Nucleic Acids Research, 2000, Vol. 28, No. 18 3551–3557

Prod is a novel DNA-binding protein that binds to the 1.686 g/cm^3 10 bp satellite repeat of *Drosophila melanogaster*

Tibor Török*, Mátyás Gorjánácz, Peter J. Bryant[1] and István Kiss

Institute of Genetics, Biological Research Center of the Hungarian Academy of Sciences, Szeged H-6701, Hungary and [1]Developmental Biology Center, University of California, Irvine, CA 92717, USA

Received May 31, 2000; Revised and Accepted July 28, 2000

DDBJ/EMBL/GenBank accession no. U83596

ABSTRACT

The *proliferation disrupter* (*prod*) gene of *Drosophila melanogaster* encodes a novel protein associated with centromeric chromosomal regions that is required for chromatin condensation and cell viability. We have examined the binding of the Prod protein to DNA *in vitro*. Co-immunoprecipitation experiments demonstrate that Prod is a DNA-binding protein that specifically recognizes the 10 bp AGAATAACAT satellite repeat of *D.melanogaster*. Footprinting experiments show that the protein interacts with a 5–8 bp target sequence in each 10 bp repeat and suggest that it can mediate condensation of this satellite into a superhelix. Gel retardation experiments indicate that Prod does not have a well defined DNA-binding domain and it binds the satellite in a co-operative manner, probably forming Prod multimers. Since Prod localizes to both heterochromatin and euchromatin *in vivo*, we discuss the possibility that the ability of pre-existing euchromatic proteins to bind DNA in a co-operative manner, might be a prerequisite of satellite compaction and satellite amplification, thereby providing a basic factor in heterochromatin evolution.

INTRODUCTION

A significant fraction of the DNA in higher eukaryotes contains long arrays of highly repetitive simple sequences known as satellite DNAs, which form a compact type of chromatin known as constitutive or α-heterochromatin. Most of this chromatin is localized near centromeres and it differs markedly in many respects from the rest of the genome (1). The constitutive heterochromatin is almost devoid of genes, replicates late in S phase, remains tightly condensed throughout the cell cycle and can vary in amount between different tissues of the same individual (2). The most distinctive feature of heterochromatin is its tightly compacted structure. Although several proteins have been described that associate with heterochromatin and might play a role in its condensation (3–7) their mode of action remains largely unknown.

The proliferation disrupter protein (Prod) of *Drosophila melanogaster* has been shown to be required for the condensation of mitotic chromosomes and, in particular, for the condensation of heterochromatic regions located close to the centromeres (8). Prod is a non-histone chromosomal protein associated with over 400 euchromatic loci which also accumulates dramatically in the constitutive heterochromatin of the second and third chromosomes suggesting preferred association with the AATAACATAG satellite repeat, that is specific to these chromosomes (8). Immunostaining of *D.melanogaster* chromosomes have indeed shown that the Prod signal in heterochromatin colocalizes with fluorescent *in situ* hybridization signals of the AATAACATAG repeat satellite DNA probes (9); however, it remains unknown whether Prod binds DNA directly or binds other chromosomal proteins specific for this region. The heterochromatic accumulation of the Prod protein has been shown to be cell cycle dependent (9); it is present on the heterochromatin during mitosis, and in interphase periods near mitosis, however, during the lengthy interphase cycles of larval brain cells the protein gradually disappears from its heterochromatic sites and is probably shifted to euchromatin.

Loss of function mutations in the *prod* gene lead to abnormal centromere condensation, anaphase defects and cell death, all of which have been interpreted as resulting from the absence of Prod from its target sites in heterochromatin (8). It appears likely that the association of the Prod protein with chromosomes directly influences chromatin condensation and segregation. If so, it is important to understand the nature of its association with discrete chromosomal sites, and of equal interest is the mechanism whereby Prod may effect chromatin condensation. Here we provide evidence, from three independent assays *in vitro*, that Prod is a novel sequence-specific DNA-binding protein whose target sequence is contained in the 1.686 g/cm^3 satellite DNA repeat with the sequence AATAACATAG (10). This satellite, which we call Prodsat, represents ~2% of the genome of *D.melanogaster*, that is localized in a single, Hoechst bright, nearly 2 Mb block in the heterochromatin of chromosome 2 at h37 and a similarly sized block in the heterochromatin of chromosome 3 at h48 (11). We propose that Prod

*To whom correspondence should be addressed at present address: Institute of Molecular Pathology, Dr. Bohr gasse 7, 1030 Vienna, Austria. Tel: +43 1 797 30; Fax: +43 1 798 7153; Email: torok@nt.imp.univie.ac.at

alone might be capable of directly folding Prodsat arrays into a superhelix, a property which might allow it to maintain the compact state of the heterochromatin. The *prod* mutant phenotype suggests that this compaction is required for the proper segregation of mitotic chromosomes.

MATERIALS AND METHODS

Subcloning of fusion protein constructs

Each fusion construct is named after the predicted molecular weight of the protein fragment without the GST domain (Fig. 1A). GST–Prod (*P*) is a 1403 bp *Eco*RI fragment; GST–25 (*25*) is a 741 bp *Eco*RI–*Bam*HI fragment; GST–13 (*13*) is a 414 bp *Eco*RI–*Sau*3AI fragment and GST–15 (*15*) is a 661 bp *Bam*HI–*Xho*I fragment subcloned from the pBluescript vector containing the Prod cDNA into pGEX-4T-2 (Pharmacia Biotech). Constructs GST–28 (*28*), GST–19 (*19*), GST–16 (*16*), GST–14 (*14*) and GST–7 (*7*) were made by PCR amplification of the appropriate Prod cDNA fragments and subcloning into pGEX-4T vectors. All final constructs were verified by sequencing.

Induction and purification of fusion proteins

Bacteria carrying plasmids with GST–Prod fusion constructs were grown overnight at 30°C in NZCYM medium containing 100 µg/ml ampicillin (Amp). The following morning, the overnight culture was diluted 1:10 in fresh NZCYM/Amp medium and grown for 3 h at 30°C before addition of IPTG (0.1 mM final) and continued incubation for 2 h at 30°C. Bacteria were harvested by centrifugation (3 min at 3000 r.p.m., Sorvall RC5C plus, GSA rotor), washed in 1× PBS and pelleted again. Cells were resuspended in 1/20 culture volume of 1× PBS and affinity purification was performed using the Bulk GST Purification Modules of Pharmacia Biotech according to the manufacturer's instructions. Purified fusion proteins were desalted and concentrated using Centricon 10 cartridges (Amicon Inc.). For co-immunoprecipitation experiments, bacterial cells were pelleted and resuspended in 1/30 culture vol of RIPA buffer (150 mM NaCl, 1.0% NP-40, 0.5% DOC, 0.1% SDS, 50 mM Tris pH 8.0) containing 200 µg/ml PMSF and 1 µg/ml aprotinin, then sonicated on ice in four 30 s bursts. The cell lysate was incubated on ice for 30 min, then centrifuged at 15 000 r.p.m. (Eppendorf table top centrifuge) for 10 min at 4°C. An equal volume of glycerol was added to the supernatant and the protein extract was stored at –20°C. Extracts were tested for the presence of fusion proteins by western blotting (Fig. 1B).

Co-immunoprecipitation

DNA binding and immunoprecipitation of Prod–DNA complexes was essentially done as described by Desplan *et al.* (12). A collection of randomized DNA fragments was created by digesting λ phage DNA with *Taq*I or *Sau*3AI and mixing the restriction fragments (300 ng) before end-labelling with T4 polynucleotide kinase. For binding assays, labeled fragments (50 ng) were incubated with 200 µg bacterial total protein extract in 100 µl binding buffer (BB) for 30 min at 4°C with constant rocking. For the general DNA-binding experiments shown in Figure 2, BB contained 50 mM NaCl, 20 mM Tris–HCl pH 7.6, 0.25 mM EDTA, 1 mM DTT, 10% glycerol and 10 mM PMSF. For the sequence-specific binding experiments shown in Figure 3B, the NaCl concentration was raised to 170 mM and cold salmon sperm DNA was added to 100 µg/ml. The binding reaction was precleared as follows: 6 µl pre-immune rabbit serum was added to the extract and incubated for 30 min before precipitation with 20 µl 50% protein-A Sepharose for 1 h and centrifugation. Goat-anti-GST antibody (1 µl) (Pharmacia) was then added to the supernatant and incubation was continued for an additional 30 min before addition of 1.5 µl rabbit-anti-goat antibody. After a 30 min incubation, the antibody complex was precipitated with 20 µl 50% protein-A Sepharose suspension and centrifuged after an additional 1 h incubation period. The pellet was washed twice in BB, phenol-extracted and ethanol-precipitated using yeast tRNA as carrier. Final pellets were dissolved in water and applied on 5% polyacrylamide or agarose gels. Gels were dried and autoradiographed.

Gel retardation

Gel retardation or gel-shift assays were performed essentially as described by Mikami *et al.* (13). The labeled probe contained 4.5 repeats of the 10 bp (AATAACATAG) Prodsat DNA subcloned into pBluescript KS+ and was prepared by PCR amplification of a 232 bp fragment from the construct, using T7 and T3 universal primers, one of them had been previously end-labeled. The probe was purified on a polyacrylamide gel as described by Sambrook *et al.* (14). The binding reaction contained the DNA probe (5000 c.p.m.), 1–4 µg sonicated salmon sperm DNA as carrier, 20 mM Tris–HCl pH 8.8, 50 mM KCl, 1 mM DTT, 0.5 mM EDTA, 7% glycerol and purified fusion protein in a final volume of 20 µl. The binding reaction was incubated for 45 min at 37°C then loaded on a gel running at maximum current. One millimeter thick 4% polyacrylamide gels (40:1 acrylamide:bis-acrylamide) were used with 0.5× TBE as running buffer. Gels were pre-run at 100 V overnight at 4°C, followed by electrophoresis of the samples for 3 h at 200 V, 4°C. Gels were dried and exposed using a Phosphorimager 445SI (Molecular Dynamics). Radioactive signals were quantitated using the ImageQuant program of the Phosphorimager.

DNase I footprinting

DNase I footprinting assays were performed essentially as described by Martino-Catt and Kay (15), using the same end-labeled probe as in the gel retardation experiments. The binding reaction was also the same as in the gel-shift experiments, except that more probe (25 000 c.p.m.) was used, no carrier was added and the reaction volume was 30 µl. DNase I digestions were performed at room temperature for 2 min by adding 2 µl of 0.1 mg/ml DNase I (Sigma) dissolved in 15 mM $MgCl_2$. Digestions were stopped by adding 5 µl 0.5 M EDTA then 85 µl stop buffer containing 5 M urea, 0.36 M NaCl, 0.5% SDS, 10 mM Tris–HCl pH 8.0, 0.2 mg/ml yeast tRNA and 1 M ammonium acetate. Samples were extracted with phenol–chloroform, precipitated with 300 µl ethanol and centrifuged. DNA pellets were dissolved in 4 µl formamide loading buffer, boiled for 3 min and loaded on 6% sequencing gels. Sequencing reactions were performed with the Sequenase 2.0 Sequencing kit (USB) using the same end-labeled T3 or T7 primers as in the footprinting reaction.

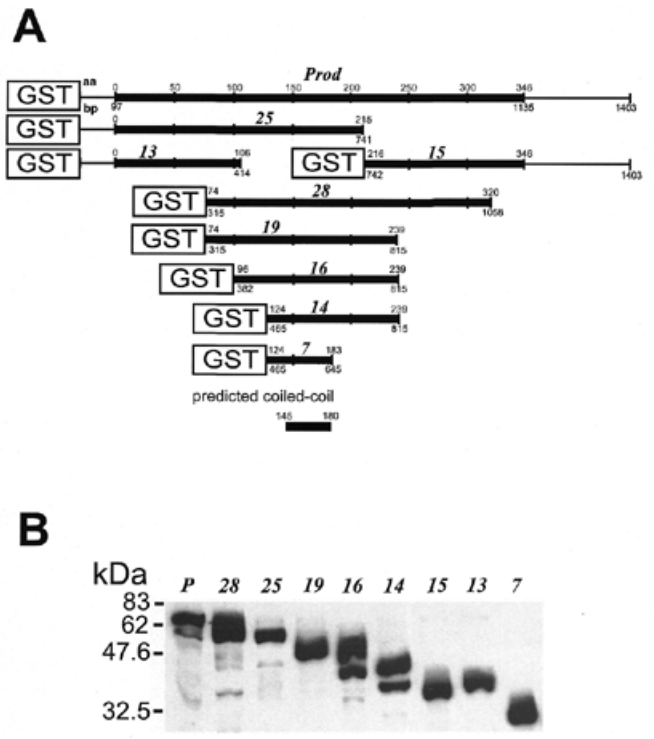

Figure 1. (**A**) The different GST–Prod fusion protein constructs used in our studies. Open rectangles, GST coding region; solid bars, Prod coding region; thin lines, non-coding regions of the Prod cDNA. Overlapping sections of the constructs are aligned above each other. Small numbers above the lines show amino acid (aa) positions in the Prod protein, numbers under the lines represent base pair (bp) positions in the prod cDNA. The localization of predicted coiled-coil is depicted at the bottom. (**B**) Western blot of the different purified GST–Prod fusion proteins illustrated in (A), revealed with an anti-GST antibody.

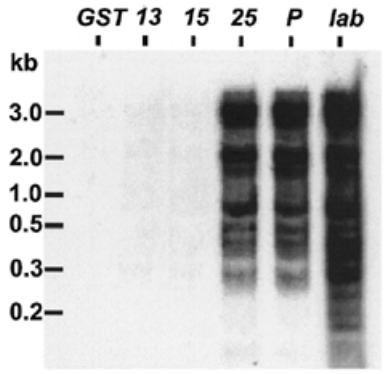

Figure 2. Co-immunoprecipitation with the different GST fusion constructs. Autoradiogram of labeled random λ DNA fragments co-immunoprecipitated with the GST–Prod fusion proteins designated above the lanes. *GST*, glutathione *S*-transferase protein without Prod; *lab*, labeled DNA before immunoprecipitation, constructs *13*, *15*, *25* and *P* (Prod) are described in Figure 1.

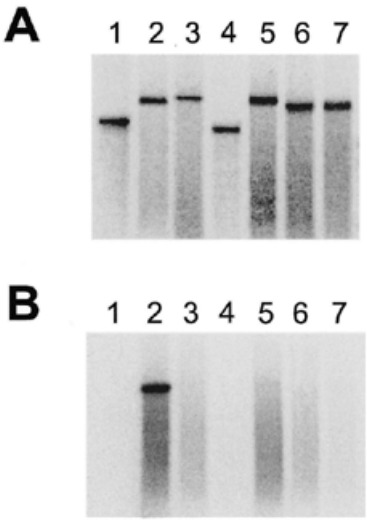

Figure 3. Co-immunoprecipitation of different satellite DNA clones. (**A**) Autoradiogram of end-labeled second and third chromosome-specific satellite DNA clones separated by agarose gel electrophoresis. Numbered lanes contain the following sequence repeats (11): AAGAC (lane 1); AATAACATAG (lane 2); 359 bp (1.688 g/cm^3) repeat (lane 3); AATAG (lane 4); AATAT (lane 5); AAGAGAG (lane 6); AAGAG (lane 7). (**B**) Autoradiogram of the same satellite clones after co-immunoprecipitation with the GST–Prod fusion protein under stringent binding conditions (170 mM NaCl, 100 µg competitor DNA).

RESULTS

Prod is a DNA-binding protein

The Prod protein does not share any homology with known proteins in the existing databases, rendering predictions of a conserved biological function difficult to make. Although coil prediction (16) suggests the presence of a coiled-coil domain in the middle of Prod (amino acids 145–180; Fig. 1), a motif known to be involved in protein–protein interactions (17), we first wanted to test whether Prod is able to bind DNA directly. Therefore we performed co-immunoprecipitation experiments based on the approach of Desplan *et al.* (12) who identified the functional homeobox within the *engrailed* gene. We constructed GST–Prod fusion proteins containing the full-length protein or different Prod sub-fragments (Fig. 1A) and expressed them in *Escherichia coli.* Radiolabeled random DNA fragments were mixed with bacterial protein extracts containing the fusion proteins and the fusion proteins were immunoprecipitated from this solution using purified anti-GST antibody. If Prod binds DNA, labeled DNA fragments should co-precipitate with the fusion protein at low salt concentration (50 mM NaCl). As shown in Figure 2, both the full-length *P* construct and construct *25*, containing the N-terminal two-thirds of the protein, co-precipitated DNA fragments with similar affinities. In contrast, constructs *13*, containing the N-terminal one-third of Prod and construct *15*, containing the C-terminal one-third of the protein, as well as the GST protein alone, did not co-precipitate labeled DNA. These results indicate that the first 215 amino acids of Prod contain a DNA-binding domain probably around the middle third of the full-length protein.

The 10 bp Prodsat provides a sequence-specific Prod binding site

As the Prod protein shows a striking accumulation in the heterochromatin of the second and third chromosomes *in vivo* (8), it seemed reasonable to assume that the DNA-binding is

sequence specific and the target sequence is the AATAA-CATAG repeat specific for these chromosomes. Sequence-specific DNA binding is resistant to relatively high salt concentrations and to competition by non-specific carrier DNA (12). As most of the major satellite sequences of *D.melanogaster* have been cloned and mapped (11), we were able to test all known satellite sequences that are present on the second and third chromosomes in a stringent Prod-binding assay to prove the above assumption and to exclude the binding to other repeats.

Plasmids containing seven different satellite repeats were linearized, end-labeled and co-immunoprecipitated with the GST–Prod fusion protein. As shown in Figure 3, when very high salt and carrier DNA concentrations were combined (170 mM NaCl, 100 µg/ml competitor DNA), only one of the satellites, the 10 bp AATAACATAG repeat, co-precipitated with the fusion protein, suggesting that this is the only satellite sequence that provides a high affinity binding site for the Prod protein. For this reason, we call this satellite Prodsat.

To test which nucleotides of the Prodsat repeat are in direct contact with the Prod protein, we performed DNase I footprinting experiments (15). A 232 bp pBluescript fragment containing 4.5 Prodsat repeats was incubated with the full-length fusion protein or the GST construct *25*. As shown in Figure 4, both proteins provided a similar protection against DNase I digestion. The footprints obtained clearly follow the periodicity of the DNA sequence, which is definitely recognized by Prod as repeated units of AGAATAACAT. This periodicity suggests that each Prodsat repeat unit is recognized and bound by one Prod protein unit (Prod monomer, dimer or other). Under these conditions, no sequences were found to be protected outside of the Prodsat repeats present in the DNA fragment tested and not all of the Prodsat bases proved to be equally protected. Every first A of the 10 bp sequence was either not or only very weakly protected and the C was not protected. In contrast, the ATAA unit was well protected, as was the A following the C. We could not clearly assess the protection of the three remaining bases (GA and the last T) because they are also poorly cleaved by DNase I in the control lane. Based on these results, we can assign the region specifically contacted by Prod from a minimum of 5 bp (ATAANA) to a maximum of 8 bp (GAATAANAT) in the Prodsat repeat.

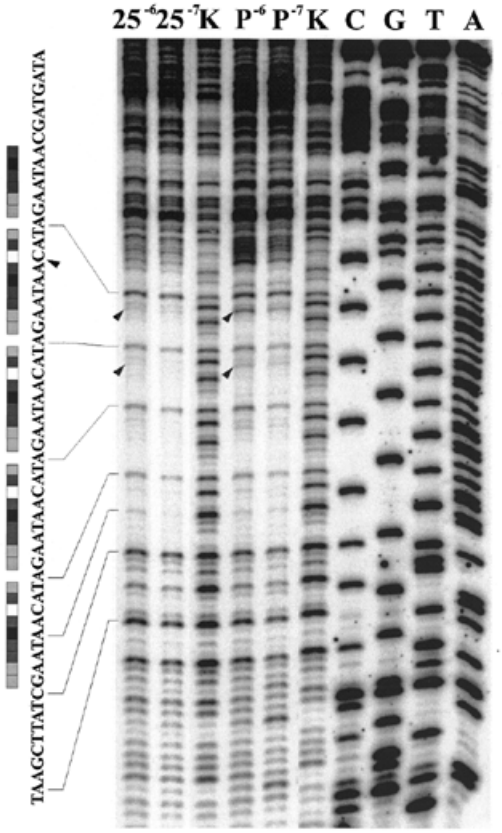

Figure 4. DNase I footprint of the Prod protein. Lanes labeled C, G, T and A are sequencing reactions, the sequence is shown on the left of the gel. Lane K is a control where no protein was added to the DNase I reaction, lane P contained the GST–Prod fusion protein and lanes labeled 25 contained the GST–25 fusion protein (Fig. 1A). Superscripts denote the protein concentration (M). Open boxes, unprotected bases; dark boxes, well protected bases; gray boxes, bases where the protection is uncertain. The arrow points at the C residue that is poorly cleaved in the control lane but well cleaved in the presence of Prod.

A large part of the protein is required for DNA binding

Next we tried to identify the DNA-binding domain, the smallest part of the protein capable of binding DNA with an affinity and a specificity comparable to that of the full-length protein. Based on the results of the co-immunoprecipitation experiments shown in Figure 2, the DNA-binding domain is contained within the first 215 amino acids of the protein, probably around the middle of Prod, since neither the N-terminal nor the C-terminal third of the protein shows DNA binding activity in itself.

In order to further localize the DNA-binding domain, we made a series of fusion constructs encoding decreasing lengths of Prod sequence around the central part of the protein (Fig. 1) and tested their ability to bind DNA in a gel-shift assay (13). In the gel-shift experiments we used the same 232 bp fragment containing 4.5 repeats of Prodsat as in the footprinting experiments shown above. We have separately tested a wide variety of binding conditions (salt, carrier and protein concentrations) for each construct in order to recognize if any remnant of DNA-binding ability could have been retained. Figure 5 shows that, out of the GST-fusion constructs tested, only four were capable of shifting the Prodsat-containing DNA fragment (data for *14* and *7*, that also did not bind DNA, are not shown). While the smallest of these, GST–19, clearly retained DNA-binding activity, the resulting weaker shift shown in Figure 5 (lane marked 19) could only be obtained if 50 times more protein was used compared to the other constructs (lanes marked P, 28 and 25), suggesting that GST–19 has a much lower affinity for Prodsat sequences. It is also notable that with construct *28* binding could be obtained only if the carrier DNA concentration was decreased. This suggests a decreased sequence specificity of *28* compared to that of *P* and *25*. Taken together, our results indicate that the smallest peptide that can bind Prodsat with an affinity comparable to that of the full-length protein is *25*.

Prod binds DNA in a co-operative manner

In order to provide a more quantitative description of the Prod–DNA interaction, we performed gel-shift titration experiments (18), in which increasing amounts of fusion peptide (*P*, *25* or *28*) are added to a constant and small amount of labeled target

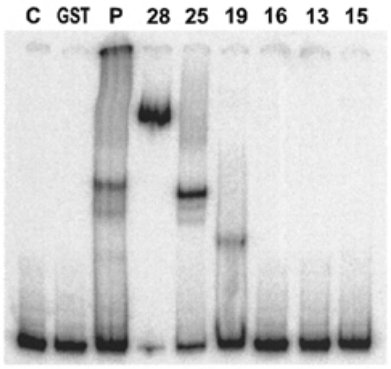

Figure 5. Capability of the different Prod fusion constructs to bind DNA. Autoradiogram of a gel-shift made with the GST–Prod fusion constructs using the 4.5× Prodsat probe. The sample in lane C corresponds to the free DNA probe (control lane, no protein added). Samples in the other lanes contained the fusion protein designated above each lane (Fig. 1A). A wide variety of different binding conditions had been tested before for each construct (also see text) and this figure shows only the optimized binding conditions achieved. Binding reactions with *GST*, *28*, *19*, *16*, *13* and *15* contained 1 µg carrier DNA while reactions with *P* and *25* contained 3 µg carrier DNA for optimal results. Lane P contains 0.5 µg protein and the relative amounts of proteins added to the other reactions were as follows: *P*:1, *25*:2, *28*:4, *19*:100, *16*:200, *13*:200, *15*:200 and *GST*:200.

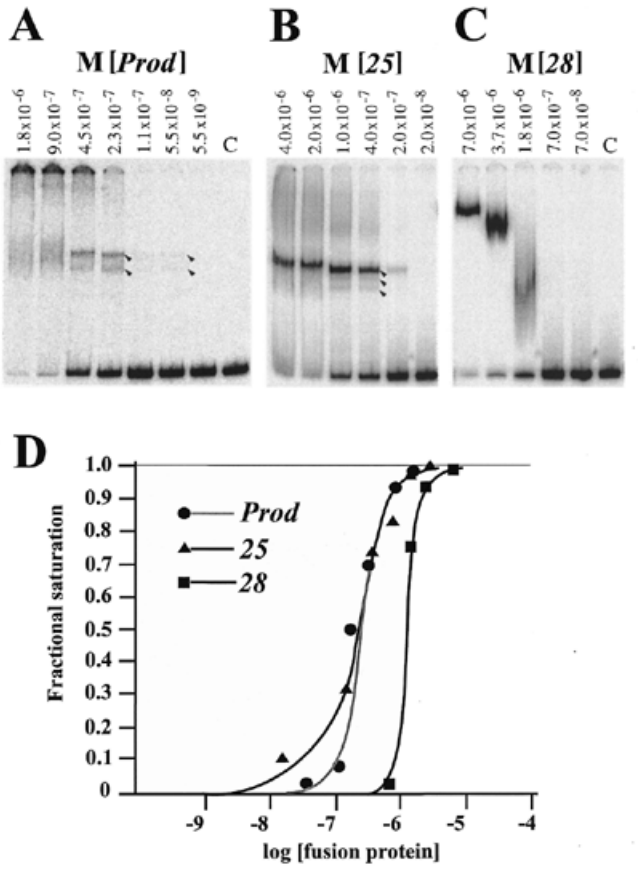

Figure 6. Gel-shift titration of the GST–Prod fusion proteins. (**A**, **B** and **C**) Gel-shifts of the 4.5× Prodsat probe by the GST–P, GST–25 and GST–28 proteins, respectively. Protein concentrations (M) are indicated above each lane. Arrows, predominant and minor faster-migrating complexes formed at low protein concentrations. Note that the reactions shown for the *P* and *25* contained 3 µg carrier DNA, while all *28* reactions contained 1 µg carrier DNA only. (**D**) The titration curves extracted from Phosphorimager quantitations of the radioactivity present as free DNA in the gels of (A), (B) and (C) (18). The fractional saturation (the ratio of DNA molecules complexed with the protein) is plotted as a function of the logarithm of the protein concentration.

DNA, until all the DNA goes into complex. In our first gel-shifts we used a target DNA that contained 12 Prodsat repeat units. However the 12× Prodsat probe constantly resulted in a sudden shift of the radioactivity at some point from the control lane to the loading well by just a 3-fold increase in the protein concentration, without having any intermediate complex size, with full-length Prod and *25* (data not shown). This was a clear indication of extremely co-operative DNA-binding (19) in case the large complexes in the loading well did not represent non-specific protein–DNA aggregates. Since these large complexes disappeared from the loading well if cold Prodsat competitor DNA (identical to the labeled probe) was added to the same reaction, they must not have represented non-specific aggregates, but huge Prod–Prodsat complexes that could not enter the gel because of their size. Nevertheless, we decided to decrease the Prodsat repeat number for having smaller complexes that can enter the gel and produce more convincing results, but are still capable of demonstrating co-operativity. For this reason all the shown gel-shift experiments were performed with a target DNA containing 4.5 Prodsat repeats only.

The behavior of *28* markedly differs from that of *25* and *P* in the gel-shift titration assay (Fig. 6A, B and C). An increasing amount of *28* resulted in a gradual decrease in the mobility of the protein–DNA complexes, which appeared to be stabilized at higher protein concentrations (Fig. 6C). In contrast, the size of the *P* and *25* (Fig. 6A and B, respectively) complexes with DNA remained constant throughout a wide range of protein concentrations. Although faster-migrating *P* and *25* complexes can also be seen as faint bands at low protein concentrations (Fig. 6A and B arrows), the apparent size of the predominant complex does not change and the faint bands disappear at higher concentrations. At high concentrations of *P* and *25*, the stability of the predominant complexes becomes less stable as suggested by their smeary appearance and at the same time very large complexes start to accumulate in the wells [Fig. 6A and B, lanes 1–3 and lane 1 (reading from left to right), respectively]. However, these very low-mobility complexes also shift back if cold Prodsat competitor DNA (identical to the labeled probe) is added to the same reaction indicating that they do not correspond to non-specific aggregates but rather represent binding of the protein to non-specific sequences outside of the Prodsat repeat within the probe if the protein concentration exceeds a certain threshold value (data not shown).

It is important to remember that the interaction of peptide *28* with DNA is much more sensitive to competition by cold carrier DNA than that of *P* or *25*. While complexes of *P* and *25* are stable in the presence of 6 µg carrier DNA, complexes of *28* are stable up to a maximum of 1 µg competitor DNA (Fig. 6 legend). This different behavior of *28* from *P* and *25* is even more obvious in footprinting experiments, where *28* does not provide a detectable protection against DNase I digestion at any of the concentrations tested (data not shown). This indicates that the *28* construct has lost most of the sequence

specificity that is seen in the full-length protein while it is retained in construct *25*. This explains the distinct *28* gel-shift pattern.

DISCUSSION

We have shown that the Prod protein of *D.melanogaster* is a DNA-binding protein that recognizes the 10 bp AGAAT-AACAT satellite repeat with a high affinity *in vitro*, and follows the periodicity of the DNA multimer. Prod does not have a defined DNA-binding domain since the major part of the protein is necessary for binding to happen. Removing smaller parts of the protein significantly alters the binding behavior; removal of the N-terminal 74 amino acids eliminates sequence specificity, while removal of more than one-third of the C-terminus eliminates binding. DNA-binding is highly co-operative, suggesting that Prod molecules not only bind to DNA, but also to each other, most probably via the coiled-coil domain present in the middle of the protein.

The co-operativity of Prod binding

The results with 12× Prodsat DNA probe were a very strong indication of co-operative DNA-binding since the binding of a protein monomer to the probe would not have been expected to form such large complexes that could not enter the gel (19). Co-operative binding is characterized by a steep sigmoidal titration curve when graphically illustrated (18) and indeed, the Prod titration curve is very steep even with the decreased (4.5) repeat number, the lowest and highest regions of the curve lying within a 10-fold range of protein concentration (Fig. 6D). The target DNA used in the shown gel-shift titration experiments contains 4.5 Prodsat tandem repeats. If Prod binding to each repeat were independent of binding to other repeats, titration of the DNA fragment with increasing protein concentrations would have been expected to produce a mobility shift in discrete steps, each step corresponding to an occupancy of a new binding site (19). This is clearly not the case for the full-length Prod and *25* constructs, as the mobility of the major complex formed at the lowest protein concentration remains unchanged and only increases in yield at higher protein concentrations (Fig. 6A and B). The simplest explanation for this observation is co-operative DNA binding, where the Prod–DNA complexes we observe through a wide range of protein concentrations, always contain the same number of protein molecules because the first stable complex to form at low protein concentrations is already fully loaded with proteins in a conformation stabilized by the co-operation of protein molecules.

Construct *28* also shows the same steep sigmoidal titration curve (Fig. 6D) suggesting that molecules of *28* also co-operate with each other in the course of DNA-binding. However, as mentioned above, *28* binding has lost most of the sequence specificity. Thus, the titration pattern of *28* complexes can be explained in the following way: multimers of Prod and *25* are restrained by the limited number of Prodsat binding sites on the probe and can only extend beyond these sites if the protein concentration exceeds a saturating threshold value; *28* multimers, in contrast, cannot effectively distinguish Prodsat sequences from vector sequences on the probe and gradually extend beyond these sites. This behavior allows the complex size to increase with increasing protein concentration. This saturating threshold value for *P* seems to be at lower protein concentration than for *25*, based on the earlier appearance of large complexes in the loading well. This can be due to the higher co-operativity of full-length Prod molecules than *25* molecules. As the *25* titration curve is also less steep than that of *P* this suggests that the C-terminal third of Prod that is missing from *25* also contains sequences contributing to co-operativity.

We were unable to define a small DNA-binding domain in Prod as smaller constructs were also altered in their binding behavior. It seems likely that the domain that attaches to DNA is not fully functional in itself and other domains of the protein are also indispensible for binding. The fact that an extensive domain analysis yielded solely co-operatively binding Prod fragments indicates that co-operativity is the prerequisite of DNA-binding. The coiled-coil domain—a potential protein binding domain via which Prod molecules might bind to each other—is the primary candidate to mediate co-operative interactions (17) and, indeed, each construct showing DNA-binding retained this domain. Results with construct *28* indicate that the N-terminal 74 amino acids are indispensable for sequence recognition.

Prod may directly condense the Prodsat repeat

There are more independent data, listed below, that support the above statement.

(i) The regularly spaced DNase I sensitive sites revealed by the Prod footprint are very reminiscent of those periodic footprint patterns described for one group of multimeric complex dsDNA-binding proteins (MC-DBPs; 20). In complexes with all members of this group of proteins the DNA forms a superhelix around the multimeric protein core, resulting in the compaction of the DNA strand (21–24). Based on its footprint pattern and co-operative DNA binding, Prod seems to be a new member of the MC-DBP group and may also compact the DNA.

(ii) Sequence multimers containing dA·dT tracts form an intrinsically curved DNA structure if the tracts are in phase with the helical screw (i.e. the repeat unit is 10 bp) because the individual curves created by each tract are summarized coherently in the multimer (25,26). Of the different sequence motifs studied by Koo *et al.* (27), the AATAA motif was shown to cause significant DNA curvature. Since AATAA is the central motif of the 10 bp Prodsat, this bending locus can promote the formation of a superhelical structure already in the absence of proteins. This bent Prodsat DNA might be stabilized and/or further curved by the Prod multimer since AT-rich DNA regions are also known to be very flexible (28). DNA bending is also inferred from the enhanced cleavage at base C in the DNA–Prod complex on the footprint (Fig. 4 arrows), because the change of DNase I sensitivity at base C can be the consequence of an alteration in the width of the minor groove of which DNase I is most sensitive (29). These arguments lead to the suggestion that the Prod protein might be able to condense the Prodsat repeat in the absence of other chromosomal proteins.

(iii) The above interpretation is in full agreement with the conclusions of Platero *et al.* (9) who, based on the Prod immunostaining pattern of *D.melanogaster, Drosophila simulans* and *Drosophila mauritiana*, have suggested that the only heterochromatic function of Prod would be to appropriately package the Prodsat repeat by a self-assembly

mechanism during mitosis, because biased composition of repeated sequences might hinder normal condensation processes.

(iv) Recently the GAGA factor, another non-histone chromosomal protein that, like Prod, binds distinct satellite repeats (30), has been shown to bind to its natural multiple euchromatic binding sites as a multimer, in a co-operative manner (19). The DNA in these complexes is wrapped around the surface of the GAGA multimer, while histone–DNA contacts are severely compromised. It is very likely that GAGA binding to its heterochromatic 'multiple binding sites' happens with the same mechanism.

The remarkable similarities between the Prod and GAGA factor heterochromatin localization, binding dynamics and mutant phenotypes have been extensively discussed (8,9,31), now we can further extend these similarities with their co-operative DNA-binding properties.

Co-operative binding might be a general requirement for satellite compaction

Based on their similar heterochromatin binding dynamics during the cell cycle in *D.melanogaster* and the absence of Prod and GAGA-factor as well as their heterochromatic target sequences from *D.simulans* and *D.mauritiana*, the closest relatives of *D.melanogaster*, Csink and Henikoff (31) proposed the 'mitotic protein borrowing model' to explain the striking variability of the eukaryotic satellite sequences at the evolutionary level. According to this model the amplification of Prodsat in *D.melanogaster* would reflect a very recent evolutionary event, possibly selected for by the pre-existence of Prod in the nucleus. It follows from the above that the primary function of this protein must have been euchromatic, but it might have acquired an additional heterochromatic function based merely on its ability to bind and compact the Prodsat repeated sequence.

Satellite domains usually appear in large blocks and are probably derived from the amplification of a simple sequence (32). If the co-evolution of Prod and Prodsat as well as GAGA factor and AG repeats illustrate a general feature of heterochromatin evolution, then simple DNA sequences would be selected for amplification into satellites if they fulfilled two basic conditions: (i) the ability to assume a bent structure and (ii) the pre-existence of DNA-binding protein(s) that can organize/stabilize the bent structure into compacted heterochromatin. Intrinsic curvature is a general characteristic of all known satellite DNAs (33,34) in agreement with the first condition of this model. Based on the results with the GAGA factor and Prod, the ability of co-operative binding might be a second general feature of those proteins that are able to compact long tandem repeats.

ACKNOWLEDGEMENTS

We thank Allan Lohe for providing us with satellite DNA clones. We are grateful to Emmanuel Kas for critical reading of the manuscript and helpful comments. This research was supported in part by Howard Hughes Medical Institute International Program, grant no. 75195-542201 and OTKA grant T 16693 to I.K.

REFERENCES

1. Brutlag,D.L. (1980) *Annu. Rev. Genet.*, **14**, 121–144.
2. Zhimulev,I.F. (1998) *Adv. Genet.*, **37**, 1–566.
3. Bhat,K.M., Farkas,G., Karch,F., Gyurkovics,H., Gausz,J. and Schedl,P. (1996) *Development*, **122**, 1113–1124.
4. Kellum,R. and Alberts,B.M. (1995) *J. Cell Sci.*, **108**, 1419–1431.
5. Levinger,L. and Varshavsky,A. (1982.) *Proc. Natl Acad. Sci. USA*, **79**, 7152–7156.
6. Radic,M.Z., Saghbini,M., Elton,T.S., Reeves,R. and Hamkalo,B.A. (1992) *Chromosoma*, **101**, 602–608.
7. Sugimoto,K., Hagishita,Y. and Himeno,M. (1994) *J. Biol. Chem.*, **269**, 24271–24276.
8. Török,T., Harvie,P.D., Buratovich,M. and Bryant,P.J. (1997) *Genes Dev.*, **11**, 213–225.
9. Platero,J.S., Csink,A.K., Quintanilla,A. and Henikoff,S. (1998) *J. Cell Biol.*, **140**, 1297–1306.
10. Lohe,A.R. and Brutlag,D.L. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 696–700.
11. Lohe,A.R., Hilliker,A.J. and Roberts,P.A. (1993) *Genetics*, **134**, 1149–1174.
12. Desplan,C., Theis,J. and O'Farrell,P.H. (1985) *Nature*, **318**, 630–635.
13. Mikami,K., Takase,H. and Iwabuchi,M. (1994) In *Plant Molecular Biology Manual*. Kluwer Academic Publishers, Belgium, pp. 1–14.
14. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. Martino-Catt,S.J. and Kay,S.A. (1994) In *Plant Molecular Biology Manual*. Kluwer Academic Publishers, Belgium, pp. 1–13.
16. Lupas,A., Van Dyke,M. and Stock,J. (1991) *Science*, **252**, 1162–1164.
17. Lupas,A. (1996) *Trends Biochem. Sci.*, **10**, 375–382.
18. Carey,J. (1991) *Methods Enzymol.*, **208**, 103–117.
19. Katsani,K.R., Hajibagheri,M.A.N. and Verrijzer,C.P. (1999) *EMBO J.*, **18**, 698–708.
20. Serrano,M., Salas,M. and Hermoso,J.M. (1993) *Trends Biochem. Sci.*, **18**, 202–206.
21. Drlica,K. and Rouviere-Yaniv,J. (1987) *Microbiol. Rev.*, **51**, 301–319.
22. Musgrave,D.R., Sandman,K.M. and Reeve,J.N. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10397–10401.
23. Serrano,M., Gutierrez,C., Salas,M. and Hermoso,J.M. (1993) *J. Mol. Biol.*, **230**, 248–259.
24. Ziegelin,G., Pansegrau,W., Lurz,R. and Lanka,E. (1992) *J. Biol. Chem.*, **267**, 17279–17286.
25. Hagerman,P.J. (1985) *Biochemistry*, **24**, 7033–7037.
26. Koo,H.S. and Crothers,D.M. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 1763–1767.
27. Koo,H.S., Wu,H.M. and Crothers,D.M. (1986) *Nature*, **320**, 501–506.
28. Akiyama,T. and Hogan,M.E. (1997) *Biochemistry*, **36**, 2307–2315.
29. Drew,H.R. and Travers,A.A. (1984) *Cell*, **37**, 491–502.
30. Raff,J.W., Kellum,R. and Alberts,B. (1994) *EMBO J.*, **13**, 5977–5983.
31. Csink,A.K. and Henikoff,S. (1998) *Trends Genet.*, **14**, 200–204.
32. Southern,E.M. (1970) *Nature*, **227**, 794–798.
33. Fitzgerald,D.J., Dryden,G.L., Bronson,E.C., Williams,J.S. and Anderson,J.N. (1994) *J. Biol. Chem.*, **269**, 21303–21314.
34. Martinez-Balbas,A., Rodriguez-Campos,A., Garcia-Ramirez,M., Sainz,J., Carrera,P., Aymami,J. and Azorin,F. (1990) *Biochemistry*, **29**, 2342–2348.

APPENDIX

Prod sequence data are available from DDBJ/EMBL/GenBank under accession number U83596. The sequence originally published in Török *et al.* (8) has been corrected. Due to a resequenced GC-compression at the end of the original sequence, the Prod protein became 46 amino acids longer. This does not affect the validity of previously published sequences and data.