 1995; Ortiz-Lombardía *et al.*, 1998). In this paper, we describe the purification and characterization of DDP1, a single-stranded DNA-binding protein of high molecular mass that binds the unstructured dodeca-satellite C-strand with high affinity and specificity. DDP1 associates with the pericentric heterochromatin and is a functional homologue of the yeast Scp160p, a multi-KH domain protein that is involved in the control of cell ploidy (Wintersberger *et al.*, 1995; Weber *et al.*, 1997).

Results

Purification and characterization of DDP1, a single-stranded DNA binding protein that interacts specifically with the dodeca-satellite C-strand

As described earlier, a strong dodeca-satellite C-strand-binding activity is detected in crude nuclear extracts obtained from cultured Schneider SL2 *Drosophila* cells (Ferrer *et al.*, 1995). The polypeptide responsible for this binding activity was purified through three chromatographic steps (Figure 1A) (see Materials and methods). Purified fractions, showing a strong binding to the dodeca-satellite C-strand (Figure 1D), are highly enriched in a polypeptide ~160 kDa in molecular mass (Figure 1B), called DDP1 (*D*rosophila ddodeca-satellite-binding protein 1). DDP1 is responsible for the binding to the dodeca-satellite C-strand as was demonstrated when the cDNA encoding DDP1 was cloned (see below). Recombinant DDP1, expressed in *Escherichia coli*, is recognized by specific αDDP1 antiserum (Figure 1C) and shows the same binding properties as the purified protein obtained from cultured cells (Figure 1D, lanes 1–4). In addition, the protein complex obtained with the purified fractions is specifically recognized by αDDP1 antiserum (not shown). The electrophoretic mobility of the DDP1–C-strand complex corresponds to that of the major complex detected in crude nuclear extracts (Figure 1D, lane NE), indicating that DDP1 is the major C-strand-binding activity present in SL2 cells.

DDP1 does not show any significant affinity for double-stranded DNA. No significant retardation of a 145 bp double-stranded dodeca-satellite DNA fragment was observed at DDP1 concentrations showing strong binding to the C-strand (Figure 2A, compare duplex and C-strand). Similar results were obtained with other cloned dodeca-satellite DNA fragments of slightly different nucleotide sequences (not shown). The very low affinity of DDP1 for double-stranded DNA was corroborated by competition experiments (Figure 2B). A significant binding of DDP1 to the C-strand is detected even in the presence of a large 3000-fold excess of the unspecific double-stranded DNA competitors, ds*E.coli*DNA or poly(dI·dC). On the contrary, the dodeca-satellite C-strand itself is a very efficient competitor. In this case, DDP1 binding is completely abolished in the presence of a 50-fold excess of competitor (Figure 2C, C-strand, lane 4). These large differences in competition efficiency are better reflected when the excess of competitor required to obtain 50% competition are compared, ~7-fold for the C-strand but >3000-fold for the double-stranded DNA competitors.

The interaction of DDP1 with the dodeca-satellite C-strand is specific. As shown in Figure 2A, DDP1 shows a very low affinity for the dodeca-satellite G-strand. However, as described previously, the single-stranded dodeca-satellite G-strand has a high tendency to form stable fold-back structures (Ferrer *et al.*, 1995; Ortiz-Lombardía *et al.*, 1998) that are likely to strongly influence DDP1 binding. The specificity of the interaction of DDP1 with the dodeca-satellite C-strand was also analysed through competition experiments using ssM13DNA and heat-denatured ss*E.coli*DNA as unspecific competitors (Figure 2C). ss*E.coli*DNA competes DDP1 binding less efficiently than the C-strand itself, a 500-fold excess of ss*E.coli*DNA being required to obtain 50% competition. Competition by ssM13DNA is also rather inefficient, a 170-fold excess of competitor is required to obtain 50%

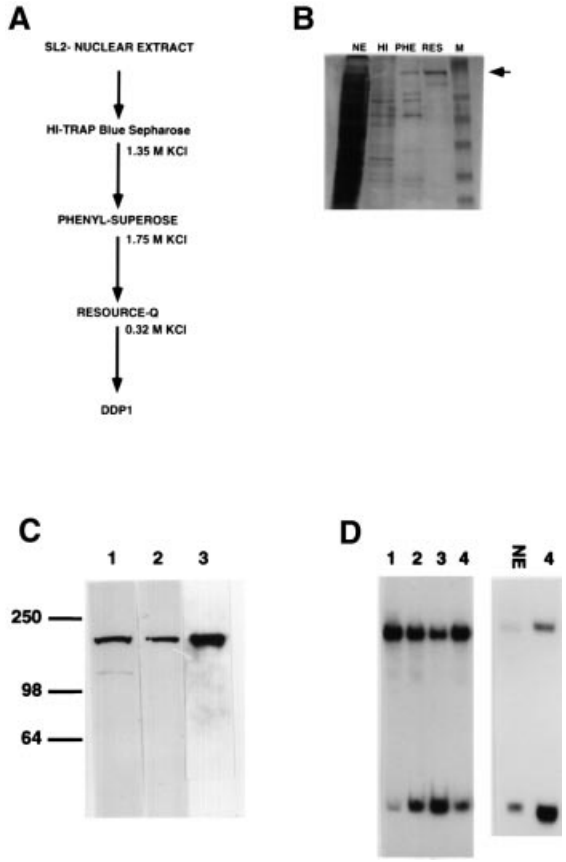


Fig. 1. Purification of DDP1. The fractionation scheme used for the purification of DDP1 is shown in (**A**) (see Materials and methods). (**B**) The protein composition of the pooled positive fractions is shown after each chromatographic step: NE, crude nuclear extract; HI, after the HI-TRAP Blue Sepharose column; PHE, after the hydrophobic Phenyl-Superose column and RES, after the anion exchange RESOURCE-Q column. Lane M corresponds to markers of known molecular weight 250, 98, 64, 50, 36 and 30 kDa from top to bottom. The arrow indicates the position of DDP1. (**C**) Western blot analysis using αDDP1 antibodies, obtained as described in Materials and methods, of: crude nuclear extract prepared from cultured SL2 cells (lane 1); total cell extract prepared from *Drosophila* embryos (0–2 h) (lane 2) and purified recombinant DDP1 expressed in *E.coli* (lane 3). The positions of markers of known M (in kDa) is indicated. (**D**) The binding to a 145-base dodeca-satellite C-strand of native DDP1 obtained from cultured cells, (lane 4) is compared to that of the recombinant protein expressed in *E.coli* (lanes 1–3). Lane NE corresponds to the binding obtained in the presence of the crude nuclear extract. The binding reactions were performed in the presence of increasing excess (w/w) of heat-denatured ss*E.coli*DNA: 0 (lanes 1 and 4); 50 (lane 2); 500 (lane 3) and 1000 (lane NE).

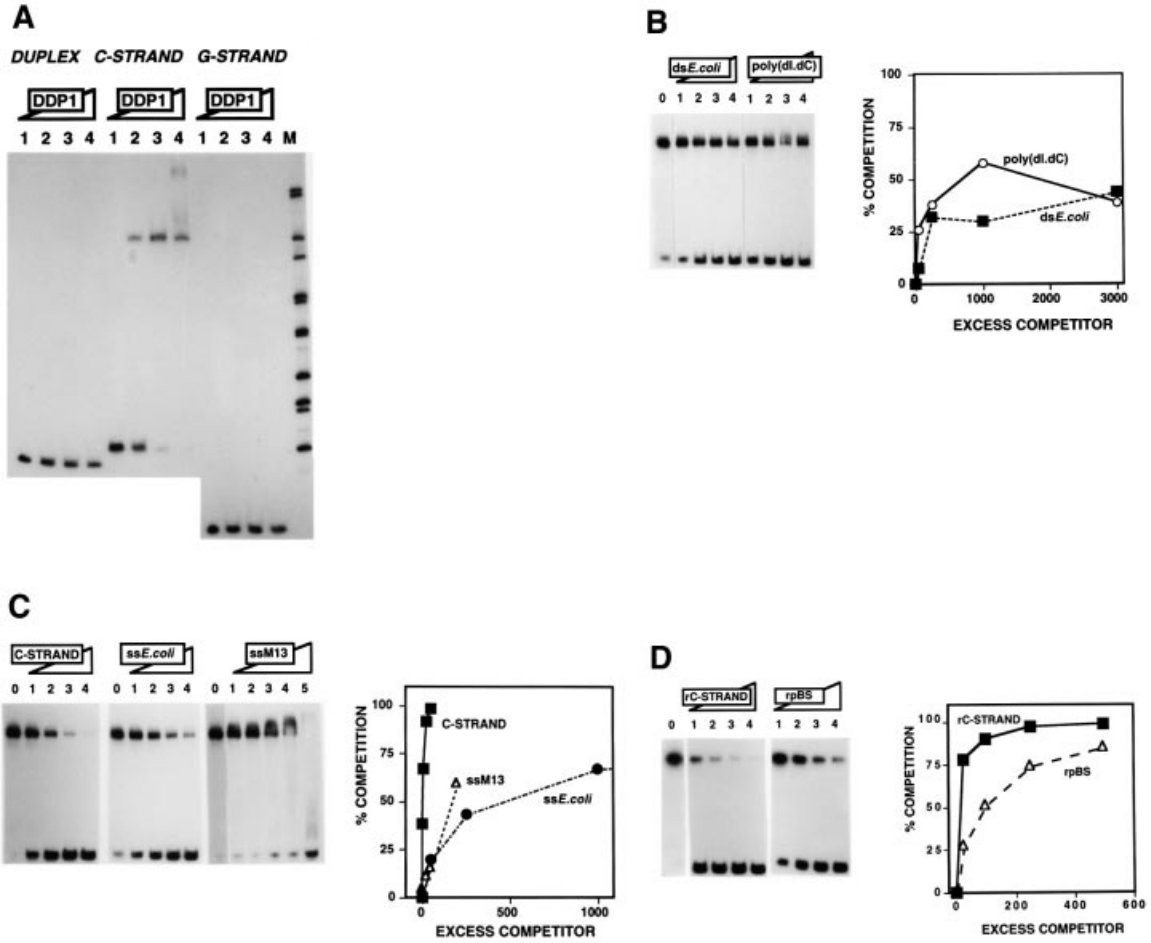


Fig. 2. The interaction of DDP1 with the dodeca-satellite. (**A**) The binding of purified DDP1, obtained from SL2 cells, to a 145 bp double-stranded dodeca-satellite sequence (DUPLEX) and to the corresponding C-strand (C-strand) and G-strand (G-strand) is analysed by EMSA as a function of increasing amounts of added protein: 0 µl (lanes 1), 0.08 µl (lanes 2), 0.3 µl (lanes 3) and 1 µl (lanes 4). Lane M corresponds to markers of known molecular mass: 2.17, 1.76, 1.23, 1.03, 0.65, 0.51, 0.45, 0.39, 0.29, 0.23/0.22 and 0.15 kb from top to bottom. (**B**) The efficiency of competition of double-stranded *E.coli* DNA (ds*E.coli*) and poly(dI·dC) is shown as a function of increasing excess (w/w) of competitor: 0 (lanes 0), 50 (lanes 1), 250 (lanes 2), 1000 (lanes 3) and 3000 (lanes 4). (**C**) The efficiency of competition of the dodeca-satellite C-strand (C-STRAND), heat denatured *E.coli* DNA (ss*E.coli*) and M13ssDNA (ssM13) is shown as a function of increasing excess (w/w) of competitor. C-strand: 0 (lane 0), 5 (lane 1), 10 (lane 2), 25 (lane 3) and 50 (lane 4). ss*E.coli*: 0 (lane 0), 50 (lane 1), 250 (lane 2), 1000 (lane 3) and 3000 (lane 4). ssM13: 0 (lane 0), 5 (lane 1), 10 (lane 2), 25 (lane 3), 50 (lane 4) and 200 (lane 5). (**D**) The efficiency of competition of RNA fragments of random sequence (rpBS) or of the dodeca-satellite C-strand sequence (rC-strand) is shown as a function of increasing excess (w/w) of competitor: 0 (lane 0), 25 (lanes 1), 100 (lanes 2), 250 (lanes 3) and 500 (lanes 4). Quantitative analysis of the results are shown on the right of each panel where the percentage of competition is shown as a function of the excess of competitor added.

competition. However, in this case the profile of the competition is different. At intermediate competitor concentrations, from 5- to 50-fold excess, a broad band of lower electrophoretic migration than that corresponding to the DDP1–C-strand complex is detected (Figure 2C, ssM13, lanes 1–4). This super-shift is not detected in the absence of DDP1 (not shown), indicating that it does not arise from the annealing of the C-strand and the ssM13DNA. At these concentrations of ssM13DNA competitor, no significant increase in the amount of unbound C-strand is observed indicating that the super-shift corresponds to the formation of a ternary complex of DDP1 and both the C-strand probe and the ssM13DNA competitor. These results suggest that DDP1 contains multiple ssDNA-binding domains capable of binding simultaneously the C-strand and the ssM13DNA. This hypothesis was confirmed when the cDNA coding for DDP1 was cloned (see below). At higher competitor concentrations, 200-fold excess, the percentage of free probe is still low but under these conditions, the band corresponding to the DDP1–DNA complex is very faint and a broad smear of higher electrophoretic mobility is detected instead (Figure 2C, ssM13, lane 5) indicating that only at this high concentration can the unspecific competitor efficiently displace the C-strand from the complex.

The affinity of DDP1 for RNA was also analysed through competition experiments using RNA fragments of the dodeca-satellite C-strand sequence (rC-strand) and of random sequence (rpBS) (see Materials and methods) (Figure 2D). The affinity of DDP1 for the RNA version of the dodeca-satellite is not significantly different than for its DNA form. A similar excess of competitor, 7-fold for the C-strand and ~10-fold for its RNA version (rC-strand), is required to obtain 50% competition. On the other hand, RNA fragments of unrelated sequence compete less efficiently (Figure 2D, rpBS). In this case, a higher 100-fold excess of competitor is required to reach 50% competition and significant binding is detected even

in the presence of a 500-fold excess. Similar results were obtained when total RNA prepared from *Drosophila* embryos was used as unspecific competitor (not shown). These results indicate that the general affinity of DDP1 for RNA is low, similar to its affinity for unspecific ssDNAs such as ssM13DNA

DDP1 associates with the pericentric heterochromatin and co-localizes with HP1

DDP1 is found associated with polytene chromosomes. When specific αDDP1 antibodies were used to immunolocalize the protein in polytene chromosomes an intense signal was detected at the chromocentre (Figure 3A). The chromocentre is a special chromosomal structure mainly consisting of repeated DNA. In addition to middle repetitive DNA sequences homologous to different transposable elements, the chromocentre also contains all the highly repetitive centromeric satellites, including the dodeca-satellite, that are, however, underreplicated. As judged by *in situ* hybridization, the dodeca-satellite is organized into three to four large blocks that are located in the pericentric heterochromatin of chromosome 3 (Figure 3B). DDP1 is present at these regions enriched in dodeca-satellite sequences. However, as seen when individual detached chromosomes are analysed, the αDDP1 signal at the chromocentre spans most of the heterochromatin of chromosome 3, extending beyond the major blocks of dodeca-satellite, as well as of chromosome 2, which does not contain detectable amounts of dodeca-satellite sequences (Figure 3B). The possibility that these other chromosomal regions could also contain low amounts of dodeca-satellite sequences, undetectable by *in situ* hybridization, cannot be ruled out totally. Nevertheless, these results strongly suggest that the distribution of DDP1 at the chromocentre is not constrained to the dodeca-satellite sequences and that DDP1 might also be associated with other heterochromatic DNAs. In addition to the strong αDDP1 reactivity of the pericentric heterochromatin, several αDDP1 signals are also detected on the chromosome arms (Figure 3A). Intriguingly, both at the chromocentre and the euchromatic chromosome arms, DDP1 shows a striking co-localization with HP1 (Figure 3C), a protein that associates in general with heterochromatin (James and Elgin, 1986; James *et al.*, 1989; Miklos and Costell, 1990; Belyaeva *et al*., 1993; Platero *et al.*, 1998). Altogether, these results indicate a general association of DDP1 with heterochromatin.

Metaphase chromosomes from larval neuroblasts show no significant αDDP1 reactivity (not shown). However, interphase nuclei react strongly with αDDP1 antibody (Figure 4). αDDP1 staining predominantly shows a uniform granular pattern (Figure 4A) which, in some cases, fuses to a single strong αDDP1 spot located at the heterochromatin region close to the nuclear membrane, lying beside the strongest DAPI signal corresponding to the heterochromatic AT-rich satellites (Figure 4B). As shown in Figure 4C, the overlap between the αDDP1 and dodeca-satellite signals is only partial. In agreement with the pattern observed in polytene chromosomes, some DDP1 locations contain no detectable dodeca-satellite sequences. Similarly, some of the dodeca-satellite rich regions appear to bind DDP1 poorly.

Consistent with its association with the chromocentre, during embryo development, DDP1 becomes nuclear at around division 13 (not shown) when, after cellularization, heterochromatin becomes more conspicuous and the chromocentre is organized (Mahowald and Hardy, 1985; Edgar *et al*., 1986; Kellum *et al*., 1995). Later, at cellular blastula and gastrula stages, DDP1 is found associated with the chromosomes (Figure 5). In synchronically dividing blastoderms, nuclei at interphase/prometaphase show a rather uniform nuclear αDDP1 reactivity (Figure 5, top), similar to that observed at interphase in larval neuroblasts. But, contrary to what is observed in larval neuroblasts, the condensed mitotic chromosomes show a strong αDDP1 reactivity at anaphase/telophase (Figure 5, centre), the αDDP1 and DAPI signals being highly superimposable. Similar results were obtained when DDP1 localization was determined at the gastrula stage (Figure 5, bottom), when mitotic domains are established (Foe, 1989). Also at this stage, interphase nuclei show a rather uniform αDDP1 reactivity and the condensed mitotic chromosomes are reactive.

The chromocentric association of DDP1, a ssDNA-binding protein, suggests that the specialized heterochromatin structure of the chromocentre contains regions of ssDNA. Consistent with this hypothesis, when fixed polytene chromosomes are treated with DNA polymerase I (Pol I) a significant nucleotide incorporation is observed at the chromocentre, which becomes labelled, while the rest of the chromosome does not (Figure 3D).

DDP1 is a multi-KH domain protein homologous to the yeast Scp160p, which is involved in the control of ploidy

No N-terminal amino acid sequence could be obtained from the direct analysis of purified DDP1. Two internal peptides were prepared by proteolytic digestion of the purified protein with either V8 protease or thrombin. The N-terminal amino acid sequences of these two peptides are shown in Table I. Oligonucleotides derived from these sequences were used to amplify a specific DNA fragment from a *Drosophila* cDNA plasmid library that was used to screen the same cDNA plasmid library. Several positive clones were obtained that, when analysed by restriction mapping and sequencing, were shown to all contain an open reading frame (ORF) coding for a polypeptide of 1301 aa (144 kDa) whose predicted sequence contains the two internal DDP1 peptides presented in Table I (Figure 6). When this cDNA was expressed in *E.coli*, the recombinant protein was recognized by specific αDDP1 antibodies (Figure 1C). Furthermore, the recombinant protein binds the dodeca-satellite C-strand with equal affinity and specificity as the native protein purified from SL2 cells, giving rise to a protein–DNA complex of the same electrophoretic mobility (Figure 1D). These results indicate that the cDNA shown in Figure 6 codes for DDP1.

A sequence homology search showed that DDP1 is homologous to the multi-KH domain proteins Scp160p from yeast (26% identity) and vigilin, either chicken or human (46 and 47% identity, respectively) (Figure 7A) (McKnight *et al*., 1992; Schimdt *et al.*, 1992; Plenz *et al.*, 1994; Wintersberger *et al.*, 1995). The KH-domain is a 65–80 aa single-stranded nucleic acid-binding motif that was originally identified in the human RNA-binding protein hnRNP K (Siomi *et al.*, 1993), and since then has been found in a number of other

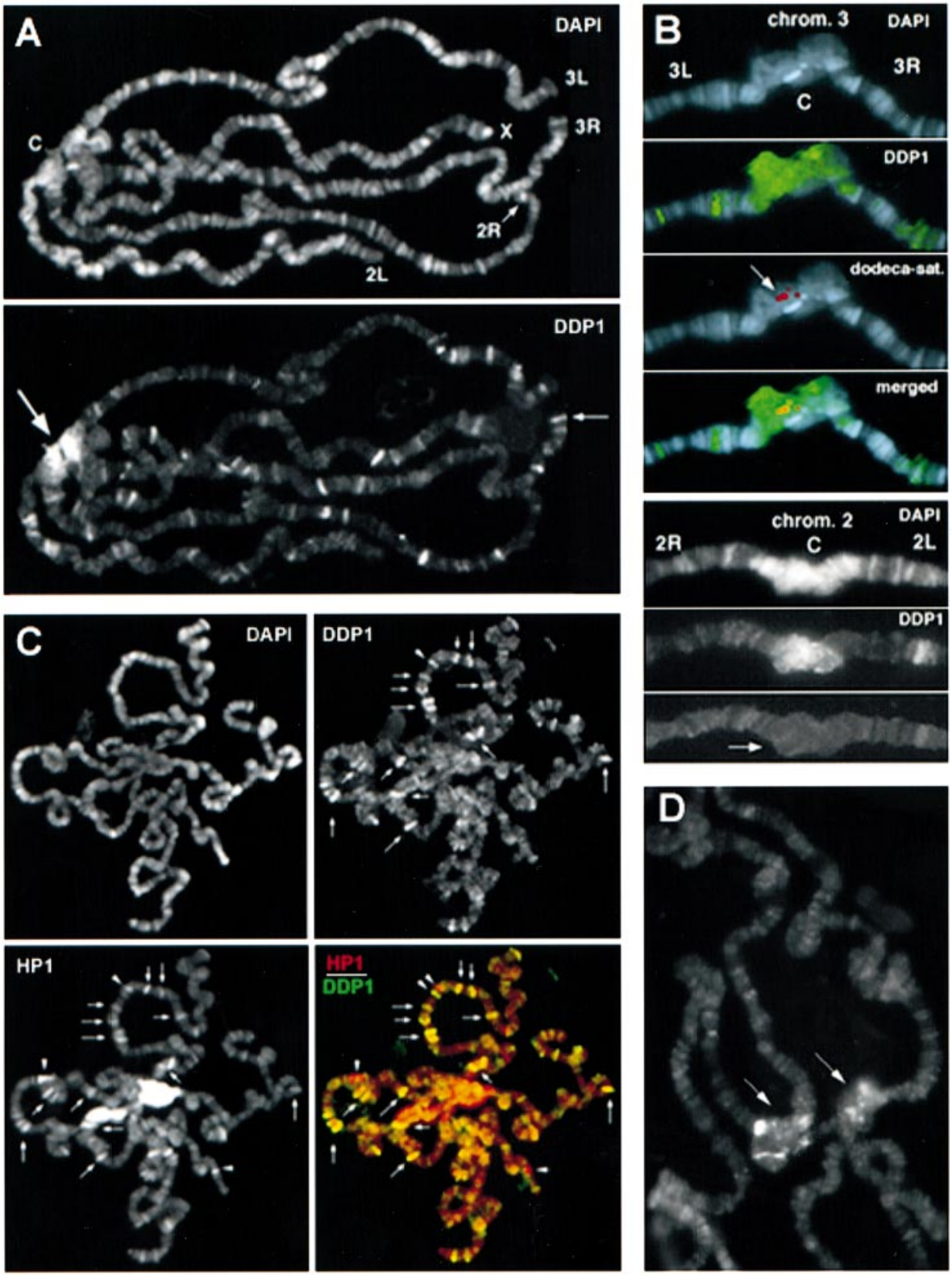


Fig. 3. DDP1 localization in polytene chromosomes. (**A**) The immunostaining with αDDP1 antiserum is shown. The protein accumulates at the chromocentre (large arrow) and is also present at multiple euchromatic locations and at the telomere of the chromosome 3 right arm (small arrow). (**B**) Sequential immunostaining with the αDDP1 antibody and *in situ* hybridization with a probe corresponding to a fragment of the dodeca-satellite DNA. (**C**) Simultaneous immunostaining with αDDP1 and αHP1 antibodies. In the merged figure, the DDP1 and HP1 immunopatterns were pseudocoloured in green and red, respectively. (**D**) *In situ* labelling by DNA Pol I. The arrows indicate the position of the chromocentre.

proteins sharing the capability to bind single-stranded nucleic acids, RNA and/or ssDNA. The KH-domain has a specific fold in which a stable three-stranded antiparallel β-sheet is packed against three α-helices on one face (Musco *et al.*, 1996). Most KH-domain proteins contain a few KH-motifs spaced by regions of variable length

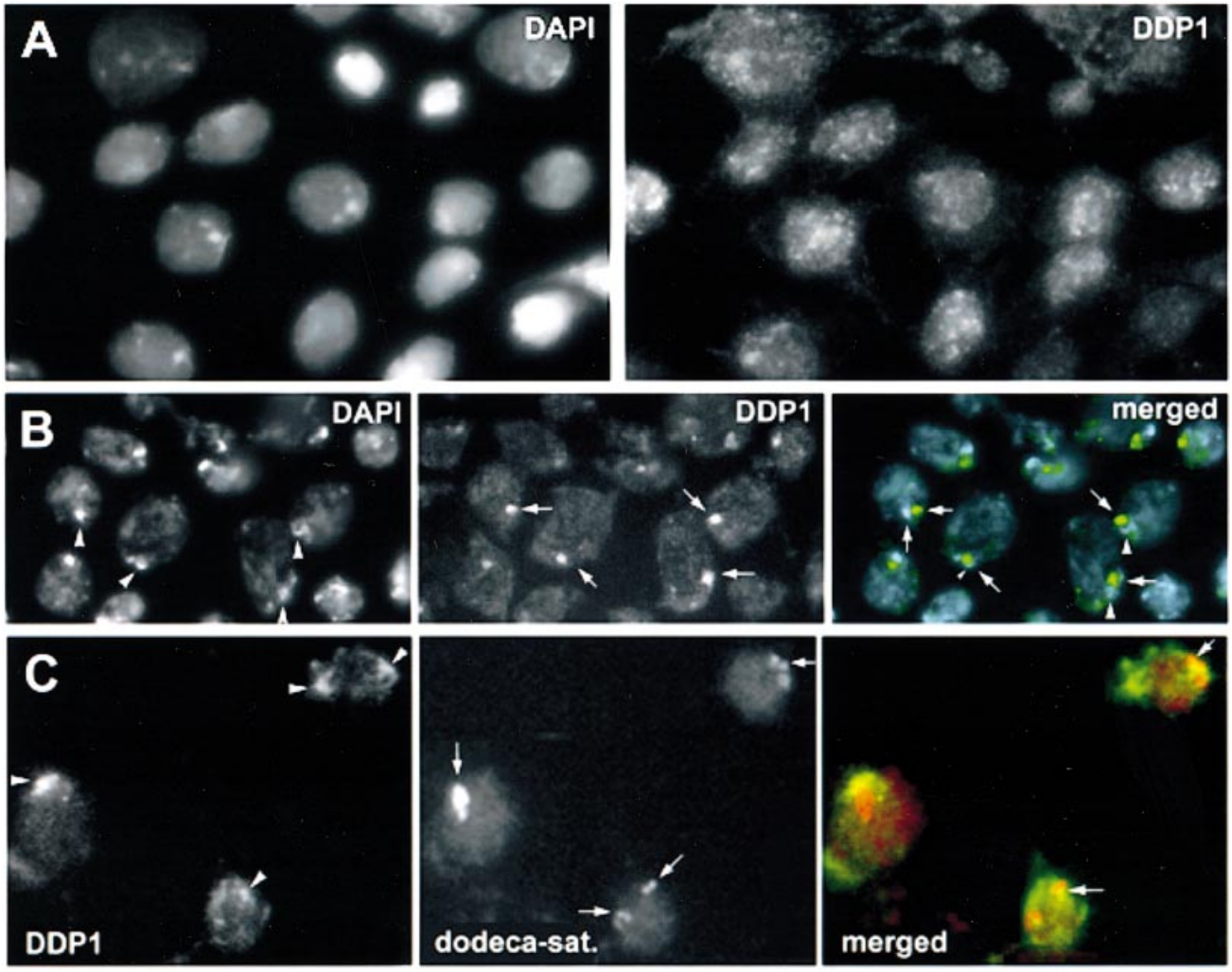


Fig. 4. DDP1 localization in nuclei from larval neuroblasts. In general, nuclei show a granular αDDP1 pattern (**A**), but in some nuclei (**B**), a single strong αDDP1 spot is observed (arrows), which localizes close to the DAPI bright chromocentre spots (arrowheads) as seen in the merged figure. (**C**) Sequential αDDP1 immunostaining (arrowheads) and *in situ* hybridization with a dodeca-satellite DNA probe (arrows). In the merged figure, the DDP1 and dodeca-satellite signals were pseudocoloured in green and red, respectively.

Table I. N-terminal amino acid sequence of two internal peptides obtained by proteolytic digestion of purified DDP1

Protease	M (kDa)	Sequence
V8	125	(A/S)AEQYKKISDRISVP(K/S)(K/I)YS
Thrombin	32	KANVRQFMSKHDVHVELPP

and sequence. What is peculiar about DDP1, Scp160p and vigilins is, on one hand, the high number of KH-domains they contain—fifteen—and, on the other hand, their tandem organization (Figure 7). Actually, these multi-KH domain proteins are formed almost exclusively by the tandem repetition of the KH-motif, preceded and followed by relatively short N- and C-terminus regions. Most of the homology is therefore found in the multi-KH region. The different KH-domains of DDP1 are themselves homologous, showing the predicted KH-fold and containing basically all the residues which are found conserved in the rest of KH-domains analysed to date (Figure 7B), and in particular, the two glycine residues flanking the α1 and α2 helices, which appear to be involved in nucleic acid binding. The twelfth KH-domain of DDP1 is longer than the rest, such as in vigilins but not in Scp160p (Figure 7). The structural organization of DDP1 confirms its strong affinity for single-stranded nucleic acids.

In yeast, Scp160p is involved in the control of ploidy (Wintersberger *et al.*, 1995). Disruption of the *SCP160* gene results in cells of decreased viability and increased number of chromosomes. Expression of Scp160p in *Δscp160* cells obtained from diploid heterozygous *SCP160*/*Δscp160* strains, but not in haploid *Δscp160* strains, rescues the mutant phenotype (Wintersberger *et al.*, 1995). Figure 8 shows a FACS analysis of the DNA content of spores from the heterozygous strain AK303 in which one copy of the *SCP160* gene is deleted by insertion of *LEU2*. Leu^- spores, which are wild type for *SCP160*, showed a normal FACS profile corresponding to a haploid DNA content (Figure 8A), identical to that of spores obtained from the wild-type strain AK300 (not shown). However, Leu^+ spores carrying the *Δscp160* deletion have a higher DNA content, mainly double, than wild-type cells (Figure 8B). In *Δscp160* cells, the peaks in the FACS analysis

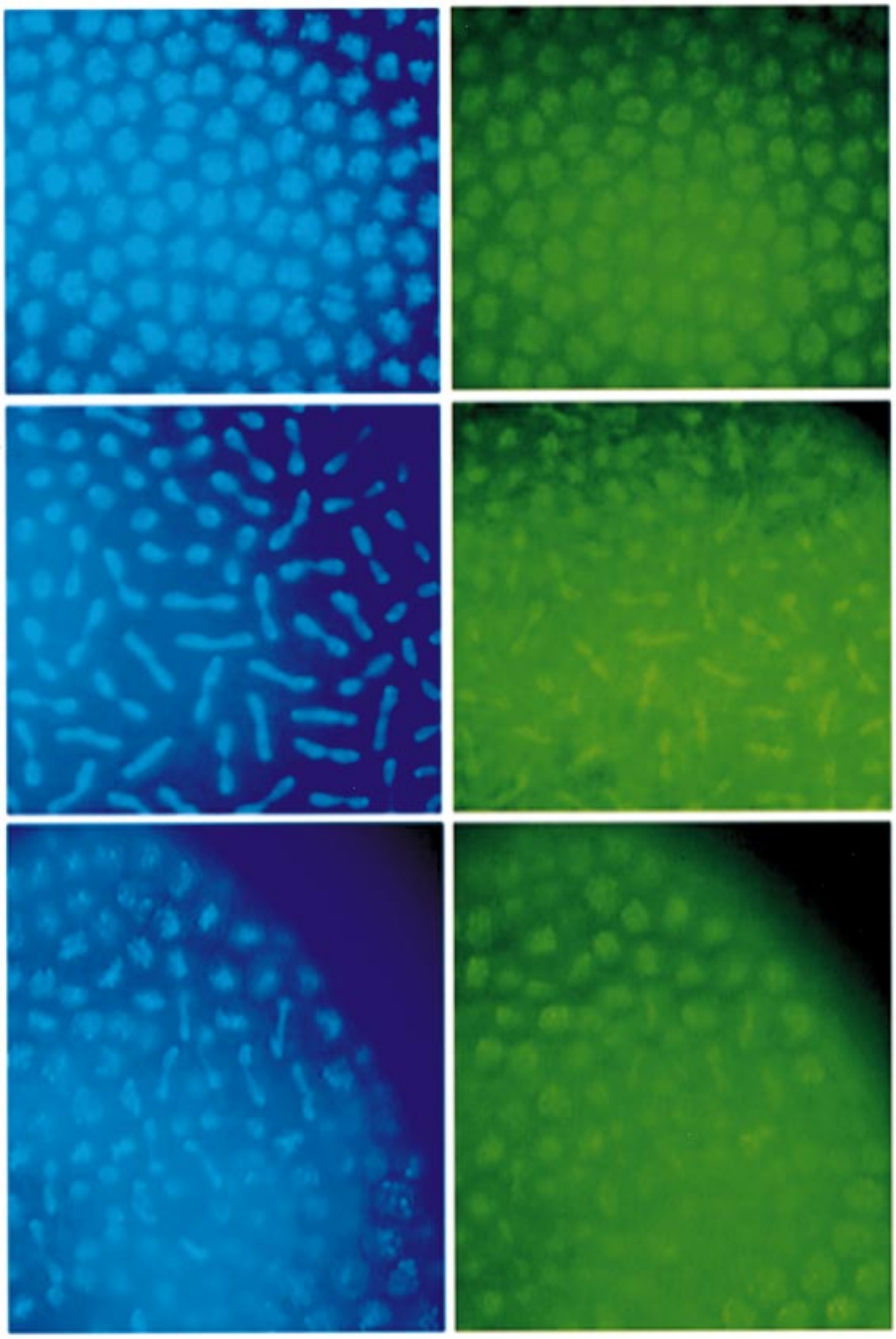


Fig. 5. DDP1 localization during embryo development. DAPI (left) and αDDP1 (right) fluorescence is shown for synchronically dividing cellular blastoderm nuclei at interphase/prometaphase (top) and anaphase/telophase (centre) stages of the mitotic cycle. Nuclei of a mitotic domain at the gastrula stage are shown at the bottom.

are broader than in wild-type cells indicating that the actual DNA content per cell is more variable in *Δscp160* cells. Expression of DDP1 rescues this mutant phenotype. The abnormal FACS profile of *Δscp160* cells was not observed in spores containing the pYES-DDP1 plasmid when germinated in galactose to allow expression of DDP1. Under these conditions, Leu^+ spores have a normal DNA content, showing a FACS profile very similar to that corresponding to wild-type cells (Figure 8C). These results indicate that DDP1 is functionally homologous to Scp160p.

Discussion

Here we have shown that DDP1, a novel *Drosophila* single-stranded nucleic acid-binding protein, is found associated with pericentric heterochromatin both in polytene chromosomes and interphase nuclei from larval neuroblasts. Like other heterochromatin associated proteins, such as HP1 and SU(VAR)3-7 (Kellum *et al.*, 1995; Cléard *et al.*, 1997), DDP1 becomes nuclear only at the cellular blastula stage when heterochromatin is fully organized (Mahowald and Hardy, 1985; Edgar

et al., 1986; Kellum *et al*., 1995). The presence of DDP1 in the condensed metaphase chromosomes could not be unambiguously demonstrated. On one hand, no significant αDDP1 reactivity was detected in metaphase chromosomes from larval neuroblasts. But on the other hand, at cellular blastula and gastrula stages condensed mitotic chromosomes are strongly reactive with αDDP1 antibodies, suggesting that the lack of reactivity observed in mitotic chromosomes from larval neuroblasts could arise from a lack of accessibility to the antibody. In polytene chromosomes, DDP1 is also found at several sites in the euchromatic arms mostly co-localizing with HP1, a protein that, like many of the known heterochromatin associated proteins, does not show any significant nucleic acid-binding activity. These results suggest that DDP1 may play a general role in organizing chromosomal domains. DDP1 might be involved in the formation of heterochromatin by recognizing specific features on the DNA (i.e. the formation of ssDNA), and perhaps facilitating the recruitment of additional protein factors.

The structural organization of DDP1, with 15 consecutive KH-domains, strongly indicates that protein–nucleic acid interactions are the major determinants of the chromosomal association of DDP1. Several indications suggest that binding of DDP1 to chromosomes probably

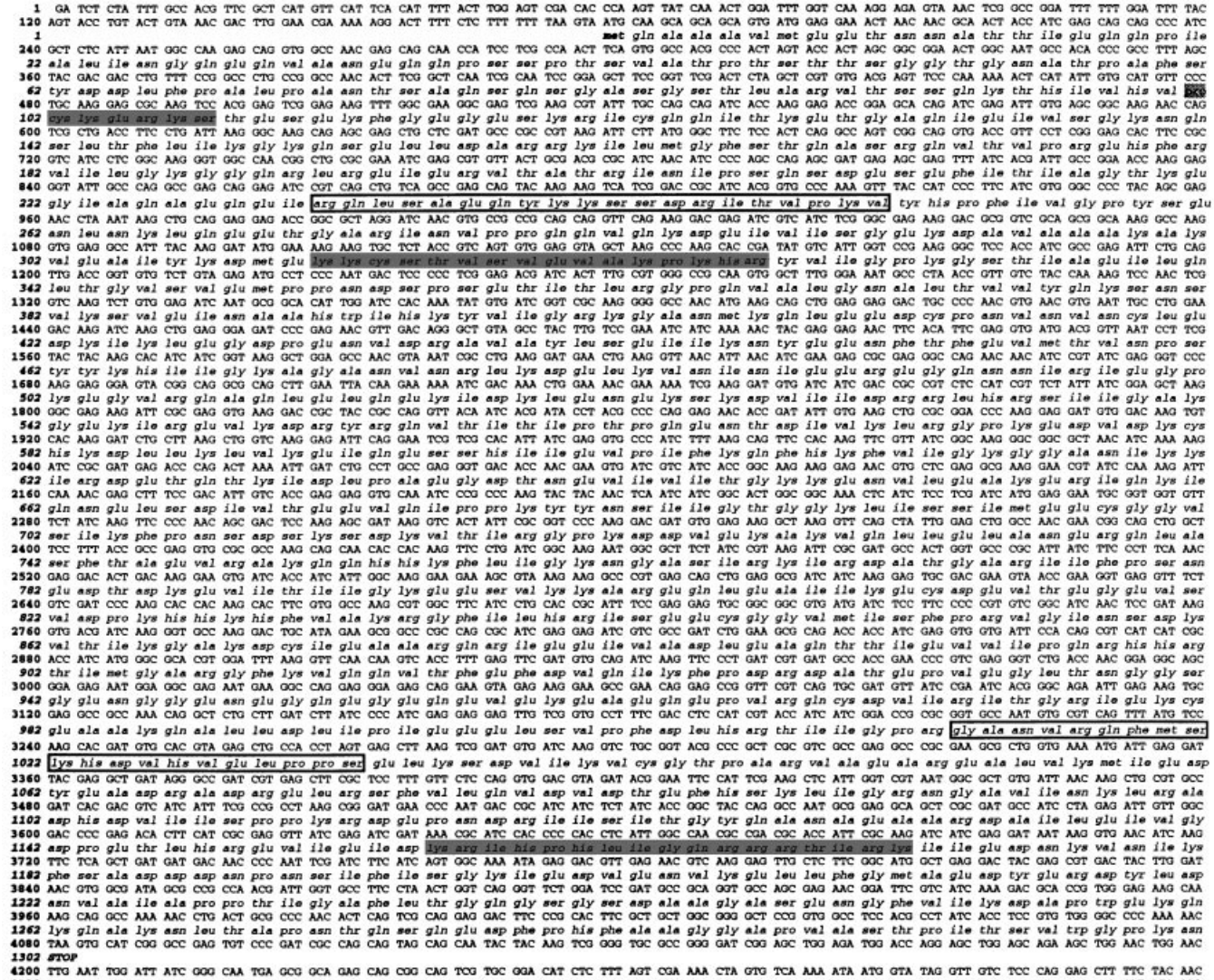
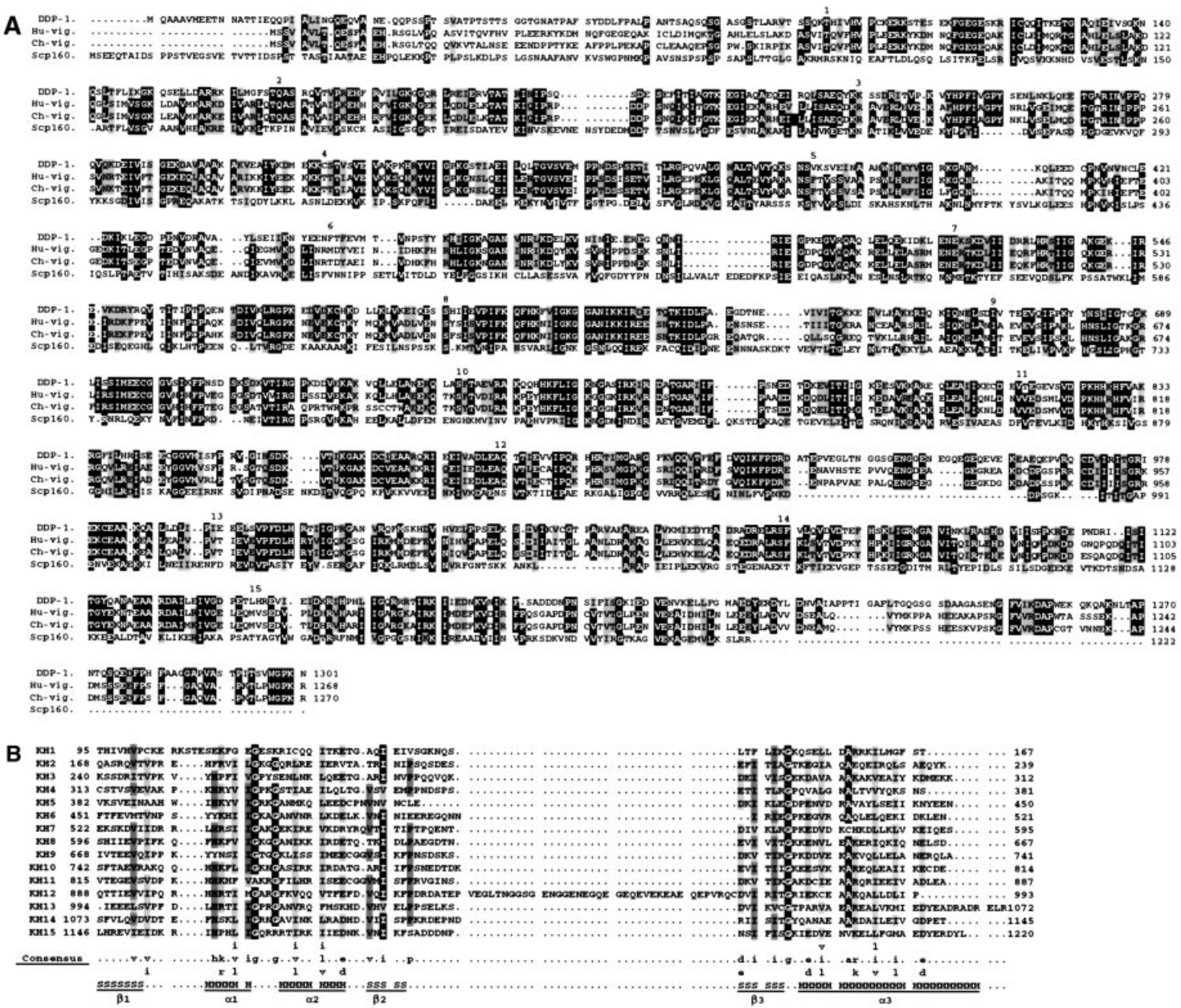


Fig. 6. Nucleotide and predicted amino acid sequence of the cDNA encoding DDP1. The deduced aa sequence begins at the first ATG codon of the ORF. Boxes indicate the positions corresponding to the two internal peptides whose sequence was determined experimentally from native purified DDP1 obtained from cultured cells (Table I). Regions in grey correspond to putative nuclear localization signals (DDBJ/EMBL/GenBank accession No. DS38396/97961).


A
DDP-1.
Hu-vig.
Ch-vig.
Scp160.
B
KH1
KH2
KH3
KH4
KH5
KH6
KH7
KH8
KH9
KH10
KH11
KH12
KH13
KH14
KH15
Consensus
β1
α1
α2
β2
β3
α3

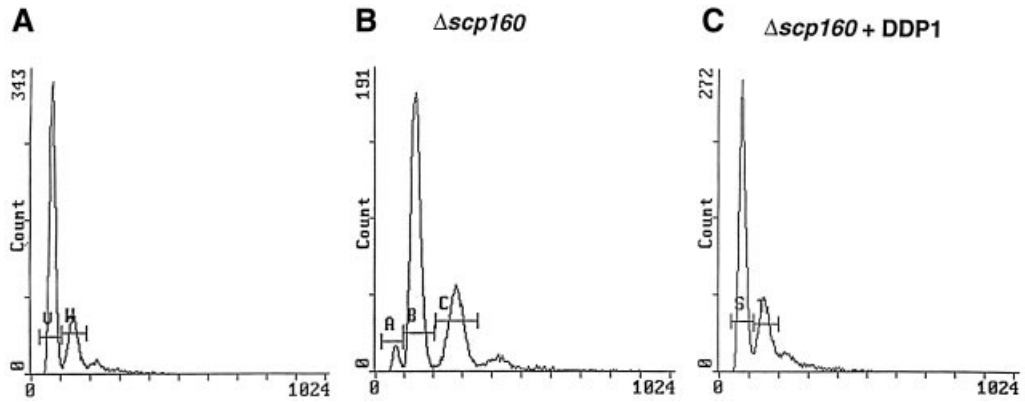


Fig. 8. Expression of DDP1 rescues a *Δscp160* deletion in yeast. FACS analysis are presented for a Leu$^+$ spore, carrying a *Δscp160* deletion, obtained from the heterozygous *SCP160/Δscp160* strain AK303 that was transformed, as indicated under Materials and methods, with a plasmid expressing DDP1 (**C**) or not (**B**). (**A**) The FACS analysis corresponding to a wild-type Leu$^-$ *SCP160* spore from AK303 strain.

occurs via ssDNA recognition. On one hand, DDP1 does not localize to regions enriched in RNA; no αDDP1 signal is observed at the nucleolus or associated to the rRNA gene cluster or other intensively transcribed regions. On the contrary, a strong αDDP1 reactivity, which is resistant to RNase A treatment (not shown), is found at the chromocentre, a region particularly poor in actively transcribed genes. In addition, the strongest DDP1–RNA interaction we have detected was with RNAs of the dodeca-satellite C-strand sequence that are not likely ever to be synthesized *in vivo*, since the dodeca-satellite is not actively transcribed. However, we cannot totally exclude the possibility that DDP1 would interact with specific heterochromatin associated RNAs. The association of RNA species with specific chromatin structures has been demonstrated in some cases, such as in the inactive X chromosome of vertebrates (reviewed in Brockdorff, 1998).

The chromosomal distribution of DDP1 is not constrained to the regions rich in dodeca-satellite sequences suggesting that DDP1 can also recognize other DNA sequences. The affinity of DDP1 for the dodeca-satellite C-strand depends on its nucleotide sequence but is largely influenced by the extent of secondary structure of the DNA substrate, being significantly reduced in C-strand DNA fragments capable of folding into secondary structures (A.Cortés and F.Azorín, in preparation). Likewise, base substitutions that reduce the degree of secondary structure of the DNA fragments result in a stronger interaction (A.Cortés and F.Azorín, in preparation). Nucleic acid recognition by other multi-KH domain proteins was found to be little sequence-specific, showing a similar strong dependence on the extent of secondary structure of the substrate (Kanamori *et al.*, 1998). Moreover, homologous KH-domains of slight different amino acid sequences show remarkable different nucleic acid-binding preferences *in vitro* (Dejgaard and Leffers, 1996). Altogether, these observations suggest that binding of DDP1 to the dodeca-satellite C-strand is largely conformational dependent and support the hypothesis that DDP1 could also bind ssDNA fragments other that the dodeca-satellite C-strand. Actually, though with lower affinity, DDP1 can also bind the pyrimidine strand of the AAGAG satellite (A.Cortés and F.Azorín, in preparation) that, being present at the pericentric heterochromatin of chromosome 2 (Lohe *et al.*, 1993; Platero *et al.*, 1998), is also capable of forming altered DNA structures in which the pyrimidine strand is single-stranded (Ortiz-Lombardía *et al.*, 1998). The *Drosophila* pericentric heterochromatin is also rich in different types of transposable elements (Pimpinelli *et al.*, 1995; Sun *et al.*, 1997) that could also contribute to DDP1 binding. Actually, DDP1 is abundant at the β-heterochromatin that is highly enriched on these repetitive elements. Interestingly, as for the dodeca-satellite, the LTRs of some retrotransposons were found to form hairpin structures (Slama-Schwok *et al.*, 1998).

The presence of DDP1 and its *in situ* labelling suggest that, at least at some stage of the cell cycle, pericentric heterochromatin contains regions of ssDNA. The presence of ssDNA at these regions could reflect the formation of altered DNA structures such as those proposed for the *Drosophila* dodeca-satellite and several other centromeric satellites, in which the purine-strand folds into very stable intramolecular hairpins leaving the complementary pyrimidine-strand single-stranded (Grady *et al.*, 1992; Catasti *et al.*, 1994; Chou *et al.*, 1994; Ferrer *et al.*, 1995; Ortiz-Lombardía *et al.*, 1998). DDP1 could play an important role on the induction and/or stabilization of these altered DNA structures. However, it is unlikely that the whole of the dodeca-satellite sequences would exist as ssDNA. On the contrary, formation of ssDNA at these regions would likely be a local event involving only relatively short stretches. Actually, the lack of DDP1 binding at some

Fig. 7. (**A**) Sequence comparison of DDP1 with the yeast Scp160p and the chicken (Ch-vig) and human (Hu-vig) vigilins. Residues present in at least three of the sequences are shaded black. Residues present in only two sequences with at least one similar residue in another sequence or being similar in at least three sequences are shaded grey. Numbers on top indicate the first residue of each of the consecutive KH-domains. Numbers on the right indicate the corresponding aa residue. (**B**) Sequence comparison of the 15 consecutive KH-domains of DDP1. Residues which are present in at least 12 or 10 domains are shaded black or dark grey, respectively. Light grey indicate similar residues when present in at least 10 domains. The consensus KH-domain sequence and the corresponding predicted secondary structure are indicated on the bottom.

dodeca-satellite regions might reflect this circumstance. DNA replication could also drive the formation of ssDNA. In this respect, it is interesting to note that HP1 was also shown to co-localize with the origin recognition complex of *Drosophila* (DmORC2) (Pak *et al.*, 1997). Therefore, it is possible that some of the chromosomal locations of DDP1 would reflect the presence of regions of ssDNA occurring at origins of replication. In the case of the pericentric heterochromatin, this possibility appears unlikely since the amount of DmORC2 present at the chromocentre is low and constrained to the α-heterochromatin (Pak *et al.*, 1997). On the contrary, DDP1 is also abundant at the β-heterochromatin.

DDP1 is homologous to the yeast Scp160p, chicken and human vigilins (McKnight *et al*., 1992; Schimdt *et al.*, 1992; Plenz *et al.*, 1994; Wintersberger *et al.*, 1995). Vigilins constitute a family of highly conserved single-stranded nucleic acid-binding proteins present in all eukaryotes analysed to date. The function(s) of vigilins is only poorly understood. It was shown that the *Xenopus* vigilin binds *in vitro* the 3′ UTR of the vitellogenin mRNA and was proposed to increase its stability in response to estrogens (Dodson and Shapiro, 1997). Binding of vigilin to the human dystrophin mRNA was also demonstrated *in vitro* (Kanamori *et al.*, 1998). However, determination of the intracellular targets and functions of this family of proteins remains elusive. Vigilins are upregulated in rapidly dividing cells (Plenz *et al.*, 1994), and in yeast, disruption of the *SCP160* gene showed a phenotype related to the control of cell ploidy (Wintersberger *et al.*, 1995). The association of DDP1 with pericentric heterochromatin points towards the possibility that some of the functions of vigilins would take place at the chromosomal level contributing, in higher eukaryotes, to the structural and functional organization of the centric heterochromatin.

Recent results strongly indicate that satellite DNAs are important functional elements of the centromere (Harrington *et al.*, 1997; Sun *et al.*, 1997; Ikeno *et al.*, 1998). In particular, the centromere of the *Dp1187* mini-chromosome of *Drosophila*, the only higher eukaryotic centromere characterized at high resolution, is primarily composed of two different satellite DNAs, the AT-rich AATAT-satellite and the 'telomere-like' AAGAG satellite (Sun *et al.*, 1997). This sequence organization appears to be common to many higher eukaryotic centromeres (reviewed in Choo, 1997) and is also present in the heterochromatic centromeres of the fission yeast *Schizosaccharomyces pombe* in which a central AT-rich region is flanked by repetitive elements of which the K″/K repeats, which contain 22 copies of the telomere-like motif TGGAAA (Takahashi *et al.*, 1992), are essential for proper chromosome segregation (Clarke *et al.*, 1993). These observations suggest that centromeres might share a common structural organization in which AT-rich and telomere-like satellites cooperate in the formation of a specific heterochromatin structure to which centromere function appears to be linked (Zinkowski *et al.*, 1991; Sunkel and Coelho, 1995). *Drosophila* centromeres are rich in transposable elements (Pimpinelli *et al*., 1995; Sun *et al.*, 1997) that are also likely to contribute to the structural organization of the pericentric heterochromatin. The formation of this centromere-specific structure might be regulated epigenetically (Karpen and Allshire, 1997; Murphy and Karpen, 1998; Wiens and Sorger, 1998). It should be noted that the formation of an altered DNA structure is in itself an epigenetic phenomena since, for any DNA sequence, the regular double-stranded B-form is the preferred structural conformation. Once such an altered DNA conformation is established it could easily propagate through successive rounds of replication provided persistence of the corresponding stabilizing factors.

Materials and methods

DNAs and RNAs

Plasmid pBK6E215 has been described elsewhere (Abad *et al.*, 1992). It is a pBluescript derivative carrying a 145-bp dodeca-satellite sequence inserted at the unique *Spe*I site of pBluescript KS. The individual dodeca-satellite strands were purified by electrophoresis as described earlier (Ferrer *et al*., 1995). Poly(dI·dC) (Boehringer Mannheim), ssM13 DNA (New England Biolabs) and *E.coli* DNA (Sigma) were used as DNA competitors. *Escherichia coli* DNA was sonicated to an average length of ~200 bp and ss*E.coli*DNA was prepared by heat denaturation at 90°C for 4 min and quickly cooling on ice. DNA concentration was determined by UV spectroscopy. RNAs used as competitors were obtained by *in vitro* transcription with T7 RNA polymerase (Promega) of plasmid pBK6E215 (rC-strand fragments) or pBluescript (rpBS fragments). Before *in vitro* transcription, plasmids were linearized with appropriate restriction endonucleases to obtain transcripts ~200 (rC-strand) or ~330 bases long (rpBS). Transcripts were analysed by agarose electrophoresis and the concentration determined by UV spectroscopy.

Purification and characterization of DDP1

DDP1 was purified from Schneider SL2 crude nuclear extracts (Figure 1A). Cells, grown to a density of $4–8\times10^6$ cells/ml, were harvested by centrifugation and lysed by freeze–thawing three times in a small volume of phosphate-buffered saline (PBS) (2 ml of PBS per ml of pellet). Nuclei were then recovered and nuclear extracts were prepared according to Digman *et al.* (1983), dialysed against buffer D [0.1 M KCl, 0.2 mM EDTA pH 8, 20 mM HEPES pH 7.9, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 0.5 mM dithiothreitol (DTT), 20% glycerol] and stored at –80°C. Nuclear extracts were then fractionated by FPLC (Pharmacia) through a Pharmacia HI-TRAP Blue Sepharose pseudoaffinity column eluted with a 0.1–2 M KCl linear gradient in buffer D. Fractions were analysed by EMSA (see below). Positive fractions, eluting at around 1.35 M KCl concentration, were pooled and concentrated by centrifugation through Millipore Biomax-5K Ultrafree filters. After concentration samples were brought to 2.5 M KCl and fractionated through a Pharmacia Phenyl-Superose hydrophobic column eluted with a 2–1 M KCl linear gradient in buffer D. Positive fractions eluting at 1.75 M KCl were desalted in Pharmacia HI-TRAP desalting columns and fractionated through a RESOURCE-Q anionic exchange column eluted with a 0.1–0.8 M KCl linear gradient. From this last column, DDP1 eluted at a 0.32 M KCl concentration and was purified almost to homogeneity (Figure 1B).

For microsequencing, internal peptides were obtained by proteolytic digestion of purified DDP1 with either V8 protease or bovine thrombin (Calbiochem). Peptides were resolved by SDS–PAGE and transferred to a PVDF membrane. Most prominent peptides were cut out from the filters and subjected to microsequencing by the EDMAN method in a Beckman LF3000 automatic microsequencer.

Obtention of αDDP1 antibodies

For the obtention of polyclonal αDDP1 antibodies, higly enriched fractions (RESOURCE-Q positive fractions) were subjected to SDS–PAGE and the band corresponding to DDP1 was cut and blended. Antibodies were raised in mice. The specificity of the αDDP1 antiserum is high, recognizing a single polypeptide of the same electrophoretic mobility as purified DDP1 in crude nuclear extracts obtained from cultured SL2 cells and *Drosophila* embryos (Figure 1D), or adult flies (not shown). In some cases, a second faint band

of lower molecular weight could also be detected which is likely to arise from the proteolytic degradation of DDP1.

EMSA experiments

For the EMSA experiments, ~1 ng of radioactively labelled DNA, prepared as described above, was incubated for 20 min at 4°C with different amounts of nuclear extract or purified DDP1 in a buffer containing 250 mM NaCl, 1 mM EDTA, 10 mM Tris, 10 mM β-mercaptoethanol, 0.1% Triton X-100 (v/v), 4% glycerol, pH 7.5. When competition experiments were performed, both labelled and competitor DNA were added to the reaction mixture prior to the addition of the protein. Formation of protein–DNA complexes was visualized by gel electrophoresis in 5% polyacrylamide non-denaturing gels run in 0.5× TBE.

Immunostaining experiments

For the immunostaining of polytene chromosomes, salivary glands were dissected from third instar larvae in saline physiological solution, transferred in a drop of fixative (45% acetic acid, 2% formaldehyde) for 10 min and squashed. Larval brains from third instar larvae were also dissected in saline physiological solution and, before fixation, were incubated in hipotonic solution (0.5% sodium citrate) for 10 min. Slides were frozen in liquid nitrogen, washed once in PBS after flipping off the coverslip, immersed in PBS, 1% Triton X-100 for 10 min and then in PBS with dried nonfat milk for 30 min. Preparations were incubated overnight at 4°C with αDDP1 antiserum at 1:400 dilution in PBS, 1% bovine serum albumin (BSA), washed three times in PBS, and incubated for 30 min at room temperature with secondary antibody (FITC-anti-mouse in PBS, 1% BSA), washed three times in PBS, stained with DAPI (0.05 μg/ml) and mounted in anti-fading medium. For the sequential immunolabelling with αDDP1 antibody and *in situ* hybridization with dodeca-satellite DNA in some of the preparations the coverslip was gently removed after microscope observation and the slides were washed five times in PBS for 5 min each, dehydrated in absolute ethanol for 3 min and air-dried for 2 days at room temperature or overnight at 37°C. The preparations were then processed for fluorescence *in situ* hybridization (FISH) as described in Gatti *et al.* (1994). For the double immunostaining with αDDP1 and αHP1 antibodies, the preparations were made according to the procedure described in James *et al.* (1989). For DNA Pol I labelling of polytene chromosomes, after freezing in liquid nitrogen, slides were washed once in PBS, treated for 10 min at 37°C with 1 mg/ml of proteinase K, washed twice in PBS and subjected to nick translation for 30 min at room temperature with 5 U/ml of DNA Pol I and 0.04 mM of the fluorescent nucleotide Cy3-dUTP in NT buffer (50 mM Tris–HCl pH 7.9, 5 mM $MgCl_2$, 50 mg/ml BSA, 10 mM β-mercaptoethanol) containing 0.15 mM dATP, dCTP and dGTP, and 0.01 mM dTTP.

Chromosome preparations were analysed using a computer-controlled Zeiss Axioplan epifluorescence microscope equipped with a cooled CCD camera (Photometrics). The fluorescent signals, recorded separately as grey scale digital images, were pseudocoloured and merged using the Adobe Photoshop program.

Embryos for immunostaining were collected at different times, decorionated in 100% bleach, rinsed in 0.1% Triton X-100 and fixed for 20 min in 2 ml PEM (0.1 M PIPES, 1 mM EGTA, 2 mM $MgSO_4$, pH 6.9), 0.25 ml formaldehyde and 2 ml heptan. The lower phase was then discarded, 2 ml of methanol added and, after 20 s of vortexing, embryos were collected from the bottom, washed three times in methanol and stored at 4°C. Embryos were then permeabilized in PBS, 0.3% Triton X-100, incubated over night at 4°C with αDDP1 antiserum at a 1:400 dilution in PBS, 0.3% Triton X-100, blocked in PBS, 0.3% Triton X-100, 2% BSA, stained with secondary antibody (FITC-anti-mouse in PBS, 0.3% Triton X-100, 2% BSA), washed in PBS, 0.3% Triton X-100, stained with DAPI and mounted in anti-fading medium.

Obtention of the full-length cDNA of DDP1 and expression in E.coli

Degenerate oligonucleotides were designed from the sequences of the two internal peptides described in Table I, and used to PCR-amplify an 800 bp specific DNA probe from a mixture of two plasmid cDNA libraries constructed from 0–4 and 4–8 h *Drosophila* embryos (Brown and Kafatos, 1988), a gift from Dr Brown. The specific probe was then used to screen 350 000 clones of the same libraries. Three independent positive clones were obtained, one of which contained the full cDNA of DDP1. For expression in *E.coli*, the full DDP1-coding sequence was cloned in a pET29-b expression vector (Novagen) and expressed in BL21 cells according to standard protocols as a fusion protein carrying a C-terminal His_6-tag. The recombinant protein was purified in a Ni^{2+}-NTA column (Qiagen), washed with 20 mM imidazol and eluted with 100 mM imidazol.

Yeast experiments

Complementation of the yeast *Δscp160* deletion by DDP1 was performed essentially as described elsewhere (Wintersberger *et al.*, 1995). Diploid strains AK300 (*MAT***a**/α *ade2-1/ ade2-1 trp1-1/trp1-1 leu2-3,112/leu2-3,112 his3-11,15/his3-11,15 ura3/ura3 SCP160/SCP160*) and AK303 (isogenic to AK303, but *SCP160/scp160::LEU2*) were kindly supplied by Dr Wintersberger. Both strains were transformed by the lithium acetate method (Sherman, 1991) with the plasmid pYES-DDP1, a derivative of pYES2 (Invitrogen) which expresses the DDP1 coding sequence under the GAL1-10 promoter. Transformants were sporulated and dissected as described elsewhere (Sherman and Hicks, 1991). Spores were isolated in a Tetrad Dissection Microscope (Micro Video Instruments, Inc., Avon, MA) after digesting the ascus wall with Lyticase (Sigma). Spores were germinated in YEP-galactose (Sherman, 1991) to allow expression of DDP1. DNA contents of relevant clones (both parental and spores) were measured at late exponential phase by flux cytometry (FACS) in a Coulter Epics Elite cytometer (Serveis Tècnics, Universitat de Barcelona, Spain) with a blue argon laser (488 nm, 15 mW). Fluorescence was detected at 665–685 nm.

Acknowledgements

We are very thankful to Dr Wintersberger for kindly providing us with yeast strains AK300 and AK303. We are also thankful to Drs Villasante, Ripoll and Casanova for helpful discussions and suggestions, and to Mr Casali for technical assistance with the flies. The contribution of Dr Ferrer at the very initial phases of the project is also acknowledged. This work was financed by grants from the Spanish DGES (PB96-812), the EU (BIO2-CT94-3069) and the CIRIT (SGR97-55). A.C. and F.X.M. acknowledge receipt of doctoral fellowships from the CIRIT. This work was carried out within the framework of the 'Centre de Referència en Biotecnologia' of the Generalitat de Catalunya.

References

Abad,J.P., Carmena,M., Baars,S., Saunders,R.D.C., Glover,D.M., Ludeña,P., Sentis,C., Tyler-Smith,C. and Villasante,A. (1992) Dodeca-satellite: a conserved G+C-rich satellite from centromeric heterochromatin of *Drosophila melanogaster. Proc. Natl Acad. Sci. USA*, **89**, 4663–4667.

Alfenito,M.R. and Birchler,J.A. (1993) Molecular characterization of a maize B chromosome centric sequence. *Genetics*, **135**, 589–597.

Belyaeva,E.S., Demakova,O.V., Umbetova,G.H. and Zhimulev,I.F. (1993) Cytogenetic and molecular aspects of position-effect variegation in *Drosophila melanogaster.* V. Heterochromatin-associated protein HP1 appears in euchromatic chromosomal regions that are inactivated as a result of position-effect variegation. *Chromosoma (Berlin)*, **102**, 583–590.

Brockdorff,N. (1998) The role of Xist in X-inactivation. *Curr. Opin. Genet. Dev.*, **8**, 328–333.

Brown,N.H. and Kafatos,F.C. (1988) Functional cDNA libraries from *Drosophila* embryos. *J. Mol. Biol.*, **205**, 425–437.

Carmena,M., Abad,J.P., Villasante,A. and González,C. (1993) The *Drosophila melanogaster* dodecasatellite sequence is closely linked to the centromere and can form connections between sister chromatids during mitosis. *J. Cell Sci.*, **105**, 41–50.

Catasti,P., Gupta,G., García,A.E., Hong,L., Yau,P., Moyzis,R.K. and Bradury,E.M. (1994) Unusual structures of tandem repetitive sequences located at human centromeres. *Biochemistry*, **33**, 3819–3830.

Choo,K.H. (ed.) (1997) *The Centromere*. Oxford University Press, Oxford, UK.

Chou,S.-H., Zhu,L. and Reid,B.R. (1994) The unusual structure of the human centromere $(GGA)_2$ motif. *J. Mol. Biol.*, **244**, 259–268.

Clarke,L., Baum,M., Marschall,L.G., Ngan,V.K. and Steiner,N.C. (1993) Structure and function of *Schizosaccharomyces pombe* centromeres. *Cold Spring Harbor Symp. Quant. Biol.*, **58**, 687–695.

Cléard,F., Delattre,M. and Spierer,P. (1997) SU (VAR)3-7, a *Drosophila*

heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. *EMBO J.*, **16**, 5280–5288.
Dejgaard,K. and Leffers,H. (1996) Characterisation of the nucleic-acid-binding activity of KH domains. Different properties of different domains. *Eur. J. Biochem.*, **241**, 425–431.
Digman,J.D., Lebowitz,R.M. and Roeder,R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.
Dodson,R.E. and Shapiro,D.J. (1997) Vigilin, a ubiquitous protein with 14 K homology domains, is the estrogen-inducible vitellogenin mRNA 3′-untranslated region-binding protein. *J. Biol. Chem.*, **272**, 12249–12252.
du Sart,D. *et al.* (1997) A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. *Nature Genet.*, **16**, 144–153.
Edgar,B.A., Kiehle,C.R. and Schubiger,G. (1986) Cell cycle control by the nucleo-cytoplasmic ratio in early *Drosophila* development. *Cell*, **44**, 365–372.
Ferrer,N., Azorín,F., Villasante,A., Gutiérrez,C. and Abad,J.P. (1995) Centromeric dodeca-satellite DNA sequences form fold-back structures. *J. Mol. Biol.*, **245**, 8–21.
Foe,V.E. (1989) Mitotic domains reveal early commitment of cells in *Drosophila* embryos. *Development*, **107**, 1–22.
Fry,K. and Salser,W. (1977) Nucleotide sequence of HS-α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell*, **12**, 1069–1084.
Gatti,M., Bonaccorsi,S. and Pimpinelli,S. (1994) Looking at *Drosophila* mitotic chromosomes. *Methods Cell Biol.*, **44**, 371–391.
Grady,D.L., Ratliff,R.L., Robinson,D.L., McCanlies,E.C., Meyne,J. and Moyzis,R.K. (1992) Highly conserved repetitive DNA sequences are present at human centromeres. *Proc. Natl Acad. Sci. USA*, **89**, 1695–1699.
Harrington,J.J., Bokkelen,G.V., Mays,R.W., Gustashaw,K. and Willard,H.F. (1997) Formation of *de novo* centromeres and construction of first-generation human artificial minichromosomes. *Nature Genet.*, **15**, 345–355.
Ikeno,M., Grimes,B., Okazaki,T., Nakano,M., Saitoh,K., Hoshino,H., McGill,N., Cooke H. and Masumoto,H. (1998) Construction of YAC-based mammalian artificial chromosomes. *Nature Biotechnol.*, **16**, 431–439.
James,T.C. and Elgin,S.C.R. (1986) Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila* and its gene. *Mol. Cell. Biol.*, **6**, 3862–3872.
James,T.C., Eissenberg,J.C., Craig,C., Dietrich,V., Hobson,A. and Elgin,S.C.R. (1989) Distribution of HP1, a heterochromatin-associated nonhistone chromosomal protein of *Drosophila*. *Eur. J. Cell Biol.*, **50**, 170–180.
Kalitsis,P. and Choo,K.H. (1997) Centromere DNA of higher eukaryotes. In Choo,K.H. (ed.), *The Centromere*. Oxford University Press, Oxford, UK, pp. 97–142.
Kanamori,H., Dodson,R.E. and Shapiro,D.J. (1998) *In vitro* genetic analysis of the RNA binding site of Vigilin, a multi-KH-domain protein. *Mol. Cell. Biol.*, **18**, 3991–4003.
Karpen,G. and Allshire,R. (1997) The case for epigenetic effects on centromere identity and function. *Trends Genet.*, **13**, 489–496.
Kellum,R., Raff,J.W. and Alberts,B.M. (1995) Heterochromatin protein 1 distribution during development and during the cell cycle in *Drosophila* embryos. *J. Cell Sci.*, **108**, 1407–1418.
Lohe,A.R., Hilliker,A.J. and Roberts,P.A. (1993) Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics*, **134**, 1149–1174.
Mahowald,A.P. and Hardy,P.A. (1985) Genetics of *Drosophila* embryogenesis. *Annu. Rev. Genet.*, **19**, 149–177.
Martínez-Balbás,A., Rodríguez-Campos,A., García-Ramírez,M., Sainz,J., Carrera,P., Aymamí,J. and Azorín,F. (1990) Satellite DNAs contain sequences that induce curvature. *Biochemistry*, **29**, 2342–2348.
McKnight,G.L. *et al*. (1992) Cloning and expression of a cellular high density lipoprotein-binding protein that is up-regulated by choresterol loading of cells. *J. Biol. Chem.*, **267**, 12131–12141.
Meyne,J. *et al*. (1990) Distribution of non-telomeric sites of the $(TTAGGG)_n$ telomeric sequences in vertebrade chromosomes. *Chromosoma*, **99**, 3–10.
Miklos,G.L.G. and Costell,J.N. (1990) Chromosome structure at interfaces between major chromatin types: α- and β-heterochromatin. *BioEssays*, **12**, 1–6.
Murphy,T.D. and Karpen,G.H. (1995) Localization of centromere function in a *Drosophila* minichromosome. *Cell*, **82**, 599–609.
Murphy,T.D. and Karpen,G.H. (1998) Centromeres take flight: alpha satellite and the quest for the human centromere. *Cell*, **93**, 317–320.
Musco,G., Stier,G., Joseph,C., Castiglione Morelli,M.A., Nilges,M., Gibson,T.J. and Pastore,A. (1996) Three-dimensional structure and stability of the KH-domain: molecular insights into the fragile X syndrome. *Cell*, **85**, 237–245.
Novak,U. (1984) Structure and properties of highly repetitive DNA sequences in sheep. *Nucleic Acids Res.*, **12**, 2343–2350.
Ortiz-Lombardía,M., Cortés,A., Huertas,D., Eritja,R. and Azorín,F. (1998) Tandem 5′-GA:GA-3′ mismatches account for the high stability of the fold-back structures formed by the centromeric *Drosophila* dodeca-satellite. *J. Mol. Biol.*, **277**, 757–762.
Pak,D.T.S., Pflumm,M., Chesnokov,I., Huang,D.W., Kellum,R., Marr,J., Romanowski,P. and Botchan,M.R. (1997) Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell*, **91**, 311–323.
Pimpinelli,S., Berloco,M., Fanti,L., Dimitri,P., Bonaccorsi,S., Marchetti,E., Caizzi,R., Caggese,C. and Gatti,M. (1995). Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. *Proc. Natl Acad. Sci. USA*, **92**, 3804–3808.
Platero,J.S., Csink,A.K., Quintanilla,A. and Henikoff,S. (1998) Changes in chromosomal localization of heterochromatin binding proteins during the cell cycle in *Drosophila*. *J. Cell Biol.*, **140**, 1297–1306.
Plenz,G., Kügler,S., Schnittger,S., Rieder,H., Fonatsch,C. and Müller,P.K. (1994) The human vigilin gene: identification, chromosomal localisation and expression pattern. *Hum. Genet.*, **93**, 575–582.
Rattner,J.B. (1991) The structure of the mammalian centromere. *BioEssays*, **13**, 51–56.
Richards,E.J., Goodman,H.M. and Ausubel,F.M. (1991) The centromere region of *Arabidopsis thaliana* chromosome 1 contains telomere-similar sequences. *Nucleic Acids Res.*, **19**, 3351–3357.
Sacchi,N., Magnani,I., Fuhrman-Conti,A.M., Monard,S.P. and Darfler,M. (1996) A stable marker chromosome with a cryptic centromere: evidence for centromeric sequences associated with an inverted duplication. *Cytogenet. Cell Genet.*, **73**, 123–129.
Schmidt,C., Henkel,B., Pöschl,E., Zorbas,H., Purschke,W.G., Gloe,T.R. and Müller,P.K. (1992) Complete cDNA sequence of chicken vigilin, a novel protein with amplified and evolutionary conserved domains. *Eur. J. Biochem.*, **206**, 625–634.
Sherman,F. (1991) Getting started with yeast. *Methods Enzymol.*, **194**, 3–21.
Sherman,F. and Hicks,J. (1991) Micromanipulation and dissection of asci. *Methods Enzymol.*, **194**, 21–37.
Siomi,H., Matunis,M.J., Michael,W.M. and Dreyfuss,G. (1993) The pre-mRNA binding protein K contains a novel evolutionarily conserved motif. *Nucleic Acids Res.*, **21**, 1193–1198.
Slama-Schwok,A., Brossalina,E., Demchenko,Y., Best-Belpomme,M. and Vlassov,V. (1998) Structural flexibility of a DNA hairpin located in the long terminal repeat of the *Drosophila* 1731 retrotransposon. *Nucleic Acids Res.*, **26**, 5142–5151.
Sullivan,B.A. and Schwartz,S. (1995) Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. *Hum. Mol. Genet.*, **4**, 2189–2197.
Sun,X., Wahlstrom,J. and Karpen,G. (1997) Molecular structure of a functional *Drosophila* centromere. *Cell*, **91**, 1007–1019.
Sunkel,C.E. and Coelho,P.A. (1995) The elusive centromere: sequence divergence and functional conservation. *Curr. Opin. Genet. Dev.*, **5**, 756–767.
Takahashi,K., Murakami,S., Chikashige,Y., Funabiki,Y. and Yanagida,M. (1992) A low copy number central sequence with strict symmetry and unusual chromatin structure in fission yeast centromere. *Mol. Cell. Biol.*, **3**, 819–835.
Taparowsky,E.J. and Gerbi,S.A. (1982) Sequence analysis of bovine satellite I DNA (1.715 gm/cm^3) *Nucleic Acids Res.*, **10**, 1271–1281.
Voullaire,L.E., Slater,H.R., Petrovic,V. and Choo,K.H. (1993) A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am. J. Hum. Genet.*, **52**, 1153–1163.
Weber,V., Wernitzing,A., Hager,G., Harata,M., Frank,P. and Wintersberger,U. (1997) Purification and nucleic-acid-binding properties of a *Saccharomyces cerevisiae* protein involved in the control of ploidy. *Eur. J. Biochem.*, **249**, 309–317.

Wiens,G.R. and Sorger,P.K. (1998) Centromeric chromatin and epigenetic effects in kinetochore assembly. *Cell*, **93**, 313–316.
Williams,B.C., Murphy,T.D., Goldberg,M.L. and Karpen,G.H. (1998) Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nature Genet.*, **18**, 30–37.
Wintersberger,U., Kühne,C. and Karwan,A. (1995) Scp160p, a new yeast protein associated with the nuclear membrane and the endoplasmic reticulum, is necessary for maintenance of exact ploidy. *Yeast*, **11**, 929–944.
Zinkowski,R.P., Meyne.,J. and Brinkley,B.R. (1991) The centromere-kinetochore complex: a repeat subunit model. *J. Cell Biol.*, **113**, 1091–1110.

Received March 18, 1999; revised and accepted May 11, 1999