11-1992

Heterochromatin protein 1, a known suppressor of position-effect variegation, is highly conserved in Drosophila

Robert F. Clark
Washington University in St. Louis

Sarah C.R. Elgin
Washington University in St. Louis, selgin@wustl.edu

Nucleic Acids Research, 1992, Vol. 20, No. 22 6067–6074

Heterochromatin protein 1, a known suppressor of position-effect variegation, is highly conserved in *Drosophila*

Robert F.Clark and Sarah C.R.Elgin
Department of Biology, Box 1137, Washington University, St Louis, MO 63130, USA

Received June 17, 1992; Revised and Accepted October 19, 1992 GenBank accession no. M88753

ABSTRACT

The *Su(var)205* gene of *Drosophila melanogaster* encodes heterochromatin protein 1 (HP1), a protein located preferentially within β-heterochromatin. Mutation of this gene has been associated with dominant suppression of position-effect variegation. We have cloned and sequenced the gene encoding HP1 from *Drosophila virilis*, a distantly related species. Comparison of the predicted amino acid sequence with *Drosophila melanogaster* HP1 shows two regions of strong homology, one near the N-terminus (57/61 amino acids identical) and the other near the C-terminus (62/68 amino acids identical) of the protein. Little homology is seen in the 5′ and 3′ untranslated portions of the gene, as well as in the intronic sequences, although intron/exon boundaries are generally conserved. A comparison of the deduced amino acid sequences of HP1-like proteins from other species shows that the cores of the N-terminal and C-terminal domains have been conserved from insects to mammals. The high degree of conservation suggests that these N- and C-terminal domains could interact with other macromolecules in the formation of the condensed structure of heterochromatin.

INTRODUCTION

In *Drosophila*, euchromatic genes placed adjacent to heterochromatin by chromosomal rearrangements may be inappropriately inactivated in somatic cells by an epistatic process. The decision to switch the gene off is variable, but once made is clonally inherited during mitotic growth, giving rise to a variegated pattern of gene expression. This phenomenon is known as position-effect variegation (PEV). Almost all *Drosophila* genes are susceptible to PEV, the extent of transcriptional inactivation reflecting proximity to the heterochromatin breakpoint in a given rearrangement. This 'spreading effect' of heterochromatin can extend over eighty bands on polytene chromosomes or approximately 1500 kb (1).

The DNA sequence of a gene exhibiting PEV is unaltered, as shown by recovery of full activity in revertants that place the gene once again in a euchromatic environment; the inactivation process reflects the position of the gene within the genome (1). In theory, PEV might be caused either by under-replication or loss of DNA, or by changes in the chromatin packaging of the gene. While it has been shown that under-replication of such gene sequences can occur in polytene cells, and such under-replication has been associated in some cases with PEV (2–3), many results point to the packaging mechanism as an essential component of PEV. Inspection of polytene chromosomes suggests that the euchromatin adjacent to the breakpoint can lose its normal banded appearance and assume the amorphous structure of β-heterochromatin, again as a variegating event (4). In several cases, direct analysis of variegating genes has found no evidence of under-replication; the cases investigated include rearranged heat shock genes and eye-color mutants (4–6).

In a stock carrying a variegating locus, second site mutations can be recovered which suppress or enhance the variegation. It has been estimated that there may be as many as 150 such suppressor and/or enhancer loci (7). Among these, a few loci showing haplo-suppressor/triplo-enhancer effects on PEV in *Drosophila* have been identified (4). It has been suggested that the proteins encoded by these loci may be structural components of heterochromatin or their modifiers (8,9). One such gene, *Suvar(3)7*, encodes a protein with five widely spaced zinc fingers (10). Another such gene, *Su(var)205*, encodes a protein preferentially associated with β-heterochromatin, heterochromatin protein 1 (HP1) (11–13).

HP1 was initially identified in *D.melanogaster* as a 19 kDa protein present in embryonic nuclei, extractable with 1–2 M potassium isothiocyanate (11). Monoclonal antibodies produced against this protein showed that it is found in polytene chromosomes preferentially at the chromocenter, which is composed of the α- and β-heterochromatin of the four chromosomes. The presence of a *D.virilis* chromosomal protein antigenically similar to that of *D.melanogaster* was established by immunofluorescent staining of *D.virilis* polytene chromosomes with antibodies raised against the *D.melanogaster* protein (14); the protein in *D.virilis* is also associated with the centric heterochromatin in the polytene chromosomes. The cDNA and genomic clones encoding this protein in *D.melanogaster* were sequenced, and the gene was mapped to cytological position 29A (11,12), a region where a dominant suppressor of PEV had earlier been mapped (15). Characterization of the DNA sequences of known mutations at this locus has confirmed that mutation in the gene encoding HP1 (including mutations that should result in reduced amounts of gene product) can result in suppression of

PEV (12,13). The stoichiometric effects observed suggest that HP1 might be one of the 'building blocks' of heterochromatin structure.

As an initial step in determining the functionally relevant domains of the HP1 protein, we have analyzed the evolutionary divergence of the genes encoding HP1 from two distantly related *Drosophila* species, *D.melanogaster* and *D.virilis*. *D.melanogaster* and *D.virilis* diverged approximately 40–60 million years ago (16,17). This is sufficiently distant for unconstrained DNA sequences to have diverged extensively, so that functionally meaningful elements may be identified by sequence conservation (18–24). As presented here, much of the non-coding DNA sequence upstream, downstream, and within the introns of the genes encoding HP1 has diverged considerably between the two species, although the locations of intron/exon boundaries are generally conserved. Two highly conserved regions in the coding DNA exist, one in the N-terminal and the other in the C-terminal portions of the predicted amino acid sequence of HP1. The high degree of conservation suggests that these regions could interact with other macromolecules in the formation of the condensed structure of heterochromatin.

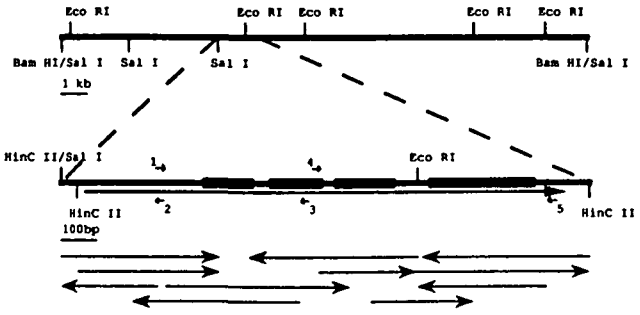

Figure 1. Restriction map and sequencing strategy for the *Drosophila virilis* HP1 genomic clone. The top portion of the figure shows the restriction map of the 18 kb *Bam* H1 fragment isolated from a λEMBL3 *D.virilis* genomic library, using the one kb full length *D.melanogaster* HP1 cDNA (11) as a probe. The bottom portion of the figure shows the restriction map of the HP1 gene and flanking DNA, with the regions sequenced represented by the thin arrows. The thick arrow indicates the RNA transcript, from the experimentally determined transcriptional start sites to the consensus polyadenylation signal site (30). The thick line represents the four exons of the HP1 gene. The small numbered lines above and below the map indicate the relative positions of oligonucleotides used in sequencing. Oligonucleotide sequences are as follows: 1, 5′-GCGCTAAAATCTGCACAC-3′; 2, 5′-GTGTGCAGATTTTAGCGC-3′; 3, 5′-ATAAGATCCTGGCA-3′; 4, 5′-CGCAAGGATGAGGTAAG-3′; 5, 5′-TTAATATGCTGAGAAAAA-3′.

MATERIALS AND METHODS

D.virilis genomic clone

Drosophila virilis HP1 genomic clones were isolated from a λEMBL3 *D.virilis* genomic library (18) using a one kb full-length *Drosophila melanogaster* HP1 cDNA (11) as a probe. The probe was labeled by random priming (25) to an activity of about 10^8 cpm/μg. Approximately 30,000 plaques were transferred to Nytran filters (Schleicher and Schuell) using standard methods (26). The filters were prehybridized in 2×SSC (1×SSC, standard saline citrate, is 0.15 M NaCl/0.015 M sodium citrate, pH 7) and 0.1% SDS, and then hybridized in 2×SSC/0.1% SDS/0.1% Ficoll/0.1% bovine serum albumin/0.1% polyvinylpyrrolidone/100 mg/ml sonicated salmon sperm DNA with labeled probe at 42°C. Filters were washed at 50°C in 2×SSC/0.1% SDS (26), and exposed to XAR-5 X-ray film (Kodak) at −80°C with a Cronex Lightning Plus intensifying screen (DuPont). Eighteen positive plaques were isolated and rescreened with the same probe, as above. One strongly hybridizing clone was isolated after this rescreen, and DNA from this phage was purified using standard methods (26).

Southern and Northern analysis

Southern analysis using high molecular weight genomic DNA from *D.virilis* adult flies (27), and Northern analysis using total RNA extracted from 30 *D.virilis* larvae or female adults (28), were carried out using standard techniques (26). The filters were prehybridized, and then hybridized with the labeled one kb full length *D.melanogaster* HP1 cDNA (11) (at 42°C for the Southern, 50°C for the Northern), as above. The filters were washed at 50°C in 2×SSC/0.1% SDS (26), and exposed to X-ray film.

DNA sequencing and sequence comparisons

Southern blot analysis of restriction digests of DNA from the isolated λEMBL3 *D.virilis* clone revealed a 1.4 kb *HinC* II fragment that hybridized with the *D.melanogaster* HP1 cDNA. This fragment and other restriction fragments bearing regions of interest were subcloned into the single-stranded bacteriophage vectors M13mp18 and M13mp19. Sequencing was done by the dideoxy chain-termination method (29) with [α-^{35}S]dATP and Sequenase (United States Biochemical); either the universal M13 primer or synthesized oligonucleotide primers complementary to sequences within the HP1 gene were used (Figure 1).

The DNA sequences of the genes encoding HP1 from *D.virilis* and *D.melanogaster* (12) were compared using the GCG programs (Wisconsin Computer Group). Gap weight was set at 5, and length weight was set at 0.3. Four hundred bp regions (with the highly conserved coding sequences as anchors) were used to align the upstream, downstream, and intronic sequences.

The deduced amino acid sequences of HP1 from *D.virilis* and *D.melanogaster* (12), and the HP1-like genes from mouse (31), human (32), and mealybug (33) were compared using the GCG programs (Wisconsin Computer Group), aligning amino acids having similar chemical sidechains (34). Each sequence was compared pairwise, and the optimal arrangement was maximized. The same program was used to align a portion of the *Polycomb* (Pc) protein (35) with the HP1 amino acid sequences.

Primer extension analysis

To determine the transcriptional start site for the *D.virilis* HP1 gene, primer extension analysis was performed using end-labeled synthetic oligonucleotides specific for sequences upstream of the open reading frame. End-labeling was carried out with T4 polynucleotide kinase (USB) and [γ-^{32}P]dATP (26). 10^5 cpm of labeled oligonucleotides were added to 10 μg of total RNA from *D.virilis* larvae and/or adults. The primer and the RNA were dried and then dissolved in 20 μl of 40 mM Pipes, pH 6.4/400 mM NaCl/1 mM EDTA/50% formamide (36), placed at 65°C for one min, and incubated at 37°C for 60 min. Hybridization products were ethanol precipitated and dissolved in 20 μl reaction buffer (36) with fresh ultrapure deoxynucleotide triphosphates and 200 U Moloney murine leukemia virus (M-MuLV) reverse transcriptase (BRL). The mixture was incubated for 2 hours at 37°C, phenol-chloroform extracted, ethanol precipitated, and the pellet dissolved in 6 ml of standard loading dye for DNA

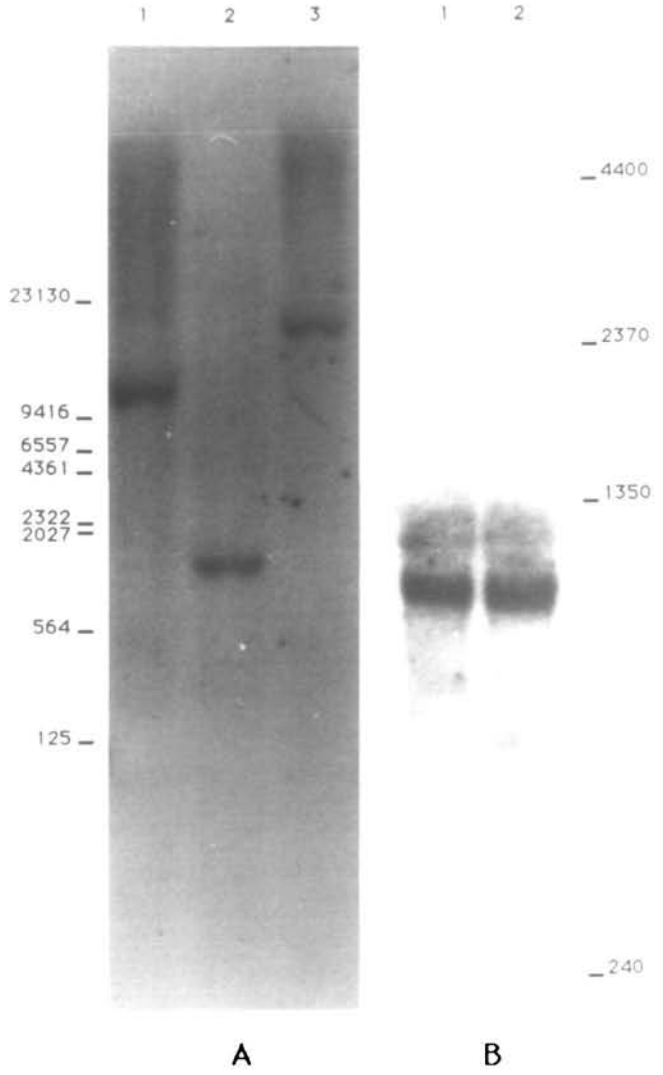

Figure 2. Southern and Northern analyses of the *D.virilis* HP1 gene (A) Autoradiograph of a genomic Southern blot of *D.virilis* high-molecular weight DNA. Genomic DNA was isolated from adult flies, digested with *Sal* I (lane 1), *HinC* II (lane 2), and *Bam* HI (lane 3), size fractionated in an agarose gel, transferred onto a filter, and probed with *D.melanogaster* HP1 cDNA (11). The same pattern is seen when the filter is probed with labeled *D.virilis* DNA fragments from the genomic HP1 clone (data not shown). (B) Autoradiograph of a Northern blot of *D.virilis* RNA. Total RNA from 30 *D.virilis* larvae (lane 1) or 30 female adults (lane 2), was size fractionated through a formaldehyde/agarose denaturing gel. The RNA was transferred onto a filter, and probed with labeled one kb full length *D.melanogaster* HP1 cDNA (11). The same pattern is seen when the filter is probed with labeled *D.virilis* DNA fragments from the genomic HP1 clone (data not shown). Size markers (in bp) are indicated for both figures.

sequencing and analyzed on an 8% sequencing gel. The same results were obtained with RNA from six independent isolations, and from the use of either larval or adult RNA.

Polymerase chain reaction

To determine whether or not an intron is present in the 5′ untranslated sequence of *D.virilis* HP1, as exhibited for the gene encoding HP1 in *D.melanogaster* (12), first-strand cDNA product was used as a template for polymerase chain reaction (PCR) (37) using the oligonucleotides #1 and #3, and #2 and #4, as shown in Figure 3. First-strand cDNA product from *D.virilis* RNA made using oligonucleotide #5 (Figure 3) was gel purified on 2% agarose; the gel band containing the product was excised, and the DNA was electrophoresed onto DEAE paper (Schleicher and Schuell), eluted from the paper, and ethanol precipitated. Two PCR reactions were carried out on this DNA using pairs of synthetic oligonucleotides (#1 and #3, and #2 and #4), such that 200 bp of the 215 bp untranslated upstream region were included in the analysis. Control experiments under identical conditions were performed using the cloned *D.virilis* genomic HP1 DNA as a template; PCR products from the cDNA and genomic templates were compared by electrophoresis on a 2% agarose gel.

RESULTS AND DISCUSSION

Isolation and sequencing of the *D.virilis* HP1 gene

The *D.virilis* HP1 gene was recovered from a λEMBL3 genomic library of *D.virilis* using the full-length *Drosophila melanogaster* HP1 cDNA as a probe (see Materials and Methods). The regions of homologous sequence were found to be completely contained within a 1.4 kb *D.virilis HinC* II fragment (Figure 1). The *D.virilis* gene was shown to be single copy by Southern analysis of restriction enzyme-digested high molecular weight genomic DNA (Figure 2a). The restriction map of the *D.virilis* HP1 clone matched that of genomic DNA from *D.virilis*. The *D.melanogaster* cDNA probe hybridized to a one kb RNA transcript in a Northern blot of *D.virilis* RNA from larvae and adults (Figure 2b). The restriction map and strategy used to sequence *D.virilis* HP1 genomic DNA are diagrammed in Figure 1. All sequencing data were derived from both strands of the HP1 genomic DNA. The DNA sequence of the 1.4 kb *HinC* II fragment containing the *D.virilis* HP1 gene, and an additional 49 bp *HinC* II fragment immediately upstream, is shown in Figure 3. This sequence contains all of the protein-coding sequences, as well as 405 bp of 5′ and 163 bp of 3′ noncoding DNA.

Transcriptional initiation of the *D.virilis* HP1 gene

Primer extension analysis was used to identify the sites of transcriptional initiation of the *D.virilis* HP1 gene. Using oligonucleotide #3 (Figure 3) as a primer, extension products of 183 bp, 205 bp, and 450 bp were observed upon extension with M-MuLV reverse transcriptase (data not shown). The 183 bp and 205 bp products indicate RNA transcripts that would initiate at the sites shown in Figure 3. Such transcriptional start sites will produce transcripts initiated at similar distances from the translational start as those found for the gene encoding *D.melanogaster* HP1 (12). However, a major product of 450 bp, inconsistent with the *D.melanogaster* transcriptional start sites, was also found. No transcript of a similar length was seen upon Northern analysis (Figure 2b) and no products were found when oligonucleotides further upstream of the first start site were used for primer extension analysis or as probes for Northern blots. Therefore, we believe the transcriptional start sites for *D.virilis* HP1 are found 321 bp and 342 bp upstream of the translational start codon (Figure 3). The site 321 bp upstream of the translational start shows a resemblance to the *D.melanogaster* transcription initiation consensus sequence ATCA(G/T)T(C/T) (38); it is CGCATTC.

The *D.virilis* gene lacks sequences for the consensus splice junction (GT-AG) in this upstream region (39), and therefore

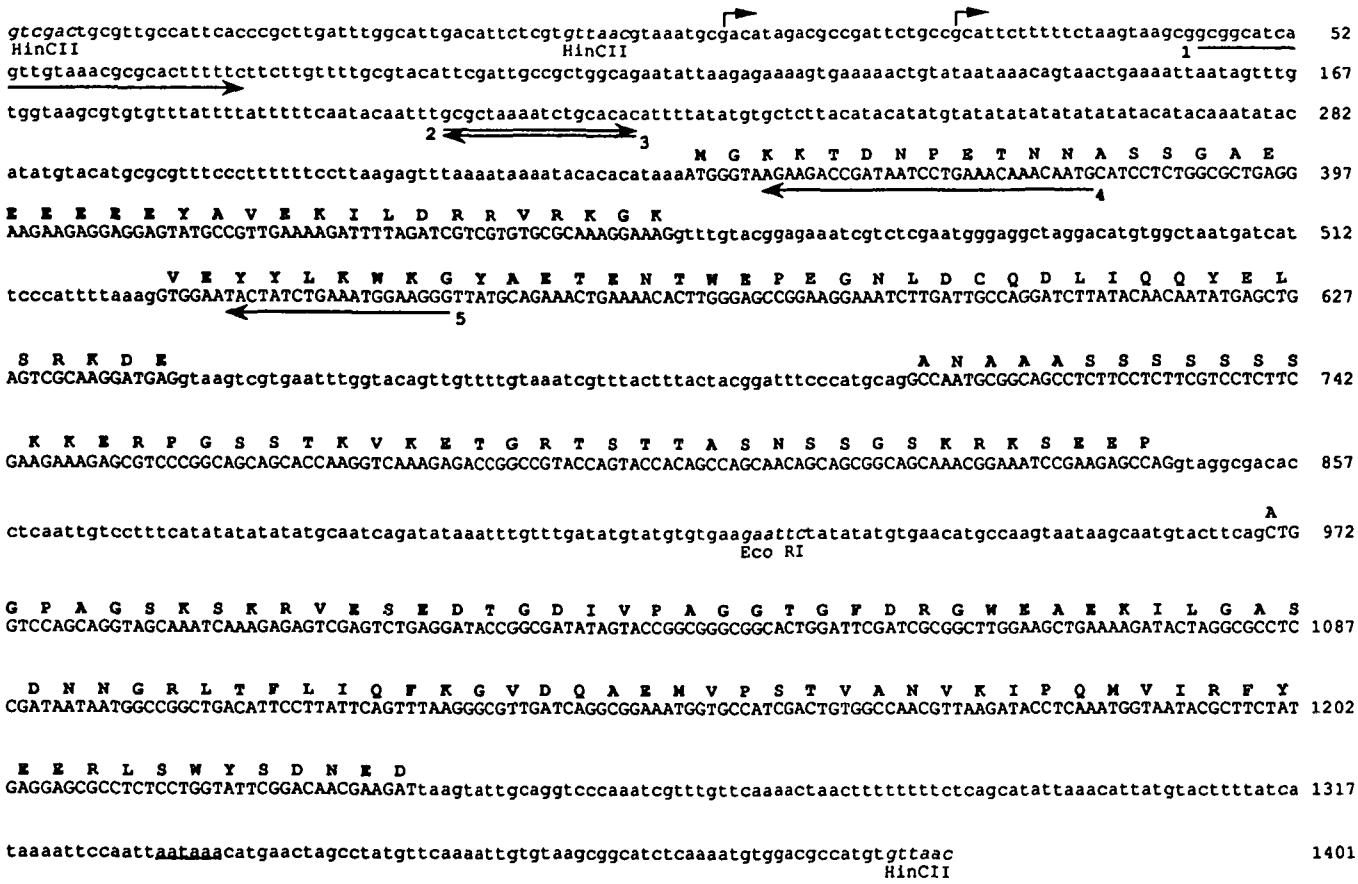

Figure 3. Sequence of the *D. virilis* HP1 gene and flanking DNA, and the deduced amino acid sequence of the HP1 protein. The predicted protein-coding DNA sequences are shown by uppercase letters; +1 indicates the first transcriptional start site; bent arrows indicate the transcriptional start sites as determined by primer extension; the underlined sequence indicates the consensus polyadenylation signal (30). The numbered arrows indicate the synthetic oligonucleotides used with PCR to search for a possible intron in the upstream untranslated region. Oligonucleotides are numbered at their 5′ ends.

might lack the intron found in the 5′ untranslated leader of the *D. melanogaster* gene. To verify this, PCR analysis was utilized. Oligonucleotide #5 (Figure 3) was used to prime first-strand cDNA from *D. virilis* RNA. Oligonucleotide pairs #1 and #3, and #2 and #4, were used to generate PCR products from this cDNA and from the genomic *D. virilis* HP1 clone. PCR products from the cDNA and genomic templates were identical in length for both pairs of oligonucleotides (data not shown). Therefore, unlike *D. melanogaster*, *D. virilis* lacks an intron in the noncoding region of the HP1 gene.

Deduced amino acid sequence and comparison of coding regions

When the DNA sequences of the genes encoding *D. virilis* and *D. melanogaster* HP1 were aligned and compared, the highest conservation was found within the coding regions of the two genes (Figure 4). The *D. virilis* HP1 protein, like that of *D. melanogaster*, is encoded by four exons. The conservation at the DNA level in the four exons, calculated as percent nucleotide identity relative to the total number of nucleotides in the *D. melanogaster* sequence, is 67.5%, 83.8%, 49.6%, and 71.1% in the first, second, third, and fourth exons, with an overall identity of 68.6% for the entire coding region.

Amino acid sequences were deduced from the coding regions of the *D. virilis* and *D. melanogaster* genes, and these sequences were aligned to optimize amino acid identities (Figure 5). The predicted *D. virilis* HP1 protein is 213 amino acids long, seven amino acids longer than the *D. melanogaster* protein. The overall identity at the amino acid level is 77% between the *Drosophila* species, and the similarity is 83% when conservative substitutions are included. Similar levels of conservation have been described for the genes *en* (19), *hunchback* (22), *period* (40), and parts of *Ubx* (41), all essential genes. All alleles of the gene encoding HP1 are both dominant suppressors of PEV and recessive lethals; clearly its product is functionally important.

The deduced products of HP1-like genes from mouse (31), human (32), and mealybug (33) were aligned and compared to the HP1 protein of *Drosophila*. A second HP1-like gene product, identical to one of the mouse gene products, has been reported in humans (31). The non-*Drosophila* amino acid sequences share a 30–45% identity with their *Drosophila* counterparts, with two regions of high homology (Figure 5). One region is a 61-amino acid domain near the C-terminal end of the HP1-like proteins that shows 28% identity among all HP1-like proteins characterized to date, with conservative replacements accounting for another 18%. Using only the HP1 proteins of *D. virilis* and *D. melanogaster*, this region can be extended to 68 amino acids, which share a 91% identity. The second conserved region near the N-terminal end of the protein is a 46-amino acid domain showing a 52% identity among all of the HP1-like proteins

Figure 4. Sequence comparison of the genes encoding HP1 in *D. virilis* and *D. melanogaster*. The sequences for *D. virilis* are displayed above those of *D. melanogaster* (12, J. Eissenberg, personal communication). Symbols and numbering are as used in Figure 3. The two dark bars over the sequence indicate the two strongly homologous regions (the upper bar is broken by an intron), and the square brackets delineate the intron in the upstream untranslated sequence of the *D. melanogaster* gene.

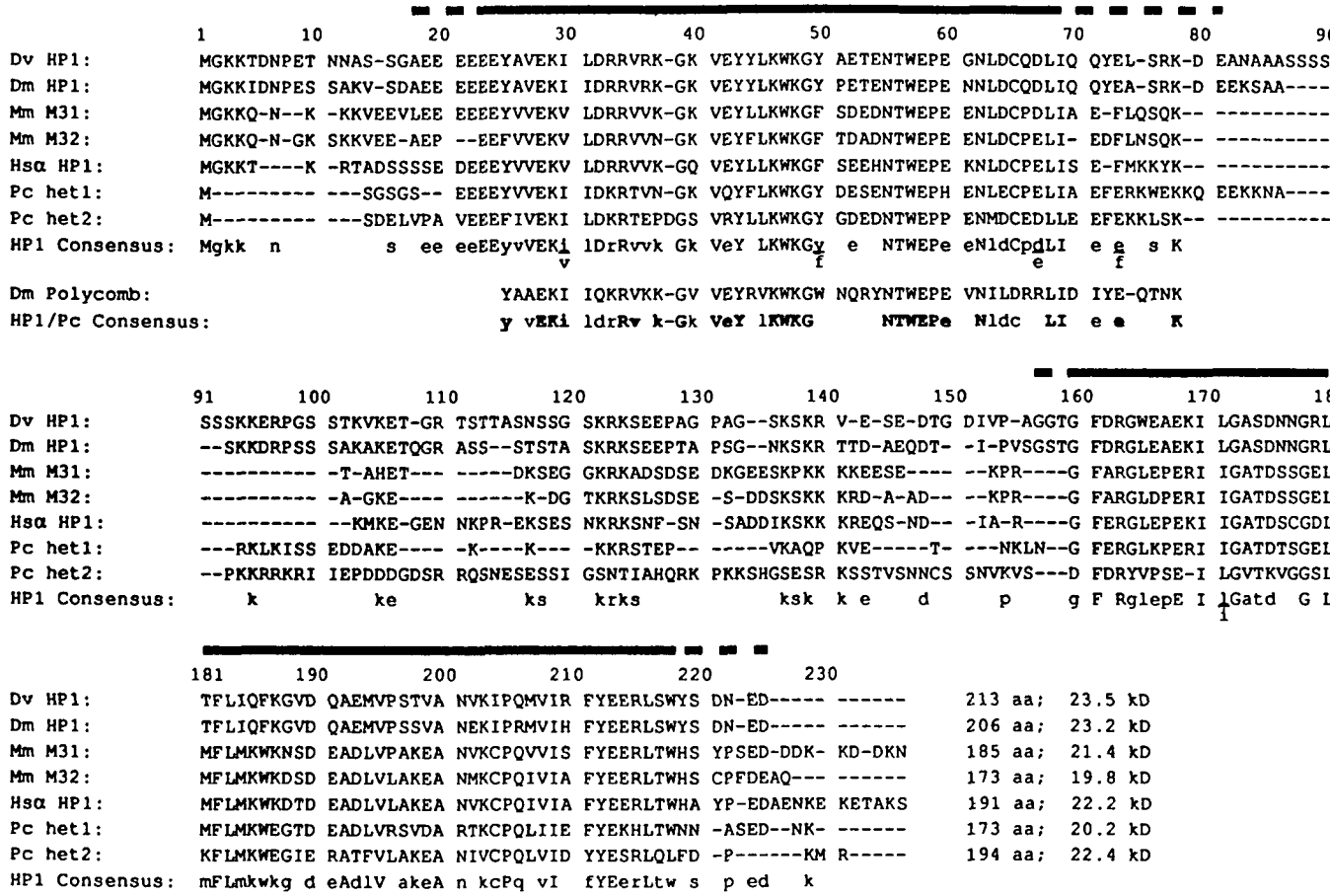

Figure 5. Comparisons of deduced amino acid sequences from the HP1 genes of *D. virilis* and *D. melanogaster*, the HP1-like genes of mouse, human, and mealybug, and the *Polycomb* gene of *D. melanogaster*. The deduced amino acid sequences of HP1 from *D. virilis* and *D. melanogaster* (12), and the HP1-like genes from mouse (31), human (32), and mealybug (33), are compared, along with a portion of the *Polycomb* (Pc) protein (35). Gaps in sequence alignments are indicated by dashes. An HP1 consensus sequence was deduced; uppercase letters indicate fully conserved amino acids, lowercase letters indicate conservation in 4 out of 7 sequences, and a doublet of lowercase letters indicates a choice of two amino acids at this location (where the distribution is 3/7 versus 4/7). The HP1/Pc consensus sequence was also deduced over the region of greatest identity between the two proteins. It extends 16 residues longer than the homologous domain discussed by Paro and Hogness (35). Bold uppercase letters indicate Pc residues identical to amino acids in the HP1 consensus, and lowercase letters indicate conservation in 5 out of 8 sequences (bold type representing those residues found in Pc). The dark bars indicate the two strongly homologous regions among all of the HP1-like proteins studied to date, and the dashed bars indicate the further extensions of these regions shared between the HP1 proteins of *D. melanogaster* and *D. virilis*.

characterized thus far, with conservative replacements accounting for another 20%. Using only the HP1 proteins of *D. virilis* and *D. melanogaster*, this region can be extended to 61 amino acids, which share a 93% identity.

This N-terminal conserved region overlaps a motif shared by HP1 and the *Polycomb* protein; this motif has been called the 'chromo domain' (chromatin modification organizer) (31). Because this designation implies a function not yet established, we prefer the term 'HP1/Pc box'. This region is a 37 amino acid domain in the *D. melanogaster* HP1 protein that is 65% identical with a domain of the *Polycomb* protein, with conservative replacements accounting for another 19% (35). More recently, this conserved domain was extended to 48 amino acids to take into account the analysis of Pc mutant alleles (42). We have extended the HP1/Pc box to 52 amino acids to account fully for all homology between the *Polycomb* protein and all of the HP1-like proteins studied thus far (Figure 5). The *Polycomb* protein, working with other gene products, serves as a repressor of some of the homeotic genes (43); it has been suggested that these proteins are responsible for the maintenance of the limited pattern of expression of these homeotic genes throughout development (35). The conservation of this HP1/Pc box suggests that HP1 and the *Polycomb* protein may use analogous mechanisms for the stable transmission of a determined state at the level of higher order chromatin structure.

The amino acid sequences of the HP1 proteins share other distinct features. All are very hydrophilic (33–40%) in nature, with a basic:acidic amino acid ratio of 1:1.2–1:1.3. These charged amino acids seem to alternate in blocks of basic residues and acidic residues along the entire protein (Figure 6); this may be important for the function of HP1. All of the HP1 proteins, but not the *Polycomb* protein, have a glutamic acid-rich region immediately upstream of the HP1/Pc box. One feature that is unique to the *D. virilis* protein is a stretch of seven serine residues immediately downstream of the HP1/Pc box.

Comparison of upstream and intron sequences

Interspecies comparisons of non-protein-coding sequences have aided in the identification of important cis-regulatory DNA sequences in several *Drosophila* genes (18,19,40,44–46).

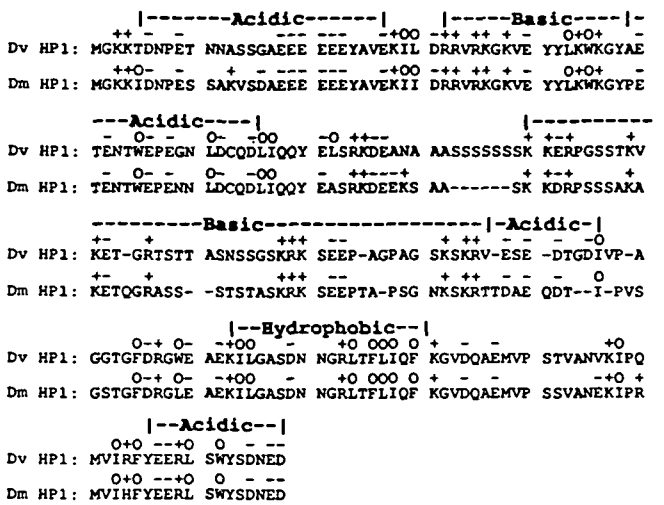

Figure 6. Comparison of the deduced amino acid sequences of the *Drosophila* HP1 proteins and their charge distributions. The amino acid sequences deduced from the *D. virilis* and *D. melanogaster* genes (12) are aligned. Minus signs above the amino acid sequence indicate acidic residues, plus signs indicate basic residues, and the circles represent highly hydrophobic residues. Charged amino acids alternate in blocks of acidic residues and blocks of basic residues along the entire protein, hydrophobic amino acids also appear to be grouped.

However, there is little sequence homology in the DNA upstream of the coding sequences of the HP1 genes from *D. virilis* and *D. melanogaster*. Although the nucleotides around the translational start codon and the first few amino acids of the open reading frame are conserved, there is little sequence homology in the non-coding sequences found immediately upstream of the transcriptional start sites, even at a gap weight of 1.0 and/or a length weight of 0.06.

Intron/exon splice sites for *D. virilis* HP1 were determined by comparison with the *D. melanogaster* gene and with the consensus splice signal sequence [(C/T)T(A/G)A(T/C)] for *Drosophila* genes (47). A putative splice signal found within the first intron is an exact fit to the consensus (CTAAT), while those within the second and third introns show a close resemblance (TTACT and GTAAT). The overall sequence conservation of the introns is very low; it can be estimated at 39%. The intron sequences and sizes have diverged substantially. The DNA sequences at half of the splice sites, however, are conserved, specifically the splice junctions of the first intron and the 5′ splice sites of the second and third introns (Figure 4). The lack of conservation of the 3′ sites of the second and third introns is surprising since splice site sequences are usually well conserved between genes of the two species (22,23). Sequence comparison with *D. melanogaster* intron/exon boundaries for the third intron in the coding sequence suggests that the *D. virilis* boundaries may be shifted by one codon (Figure 4). A similar shift in the boundaries of the first intron was reported in the initial comparison of the genes encoding HP1 and *Polycomb* (35).

Conclusions

We have shown that HP1 is a highly conserved protein, particularly in two regions near the N-terminal and the C-terminal ends. One can conjecture that many conserved motifs are necessary for the function of a structural heterochromatin protein such as HP1. Since neither HP1 (T. C. James, personal communication) nor the *Polycomb* protein (48) appears to bind to DNA directly, the HP1/Pc box is unlikely to be involved in protein-DNA interactions. Possible functions for either of the two conserved motifs in HP1 are as a nuclear-targeting signal, a domain necessary for the repressor activity of heterochromatin, or a protein-binding domain providing interaction with other heterochromatin-specific protein(s), possibly including a DNA-binding protein.

It has been suggested that 20–30 dominant suppressors of position-effect variegation exist (8), many of which could be structural proteins of heterochromatin. Two HP1-like proteins have been inferred for mouse (31), human (31–32), and mealy bug (33) from cDNA sequences. Probing a Southern blot of a restriction digest of genomic DNA from plants or animals using the HP1 sequence from the HP1/Pc box reveals multiple fragments (2–10) with homology (31). It is tempting to suggest that there might be other proteins containing this motif, perhaps involved in packaging other sets of genes. We are currently screening, using a variety of DNA techniques, to determine whether or not other proteins sharing this conserved region and common functions can be identified.

ACKNOWLEDGMENTS

We thank R.Blackman for the *D. virilis* genomic library, and the Elgin lab for critical review of this manuscript. This work was supported by NIH Postdoctoral Fellowship GM 13347 to R.F.C., and by NIH grant HD 23844 to S.C.R.E.

REFERENCES

1. Spofford, J. B. (1976) In Ashburner, M. and Novitski, E. (eds.), The Genetics and Biology of *Drosophila*, Academic Press, New York, pp. 955–1018.
2. Karpen, G. H. and Spradling, A. C. (1990) *Cell* **63,** 97–107.
3. Spradling, A. C. and Karpen, G. H. (1990) *Genetics* **126,** 779–784.
4. Henikoff, S. (1981) *Chromosoma* **83,** 381–393.
5. Rushlow, C. A. and Chovnick, A. (1984) *Genetics* **108,** 589–602.
6. Hayashi, S., Ruddell, A., Sinclair, D., and Grigliatti, T. (1990) *Chromosoma* **99,** 361–400.
7. Wustmann, G., Szidonya, J., Taubert, H. and Reuter, G. (1989) *Mol. Gen. Genet.* **217,** 520–527.
8. Locke, J., Kotarski, M. A. and Tartof, K. D. (1988) *Genetics* **120,** 181–198.
9. Grigliatti, T. (1991) *Methods in Cell Biology,* **Vol. 35,** pp. 588–627.
10. Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A. and Spierer, P. (1990) *Nature (London)* **344,** 219–223.
11. James, T. C. and Elgin, S. C. R. (1986) *Mol. Cell. Biol.* **6,** 3862–3872.
12. Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V. and Elgin, S. C. R. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9923–9927.
13. Eissenberg, J. C., Morris, G. D., Reuter, G. and Hartnett, T. (1992) *Genetics* **131,** 345–352.
14. James, T. C., Eissenberg, J. C., Craig, C., Dietrich, V., Hobson, A. and Elgin, S. C. R. (1989) *Eur. J. Cell Biol.* **50,** 170–180.
15. Sinclair, D. A. R., Mottus, R. C. and Grigliatti, T. A. (1983) *Mol. Gen. Genet.* **191,** 326–333.
16. Beverly, S. M. and Wilson, A. C. (1984) *J. Mol. Evol.* **21,** 1–13.
17. Moriyama, E. N. (1987) *Jpn. J. Genet.* **62,** 139–147.
18. Blackman, R. K. and Meselson, M. (1986) *J. Mol. Biol.* **188,** 499–515.
19. Kassis, J. A., Poole, S. J., Wright, D. K. and O'Farrell, P. H. (1986) *EMBO J.* **5,** 3583–3589.
20. Tautz, D., Tautz, C., Webb, D. A. and Dover, G. A. (1987) *J. Mol. Biol.* **195,** 525–542.
21. Martinez-Cruzado, J. C., Swimmer, C., Fenerjian, M. G. and Kafatos, F. C. (1988) *Genetics* **119,** 663–677.
22. Treier, M., Pfeifle, C. and Tautz, D. (1989) *EMBO J.* **8,** 1517–1525.
23. Michael, W. M., Bowtell, D. D. L. and Rubin, G. M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5351–5353.
24. Heberlein, U. and Rubin, G. M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5916–5920.
25. Feinberg, A. P. and Vogelstein, B. (1984) *Anal. Biochem.* **137,** 266–267.

26. Maniatis, T., Fritsch, E. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor.
27. Lis, J. T., Simon, J. A. and Sutton, C. A. (1983) *Cell* **35,** 403–410.
28. Wallrath, L. L., Burnett, J. B. and Friedman, T. B. (1990) *Mol. Cell. Biol.* **10,** 5114–5127.
29. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
30. Proudfoot, N. J. and Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
31. Singh, P. B., Miller, J. R., Pearce, J., Kothary, R., Burton, R. D., Paro, R., James, T. C. and Gaunt, S. J. (1991) *Nucleic Acids Res.* **19,** 789–794.
32. Saunders, W. S., Chue, C., Goebl, M., Craig, C., Clark, R. F., Powers, J., Eissenberg, J. C., Elgin, S. C. R., Rothfield, N. F., and Earnshaw, W. C. (1992), manuscript submitted.
33. Epstein, H., James, T. C., and Singh, P. B. (1992) *J. Cell Sci.* **101,** 463–474.
34. Gribskov, M. and Burgess, R. R. (1986) *Nucleic Acids Res.* **14,** 6745–6763.
35. Paro, R. and Hogness, D. S. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 263–267.
36. Hirsh, J., Morgan, B. A. and Scholnick, S. B. (1986) *Mol. Cell. Biol.* **6,** 4548–4557.
37. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988) *Science* **239,** 487–491.
38. Hultmark, D., Klemenz, R. and Gehring, W. J. (1986) *Cell* **44,** 429–438.
39. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–471.
40. Colot, H. V., Hall, J. C. and Rosbash, M. (1988) *EMBO J.* **7,** 3929–3937.
41. Wilde, C. D. and Akam, M. (1987) *EMBO J.* **6,** 1393–1401.
42. Messmer, S., Franke, A., and Paro, R. (1992) *Genes Dev.* **6,** 1241–1254.
43. Lewis, E. B. (1978) *Nature (London)* **276,** 565–570.
44. Bray, S. J. and Hirsh, J. (1986) *EMBO J.* **5,** 2305–2311.
45. Kassis, J. A., Desplan, C., Wright, D. K. and O'Farrell, P. H. (1989) *Mol. Cell. Biol.* **9,** 4304–4311.
46. Fortini, M. E. and Rubin, G. M. (1990) *Genes Dev.* **4,** 444–463.
47. Keller, E. B. and Noon, W. A. (1985) *Nucleic Acids Res.* **13,** 4971–4981.
48. Paro, R. (1990) *Trends Genet.* **16,** 416–421.