THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 44, pp. 31581–31591, November 1, 2013

Silk Gland Factor-2, Involved in Fibroin Gene Transcription, Consists of LIM Homeodomain, LIM-interacting, and Single-stranded DNA-binding Proteins*

Received for publication, August 29, 2013 Published, JBC Papers in Press, September 10, 2013, DOI 10.1074/jbc.M113.514471

Kaoru Ohno[‡], Jun-ichi Sawada[§1], Shigeharu Takiya[‡¶2], Mai Kimoto[¶], Akiko Matsumoto[§], Takuya Tsubota[‖], Keiro Uchino[‖], Chi-chung Hui[‡3], Hideki Sezutsu[‖], Hiroshi Handa[§], and Yoshiaki Suzuki[‡4]

From the [‡]National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki 444-8585, the [§]Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501, the [¶]Hokkaido University, North 10 West 8, Kita-ku, Sapporo 060-0810, Japan, and the [‖]National Institute of Agrobiological Sciences, Owashi 1-2, Tsukuba 305-8634, Japan

Background: Silk gland factor-2 (SGF-2) is a key factor regulating tissue-specific expression of the fibroin gene.
Results: SGF-2 is a 1.1-MDa heteromeric complex containing Awh, Ldb, Lcaf, and fibrohexamerin proteins.
Conclusion: Awh, Ldb, and Lcaf interact functionally in SGF-2 to control fibroin gene expression.
Significance: This study provides new insight into the functional role of single-stranded DNA-binding proteins in protein-protein interaction and transcriptional regulation.

SGF-2 binds to promoter elements governing posterior silk gland-specific expression of the fibroin gene in *Bombyx mori*. We purified SGF-2 and showed that SGF-2 contains at least four gene products: the silkworm orthologues of LIM homeodomain protein Awh, LIM domain-binding protein (Ldb), a sequence-specific single-stranded DNA-binding protein (Lcaf), and the silk protein P25/fibrohexamerin (fhx). Using co-expression of these factors in Sf9 cells, Awh, Ldb, and Lcaf proteins were co-purified as a ternary complex that bound to the enhancer sequence *in vitro*. Lcaf interacts with Ldb as well as Awh through the conserved regions to mediate transcriptional activation in yeast. Misexpression of *Awh* in transgenic silkworms induces ectopic expression of the fibroin gene in the middle silk glands, where *Ldb* and *Lcaf* are expressed. Taken together, this study demonstrates that SGF-2 is a multisubunit activator complex containing Awh. Moreover, our results suggest that the Ldb·Lcaf protein complex serves as a scaffold to facilitate communication between transcriptional control elements.

Expression of the silk genes is a trait of terminal differentiation of the silk gland of *Bombyx mori* (1). The fibroin gene encoding the silk fiber protein is expressed only in cells of the posterior silk gland (PSG)[5] (1, 2), whereas the genes for glue proteins, sericins, are expressed only in cells of the middle silk gland (MSG) (3–5). The cell-free transcription systems using silk genes and crude nuclear extracts derived from silk gland tissues (6, 7) led to the *in vitro* reconstitution of tissue-specific transcription of the fibroin gene and the identification of *cis*-elements important for its transcriptional activity (1, 7–9). In particular, the region between −214 and −180 in the upstream promoter element En I, designated as the E site, is essential for tissue-specific transcriptional enhancement (7, 8, 10) (see Fig. 1*A*). Recently, Shimizu *et al.* (11) demonstrated that additional enhancer elements further upstream, which contain similar sequences of the E site, are also necessary for full activation of the fibroin gene *in vivo*.

We have demonstrated the presence of factors that bind specifically to these elements (10, 12). Silk gland factor-1 (SGF-1) is a forkhead (Fkh) protein binding to the proximal upstream region (12, 13). Fibroin modulator-binding protein-1 (FMBP-1), which contains a novel DNA-binding domain, binds to both En I and the intronic element En II (12, 14). Silk gland factor-2 (SGF-2) binds to the E site with two AT-rich repeat sequences, which resemble the consensus sequence recognized by homeodomain proteins. Of these factors, SGF-2 is specifically detected in PSG (10).

In this study, SGF-2 was purified, and its composition was determined. SGF-2 contains at least four components: the silk protein P25/fhx, a LIM homeodomain (LIM-HD) protein *Bombyx* Arrowhead (Awh), LIM domain-binding (Ldb) pro-

* This work was partially supported by a grant-in-aid for scientific research on priority areas from The Ministry of Education, Culture, Sports, Science and Technology (MEXT), and by a grant from the Ministry of Agriculture, Forestry, and Fisheries (MAFF) of Japan, a research grant from Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Corporation (JST), research fellowships from the Japan Society for the Promotion of Science for Young Scientists, and a grant for R&D projects in cooperation with academic institutions from the New Energy and Industrial Technology Development Organization (NEDO).
The nucleotide sequence reported in this paper has been deposited in the DNA Data Bank of Japan (accession numbers AB687553–AB687557).
[1] Present address: University of Shizuoka, 52-1 Yada, Suruga-ku, Shizuoka 422-8526, Japan.
[2] To whom correspondence should be addressed: Division of Biological Sciences, Faculty of Science, Hokkaido University, North 10 West 8, Kita-ku, Sapporo 060-0810, Japan. Tel.: 81-11-706-3590; Fax: 81-11-706-3588; E-mail: takiya@sci.hokudai.ac.jp.
[3] Present address: The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada.
[4] Present address: 5651-451 Matsukawa Village, North Azumi County, Nagano 399-8501, Japan.

[5] The abbreviations used are: PSG, posterior silk gland; MSG, middle silk gland; SGF-2, silk gland factor 2; LIM-HD, LIM homeodomain; Ldb, LIM domain binding; SSDP, sequence-specific single-stranded DNA-binding protein; Awh, Arrowhead; 3-AT, 3-amino-1, 2, 4-triazole; GAL4, galactose 4; UAS, upstream activation sequence; Ni-NTA, nickel-nitrilotriacetic acid; AD, activation domain; SD, synthetic defined.

tein, and a member of the sequence-specific single-stranded DNA-binding protein (SSDP) family. By misexpression of *Awh* in transgenic silkworms, expression of the fibroin gene was induced in the middle silk glands, demonstrating that SGF-2 is a tissue-specific activator complex of the fibroin gene.

EXPERIMENTAL PROCEDURES

Electrophoretic Mobility Shift Assay (EMSA)—Each protein-DNA binding reaction contained 5–10 fmol of probe (39-mer oligonucleotide with the E site sequence), 1 μg of poly(dI-dC) (GE Healthcare), and protein samples in a volume of 10 μl (15). After incubation on ice for 30 min, samples were analyzed by electrophoresis on 3.2% acrylamide gel containing 2.5% (v/v) glycerol at 4 °C in 0.25× TBE (22.5 mM Tris borate (pH 8.0), 0.5 mM EDTA) buffer.

Purification of SGF-2—Commercial silkworm strains (Kin-Shu x Sho-Wa or Shun-Rei x Sho-Getsu from Kanebo Silk Co., Kasugai City, Japan) of *B. mori* were reared at 27 °C on an artificial diet from Kyodo Shiryo Co. (Yokohama, Japan). SGF-2 was purified from crude nuclear extracts of PSG from V2 instar larvae through six column chromatographic steps (see Fig. 1*C*). Crude nuclear extract (protein, 80.0 g; volume, 2,040 ml) from 40,000 pairs of PSG from V2 instar larvae was prepared as described previously (6, 7, 9) and subjected to following purification steps (see Fig. 1*C*). The nuclear extract was diluted 5-fold by adding $TEMGTK_0$ buffer. TEMGTK buffers contain 20 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 12.5 mM $MgCl_2$, 10% glycerol, 0.1% Tween 20, 0.1% PMSF; the number following TEMGTK denotes the concentration (in mM) of KCl. After 30 min of stirring at 4 °C, the sample was centrifuged with a JA-10 rotor (Beckman) at 10,000 rpm for 1 h at 4 °C. The supernatant was filtered through a Y020A047A membrane filter (ADVANTEC, Tokyo, Japan) and sequentially applied to a 500-ml SP Sepharose Fast Flow resin column (GE Healthcare) equilibrated in $TEMGTK_{20}$. The column was washed with 1,000 ml of $TEMGTK_{40}$ and then eluted with $TEMGTK_{120}$. The eluate (volume, 1,000 ml) was loaded onto a 40-ml Source 30Q resin column (GE Healthcare). The column was washed with 200 ml of $TEMGTK_{140}$, and the SGF-2 activity was eluted with $TEMGTK_{190}$. The SGF-2 fraction (protein, 511 mg; volume, 983 ml) was diluted with $TEMGTK_0$ to adjust the KCl concentration to 100 mM. The diluted eluate underwent DNA affinity purification using Dynabeads (Dyna-l) on which 4 μg of poly(EW) DNA was immobilized per mg of beads. Poly(EW) DNA contained tandem repeats of SGF-2 binding sequence derived from the E site of fibroin promoter and was 0.2–1.0 kbp in length. For one round of DNA affinity purification, 190 ml of diluted eluate was incubated with 10 mg of heat-denatured salmon sperm DNA and 10 mg of poly(dI-dC) on ice for 10 min and then mixed with 10 mg of the beads. After the binding reaction at 4 °C for 30 min, the beads were collected using a magnetic stand, and the supernatant was removed. The beads were washed eight times (50, 20, 10, 5, 2, 1, 0.5, and 0.2 ml, respectively) with $TEMGTK_{100}$ batch-wise, and bound proteins were eluted with 1 ml of $TEMGTK_{1000}$ twice. The total eluate (0.15 mg, 2.1 ml) was dialyzed against $TEMGTK_{100}$ and loaded onto a column of Bio-Silect 250 (Bio-Rad) equilibrated with $TEMGTK_{100}$. Fractions containing SGF-2 activity (71 μg, 11.8 ml) were pooled and applied to a 0.1-ml Mini S column (GE Healthcare, SMART system) equilibrated in $TEMGTK_0$. The column was washed with $TEMGTK_{40}$, and proteins were eluted with $TEMGTK_0$ containing 6 M urea and then with $TEMGTK_0$ containing 6 M guanidine HCl. SGF-2 activity was not detected in the $TEMGTK_{40}$ wash fraction. The eluate with 6 M urea was dialyzed against $TEMGTK_{40}$ and applied to a 0.1-ml Mini Q column (GE Healthcare, SMART system) equilibrated in $TEMGTK_{40}$. The column was washed with $TEMGTK_{40}$, and bound proteins were eluted with $TEMGTK_0$ containing 6 M guanidine HCl.

Amino Acid Sequencing of SGF-2—Purified proteins were resolved by SDS-PAGE, and the bands of interest were subjected to in gel tryptic digestion, as described previously (13). The generated tryptic peptides were fractionated with a reverse-phase column, and the resolved peptide peaks were subjected to automated Edman degradation on an ABI Procise 477A protein sequencer (Applied Biosystems).

Isolation of cDNAs for SGF-2 Subunits—cDNA library prepared using poly(A)$^+$ RNA from V2 PSG was screened with a random primed probe made from RT-PCR products amplified using primer sets designed on the basis of the results of amino acid sequencing (see Table 1). The positive clones were sequenced. The accession numbers of these cDNA clones of SGF2 subunits, p36 (Awh), p47B (Ldb), and p48/p47G/p45 (Lcaf), are AB687553, AB687554, and AB687556, respectively. During the cDNA cloning of SGF2 p47B and p48/p47G/p45, the other clones, named Ldbβ (AB687555) and Lcafβ (AB687557), which are probably derived from alternatively spliced mRNA, were also obtained.

Preparation of Recombinant Proteins—Recombinant proteins were produced in Sf9 cells using the baculovirus expression system. DNA fragments encoding proteins of interest were cloned into pFastBac donor plasmids, and recombinant baculoviruses were obtained using Bac-to-Bac baculovirus expression systems (Invitrogen). For expression of recombinant proteins, Sf9 cells were infected by the virus(es) and cultured at 27 °C for 60 h. The infected cells were lysed in 10 mM HEPES-KOH (pH 7.9), 10 mM KCl, 1.5 mM $MgCl_2$, 0.05 mM $ZnSO_4$, 1 mM DTT, 0.5% Nonidet P-40, and protease inhibitors (Roche Applied Science). The lysates were centrifuged at 15,000 × *g* for 10 s at 4 °C. The precipitate was resuspended in 20 mM HEPES-KOH (pH 7.9), 0.38 M $(NH_4)_2SO_4$, 17% glycerol, 0.2 mM EDTA, 4 mM $MgCl_2$, 0.05 mM $ZnSO_4$, 1 mM DTT, and protease inhibitors and centrifuged at 120,000 × *g* for 1 h at 4 °C. The recombinant proteins with His tag were purified using Ni-NTA-agarose (Qiagen). Each fraction was analyzed by SDS-PAGE and Western immunoblotting with anti-His_6 (Covance), anti-HA 12CA5, and anti-FLAG M2 (Covance) monoclonal antibodies.

Plasmid Construction—Expression plasmids for the yeast two-hybrid assay and yeast one-hybrid assay were constructed using pLexA/NLS, which were LexA-fused protein expression vectors carrying the *TRP1* gene, and using pGAD424 for GAL4-AD-fused protein expression vector with the *LEU3* gene (16).

Interaction Assay by Yeast Two-hybrid System—A qualitative interaction assay was performed to measure *HIS3* gene expression (16). A pair of fusion gene plasmids was introduced into yeast strain L40 by the standard lithium acetate transformation

procedure. Transformants were plated on an SD agar plate containing 10 mM 3-amino-1,2,4-triazole (3-AT) without histidine, leucine, tryptophan, lysine, and uracil and incubated overnight at 30 °C.

In Vivo Transcriptional Activity Assay by Yeast One-hybrid System—The quantitative yeast one-hybrid assay was performed to measure the expression of the β-galactosidase gene under the control of four tandem repeated LexA-binding sequences. Equal amounts of logarithmically growing yeast transformants expressing each LexA hybrid protein were subjected to β-galactosidase activity assay (17).

Preparation of Transgenic Silkworms—The Awh ORF was amplified by using primers 5′-agtctagaatgaagacggagcaccgcac-3′ and 5′-agtctagatcagacttcactctgcatgc-3′ and inserted into the BlnI site of the pBacUASMCS vector (18), which has CFP gene as a screening marker. The plasmid was injected into *w1-pnd* embryos to obtain the UAS-Awh strains. The established strains were crossed with the hs-GAL4 strain (19).

RESULTS

AT-rich Sequences of E Site in En I Are Essential for SGF-2 Binding—Our previous results showed that SGF-2 binds to both C and E sites in the upstream enhancer element En I of the fibroin gene, with stronger preference for the E site (10) (Fig. 1*A*). For further investigation of the sequence important for binding to SGF-2, EMSA was performed using a series of mutant E sites (Fig. 1*B*). Two AT-rich sequences (Fig. 1*B*, *boxed*) are critical for SGF-2 binding, overlap with the protected sequences in an *in vitro* footprint assay using V2 PSG extract, and contain homeodomain protein-binding sequences. A similar AT-rich sequence is found in the C site. The importance of these regions for preferential transcription of the fibroin gene in the PSG extracts has been demonstrated repeatedly previously (7–9).

Purification of SGF-2—SGF-2 was purified from V2 PSG extract through six chromatographic steps (Fig. 1*C*). Fig. 1*D* depicts a silver-stained SDS-PAGE gel containing active fractions from the third step of the purification using DNA-immobilized beads. In this step, we used not only EW oligonucleotide with an intact E site but also with two mutant oligonucleotides, EgcW and EgcM. EgcW contains mutations but maintains SGF-2 binding activity, whereas mutations in EgcM completely abolish SGF-2 binding. The elution profile of SGF-2 activity in the fourth step of purification (size exclusion chromatography using Bio-Silect 250) correlated with EW oligonucleotide-specific polypeptides visualized on silver-stained SDS-PAGE (Fig. 2). The native molecular mass of SGF-2 activity was estimated as about 1.1 MDa by gel filtration chromatography.

In the fifth step of Mini S ion exchange chromatography purification, the bound proteins were eluted by 6 M urea and 6 M guanidine HCl (Fig. 1*C*). The urea-eluted fraction contained six polypeptides of 48, 47, 45, 33, 32, and 30 kDa (Fig. 2*C*, *lanes 5–8* and *7′*), and the guanidine HCl-eluted fraction consisted of five peptides with molecular masses of 73, 55, 50, 47, and 36 kDa (Fig. 2*C*, *lanes 9* and *10*). In the sixth step, the urea-eluted fraction was further fractionated by Mini Q ion exchange chromatography into three polypeptides of 48, 47, and 45 kDa in the guanidine HCl-eluted fraction and 3 polypeptides of 33, 32, and 30 kDa in the flow-through (Fig. 1*C*).

cDNA Cloning of SGF-2 Components—To identify the components of SGF-2, amino acid sequence analysis of the purified polypeptides was performed (Table 1). This analysis revealed that p33, p32, and p30 are derived from the silk protein P25/fhx, which was identified as a fibroin-associated protein (20, 21), and other proteins represent novel *Bombyx* gene products. The peptide sequences from p55, p50B ("B" indicates light brown protein bands in the silver-stained gel), and p47B were mostly identical, and so were those from p48, p47G ("G" indicates gray protein bands), and p45. These results suggest that p55/p50B/p47B and p48/p47G/p45 might represent products from two distinct genes by alternative splicing, respectively.

A 2.2-kb cDNA clone for the 36-kDa protein encodes an LIM-HD protein of 274 amino acids. Because the deduced amino acid sequence is highly homologous to that of the *Drosophila* Arrowhead protein (22) and orthologues in other species (Fig. 3*A*), we named the protein as *Bombyx* Arrowhead (Awh).

Next, we isolated the cDNA clones for p47B. The predicted protein product, which contains 357 amino acid residues, is highly homologous to mouse Ldb1/NLI/CLIM-2, CLIM-1, and *Xenopus* XLdb1 (23–25). It possesses a LIM domain-interacting domain (LID), which is identical among all Ldb proteins (Fig. 3*B*) (26, 27). We designated this protein as *Bombyx* LIM domain-binding protein (Ldb). Supporting our notion that p47B, p50B, and p55 are products derived from the same gene, their peptide sequences are found in the predicted amino acid sequence of Ldb.

Finally, we isolated the cDNA for the SGF-2 components p48, p47G, and p45. A 3.0-kb cDNA clone, which encodes a novel protein of 357 amino acid residues containing all peptide sequences from p48, p47G, and p45, was isolated and designated as Lcaf (LIM-HD and Ldb complex-associated factor). We searched the DNA database for molecules related to Lcaf and identified SSDP as the closest relative in the vertebrate (Fig. 3*C*). The amino acid sequence of Lcaf shows high similarity to that of various vertebrate SSDPs, especially in the N-terminal 92-amino acid sequence. Interestingly, although SSDP was reported originally as a factor binding to the DNase I hypersensitive region of chicken α2(I) collagen gene promoter (28), it was also identified as a factor interacting with Ldb proteins (29, 30).

SGF-2 Subunits Show Restricted or Preferential Expression in PSG—SGF-2 is detected in the extract of PSG, but not of MSG (10). Northern blot analysis using total RNA derived from the posterior or middle portion of the fifth instar silk glands showed that *Awh* and *P25/fhx* transcripts were only detected in PSG (Fig. 3*D*). On the other hand, *Ldb* and *Lcaf* transcripts were found in both regions of the silk gland, but preferentially in the posterior portion.

Lcaf Forms a DNA-binding Protein Complex with Awh and Ldb—To examine whether Lcaf forms a complex with other SGF-2 subunits Awh, Ldb, and P25/fhx, all four proteins were co-expressed in Sf9 insect cells by using the baculovirus expression system. We constructed recombinant baculoviruses expressing each of the following: HA-tagged Awh (ha:Awh),

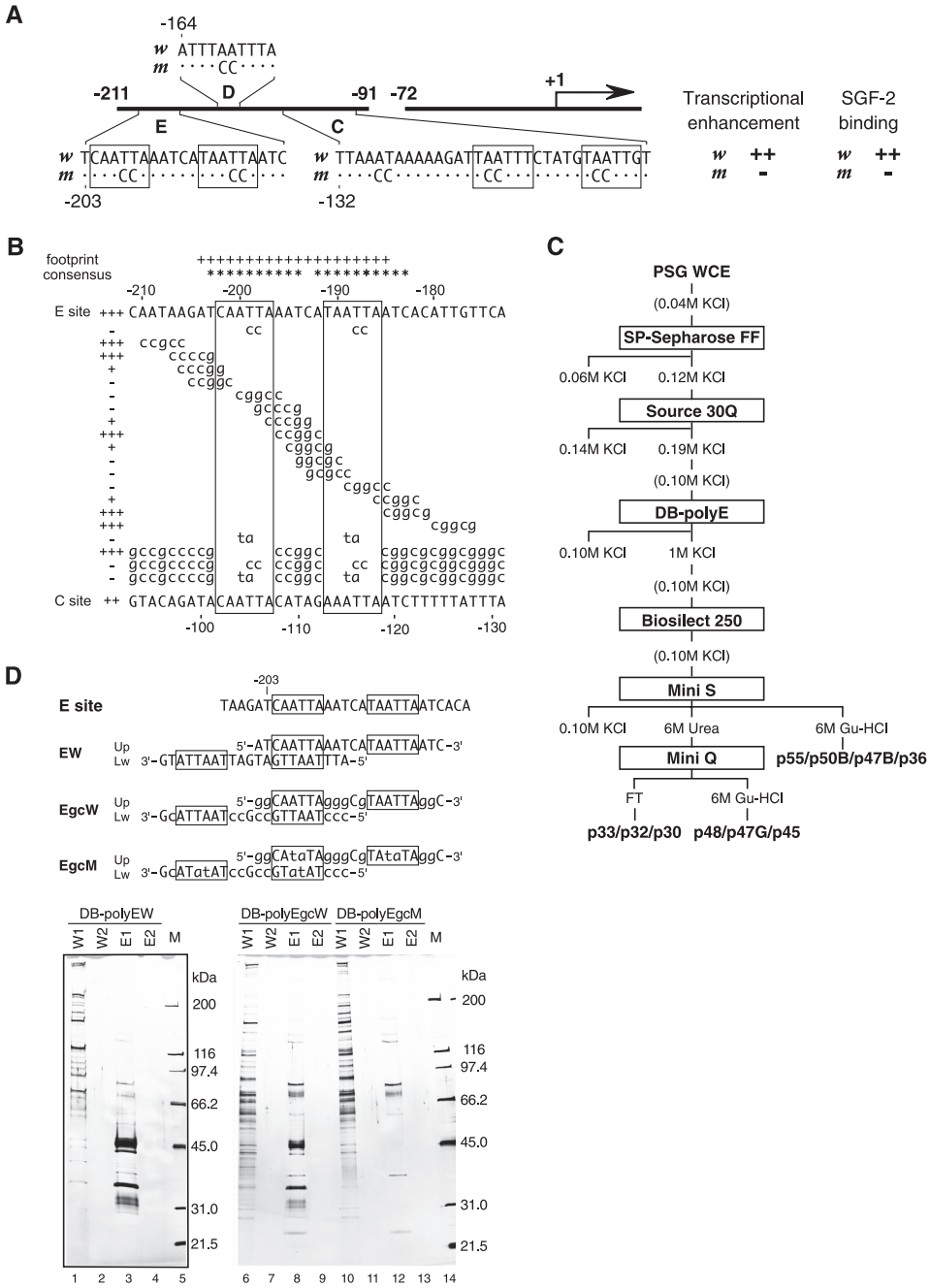


FIGURE 1. **Purification of SGF-2.** *A*, effects of mutations in AT-rich regions of En I element (−211 to ~−91) on PSG-specific transcriptional enhancement and SGF-2 binding are summarized. The AT-rich sequences (E box) recognized by SGF-2 are *boxed*. *B*, the binding activity of each mutation construct of the E site to SGF-2 in EMSA was indicated as + and − on the *left side*. The result of *in vitro* footprint assay performed by Hui *et al.* (10) was also summarized. + at the top indicates positions protected by SGF-2 in footprint assay. The "consensus" at the *top* indicates the homeodomain protein-binding sequences. The nucleotide sequence of the C site is shown at the *bottom line*. *C*, column chromatography procedures of SGF-2 purification. *PSG WCE*, whole cell extract from the posterior silk glands. *D*, comparison of peptides purified by DNA-immobilizing beads (*DB*) with intact and mutated E box from Source 30Q-purified SGF-2 fraction. The *top panel* represents the sequence of one unit of DNA. EW and EgcW contain the intact sequence of the E box, whereas EgcM has mutated sequence. These DNAs were polymerized and immobilized onto beads. The *bottom panel* represents silver-stained SDS-PAGE gels of fractions bound to each DNA-immobilized bead. *W1*, wash fraction 1 with TEMGTK100; *W2*, wash fraction with TEMGK100; *E1*, eluate fraction 1 with TEMGTK1000; *E2*, eluate fraction 2 with TEMGTK1000; *M*, size marker.

FLAG-tagged Ldb (f:Ldb), His-tagged Lcaf (h:Lcaf), and Myc-tagged P25/fhx (m:P25/fhx). When cells were infected with baculovirus expressing ha:Awh, f:Ldb, or m:P25/fhx individually, the recombinant proteins were insoluble and not recovered well. When cells were infected with the h:Lcaf baculovirus, a 45-kDa protein band together with a minor protein band just above it were detected in the affinity-purified fraction (Fig. 4*A*, *left panel*, *lane 8*). On the other hand, when cells were co-infected with ha:Awh, f:Ldb, and h:Lcaf baculoviruses and h:Lcaf protein was purified with nickel affinity chromatography, the 45-kDa protein was co-purified with several proteins in an almost stoichiometric manner. Immunoblotting analysis using anti-His_6, anti-HA, and anti-FLAG antibodies showed that the proteins co-purified with h:Lcaf were ha:Awh and f:Ldb (Fig.

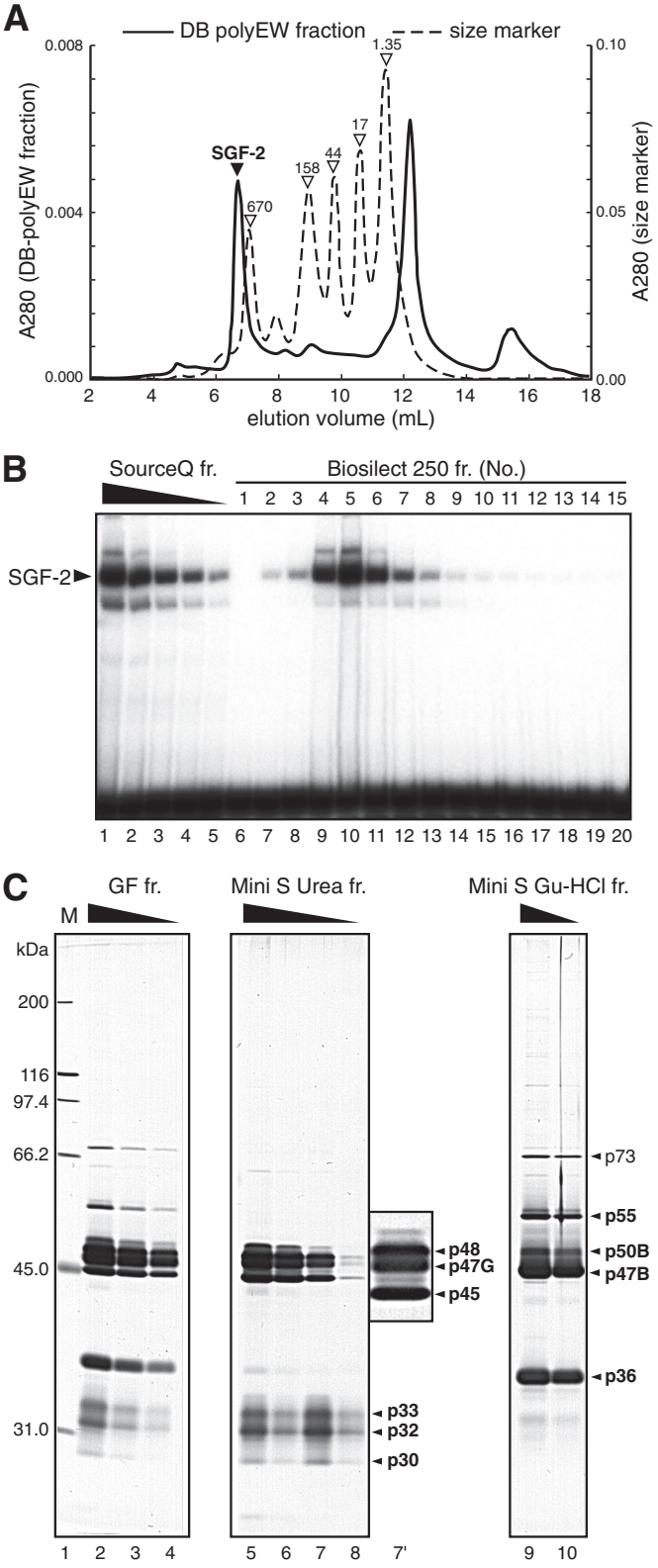


FIGURE 2. **SGF-2 is a 1.1-MDa heteromeric complex.** *A*, chromatography chart of SGF-2 activity passed through a Bio-Silect 250 gel filtration column. *B*, EMSA was performed using the E site probe, and each fraction passed through a Bio-Silect 250 column. *Source Q fr* is the peak fraction from the Source 30Q column, and the protein amount changed in *lanes 1–5*. *C*, silver-stained SDS-PAGE gel of Bio-Silect 250 (*lanes 2–4*), Mini S 6 M urea (*lanes 5–8*), and Mini S 6 M Gu-HCl (*lanes 9* and *10*) fractions. *Lane 7′* shows a magnified image of *lane 7*. *M* in *lane 1* is a size marker; *fr.*, fraction; *Gu-HCl*, guanidine HCl.

TABLE 1

Amino acid sequences of digested peptides from SGF-2 components

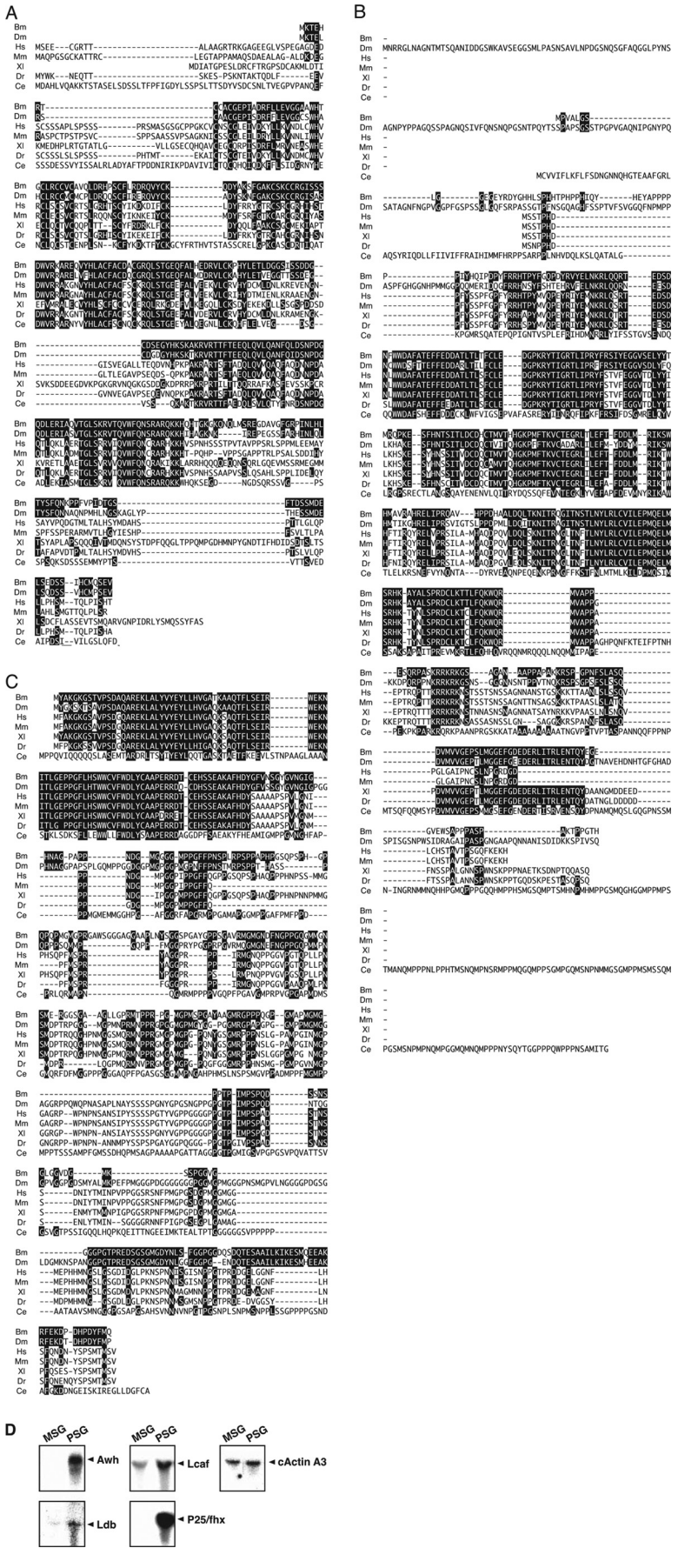
Oligonucleotides encoding underlined sequences were used for RT-PCR to amplify partial cDNA fragments of the corresponding SGF-2 components. a, sequence was from undigested p36 protein. a′, sequence was included in a. b and d–f were the same sequence. b includes the c′ sequence. g, sequences were found in every proteins of p48, p47G, and p45.

p48	#1	LALYVYEYLLHVGAT [(g)]
p47G	#2	AAQTFLSEIR [(g)]
p45	#3	NITLGEPPGF [(g)]
	#4	SSPGGVGGGGPGTP [(g)]
	#5	EDSGSGMGDYNLSFGG-PGGDQSDQTESAAIL [(g)]
	#6	IKESMQEER [(g)]
	#7	FEKDPDHPDYFM [(g)]
p55	#1	ELIPR [(e)]
	#2	HTPYFGQPDYR [(b)]
p50B	#1	HTPYFGQPDYR [(b)]
	#2	TLIPR
	#3	LILEFTFDDLMR [(c′)]
	#4	SWHMAVR [(d)]
	#5	ELIPR [(e)]
	#6	LCVILEPMQELMSR [(f)]
	#7	TTLFQK
p47B	#1	VALGSL
	#2	HTPYFGQPDYR [(b)]
	#3	VYELNK
	#4	VCTEGRLILEFTFDDLMR [(c)]
	#5	SWHMAVR [(d)]
	#6	ELIPR [(e)]
	#7	LCVILEPMQELMSR [(f)]
	#8	WQR
	#9	KGSAGANAAP
	#10	VALG
p36	#1	MLTEHRTCCACGEPIAD-RFLLEVGGAAWHT [(a)]
	#2	TCCACGEPIADR [(a′)]
	#3	FLLEVGGAAWHTGCLR
	#4	CCVCAVQLDR
	#5	HPSCFLR
	#6	QVYCK
	#7	GISSSDWVR
	#8	EQVYHLACF
	#9	QLSTGEQFALHED
	#10	VLCKP
	#11	VTQVWFQNR
	#12	QNQLMSR
	#13	PINLHLTY
p33	#1	GQIPSQYEIPVFQFEIPYF
	#2	ATYVD
	#3	NLI
p32	#1	EDVVLSFYIDGSYS
p30	#1	MLAR
	#2	PCYLDDYK
	#3	CIPGR
	#4	NHDQCR
	#5	TLAQHMSFK
	#6	LTTVFDK
	#7	TAQWLSK
	#8	EHIFGK
	#9	NWLAR
	#10	TLCDFGCQH

4*A*, *right panel*). Co-infection of Sf9 cells with m:fhx/P25 baculovirus and the other three baculoviruses was performed, but we did not detect integration of P25/fhx protein into the Awh·Ldb·Lcaf complex.

To examine the possible DNA binding activity of the h:Lcaf complex, EMSA was performed. As shown in Fig. 4*B*, DNA binding activity to the E site was detected in the 60 and 80 mM imidazole fractions of nickel affinity chromatography. The complex migrated slightly faster than native SGF-2 purified from PSG extract (Fig. 4*B*, compare *lane 14* with *lane 15*). This DNA·protein complex is supershifted by the addition of antibodies against HA, FLAG, and His epitopes, but not by anti-Myc antibody (Fig. 4*C*). Most importantly, similar to SGF-2, this complex was specifically abolished by the addition of anti-SGF-2 antibody (Fig. 4*C*, compare *lanes 1* and *2* with *lanes 4* and *9*). These results clearly illustrate that SGF-2-like complex with specific DNA binding activity to the E site can be reconstituted by recombinant proteins encoded by *Awh*, *Ldb*, and *Lcaf* cDNA.

We also purified the protein complex from Sf9 cells co-infected with f:Ldb- and ha:Awh-expressing baculoviruses using FLAG tag and examined its DNA binding activity. The f:Ldb complex could bind specifically to the E site in EMSA, but the DNA·protein complex migrated much faster than that of the h:Lcaf complex (Fig. 4*D*, *lanes 1* and *3*). The DNA·protein complex was confirmed to contain both f:Ldb and ha:Awh by supershift migration using anti-HA and anti-FLAG antibodies, but was not affected by anti-SGF-2 antibody. These results demonstrate that the complex of Awh and Ldb is sufficient for the specific binding to the E site, but is not equivalent to the purified SGF-2.


A
B
C
D
MSG
PSG
Awh
Lcaf
cActin A3
Ldb
P25/fhx

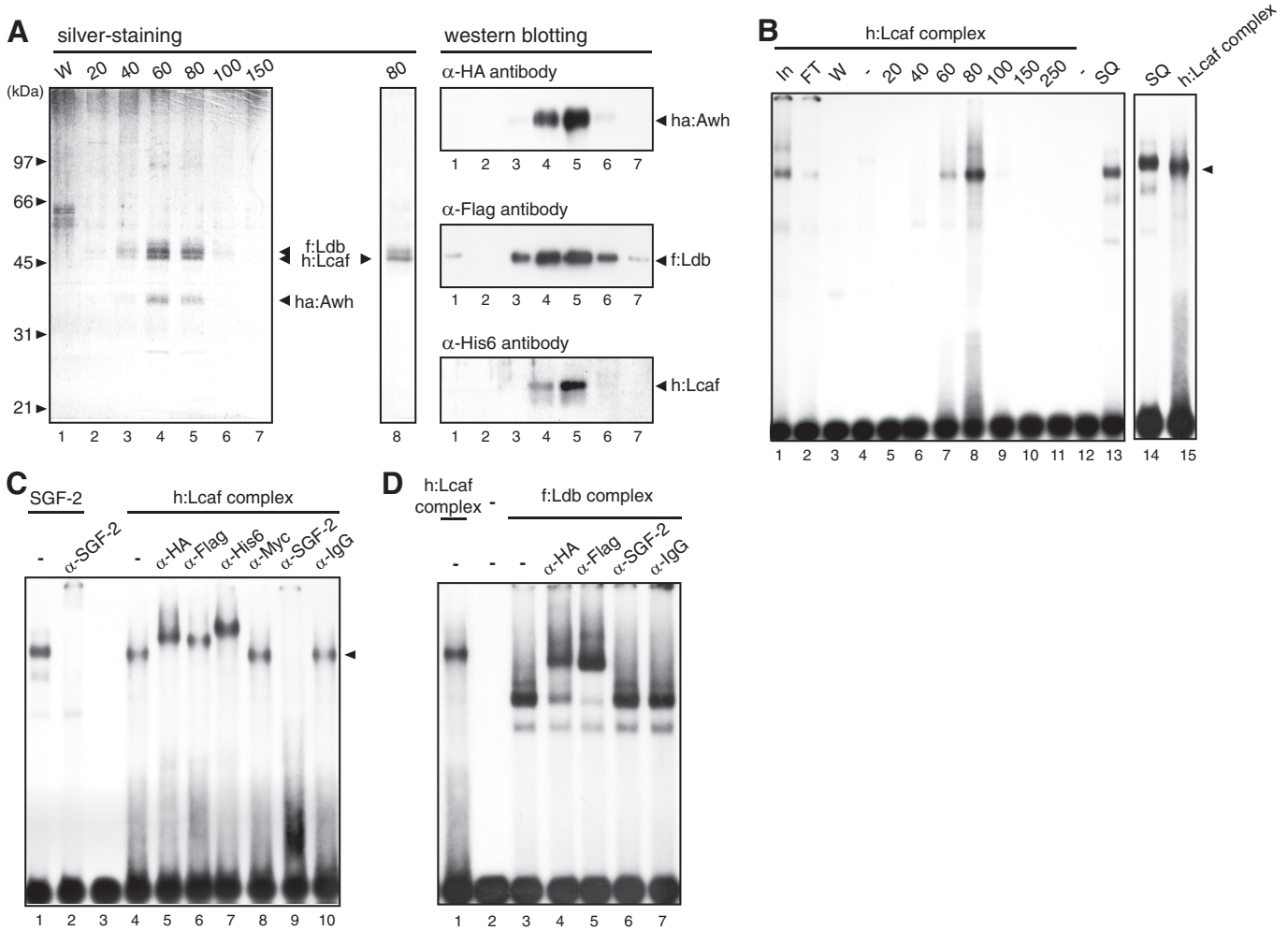


FIGURE 4. **Lcaf forms a DNA-binding protein complex with Awh and Ldb.** *A*, *left*, silver-stained SDS-PAGE gel of Ni-NTA-agarose purification fractions from the extract of Sf9 cells co-expressed with ha:Awh, f:Ldb, and h:Lcaf by the baculovirus expression system. *W* in *lane 1*: wash fraction of Ni-NTA-agarose. *Numbers at the top* indicate imidazole concentration (in mM) for elution. *Lane 8*, the purified fraction from Sf9 cell extract expressing only h:Lcaf. The character of native Ldb protein with light-brown color in the silver-stained gel was also observed for recombinant Ldb protein, but the slightly slower migrating band than the h:Lcaf band in *lane 8* was gray-colored, not light brown; therefore it seemed not to be Ldb. *Right*, Western blot analysis using Ni-NTA-agarose purification fractions. Each lane corresponds to that of the silver-stained SDS-PAGE gel shown on the *left*. *B*, EMSA using Ni-NTA-agarose fractions of h:Lcaf complex with the E box DNA probe. *Numbers* and *SQ* at the *top* indicate imidazole concentration (in mM) for elution and partial purified native SGF-2 fraction after Source 30Q column, respectively. *Lanes In* and *FT* indicate the extracts (*In*) and flow-through (*FT*) fraction of Ni-NTA-agarose purification, respectively. *C*, EMSA of h:Lcaf complex with antibodies. The h:Lcaf complex eluted by 80 mM imidazole was used with each antibody indicated at the *top*. Anti-IgG and anti-c-Myc antibodies were used as negative controls. *D,* EMSA of f:Ldb complex with antibodies. The f:Ldb complex eluted by 80 mM imidazole was used with each antibody indicated at the top.

Induction of Ectopic Expression of the Fibroin Gene by Awh in MSG—To investigate whether SGF-2 is a tissue-specific transcriptional activator of the fibroin gene, we generated transgenic silkworms that possess a *UAS-Awh* transgene in which *Bombyx Awh* was under the control of a UAS promoter. *UAS-Awh* silkworms were crossed with *hs-GAL4* transgenic silkworms, and *hs-GAL4/UAS-Awh* offspring were selected. These transgenic worms were kept at 42 °C for 2 h on day 1 of the fourth instar, and the expression of the fibroin gene was then analyzed. Strikingly, by misexpression of the *Awh* gene in transgenic worms, the fibroin gene was induced in MSG (Fig. 5), where *Ldb* and *Lcaf* genes are expressed (Fig. 3*D*), indicating that Awh protein is a PSG-specific activator of the fibroin gene.

Self-association of Lcaf—To compare the size of the h:Lcaf complex with native SGF-2, gel filtration chromatography was performed. The elution profiles of the h:Lcaf complex are shown in the *top panel* of Fig. 6*A*, in which the majority was eluted in a peak corresponding to a molecular mass of ~800 kDa. Although previous studies showed that the LIM-HD and Ldb proteins bound to each other to form a heterotetrameric complex *in vitro* (24, 31), it is still possible that Lcaf could oligomerize by itself. When h:Lcaf protein, expressed by the bacu-

FIGURE 3. **Deduced amino acid sequences of SGF-2 subunits and their restricted expression in the posterior silk gland.** *A*, deduced amino acid sequence of p36 protein, *Bombyx* Awh, is compared with orthologues of other animals. Identical amino acids are indicated with *white letters* on *black*. *Bm*, *B. mori*; *Dm*, *Drosophila melanogaster*; *Hs*, *Homo sapience*; *Mm*, *Mus musculus*; *Xl*, *Xenopus laevis*; *Dr*, *Danio rerio*; *Ce*, *Caenorhabditis elegans*. *B*, deduced amino acid sequence of p48 protein, *Bombyx* Ldb, is compared with orthologues of other animals. *C*, deduced amino acid sequence of p45 protein, *Bombyx* Lcaf, is compared with orthologues of other animals. *D*, Northern blot analysis was performed using total RNA derived from middle and posterior silk gland at the fifth instar. *cActin A3*: cytoplasmic actin A3, used as a control.

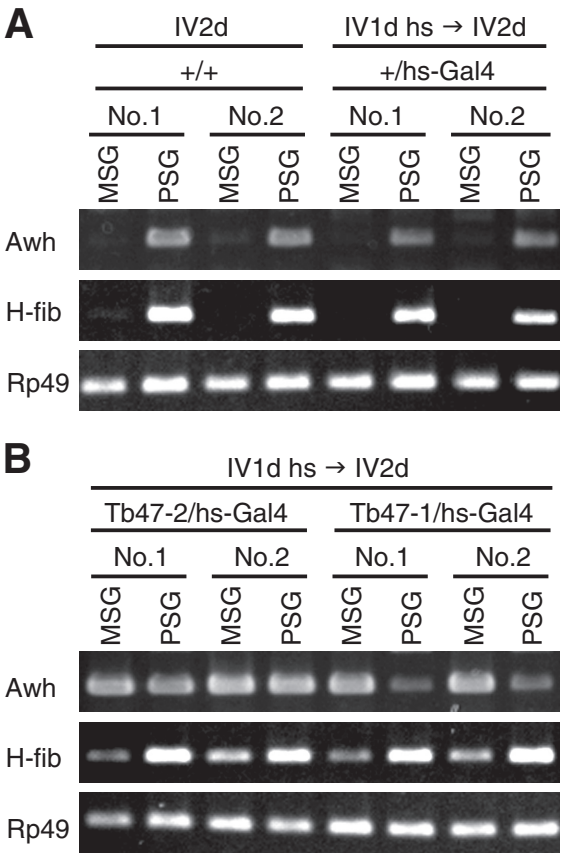


FIGURE 5. **Misexpression of *Awh* induced expression of the fibroin gene in MSG.** *A*, total RNA of MSG or PSG was prepared individually from two larvae (*No. 1* and *No. 2*) of the wild type (+/+) silkworms or the silkworms carrying *hs-Gal4* transgene alone (+/hs-Gal4), with (*IV1d hs → IV2d*) or without (*IV2d*) heat shock at 42 °C. The cycle numbers PCR are 35 for *Awh* and 21 for *H-fib* and *Rp49* control. *B*, RT-PCR was performed with total RNA from two strains of transgenic worms (Tb47-1/hs-Gal4 and Tb47-2/hs-Gal4) with *UAS-Awh* transgene in different locus.

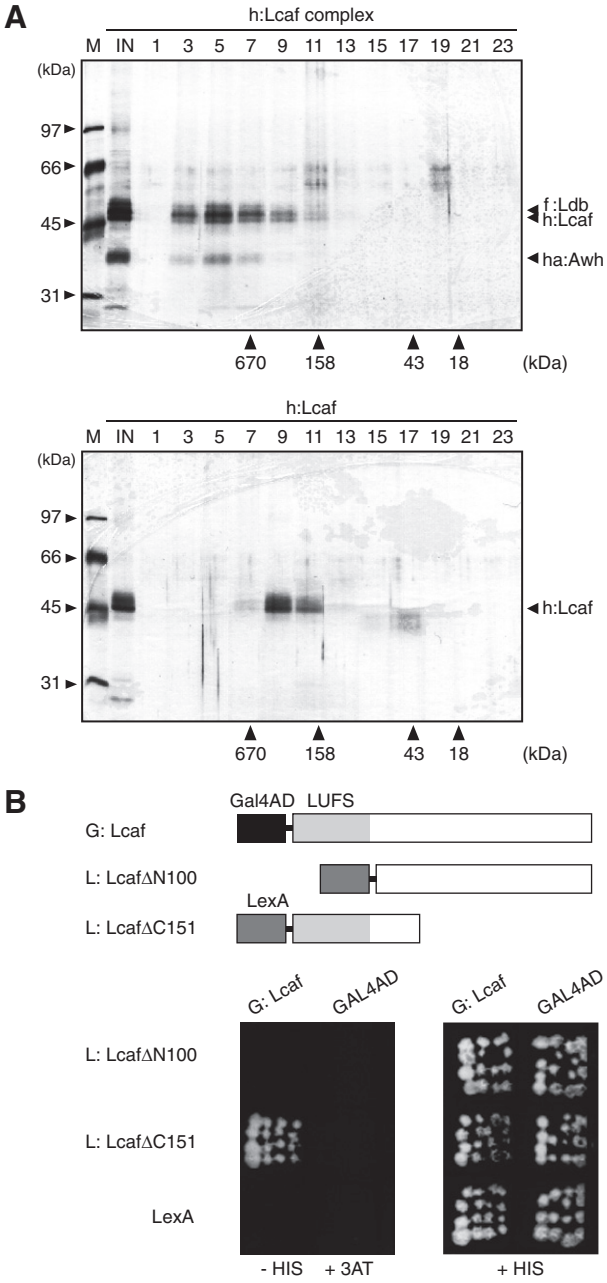


FIGURE 6. **Self-association of Lcaf protein.** *A*, elution profiles of the h:Lcaf complex (*top*) and h:Lcaf protein (*bottom*) in gel filtration chromatography. *Numbers at the top* indicate fraction numbers of the chromatography. Elution of markers is shown on the horizontal axis. *M* and *IN* at the *top* are the molecular mass marker and input protein fraction on chromatography, respectively. *B*, self-association of Lcaf in yeast two-hybrid analysis. *HIS3* gene was used as a reporter gene. *Top*, schematic structures of Lcaf hybrid proteins. *LUFS*, the conserved domain, named after the four founding members LUG, LUH (for LEUNIG_HOMOLOG), yeast Flo8, and human SSDP. *Bottom*, each yeast transformant was inoculated on an SD agar plate containing 10 mM 3-AT without histidine or on SD agar with histidine.

lovirus system and purified by nickel affinity chromatography, was subjected to gel filtration chromatography, the majority of h:Lcaf was eluted in a peak corresponding to a molecular mass of ~300 kDa (Fig. 6*A*, *bottom panel*). This is equivalent to almost six times the predicted molecular mass of a sole h:Lcaf molecule (48 kDa). The self-association ability of the Lcaf protein was confirmed by a yeast two-hybrid system using the GAL4 activation domain (GAL4-AD) and LexA as a DNA-binding portion. We prepared expression plasmids for Lcaf fused to an N-terminal Gal4-AD, called G:Lcaf, along with two Lcaf truncated mutants fused to an N-terminal LexA, which contained the N-terminal region (1–150 amino acids) or the C-terminal region (101–357 amino acids) of Lcaf, named L:LcafΔC151 and L:LcafΔN100, respectively. As shown in Fig. 6*B*, G:Lcaf interacted with L:LcafΔC151, but not with L:LcafΔN100, in yeast. These findings indicate that Lcaf protein can form a homo-oligomer through its N-terminal 100-amino acid sequence, which may contribute to the formation of a huge h:Lcaf complex with ha:Awh and f:Ldb.

Lcaf Interacts with Ldb—Co-purification of ha:Awh and f:Ldb proteins with h:Lcaf suggested possible direct interactions of Lcaf with Awh and Ldb. The yeast two-hybrid system was used to examine this possibility. We constructed expression plasmids for Ldb as LexA fusion L:Ldb and Awh as GAL4-AD fusion G:Awh (Fig. 7*A*). Yeast transformants co-expressing L:Ldb and GAL4-AD did not grow on an SD agar plate containing 10 mM 3-AT without histidine. When L:Ldb was co-expressed with G:Awh or G:Lcaf in the reporter yeast strain, both transformants were able to grow under the same conditions (Fig. 7*B*).

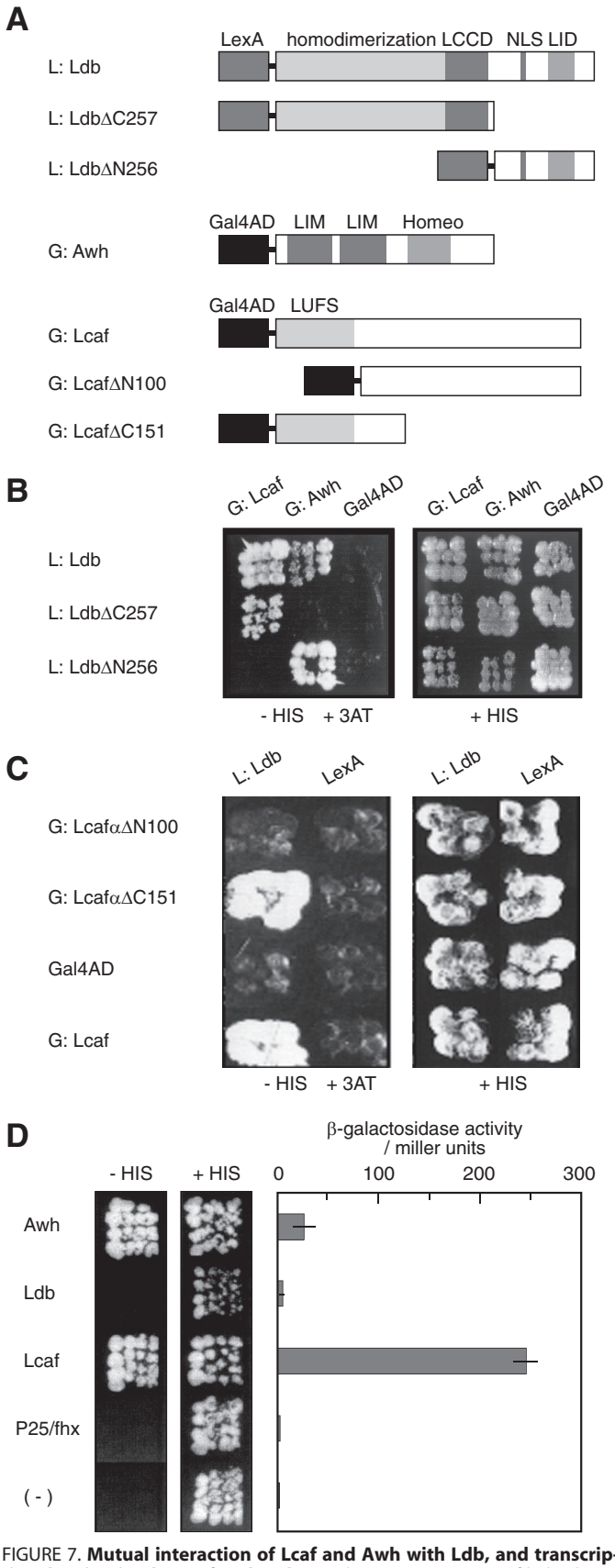


FIGURE 7. **Mutual interaction of Lcaf and Awh with Ldb, and transcriptional activation by Lcaf and Awh.** *A*, schematic structures of hybrid proteins used in the yeast two-hybrid assay. *LCCD,* Ldb/CHIP conserved domain; *NLS*, nuclear localization sequence; *LID*, LIM domain-interacting domain. *B*, Lcaf and Awh interact with the amino- (1–256 amino acids) and C-terminal (257–376 amino acids) portion of Ldb, respectively, in the yeast two-hybrid assay. *C*, Ldb interacts with the N-terminal portion (1–151 amino acids) of Lcaf. *HIS3* gene was used as a reporter gene. Each yeast transformant was inoculated onto an SD agar plate containing 10 mM 3-AT without histidine or on SD agar with histidine. *D*, yeast one-hybrid assay was performed using each SGF-2 subunit fused with LexA and nuclear location signal. *HIS3* gene and β-galactosidase gene were used as reporters. *Left panel* shows that yeast transformants expressing Awh and Lcaf grew well in medium containing 10 mM 3-AT without histidine. *Right panel* represents β-galactosidase activity of each yeast extract. All results are the mean ± S.E. of at least four independent transformants.

To define the regions on Ldb involved in the interactions with Awh and Lcaf, two LexA hybrid proteins called L:LdbΔC257 and L:LdbΔN256 that contain the N-terminal region (1–256) and the C-terminal region (257–376) of Ldb, respectively, were examined. Our results showed that G:Lcaf interacts with L:LdbΔC257 (*i.e.* the N-terminal portion of Ldb), whereas G:Awh binds to L:LdbΔN256 (*i.e.* the C-terminal portion of Ldb) in yeast (Fig. 7*B*). These observations suggest that Ldb interacts with Awh and Lcaf through distinct binding domains. To determine which portion of Lcaf is necessary for the Lcaf-Ldb interaction, we used the two truncation mutants of Lcaf. Although G:LcafΔC151 hybrid protein maintained its interaction with L:Ldb in the reporter yeast strain, G:LcafΔN100 lost this ability under the same condition (Fig. 7*C*), suggesting that the N-terminal 100 amino acid sequence of Lcaf is necessary for Lcaf-Ldb interaction.

Awh and Lcaf Contribute to Transcriptional Activation—To examine which SGF-2 subunits identified here play a significant role in transcriptional activation, we performed experiments using the yeast one-hybrid system. Yeast transformants expressing Awh fused to the N-terminal LexA, L:Awh, or L:Lcaf grew well on medium containing 10 mM 3-AT without histidine and were also positive for β-galactosidase activity (Fig. 7*D*). The β-galactosidase activity of the L:Lcaf-expressing transformant was about 8-fold stronger than those expressing L:Awh. On the other hand, yeast transformants expressing L:Ldb or L:P25/fhx did not grow under the same condition and exhibited little or no β-galactosidase activity. These results indicate that both Awh and Lcaf possess intrinsic transcriptional activation ability.

DISCUSSION

SGF-2 Is a Transcriptional Activator Complex of the Fibroin Gene—SGF-2 was originally identified by EMSA in extracts of PSG on the basis of the binding activity to the fibroin En I element and is thought to be a key transactivator for the fibroin gene (6–10). We identified four proteins, Awh, Ldb, Lcaf, and P25/fhx, as components of SGF-2. Several lines of evidence support that these proteins constitute SGF-2 and promote transcriptional activation of the fibroin gene. 1) Recombinant Awh, Ldb, and Lcaf proteins formed a complex with specific DNA binding activity to the E site. 2) The protein complex of recombinant Awh, Ldb, and Lcaf was recognized by anti-SGF-2 antibody. 3) *Awh*, *Ldb*, *Lcaf*, and *P25*/*fhx* are specifically or preferentially expressed in PSG. 4) The purified complex of recombinant Awh, Ldb, and Lcaf exhibited almost the same mass as purified SGF-2. 5) Awh and Lcaf showed transcriptional activation activity in yeast one-hybrid system. 6) Misexpression of

Awh, which is normally restricted to PSG, induced ectopic expression of the fibroin gene in MSG of transgenic silkworms. Another indirect evidence supporting the possible contribution of Awh to tissue- and developmental stage-specific transcriptional activation of the fibroin gene is that in *Drosophila* transgenic lines carrying the fibroin promoter fused to the *β-galactosidase* gene, reporter gene expression is restricted to anterior cells of the larval salivary gland, where the *Drosophila Awh* gene is specifically expressed (22, 32).

The silkworm Awh and Ldb are members of the LIM-HD family of transcription factors and the Ldb protein family, respectively. The LIM-HD·Ldb complex appears to be a critical regulator during development and functions as a transcriptional activator (23, 24, 33, 34). Our finding that SGF-2 contains Awh and Ldb is consistent with the notion that Ldb proteins are a requisite component of many transcriptional regulatory complexes involving LIM-HD factors.

P25/fhx is known to be a component of the 2.3-MDa secretory elementary unit of silk fibroin (35). Because *P25/fhx* is not expressed in MSG and misexpression of *Awh* can induce fibroin gene expression in MSG, P25/fhx appears to be a nonessential component in the SGF-2 complex. However it will be intriguing to speculate that P25/fhx might play a role in fine-tuning the molecular ratio of fibroin to P25/fhx by regulating the transcription of the fibroin gene through SGF-2. P25/fhx is a glycoprotein (36). The *N*-linked oligosaccharide chains of P25/fhx are important for maintaining the 2.3-MDa complex of fibroin elementary unit (35). However, recombinant P25/fhx protein in Sf9 cells seemed not to be glycosylated on the mobility on SDS-PAGE. It is possible that glycosylation of P25/fhx is important for its integration into Awh·Ldb·Lcaf complex.

Lcaf Is an Additional Component of LIM-HD·Ldb Complex—The present study identified another protein, Lcaf, a member of the SSDP family, which could transform the Awh·Ldb protein complex into a larger protein complex due to its oligomerization activity. The N-terminal amino acid sequence (~100 amino acids) of Lcaf is almost identical to that of SSDP, which was originally identified as a nuclear protein that binds to the single-stranded pyrimidine-rich element in the chicken α2(I) collagen gene promoter (28). This promoter element is well conserved among different mammalian species and located in a region that is DNase I-hypersensitive only when the promoter is active. These observations have led investigators to believe that SSDP might be involved in the transcriptional regulation of the α2(I) collagen gene. The remarkable conservation of the N terminus between Lcaf and mammalian SSDP (37) suggests that silkworm Lcaf protein is a functional homologue of vertebrate SSDP. The conserved domain of Lcaf was necessary not only for its self-oligomerization ability but also for Ldb interaction.

Previous studies in flies and vertebrates have revealed the importance of LIM-HD and Ldb proteins in tissue patterning and differentiation (30). However, the molecular mechanisms in which LIM-HD·Ldb protein complex functions as transcriptional regulator are not fully understood. Genetic experiments imply that Chip, a *Drosophila* homologue of Ldb proteins, may mediate communication between enhancers and promoters (34, 38, 39). Given that the SSDP/Lcaf protein family is a requisite interaction partner for Ldb proteins, they might play a role in long range enhancer-promoter communication by cooperating with Ldb proteins. In this scenario, the potential sequence-specific single-stranded DNA binding activity, interaction activity with Ldb proteins, and self-oligomerization activity of the SSDP/Lcaf protein family could facilitate enhancer-promoter communication by gathering together sequence- and structure-specific *cis*-elements scattered throughout certain gene loci and by organizing transcriptional regulatory elements on chromatin to form particular higher order structures that support transcriptional regulation. From this point of view, it is important to stress that besides the En I region of the fibroin gene, a key region in the farther upstream enhancer element from −1659 to −1590 detected *in vivo* and localized near a DNase-hypersensitive site, also possesses an SGF-2-binding sequence (11). The mouse Ldb protein was found to occupy numerous DNase I-hypersensitive sites on chromatin across a region of ~130 kb in the mouse α-globin locus (40). Long range genomic interaction via Ldb1 and GATA1 was also reported in mammalian β-globin gene locus (41). It would be interesting to investigate whether the SSDP/Lcaf protein family can co-occupy the same positions as the Ldb proteins, and if so, how it could contribute to the regulation of developmental gene expression, such as the formation of intrachromosomal loops and histone modification (42). Recently, Brandt and co-workers (43, 44) reported that SSDPs regulate the activity of Ldb-containing complex through stabilization of Ldb proteins by interfering with proteasomal degradation. SSDPs may be a multifunctional component in transcriptional regulation.

Acknowledgments—We are grateful to the members of the Y. Suzuki, H. Handa, and H. Sezutsu laboratories and associated groups of Drs. H. Kokubo, C. Sawa, K. Morohashi, and Y. Nagahama and to E. Suzuki, M. Sasaki ,and M. Masuda for technical assistance and encouragement throughout this study. We also thank Dr. Kikawada for the preparation of total RNA from silk glands and Dr. Hollenberg for providing the yeast two-hybrid system.

REFERENCES

1. Suzuki, Y., Takiya, S., Suzuki, T, Hui, C.-C., Matsuno, K., Fukuta, M., Nagata, T., and Ueno, K. (1990) *Molecular Insect Science* (Hagedorn, H. H., Hildebrand, J. G., Kidwell, M. G., and Law, J. H., eds) pp. 83–89, Plenum Press, New York
2. Suzuki, Y., and Suzuki, E. (1974) Quantitative measurements of fibroin messenger RNA synthesis in the posterior silk gland of normal and mutant *Bombyx mori*. *J. Mol. Biol.* **88,** 393–407
3. Okamoto, H., Ishikawa, E., and Suzuki, Y. (1982) Structural analysis of sericin genes: homologies with fibroin gene in the 5′ flanking nucleotide sequences. *J. Biol. Chem.* **257,** 15192–15199
4. Couble, P., Michaille, J.-J., Garel, A., Couble, M.-L., and Prudhomme, J.-C. (1987) Developmental switches of sericin mRNA splicing in individual cells of *Bombyx mori* silkgland. *Dev. Biol.* **124,** 431–440
5. Takasu, Y., Hata, T., Uchino, K., and Zhang, Q. (2010) Identification of Ser2 proteins as major sericin components in the non-cocoon silk of *Bombyx mori*. *Insect Biochem. Mol. Biol.* **40,** 339–344
6. Tsuda, M., and Suzuki, Y. (1981) Faithful transcription initiation of fibroin gene in a homologous cell-free system reveals an enhancing effect of 5′ flanking sequence far upstream. *Cell* **27,** 175–182
7. Suzuki, Y., Tsuda, M., Takiya, S., Hirose, S., Suzuki, E., Kameda, M., and Ninaki, O. (1986) Tissue-specific transcription enhancement of the fibroin gene characterized by cell-free systems. *Proc. Natl. Acad. Sci. U.S.A.*

83, 9522–9526
8. Tsuda, M., and Suzuki, Y. (1983) Transcription modulation *in vitro* of the fibroin gene exerted by a 200-base-pair region upstream from the "TATA" box. *Proc. Natl. Acad. Sci. U.S.A.* **80,** 7442–7446
9. Takiya, S., Hui, C.-C., and Suzuki, Y. (1990) A contribution of the core-promoter and its surrounding regions to the preferential transcription of the fibroin gene in posterior silk gland extracts. *EMBO J.* **9,** 489–496
10. Hui, C.-C., Matsuno, K., and Suzuki, Y. (1990) Fibroin gene promoter contains a cluster of homeodomain binding sites that interact with three silk gland factors. *J. Mol. Biol.* **213,** 651–670
11. Shimizu, K., Ogawa, S., Hino, R., Adachi, T., Tomita, M., and Yoshizato, K. (2007) Structure and function of 5′-flanking regions of *Bombyx mori* fibroin heavy chain gene: identification of a novel transcription enhancing element with a homeodomain protein-binding motif. *Insect Biochem. Mol. Biol.* **37,** 713–725
12. Takiya, S., Kokubo, H., and Suzuki, Y. (1997) Transcriptional regulatory elements in the upstream and intron of the fibroin gene bind three specific factors POU-M1, Bm Fkh and FMBP-1. *Biochem. J.* **321,** 645–653
13. Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T., and Suzuki, Y. (1995) Silk gland factor-1 involved in the regulation of *Bombyx* sericin-1 gene contains fork head motif. *J. Biol. Chem.* **270,** 9340–9346
14. Takiya, S., Ishikawa, T., Ohtsuka, K., Nishita, Y., and Suzuki, Y. (2005) Fibroin-modulator-binding protein-1 (FMBP-1) contains a novel DNA-binding domain, repeats of the score and three amino acid peptide (STP), conserved from *Caenorhabditis elegans* to humans. *Nucleic Acids Res.* **33,** 786–795
15. Suzuki, T., and Suzuki, Y. (1988) Interaction of composite protein complex with the fibroin enhancer sequence. *J. Biol. Chem.* **263,** 5979–5986
16. Vojtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993) Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* **74,** 205–214
17. Savoysky, E., Mizuno, T., Sowa, Y., Watanabe, H., Sawada, J., Nomura, H., Ohsugi, Y., Handa, H., and Sakai, T. (1994) The retinoblastoma binding factor 1 (RBF-1) site in RB gene promoter binds preferentially E4TF1, a member of the Ets transcription factors family. *Oncogene* **9,** 1839–1846
18. Sakudoh, T., Sezutsu, H., Nakashima, T., Kobayashi, I., Fujimoto, H., Uchino, K., Banno, Y., Iwano, H., Maekawa, H., Tamura, T., Kataoka, H., and Tsuchida, K. (2007) Carotenoid silk coloration is controlled by a carotenoid-binding protein, a product of the *Yellow blood* gene. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 8941–8946
19. Uchino, K., Imamura, M., Sezutsu, H., Kobayashi, I., Kojima, K., Kanda, T., and Tamura, T. (2006) Evaluating promoter sequences for trapping an enhancer activity in the silkworm *Bombyx mori*. *J. Insect Biotech. Sericol.* **75,** 89–97
20. Couble, P., Moine, A., Garel, A., and Prudhomme, J.-C. (1983) Developmental variations of a nonfibroin mRNA of *Bombyx mori* silkgland, encoding for a low-molecular-weight silk protein. *Dev. Biol.* **97,** 398–407
21. Chevillard, M., Couble, P., and Prudhomme, J.-C. (1986) Complete nucleotide sequence of the gene encoding the *Bombyx mori* silk protein P25 and predicted amino acid sequence of the protein. *Nucleic Acids Res.* **14,** 6341–6342
22. Curtiss, J., and Heilig, J. S. (1997) *Arrowhead* encodes a LIM homeodomain protein that distinguishes subsets of *Drosophila* imaginal cells. *Dev. Biol.* **190,** 129–141
23. Agulnick, A. D., Taira, M., Breen, J. J., Tanaka, T., Dawid, I. B., and Westphal, H. (1996) Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins. *Nature* **384,** 270–272
24. Jurata, L. W., Kenny, D. A., and Gill, G. N. (1996) Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. *Proc. Natl. Acad. Sci. U.S.A.* **93,** 11693–11698
25. Bach, I., Carrière, C., Ostendorff, H. P., Andersen, B., and Rosenfeld, M. G. (1997) A family of LIM domain-associated cofactors confer transcriptional synergism between LIM and Otx homeodomain proteins. *Genes Dev.* **11,** 1370–1380
26. Breen, J. J., Agulnick, A. D., Westphal, H., and Dawid, I. B. (1998) Interactions between LIM domains and the LIM domain-binding protein Ldb1. *J. Biol. Chem.* **273,** 4712–4717
27. van Meyel, D. J., O'Keefe, D. D., Thor, S., Jurata, L. W., Gill, G. N., and Thomas, J. B. (2000) Chip is an essential cofactor for apterous in the regulation of axon guidance in *Drosophila*. *Development* **127,** 1823–1831
28. Bayarsaihan, D., Soto, R. J., and Lukens, L. N. (1998) Cloning and characterization of a novel sequence-specific single-stranded-DNA-binding protein. *Biochem. J.* **331,** 447–452
29. Chen, L., Segal, D., Hukriede, N. A., Podtelejnikov, A. V., Bayarsaihan, D., Kennison, J. A., Ogryzko, V. V., Dawid, I. B., and Westphal, H. (2002) Ssdp proteins interact with the LIM-domain-binding protein Ldb1 to regulate development. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 14320–14325
30. van Meyel, D. J., Thomas, J. B., and Agulnick, A. D. (2003) Ssdp proteins bind to LIM-interacting co-factors and regulate the activity of LIM-homeodomain protein complexes *in vivo*. *Development* **130,** 1915–1925
31. Hobert, O., and Westphal, H. (2000) Functions of LIM-homeobox genes. *Trends Genet.* **16,** 75–83
32. Nony P., Prudhomme J.-C., and Couble P. (1995) Regulation of the P25 gene transcription in the silk gland of *Bombyx*. *Biol. Cell* **84,** 43–52
33. Rincon-Limas, D. E., Lu, C.-H., Canal, I., and Botas, J. (2000) The level of DLDB/CHIP controls the activity of the LIM homeodomain protein apterous: evidence for a functional tetramer complex *in vivo*. *EMBO J.* **19,** 2602–2614
34. Morcillo, P., Rosen, C., Baylies, M. K., and Dorsett, D. (1997) Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in *Drosophila*. *Genes Dev.* **11,** 2729–2740
35. Inoue, S., Tanaka, K., Arisaka, F., Kimura, S., Ohtomo, K., and Mizuno, S. (2000) Silk fibroin of *Bombyx mori* is secreted, assembling a high molecular mass elementary unit consisting of H-chain, L-chain, and P25, with a 6:6:1 molar ratio. *J. Biol. Chem.* **275,** 40517–40528
36. Tanaka, K., Inoue, S., and Mizuno, S. (1999) Hydrophobic interaction of P25, containing Asn-linked oligosaccharide chains, with the H-L complex of silk fibroin produced by *Bombyx mori*. *Insect Biochem. Mol. Biol.* **29,** 269–276
37. Bayarsaihan, D. (2002) SSDP1 gene encodes a protein with a conserved N-terminal FORWARD domain. *Biochim. Biophys. Acta* **1599,** 152–155
38. Dorsett, D. (1999) Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr. Opin. Genet. Dev.* **9,** 505–514
39. Morcillo, P., Rosen, C., and Dorsett, D. (1996) Genes regulating the remote wing margin enhancer in the *Drosophila* cut locus. *Genetics* **144,** 1143–1154
40. Anguita, E., Hughes, J., Heyworth, C., Blobel, G. A., Wood, W. G., and Higgs, D. R. (2004) Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J.* **23,** 2841–2852
41. Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory P. D., and Dean, A. (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149,** 1233–1244
42. Dean, A. (2006) On a chromosome far, far away: LCRs and gene expression. *Trends Genet.* **22,** 38–45
43. Xu, Z., Meng, X., Cai, Y., Liang, H., Nagarajan, L., and Brandt S. J. (2007) Single-stranded DNA-binding proteins regulate the abundance of LIM domain and LIM domain-binding proteins. *Genes Dev.* **21,** 942–955
44. Wang, Y., Klumpp, S., Amin, H. M., Liang, H., Li, J., Estrov, Z., Zweidler-McKay, P., Brandt S. J., Agulnick, A., and Nagarajan, L. (2010) *SSBP2* is an *in vivo* tumor suppressor and regulator of LDB1 stability. *Oncogene* **29,** 3044–3053