The closely related *Drosophila* *sry* β and *sry* δ zinc finger proteins show differential embryonic expression and distinct patterns of binding sites on polytene chromosomes

FRANÇOIS PAYRE, STÉPHANE NOSELLI, VALÉRIE LEFRÈRE and ALAIN VINCENT

Centre De Recherche de Biochimie et de Genetique Cellulaires du C.N.R.S., 118, route de Narbonne, 31062 Toulouse, Cedex, France

Summary

Serendipity (sry) beta (β) and *delta* (δ) are two finger protein genes resulting from a duplication event. Comparison of their respective protein products shows interspersed blocks of conserved and divergent amino-acid sequences. The most extensively conserved region corresponds to the predicted DNA-binding domain which includes 6 contiguous fingers; no significant sequence conservation is found upstream and downstream of the protein-coding region. We have analysed the evolutionary divergence of the *sry* β and δ proteins on two separate levels, their embryonic pattern of expression and their DNA-binding properties *in vitro* and *in vivo*. By using specific antibodies and transformant lines containing β -galactosidase fusion genes, we show that the *sry* β and *sry* δ proteins are maternally inherited and present in embryonic nuclei at the onset of zygotic transcription, suggesting that they are transcription factors involved in

this process. Zygotic synthesis of the *sry* β protein starts during nuclear division cycles 12–13, prior to cellularisation of the blastoderm, while the zygotic *sry* δ protein is not detectable before germ band extension (stage 10 embryos). Contrary to *sry* δ , the zygotic *sry* β protein constitutes only a minor fraction of the total embryonic protein. The *sry* β and δ proteins made in *E. coli* bind to DNA, with partly overlapping specificities. Their *in vivo* patterns of binding to DNA, visualised by immunostaining polytene chromosomes, differ both in the number and position of their binding sites. Thus changes in expression pattern and DNA-binding specificity have contributed to the evolution of the *sry* β and δ genes.

Key words: *Drosophila*, embryogenesis, finger protein, polytene chromosomes, serendipity genes.

Introduction

Identification and characterization of an increasing number of proteins regulating gene expression in eucaryotes has led to the discovery that transcription factors make up a remarkably diverse group of proteins constructed from a variety of combinatorially arranged functional domains. Consequently, families of eucaryotic regulators with related DNA-binding specificities and varying levels of functional homology could have been created by gene duplication and subsequent divergence during evolution (reviews by Ptashne, 1988; Mitchell and Tjian, 1989).

The Cys2/His2 zinc finger first identified in the *Xenopus* transcription factor *TFIIIA* (Brown *et al.* 1985; Miller *et al.* 1985) is one well-characterized DNA-binding motif that defines large multigene families, particularly in vertebrates (Bellefroid *et al.* 1989; Nietfeld *et al.* 1989). Characteristic of this finger is its tandem repetition occurring a variable number of times per protein (reviews by Evans and Hollenberg, 1988; Payre and Vincent, 1988). The existence of subfamilies

of Cys2/His2 finger proteins, characterised by the presence, at their *N*-terminal end, of common conserved modules unrelated to the finger, has also been shown in *Xenopus* (Knöchel *et al.* 1989). Furthermore, comparison of the gap segmentation gene *hunchback* between two *Drosophila* species has revealed that divergence of expression patterns may also contribute to the functional evolution of Cys2/His2 finger protein genes (Treier *et al.* 1989).

To study structural relative to functional parameters contributing to the evolutionary divergence of Cys2/His2 finger proteins, we started a few years ago the characterisation of the *Drosophila* *sry* β and *sry* δ genes. These two finger protein genes map at a single locus and probably result from a duplication of an ancestral gene containing at least six fingers (Vincent *et al.* 1985). Mutants for *sry* δ , which display a late embryonic lethal phenotype, have been isolated. These mutants can be rescued by transformation with a DNA fragment that contains the *sry* δ but not the *sry* β gene (Crozatier *et al.* manuscript in preparation). Therefore, it is likely that the *sry* β and *sry* δ genes are functionally distinct. The

physical linkage and evolutionary conservation of the *sry* β and *sry* δ coding sequences in distant *Drosophila* species (Aguadé, 1988; Vincent *et al.* 1988 and unpublished data) confirms the value of the *sry* β and δ genes as a potential model for the study of the functional evolution of finger proteins relative to their DNA-binding specificities.

We have recently shown that the *sry* δ protein is maternally inherited by the embryo and present in nuclei prior to the onset of zygotic transcription. Later in embryogenesis and up to the adult stage, the *sry* δ protein is present in nuclei of transcriptionally active cells, suggesting that it is a transcription factor with a role in zygotic activation and maintenance of expression of general cellular functions (Payre *et al.* 1989).

In this paper, we have determined the pattern of embryonic expression of *sry* β in order to compare it to that of *sry* δ , by using immunostaining and transgenic approaches that make it possible to distinguish between maternal and zygotic contributions to these patterns. We then show that both proteins made in *E. coli* display partly overlapping *in vitro* DNA-binding specificities and *in vivo* patterns of binding sites on polytene chromosomes.

Materials and methods

Plasmid constructions

(a) pSBL1 construct

A *sry* β DNA fragment bordered by a *Bam*H1 site 458 bp upstream of the *sry* β transcription start site, and a *Sal*I site created in the *sry* β coding region at codon position 340, after the 6th finger (Vincent *et al.* 1985), was fused in frame to *E. coli lacZ* gene using the pSDL1 plasmid (Payre *et al.* 1989) opened with *Sal*I and *Xba*I. The resulting construct (pSBL1) contained the *sry* β -*lacZ* fusion placed upstream of the *sry* δ polyadenylation site into the plasmid pDm23 (*rosy* containing P-element transformation vector, a gift of D. Messmer and G. Rubin) derived from the plasmid Carnegie 20 (Rubin and Spradling, 1983). This construct was used to transform flies from the *ry*⁵⁰⁶ stock according to standard methods (Rubin and Spradling, 1982). Six independent transformant lines, identified as being different from one another by chromosomal linkage and Southern blot analysis, were made homozygous and are referred to as SBL lines.

(b) pSBA1 construct

A *sry* β DNA fragment, starting at an *Xmn*I site at codon position 41 and ending at the *Sal*I site created at codon position 340, was inserted into the plasmid pAR3039 (Rosenberg *et al.* 1987) to give the pSBA1 plasmid.

The pSDL1 and pSDA1 plasmids are described in Payre *et al.* (1989).

(c) pRLXS1 plasmid

This was generated by inserting a λ fragment from position 31617 to 33499 on the λ map (Daniels *et al.* 1983) in the plasmid pTZ18R.

Expression in *E. coli* and purification of the *sry* β protein

Overexpression of *sry* β or *sry* δ in *E. coli* was achieved by

using the pSBA1 and pSDA1 plasmid as described by Tabor and Richardson (1985). Purification of the inclusion bodies was carried out according to Nagai *et al.* (1988). After renaturation of the proteins (Hoey and Levine, 1988), extracts were chromatographed on a monoQ HR 5/5 column (Pharmacia). Fractions corresponding to either *sry* β or *sry* δ (purity of at least 90–95% as estimated by Coomassie staining of PAGE-SDS) were pooled and loaded onto a heparine sepharose column. The fractions eluted at 0.4 M KCl were dialysed against HK buffer (25 mM Hepes, 20% Glycerol, 120 mM KCl, 1 mM DTT, 50 μ M ZnSO₄) and used for DNA-binding experiments.

Production of anti-*sry* β antibodies

20 μ g of pSBA1 encoded *sry* β purified protein mixed together with 4 mg of poly(A)-poly(U) in 2 ml of NaCl (9%) was injected intravenously into young female rabbits. The injections were repeated five times at ten day intervals. Blood was collected weekly from rabbits boosted every month. Crude sera clarified by a 17 000 g centrifugation were passed over a column of affigel 10 sepharose (Biorad) coupled with total bacterial protein from an induced culture of pAR3039. After several passages over the same column the negatively purified serum was loaded onto a similar column containing 2 mg ml⁻¹ coupled purified *sry* β protein. Following washes in conditions described by Carroll and Scott (1985), the retained fraction containing the affinity purified antibody was eluted by 2 M guanidine HCl and extensively dialysed against 1 $\times$ PBS containing 0.1% sodium azide.

Immunological detections

Western blotting was done according to standard procedures using 1/1000 dilution for both anti-*sry* β and *sry* δ and 1/3000 for anti- β -gal (Promega). Nuclei were prepared from 2 to 12 h old embryos according to Jowett (1986) and extracted in conditions described by Dignam *et al.* (1983).

Immunostaining of whole-mount embryos was performed essentially as described by Carroll and Scott (1985). Embryonic stages referred to in the text are those of Campos-Ortega and Hartenstein (1985).

Salivary glands of third instar larvae were dissected in PBS and rapidly transferred to ALF solution (acetic acid 45%, lactic acid 1:6, formaldehyde 3.7%) for 30 min. After fixation, glands were squashed in ALF. Antibody reactions and Giemsa staining were done according to Zink and Paro (1989). Procedures involving the secondary antibody were as indicated by the manufacturer (Vector Lab).

Immunoprecipitation of DNA fragments

In a typical experiment, 20 ng of end labelled DNA and 1.5 μ g of yeast tRNA, complemented or not with 200 ng to 1 μ g of sonicated calf thymus DNA, were mixed with or without 50–100 ng of protein in 25 μ l of HK buffer. Following an incubation of 30 min in wet ice, 10 μ l of 1/10 *Staphylococcus aureus* cells were first preincubated for 30 min at 4°C with 1 μ l of specific antiserum before addition to each sample. After incubation for 30 min at 4°C, 1 ml HK buffer was added, and samples centrifugated for 5 min at 4°C. After a second wash in 1 ml HK buffer containing 0.5% NP40, pellets were resuspended in formamide. Immunoprecipitated DNA fragments were then analyzed by electrophoresis on 5% acrylamide sequencing gels followed by autoradiography.

Photography

Views were taken at varied magnification using bright-field or NOMARSKI optics Nikon (see legends of figures). We used

Kodak Ektachrome 50 ASA tungsten or Ektar 25 ASA films for colour photography.

Results

Protein sequence comparison

A preliminary comparison of the sry β and sry δ proteins revealed the presence of a 28–29 amino-acid motif repeated several times in each protein, the Cys2/His2 finger motif (Miller *et al.* 1985; Vincent *et al.* 1985; Vincent, 1986). We extend here the comparison to the entire protein length (Fig. 1), taking into account minor corrections of the published sequence that make the sry β and δ proteins 356 and 434 amino-acids long, respectively (Vincent *et al.* 1985, see legend of Fig. 1). Along the sequence comparison, the amino-acid numbers refer to the sry β protein sequence. Consequently, the places where the sry δ protein has more amino-acids are considered as insertions, and where it has less as deletions without any implication of directionality in the change seen.

The sequence alignment of sry β and sry δ proteins (Fig. 1) requires only two insertions longer than one amino-acid, one of which, 5 amino-acids long, is at the position of the sry β intron which has no counterpart in sry δ . It is readily apparent that the highest degree of conservation between sry β and sry δ is found in two separate regions: (i) the *N*-terminal part of the protein between amino-acids 8 and 88, showing 44 % homology (ii) the finger domain, amino-acids 172 to 337, encompassing 6 adjacent fingers, which displays the highest degree of conservation. The degree of relatedness of individual fingers found at homologous positions in the two proteins is not, however, uniform. Fingers 1, 5 and 6 display 38 %, 55 % and 52 % identical residues, respectively, when sry β and δ are compared, while the values rise to 75 %, 65 %, and 75 %, respectively, for fingers 2, 3 and 4. The maximal homology is in the carboxy terminal part of the fingers which includes the His–His putative α helical loop (X¹-L²-X³-X⁴-H⁵-X⁶-X⁷-X⁸-H⁹) and the linker that connects adjacent fingers

(Berg, 1988). It must be noted that these linkers are not of the widespread form Thr-Gly-Glu-Lys-Pro (Schuh *et al.* 1986). Remarkably, sry δ displays 55 additional carboxy-terminal residues, including a 7th finger separated from the 6 others by 45 amino-acids (see Fig. 1).

The conserved *N*-terminal and finger regions of sry β and sry δ are separated by regions rich in negatively charged amino-acids. These acidic domains, sry β amino-acids 124 to 172 (negative charge –19) and sry δ amino-acids 90 to 183 (negative charge –36) show very divergent primary sequences; there is an extra segment of 21 amino-acids in sry δ , which includes 14 consecutive Glu-Asp residues.

Anti-sry β antibodies and transformant *Drosophila* SBL lines

To compare the expression of sry β and sry δ during embryogenesis we prepared antibodies directed against the sry β gene product. A slightly truncated form of the sry β protein was expressed in *E. coli* using the pSBA1 plasmid (see Materials and methods). This polypeptide was purified and used to immunize rabbits. The specificity of the purified anti-sry β antibodies was demonstrated by immunoblotting assays against both bacterial and *Drosophila* protein extracts (Fig. 2B). In an *E. coli* lysate, the antibody recognizes only the pSBA1 encoded sry β protein. In *Drosophila* nuclear extracts from wild-type embryos, the anti-sry β antibody reacts with a single protein of $39 \times 10^3 M_r$.

In order to distinguish between the maternal and zygotic contributions to the embryonic sry β protein, we made transformant *Drosophila* lines containing a sry β /galactosidase reporter gene. This reporter gene encodes a fusion protein that contains 340 (out of 356) NH2 terminal sry β residues, including the 6 zinc fingers (Fig. 2A, see Materials and methods). The SBL construct was introduced into the *ry*⁵⁰⁶ *Drosophila* genome to establish transformant lines referred to below as SBL lines. Western blot analysis of embryo extracts using an anti- β -galactosidase monoclonal antibody revealed a nuclear protein specific to SBL lines with an apparent

Fig. 1. Conservation and divergence within the sry β compared to the sry δ protein. The representation is adapted from that chosen in Treier *et al.* (1989). Conserved residues are indicated by black lines. Deletions/insertions, as defined in the text, are represented by triangles facing downwards and upwards, respectively. The open, and small and large black triangles correspond to 21, 1 and 5 bp insertions, respectively. The following exchanges were considered as conservative ones: L $\rightleftharpoons$ V $\rightleftharpoons$ I; E $\rightleftharpoons$ D; S $\rightleftharpoons$ T; Y $\rightleftharpoons$ F; N $\rightleftharpoons$ Q. The position of the sry β intron is marked. Open arrows indicate the sites of fusion of the sry β or δ and *E. coli* β -galactosidase open reading frames in the pSBL1 and pSDL1 reporter genes. The position of each finger (numbered from 1 to 7) is represented by an open box; the shadowed area represents the putative His–His α helical loop according to Berg (1988). Revision of the structure of the sry β gene intron (Vincent *et al.* 1985), as determined by S1 mapping and cDNA sequencing (data not shown) added 5 residues to the sry β protein sequence: Gly, Cys, Ile, Val, Pro after the Glu residue position 25. Corrections in the sry δ protein sequence include additional residues: Pro at position 62, Asp at position 162 and Val (in place of Ser)-Leu-Lys at the C-terminal end, and a.a replacements: Ser $\rightarrow$ Met position 273, Phe $\rightarrow$ Val position 394, Thr $\rightarrow$ Asn position 414.

Fig. 2. Specificity of the purified anti-*sry* β and anti-*sry* δ antibodies. (A) Diagrammatic representation of the pSBL1 and pSDL1 reporter genes introduced into the *Drosophila* genome to give rise to SBL and SDL transformant lines, respectively. The open reading frames are drawn as boxes with the shaded and crippled areas representing the acidic and finger domains, respectively. (B) Total proteins extracted from induced *E. coli* cells carrying the pSBA1 (pSBA) and pAR3039 (pAR) control plasmids, respectively, were immunoblotted with purified anti-*sry* β antibodies; 10 μ g of proteins from nuclear extracts of 0 to 12 h old embryos from control *ry*⁵⁰⁶ (*ry*), SBL and SDL transformant lines were separated by electrophoresis on 10% SDS PAGE, transferred to nitrocellulose and incubated with the purified antibodies: anti-*sry* β , (left); anti-*sry* δ (center); anti- β -galactosidase (right). The position of molecular weight markers is indicated by arrows on the left.

relative molecular mass of 155×10^3 , the predicted size of the *sry* β / β -galactosidase fusion protein (Fig. 2). The anti-*sry* β antibody also recognises the 155×10^3 M_r fusion protein; this strongly suggests that the 39×10^3 M_r protein is *sry* β .

Comparison of results obtained with extracts prepared from *ry*⁵⁰⁶ control, SBL and SDL lines and reacted with either anti-*sry* β or anti-*sry* δ antibodies (see Payre *et al.* 1989 for description of the anti-*sry* δ antibodies and SDL transformant lines) confirmed that these antibodies are highly specific for *sry* β and *sry* δ , respectively, and do not show any detectable cross-reactivity. The data show that, like *sry* δ , the *sry* β and β -galactosidase fusion proteins are nuclear proteins.

The *sry* β , *sry* δ and *sry*/ β -galactosidase fusion proteins were localised in embryos by an immunostaining procedure (Carroll and Scott, 1985) using anti-*sry* β , anti-*sry* δ (Fig. 3A), and anti- β -galactosidase (Fig. 3B) antibodies, respectively.

Expression of the sry β protein during embryogenesis; comparison with sry δ expression

The *sry* β protein is present in early nuclear cleavage stage embryos. Staining of nuclei above the uniform ooplasmic staining seen prior to fertilisation (data not shown) is detectable as early as the four nuclei stage. During intravittelline divisions and migration of nuclei to the periphery of the embryo (up to nuclear division cycle 8, but prior to pole bud formation, stage 2 of Campos-Ortega and Hartenstein, 1985) all nuclei, including the precursors of pole cell nuclei are stained (Fig. 3A, a). At the blastoderm stage, the *sry* β protein is present in all somatic nuclei but is detectable only in residual trace amounts in pole cell nuclei (Fig. 3A, b). During germ band elongation (stage 8, Fig. 3A, c) and germ band shortening (stage 13, Fig. 3A, d), the spatial patterns of expression of the *sry* β and *sry* δ proteins are virtually indistinguishable (see Payre *et al.* 1989 for a detailed description of *sry* δ expression), although the peak of accumulation of the *sry* β protein markedly precedes that of *sry* δ (see Fig. 3A a,b and c). The predominant localisation of the *sry* δ protein in the sub and supraoesophageal ganglia and the ventral nerve chord after dorsal closure (stages 16–17) is not so distinct in the case of *sry* β (Fig. 3A, e, see also Fig. 3B).

Maternal and zygotic contribution to the embryonic sry β/ and sry δ/β-galactosidase proteins

A pattern similar to that of *sry* β was obtained for the *sry* β / β -galactosidase protein in SBL homozygous embryos. The maternal and zygotic contributions to the embryonic *sry* β / β -galactosidase were distinguished by comparing the time course of accumulation of the fusion protein in embryos from the crosses ♀ *ry*⁵⁰⁶ × ♂ SBL and ♀ SBL × ♂ *ry*⁵⁰⁶. In ♀ SBL × ♂ *ry*⁵⁰⁶ embryos, the *sry* β / β -galactosidase is already detectable during intravittelline nuclear divisions (Fig. 3B,a). In ♀ *ry*⁵⁰⁶ × ♂ SBL embryos, the zygotic protein encoded by the paternal transgene is detected, starting at the 13th nuclear cleavage (Fig. 3B, b). From this stage onto the end of embryogenesis, the patterns of zygotic and maternal plus zygotic accumulation of the *sry* β / β -galactosidase protein remain virtually identical (Fig. 3B, c–e). Comparison of *sry* β / β -gal and *sry* δ / β -gal expression confirmed that *sry* β is turned on earlier than *sry* δ in embryos since expression of the *sry* δ /lacZ gene in embryos from the ♀ *ry*⁵⁰⁶ × ♂ SDL was not detected before germ band extension (stage 9–10 embryos) (Fig. 3B, d and Payre *et al.* 1989). It also confirmed that, contrary to *sry* β , *sry* δ predominantly accumulates in the central nervous system after dorsal closure of the embryo (Fig. 3B, e), and that this accumulation is due to the zygotic expression of the protein.

A anti-*sry* β anti-*sry* δ
 *ry*⁵⁰⁶ embryos

COB 10-7002L 24.07.90 L01

Fig. 3. Compared expression of *sry* β and *sry* δ in embryos. Whole-mount preparations of fixed embryos were immunostained according to Carroll and Scott (1985). Embryos are oriented so that anterior is to the left and ventral is down. (A) *ry*⁵⁰⁶ embryos stained with anti-*sry* β (left), or anti-*sry* δ (right) antibodies. (B) embryos from ♀ *ry*⁵⁰⁶ × ♂ SBL or ♂ SDL and reciprocal crosses, as indicated at the top, stained with anti- β -galactosidase antibody. Similar stages are shown in panels A and B. (a) Embryos undergoing the eight nuclear cleavage; (b) blastoderm stage (interphase of the 14th nuclear division cycle) embryos. The stronger *sry* β staining, compared to *sry* δ (panel A), is contributed by both maternal and zygotic synthesis of the *sry* β protein present at that stage (see panel B); (c) embryos during germ band elongation (stage 8) and (d) during (stage 12, panel B) or after (stage 13, panel A) germ band shortening show similar patterns of staining for *sry* β and *sry* δ . First zygotic expression of *sry* δ / β -galactosidase is seen during germ band retraction (panel B); (e) embryos after organogenesis is almost completed (stage 17): maximal accumulation of the *sry* δ protein in the brain and ventral nerve cord occurring at that stage is not so distinct in the case of *sry* β .

B

anti-B-galactosidase

♀ *ry*⁵⁰⁶ × ♂ SBL

♀ SBL × ♂ *ry*⁵⁰⁶

♀ *ry*⁵⁰⁶ × ♂ SDL

♀ SDL × ♂ *ry*⁵⁰⁶

Fig. 5. *sry* β and *sry* δ proteins bind to different sets of chromosomal sites. Protein binding sites were visualised with the peroxidase substrate DAB (brown) and chromomeric bands with Giemsa stain (blue). (A) Complete set of wild-type chromosomes stained with anti-*sry* β (a) or anti-*sry* δ (b) antibodies. (B) Magnification of sections 1 to 7 of the X chromosome; c and d show the binding sites of *sry* β and *sry* δ , respectively; c', d' show the same chromosomes counterstained with Giemsa; e, f: tip of the chromosomal arm 2L stained with anti-*sry* β (e) or anti-*sry* δ (f) antibodies, and counterstained with Giemsa. (C) Tip of the X chromosome from SBL (g) and SDL (h) transformant lines immunostained with anti- β -galactosidase antibody and counterstained with Giemsa. Black arrows point to selected *sry* β binding sites (at positions 3B3-4, 21C8 and 21D3) that do not correspond to *sry* δ binding sites (open arrows).

sry β and sry δ bind selectively to DNA in vitro
As the target sequences of sry β and sry δ are not known, we used the empirical method previously used by Desplan *et al.* (1985) to detect DNA sequence specific recognition. As a first step *Sau*3A1 digested λ

phage DNA was incubated with the purified sry proteins produced in *E. coli*. The protein-DNA complexes formed upon incubation were immuno-precipitated. The results obtained for sry β are shown in Fig. 4A. In the absence of protein, or with anti-sry- δ antibodies, no

Fig. 4. Sequence specific interaction of the sry β and sry δ proteins with bacteriophage λ and plasmid pRLXS1 DNA fragments. Immunoprecipitated DNA fragments (120 mM KCl salt conditions, see Materials and methods) were separated on a 5% denaturing polyacrylamide gel. (A) *Sau*3A1 digested λ phage DNA; (B) *Sau*3A1 digested pRLXS1 plasmid DNA. Lanes a, a': input DNA; b, b' immunoprecipitated material in the absence of added sry protein in the assay; c, c' immunoprecipitates in the presence of protein and tRNA without competitor DNA; lanes d, d', e, e' and f, f' show immunoprecipitates obtained in the presence of 200 ng, 400 ng and 1 μ g, of calf thymus competitor DNA respectively; lane g, immunoprecipitate in the presence of sry β protein but with anti-sry δ antibody. The size (bp) of the pRLXS1 *Sau*3A1 fragments is indicated. The 66, 78, 88, 48, 126, 188 and 279 bp fragments are derived from the λ phage DNA. The 75 and 258 bp fragments derive from the pTZ18R plasmid vector, as indicated by an asterisk. The open and black arrowheads point to fragments bound by either sry β or sry δ , but not by both.

DNA was detected in the precipitate (Fig. 4A, lane b and g). Without competitor DNA, many λ DNA fragments were bound nonselectively (lane c). With increasing concentrations of competitor DNA, the binding became selective with 9 out of 80 resolved DNA fragments specifically precipitated (lanes d–f). To be able to compare precisely the DNA-binding specificities of *sry* β and *sry* δ , we subcloned a λ phage fragment which contained 4 of the immunoprecipitable fragments (see Fig. 4B). *Sau*3A1 digestion of DNA from the resulting plasmid (pRLXS1) yielded 14 resolvable fragments ranging from 46 to 341 bp in size (Fig. 4B lanes a and a'). In the absence of competitor DNA, 6 and 7 fragments were quantitatively immunoprecipitated by *sry* β and *sry* δ , respectively (lanes c and c'). Only 4 of these were common to both proteins, the others being specifically bound by one or the other. It must be noted that the 258 and 75 bp fragments derive from the plasmid vector. Increasing concentrations of competitor DNA in the binding assay drastically decreased the amount of all immunoprecipitated DNA fragments in the case of *sry* δ , while specific binding of *sry* β to the 258, 126 and 66 bp fragments was partly resistant (Fig. 4B, lanes d–f and d'–f').

In vivo binding patterns of sry β and sry δ on polytene chromosomes are specific of each protein

Nuclei of late third instar larvae salivary glands contain large polytene chromosomes allowing precise mapping of protein–DNA interaction sites (see, e.g., Sass, 1982; Kabisch and Bautz, 1983; Pirrotta *et al.* 1988; Zink and Paro, 1989). Since both *sry* β and *sry* δ are expressed in the nuclei of larval salivary glands (Payre *et al.* 1989 and data not shown), we looked for their potential binding sites on the polytene chromosomes, using an indirect immunostaining procedure (see Materials and methods).

When squashed fixed chromosomes from wild-type larvae were reacted with purified anti-*sry* β or anti-*sry* δ antibodies, a discrete number of chromosomal sites of various intensities, but neither the nucleolus nor the chromocenter, were stained (Fig. 5A). The *sry* β protein was localised at about 60 specific sites (Fig. 5A,a), 54 of which were cytologically mapped (Table 1). A much larger number of *sry* δ binding sites, about 200, were revealed with anti-*sry* δ antibodies (Fig. 5A,b). In either case, the staining of any particular region appeared constant when chromosome spreads from separate larvae were compared. Observation of immunostained chromosomes after counterstaining with Giemsa revealed that the stained sites are mostly localised in interbands or at the edge of strong Bridges bands (Fig. 5B). Detailed comparison of the *sry* β and δ binding sites focused on the X chromosome and selected regions of the autosomes containing strong *sry* β binding sites (Fig. 5B and Table 1). This comparison showed that not only some weak but also some of the strong *sry* β binding sites do not correspond to *sry* δ binding sites (e.g. sites at 3B3–4, 21C8 and 21D3, Fig. 5B) while others could be common to both proteins.

Table 1. Location of *sry* β binding sites on salivary gland polytene chromosomes of Oregon R larvae

X	2L	2R	3L	3R
2B16,18+	*21C8+	42F	63D	*82B4
*3B3,4+	*21D3+	43F7,9	64F-65A+	82F4,7
4F9,10+	23C	46EF	65D+	*83A3,4+
5B	23F1,4	49A+	66AB+	85C+
5E8	24C6,9	49F10,15+	66C1,4+	85D24,27+
6CD	24E	50C+	67B6,9+	85F14,16
7CD+	25B	52F7,11	73B9,11+	86E3,6
10C1,6+	28C4,7+	55B4,9	75F6,10+	95DE+
11B7,11+	30C1,2	59F	*76C4,5	96B15,21
*11C1,2+	30F1,3			96C6,9
*11C3,4+	31B1,3+			
13A7,12				
13E7,9+				
*17B				
*18B1,2				

The cytological positions of the binding sites were determined by comparison to the map of Bridges (Lindsley and Grell, 1986).

* indicate *sry* β sites that do not overlap with *sry* δ binding sites.

Comparison was done on the entire X chromosome and autosomal regions 21, 74 to 76 and 81 to 83, which contain the strongest *sry* β binding sites, as indicated by boldface letters. † indicate chromosomal sites where DNA clones have been localised by *in situ* hybridization (Merriam, 1988).

Binding of sry β /lacZ and sry δ /lacZ fusion proteins on polytene chromosomes mimics that of the native proteins

In view of the marked differences between the *sry* β and *sry* δ patterns of chromosomal sites, it was of interest to look at the binding of the fusion proteins. We therefore immunostained polytene chromosomes from SBL and SDL homozygous larvae, (3 and 2 independent transformed lines, respectively) with a monoclonal anti- β -galactosidase antibody (Fig. 5C). The use of a single antibody for this comparison rules out the possibility that a cross-reaction between anti-*sry* β or anti-*sry* δ antibodies and other chromosomal proteins contributed to the observed patterns.

The patterns of binding sites observed for the *sry* β / β -galactosidase and *sry* δ / β -galactosidase proteins were similar to those of the wild-type *sry* β and *sry* δ proteins, respectively, as exemplified by a detailed view of the tip of the X chromosome (compare Fig. 5B,c' and 5C,g; 5B,d' and 5C,h). As a control, no staining of *ry*⁵⁰⁶ chromosomes was seen (not shown). Detailed comparison of the *sry* β and *sry* β / β -galactosidase patterns showed that the number and position of sites were conserved, with, however in some cases, a decrease in staining intensity of the fusion protein sites.

Discussion

Eucaryotic transcription factors can be broadly divided into two classes (Biggin and Tjian, 1989): one class is expressed in highly regulated spatial and temporal patterns correlating with gene-specific regulatory functions. The second class is expressed in a variety of cells and may regulate a wide range of genes. The *Drosophila sry* β , *sry* δ and *sry h-1* genes encoding Cys2/

His2 finger proteins maternally provided to the embryo could belong to this second class (Vincent *et al.* 1988).

We have chosen the sry β and δ genes – so far the only identified set of two closely related Cys2/His2 finger protein genes in *Drosophila* – as a model for evolutionary studies on the structure and function of this class of DNA-binding proteins.

The presence of acidic domains in sry proteins supports the prediction that they are transcription factors, since a significant net negative charge is characteristic of the transcription activating region of various gene control proteins (Ptashne, 1988). The divergence of the primary sequence of the acidic domain between sry β and δ contrasts with the degree of conservation of the rest of the protein (Fig. 1). In a perhaps related vein, the transcription activating domain of the yeast Gal4 transcription activator can be replaced by a number of other amino-acid sequences, related only by their negative charge (Ma and Ptashne, 1987).

The predicted DNA-binding domain of the sry β and δ proteins comprises 6 contiguous fingers, an additional carboxy terminal 7th finger being found in sry δ . The degree of relatedness between individual fingers located at homologous positions in the two proteins is variable (Fig. 1). This suggests that all the fingers were not all subject to equal evolutionary constraints and may have taken on specific individual DNA-binding functions. Those that have been best conserved between sry β and sry δ are the contiguous fingers 2,3 and 4 (Fig. 2). The conservation includes the amino-acids at positions 3 and 4 of the His–His loop, which generally correspond to highly variable positions (Bellefroid *et al.* 1989; Nietfeld *et al.* 1989) and have been proposed as being essential to specific contacts between DNA and fingers (Berg, 1988). Conservation of these 3 fingers may be responsible for the binding of the two proteins to some common DNA fragments *in vitro* (Fig. 4). Divergence of the other fingers may in turn be related to the differences in sry β and δ DNA recognition specificities *in vitro* and *in vivo* (Fig. 4, 5 and Table 1). Assuming that each Cys2/His2 finger may potentially bind ≈ 5 bp of DNA, (see discussion by Klug and Rhodes, 1987), the DNA target size of sry β and sry δ proteins could be up to ≈ 30 bp long. Even considering that possibly not all the fingers are required for specific DNA-binding (Vrana *et al.* 1988; Payre and Vincent, 1988) and not all base pairs of the recognized sequence are essential, it is statistically unlikely that *bona fide* binding sites of sry β or sry δ would be present within the 48 kbp of λ phage DNA. Binding of the sry β and sry δ proteins to several λ DNA fragments correlated with the sensitiveness of this binding to low concentrations of competitor DNA suggests that it reveals low affinity sites involving only limited specific contacts between DNA and the proteins.

The binding specificity *in vivo* might be influenced by protein modifications not occurring in *E. coli* or by interactions with other *Drosophila* factors which were lacking in our *in vitro* assay. We therefore compared the DNA-binding patterns of sry β and sry δ *in vivo*, on polytene chromosomes. Around 60 sites (Table 1) were

observed for sry β , while a much greater number of sites were reproducibly observed for sry δ (Fig. 5A). Major puffs are not stained, indicating that neither sry β or sry δ proteins are general RNA polymerase II associated factors (Kabisch and Bautz, 1983). Cytogenetic mapping of sry β (and part of sry δ , not shown) binding sites did not show up any comprehensive gene family, partly due to the absence of precisely mapped genes in the regions corresponding to mapped sites (Merriam, 1988). The number of binding sites reported for *Polycomb* and *Zeste*, two proteins that do not possess fingers and have completely unrelated structures, is similar to that of sry β (Zink and Paro, 1988; Pirrotta *et al.* 1988). A good correlation was noticed between the cytological positions of a dozen *Polycomb* binding sites and the position of genes with which *Polycomb* genetically interacts (Zink and Paro, 1988). We have no direct evidence that the binding of sry proteins at all the sites seen on chromosomes is functional (Berg and Von Hippel, 1988). Nevertheless, the different patterns of sry β and sry δ suggest that they interact with specific, although probably partly overlapping, sets of genes. The marked difference in their number of chromosomal binding sites was rather unexpected and could have been contributed to by the at least 5 fold difference in relative amounts of these two proteins in salivary gland nuclei (data not shown). For example, overproduction of *zeste* in salivary glands results in a vastly greater number of bands binding the *zeste* protein than is observed in normal conditions (Pirrotta *et al.* 1988).

The patterns of binding sites of the sry β / β -galactosidase and sry δ / δ -galactosidase fusion proteins are similar to those of sry β and sry δ , respectively (Fig. 5B and C). It both confirms the sry β and sry δ DNA-binding specificities and shows that these specificities are maintained when the proteins are fused to the enzymatically active β -galactosidase. Decrease in staining intensity of some fusion protein binding sites could reflect a decreased accessibility of these sites to the larger proteins for steric hindrance reasons. Alternatively, it could reflect site-specific competition effects between the wild-type and the chimeric proteins for binding to DNA. The chimeric protein approach used here should prove valuable for the identification of potential target genes for regulatory factors with no need for specific antibodies or the *in vivo* dissection of their DNA-binding domain. The present data indicate that the seventh carboxy-terminal sry δ finger (which is deleted in the fusion protein) does not account for the differences in the binding patterns of sry β and sry δ to polytene chromosomes. This can be brought together with data on *in vitro* mapping of TFIIA DNA-binding domains, which showed that progressive deletions of its 4 carboxy-terminal fingers do not suppress its DNA-binding specificity and result in only a slight decrease in overall binding affinities of the mutant proteins (Vrana *et al.* 1988).

The presence of the sry β and δ proteins in embryonic nuclei at the onset of zygotic transcription (Edgar and Schubiger, 1986) suggests that both proteins play a role in the activation of the zygotic genome. Like sry δ , the

sry β protein present in embryos originates from both maternal and zygotic expression of the gene (Fig. 3). Detection of the zygotic *sry β/β*-galactosidase protein as early as the syncytial blastoderm stage correlates with the pattern of transcription from a transposed modified *sry β* gene (Vincent *et al.* 1986). It confirms that 459 bp of upstream DNA are sufficient for maternal and zygotic expression of *sry β* and that the embryonic protein is mainly of maternal origin; *sry β* expression is detected, but at a much lower level, at every other period of the fly life cycle examined. (Vincent *et al.* 1985 and data not shown). The zygotic *sry δ/β*-galactosidase protein is expressed starting at the germ band extension stage (stage 9–10 embryos) but, at later stages, predominates over the maternal product. The *sry δ* mRNA and protein are also present in various amounts in larvae, pupae and adults (Vincent *et al.* 1985; Payre *et al.* 1989). How the temporal difference between the maximal levels of expression of *sry β* and *δ* proteins during embryogenesis contributes to their functional specificity is not known. Several *δ* mutant alleles but no *β* mutations have been isolated from three independent screens based on zygotic lethality (Crozatier *et al.* in preparation). It is therefore possible that the low level of zygotic *sry β* expression is dispensable for survival. In the same way, it may be noticed that the only known phenotype of mutations in *SuHw*, another Cys2/His2 finger protein expressed in various tissues (Spana *et al.* 1988), is female sterility. In the near future, a combination of genetic and biochemical approaches should enable us to understand in details of the respective functions of *sry β* and *sry δ* during *Drosophila* development.

We are grateful to François Amalric and Michèle Crozatier for critically reviewing the manuscript, Françoise Lemeunier for suggestions and Jean Antoine Lepasent for his continuous interest. We wish to thank Edmond Barbey for photography and Joëlle Maurel for editorial assistance. This work was supported by CNRS, INSERM (grant n° 881019) and the Fondation pour la Recherche Médicale. F.P., S.N. and V.L. were supported by fellowships from Ligue Nationale Contre le Cancer and Ministère de la Recherche et de l'Enseignement Supérieur.

References

- AGUADÉ, M. (1988). Nucleotide sequence comparison of the rp49 gene region between *Drosophila subobscura* and *D. melanogaster*. *Mol. Biol. Evol.* **5**, 433–441.
- BELLEFRÖID, E. J., LECOCQ, P. J., BENHIDA, A., PONCELET, D. A., BELAYEW, A. AND MARTIAL, J. A. (1989). The human genome contains hundreds of genes coding for finger proteins of the *krüppel* type. *DNA* **8**, 377–387.
- BERG, J. M. (1988). Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. *Proc. natn. Acad. Sci. U.S.A.* **85**, 99–102.
- BERG, O. G. AND VON HIPPEL, P. H. (1988). Selection of DNA binding sites by regulatory proteins. *Trends Biochem. Sci.* **13**, 207–211.
- BIGGIN, M. D. AND TJIAN, R. (1989). Transcription factors and the control of *Drosophila* development. *Trends Genet.* **5**, 377–383.
- BROWN, R. S., SANDER, C. AND ARGOS, P. (1985). The primary structure of transcription factor TFIID has 12 consecutive repeats. *FEBS Lett.* **186**, 271–274.
- CAMPOS-ORTEGA, J. A. AND HARTENSTEIN, V. (1985). *The Embryonic Development of Drosophila melanogaster*. Springer Verlag, Berlin/New York.
- CARROLL, S. B. AND SCOTT, M. P. (1985). Localisation of the *fushi tarazu* protein during *Drosophila* embryogenesis. *Cell* **43**, 47–57.
- DANIELS, D., SCHROEDER, J., SZYBĄLSKI, W., BLATTNER, F., SANGER, F., COULSON, A., HONG, G., HILL, D. AND PETERSON, G. (1982). In *LambdaII* (eds Hendrix, R., Roberts, J., Stahl, F. and Weisber, R.) Cold Spring Harbor Laboratory, New York.
- DESPLAN, C., THIES, J. AND O'FARRELL, P. H. (1985). The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. *Nature* **318**, 630–635.
- DIGNAM, J. D., LEBOVITZ, R. M. AND ROEDER, R. G. (1983). Accurate initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucl. Acids Res.* **11**, 1475–1489.
- EDGAR, B. A. AND SCHUBIGER, G. (1986). Parameters controlling transcriptional activation during early *Drosophila* development. *Cell* **44**, 871–877.
- EVANS, R. M. AND HOLLENBERG, S. M. (1988). Zinc fingers: Gilt by association. *Cell* **52**, 1–3.
- HOEY, T. AND LEVINE, M. (1988). Divergent homeobox proteins recognise similar DNA sequences in *Drosophila*. *Nature* **332**, 858–861.
- JOWETT, T. (1986). Preparation of nucleic acids. In *Drosophila, A Practical Approach* (Roberts, D. B. ed.) IRL Press. pp. 139–157.
- KABISCH, R. AND BAUTZ, E. K. F. (1983). Differential distribution of RNA polymerase B and non histone chromosomal proteins in polytene chromosomes of *Drosophila melanogaster*. *EMBO J.* **2**, 395–402.
- KLUG, A. AND RHODES, D. (1987). 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12**, 464–469.
- KNÖCHEL, W., PÖTING, A., KÖSTER, M., EL-BARADI, T., NIETFELD, W., BOUWMEESTER, T. AND PIELER, T. (1989). Evolutionary conserved modules associated with zinc fingers in *Xenopus laevis*. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6097–6100.
- LINDSLEY, D. L. AND GRELL, E. H. (1968). Genetic variations of *Drosophila melanogaster*. *Carnegie Inst. Wash. Publ.* 627.
- MA, J. AND PTASHNE, M. (1987). A new class of yeast transcriptional activators. *Cell* **51**, 113–119.
- MERRIAM, J. (1988). The *Drosophila* clone list by chromosome location. *D.I.S.* **67**, 11–135.
- MILLER, J., MCLACHLAN, A. D. AND KLUG, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609–1614.
- MITCHELL, P. J. AND TJIAN, R. (1989). Transcriptional regulation in mammalian cells by sequence specific DNA binding proteins. *Science* **245**, 371–378.
- NAGAI, K., NAKASEKO, Y., NASMYTH, K. AND RHODES, D. (1988). Zinc-finger motifs expressed in *E. Coli* and folded *in vitro* direct specific binding to DNA. *Nature* **332**, 284–286.
- NIETFELD, W., EL-BARADI, T., MENTZEL, H., PIELER, T., KÖSTER, M., PÖTING, A. AND KNÖCHEL, W. (1989). Second-order repeats in *Xenopus laevis* finger proteins. *J. molec. Biol.* **208**, 639–659.
- PAYRE, F. AND VINCENT, A. (1988). Finger proteins and DNA-specific recognition: distinct patterns of conserved amino acids suggest different evolutionary modes. *FEBS Lett.* **234**, 245–250.
- PAYRE, F., YANICOSTAS, C. AND VINCENT, A. (1989). Serendipity delta, a *Drosophila* zinc finger protein present in embryonic nuclei at the onset of zygotic gene transcription. *Devl Biol.* **136**, 469–480.
- PIRROTTA, V., BICKEL, B. AND MARIANI, C. (1988). Developmental expression of the *Drosophila zeste* gene and localisation of the *zeste* protein on polytene chromosomes. *Genes and Development* **2**, 1839–1850.
- PTASHNE, M. (1988). How eucaryotic transcriptional activators work. *Nature* **335**, 683–689.
- ROSENBERG, A. H., LADE, B. N., CHUI, D. S., LIN, S. W., DUNN, J. J. AND STUDIER, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125–135.
- RUBIN, G. M. AND SPRADLING, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353.
- RUBIN, G. AND SPRADLING, A. (1983). Vectors for P element

- mediated gene transfer in *Drosophila*. *Nucl. Acids Res.* **11**, 6341–6351.
- SASS, H. (1982). RNA polymerase B in polytene chromosomes: immunofluorescent and autoradiographic analysis during stimulated and repressed DNA synthesis. *Cell* **28**, 269–278.
- SCHUH, R., AICHER, W., GAUL, U., COTÉ, S., PREISS, A., MAIER, D., SEIFERT, E., NAUBER, U., SCHRÖDER, C., KEMLER, R. AND JÄCKLE, H. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by *Krüppel*, a *Drosophila* segmentation gene. *Cell* **47**, 1025–1032.
- SPANNA, C., HARRISON, A. AND CORCES, V. G. (1988). The *Drosophila melanogaster* suppressor of *Hairy-wing* protein binds to specific sequences of the gypsy retrotransposon. *Genes and Development* **2**, 1414–1423.
- TABOR, S. AND RICHARDSON, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1074–1078.
- TREIER, M., PFEIFLE, C. AND TAUTZ, D. (1989). Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. *EMBO J.* **8**, 1517–1525.
- VINCENT, A. (1986). *TFIIIA* and homologous genes. The 'finger' proteins. *Nucl. Acids Res.* **14**, 4385–4391.
- VINCENT, A., COLOT, H. V. AND ROSBASH, M. (1985). Sequence and structure of the *serendipity* locus of *Drosophila melanogaster*. A densely transcribed region including a blastoderm-specific gene. *J. molec. Biol.* **186**, 149–166.
- VINCENT, A., COLOT, H. V. AND ROSBASH, M. (1986). Blastoderm-specific and read-through transcription of the *sry* alpha gene transformed into the *Drosophila* genome. *Devl Biol.* **118**, 480–487.
- VINCENT, A., KEIJZLAROVA-LEPESANT, J., SEGALAT, L., YANISCOSTAS, C. AND LEPESANT, J. A. (1988). *sry h-1*, a new *Drosophila* multifingered protein gene showing maternal and zygotic expression. *Mol. Cell Biol.* **8**, 4459–4468.
- VRANA, K. E., CHURCHILL, M. E. A., TULLIUS, T. D. AND BROWN, D. D. (1988). Mapping functional regions of transcription factor *TFIIIA*. *Mol. Cell Biol.* **8**, 1684–1696.
- ZINK, B. AND PARO, R. (1989). *In vivo* binding pattern of a trans-regulator of homeotic genes in *Drosophila melanogaster*. *Nature* **337**, 468–471.

(Accepted 6 June 1990)