The *Drosophila* Stubble-stubbloid gene encodes an apparent transmembrane serine protease required for epithelial morphogenesis

(imaginal discs/bristles/microfilaments/20-hydroxyecdysone)

Laurel F. Appel*, Mary Prout, Robin Abu-Shumays, Ann Hammonds, James C. Garbe, Dianne Fristrom, and James Fristrom^{\dagger}

Division of Genetics, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Communicated by Gerald M. Rubin, February 24, 1993

ABSTRACT The Stubble-stubbloid (Sb-sbd) gene is required for hormone-dependent epithelial morphogenesis of imaginal discs of Drosophila, including the formation of bristles, legs, and wings. The gene has been cloned by using Sb-sbd-associated DNA lesions in a 20-kilobase (kb) region of a 263-kb genomic walk. The region specifies an \approx 3.8-kb transcript that is induced by the steroid hormone 20hydroxyecdysone in imaginal discs cultured in vitro. The conceptually translated protein is an apparent 786-residue type II transmembrane protein (N terminus in, C terminus out), including an intracellular N-terminal domain of at least 35 residues and an extracellular C-terminal trypsin-like serine protease domain of 244 residues. Sequence analyses indicate that the Sb-sbd-encoded protease could activate itself by proteolytic cleavage. Consistent with the cell-autonomous nature of the Sb-sbd bristle phenotype, a disulfide bond between cysteine residues in the noncatalytic N-terminal fragment and the C-terminal catalytic fragment could tether the protease to the membrane after activation. Both dominant Sb and recessive sbd mutations affect the organization of microfilament bundles during bristle morphogenesis. We propose that the Sb-sbd product has a dual function. (i) It acts through its proteolytic extracellular domain to detach imaginal disc cells from extracellular matrices, and (ii) it transmits an outside-to-inside signal to its intracellular domain to modify the cytoskeleton and facilitate cell shape changes underlying morphogenesis.

The attachment of cells to extracellular substrates-for example, by integrins (1)-plays an important role in determining cell shape and the intracellular organization of the cytoskeleton. Likewise, detachment of cells from substrates also leads to profound changes in cell shape and cytoskeletal organization. In particular, cell surface-associated proteases have been shown to mediate cell shape changes by local degradation of extracellular matrices and by signaling the reorganization of the actin cytoskeleton (2, 3). Proteases are also implicated in morphogenesis of imaginal discs to form adult appendages in Drosophila (4-7). The formation of legs and wings from discs results in part from actin- and myosindependent cell shape changes in the disc epithelium elicited by the steroid hormone 20-hydroxyecdysone (20HE) (8). Later in development, specific disc cells undergo actinmediated cell shape changes to form bristles (9, 10). Stubblestubbloid (Sb-sbd) mutants cause failures in cell-shape changes required for both disc and bristle morphogenesis. Gain-of-function Sb mutations affect bristle morphogenesis in a dominant manner. Both Sb and allelic loss-of-function sbd mutations act recessively to affect disc morphogenesis, producing characteristically malformed legs and wings (11).

We show here that the product of the *Sb-sbd* gene is an apparent transmembrane protein with an extracellular serine protease domain. The structure of the conceptual protein and phenotypes of mutants suggest that *Sb* has a dual role in degrading extracellular proteins and modifying the cytoskeleton.[‡]

MATERIALS AND METHODS

Fly Stocks and Culture. A wild-type Oregon-R stock maintained in our laboratory since 1965 was used. Stubble and stubbloid stocks are described in detail elsewhere (12). Embryonic and pupal development at 25°C were staged from egg laying and pupariation, respectively. The third instar was subdivided into four stages (E, ≥ 18 ; M1, ≥ 12 ; M2, ≈ 7 ; L, ≈ 3 hr before pupariation) by the blue-food technique (13). Imaginal discs were mass-isolated and cultured as described (14).

Chromosome Walking. A library made in phage λ vector EMBL3 from *sbd*¹⁰⁵/*TM2* genomic DNA was used with clones from the 88F region (gift of E. Fyrberg, Johns Hopkins University) to jump into 89B9-10. Chromosomal walking in a *Sb*⁺ library in λ FIX (gift of K. Moses, University of Southern California), cloning, and Southern and Northern analyses were done by standard methods. cDNA clones were isolated from a library made in λ ZAPII (Stratagene) with RNA from imaginal discs that had been cultured for 10 hr with 20HE at 1 μ g/ml (in collaboration with K. Bassler and E. Hafen, University of Zürich).

Sequencing. The sequence was determined by the dideoxy chain-termination method using Sequenase version 2.0 (United States Biochemical) and specific oligonucleotide primers that were synthesized as the sequence was determined. The sequence was assembled and analyzed with MACVECTOR 3.5 software. The predicted protein sequence was compared to sequences in the GenBank 71 and Swiss-Prot 23 data bases by using the IntelliGenetics FastDB program (15).

Northern Blot Analysis. Total RNA was separated in formaldehyde/agarose gels and transferred to Nytran (Schleicher & Schuell). Primers (nt 1753–1770 and 2235–2218) were used to generate a PCR fragment spanning nt 1753–2235 of the cDNA that is separate from sequences corresponding to the protease domain. This fragment was isolated by gel electrophoresis and ³²P-labeled with a Boehringer Mannheim random-primer labeling kit. The probe was used in standard 50% formamide hybridizations at 42°C. The final wash following hybridization was in 0.1× SSPE (1× SSPE is 0.15 M NaCl/10 mM NaH₂PO₄, pH 7.4/1 mM EDTA) with 0.1% SDS at 65°C.

Abbreviation: 20HE, 20-hydroxyecdysone.

^{*}Present address: Department of Biology, Yale University, New Haven, CT 06536.

[†]To whom reprint requests should be addressed.

vertisement" [‡]The sequence reported in this paper has been deposited in the fact. GenBank data base (accession no. L11451).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

In Situ RNA Hybridization. Single-stranded digoxygeninlabeled DNA probe was made by using a single M13 primer with a linearized pBluescript (Stratagene) cDNA subclone in a thermocycling reaction and hybridized to discs (16). The subclone used contained the 5' end of the cDNA (nt 1–1116) that lacks sequences similar to those of other serine protease genes.

Confocal Microscopy of Bristles. Staged pupae were fixed for at least 24 hr in 4% formaldehyde in phosphate-buffered saline (150 mM NaCl/5 mM KCl/10 mM NaH₂PO₄, pH 7.4). Unless otherwise indicated, images are optical sections of phalloidin-labeled, whole-mounted dorsal thoraces (8). Observations were made on a Bio-Rad 600 confocal microscope.

RESULTS AND DISCUSSION

Identification of the Stubble Gene. The Sb-sbd gene (hereafter the Sb gene) was identified by restriction fragment size alterations associated with 11 Sb and sbd mutations between map positions -50 to -70 kb in a chromosomal walk, 9 of which are shown in Fig. 1. Deficiencies that remove the entire -50- to -70-kb region—e.g., $Df(3R)sbd^{105}$ —are all associated with sbd mutations. Six sbd mutations, including 2 revertants of Sb^{63b} , are respectively associated with an insertion, an inversion, two translocations, and two partial deficiencies in the region. Sb^{1} is associated with an insertion between -55.5and -57 kb present in 13 Sb¹ strains, including the TM3 balancer marked with Sb^{1} , 5 Sb^{1} revertants, and a $sbd^{2}-Sb^{1}$ intragenic recombinant (17). Sb^{63b} , a spontaneous mutation, also has an insertion between -55.5 and -57 kb and one at approximately -61 kb. Analysis of Sb^{63b} transcripts suggests that the insertion at -61 kb is successfully spliced out of the transcript (data not shown). Taken together, these data indicate that the insertion between -55.5 and -57 kb causes the Sb^{63b} mutation. (Sb^{70} also potentially has an insertion in the region of -55.5 to -57 kb.) sbd^{202r}, a diepoxybutane-induced sbd revertant of Sb^{63b} , has a small deletion of about 200 bp

T(2,3)sbd GT11 Df(3R)sbd104 In(3R)sbdVX1 T(2.3)sbd106 Sb^{63b} sbd 205r Sb^{63b} sbd 202r -50 kb -75 kb 3 Membrane Cleavage Spanning (Activation) Domain Site

FIG. 1. The structure of the Sb gene. Sb was cloned in a 263-kb genomic walk. Coordinates of the map are based on distance (in kilobases) from the EcoRI site nearest the distal breakpoint of $Df(3R)sbd^{105}$ at 89B9-10, from which the walk was started. Eleven lesions (9 are shown) associated with Sb-sbd mutants were mapped to the -50- to -75-kb region (for details see ref. 12). The depicted positions of chromosomal breakpoints associated with sbd mutations are localized to restriction fragments. Restriction fragments between -71 and -53 kb hybridize to a major 20HE-dependent transcript of \approx 3.8 kb in imaginal discs. These fragments were used to isolate cDNA clones from an imaginal disc library. Exons corresponding to a 3.8-kb cDNA have been mapped to the genomic region. The precise positions of intron/exon boundaries have not been determined. Small introns interrupt the cDNA between nt 1753 and 2235 (0.7-kb intron) and between nt 2521 and 2998 (1-kb intron). Hatched boxes delimit the positions of insertions.

between -57.5 and -62 kb. A second *sbd* revertant of Sb^{63b} , sbd^{205r} , which was induced by mobilization of *P* elements, has an insertion at approximately -66 kb.

Expression of Sb Transcripts. Sb cDNAs were isolated and mapped to the genomic region (Fig. 1). A major \approx 3.8-kb Sb transcript that hybridizes to the cDNA is present in 12- to 18-hr embryos, in early prepupae when disc morphogenesis begins, and in 36-hr pupae when bristles are forming (Fig. 2 a and b). The Sb transcript is reduced in size in Sb¹ and Sb^{63b} (\approx 2.8 kb), but not in Sb^{Spike} prepupae (data not shown), confirming that the identified transcript is the product of the Sb gene. In discs cultured in vitro, Sb transcripts are evident after between 1 and 3 hr of incubation with 20HE (Fig. 2c). Each period of Sb expression in vivo coincides with one of elevated levels of 20HE (for review, see ref. 9). Sb transcripts were localized in tissues by in situ hybridization. Transcripts are absent in larval discs (Fig. 2d) and accumulate in prepupal discs primarily in regions that undergo profound 20HE-dependent shape changes to form appendages such as legs, wings, halteres, and antennae (Fig. 2 e and f). In contrast, the ommatidial precursors in eye discs do not accumulate transcripts (Fig. 2g). Several larval tissues (brain, salivary glands, epidermis, foregut, muscle) also lack transcripts at pupariation. In 29-hr pupae, Sb transcripts are present in wing and leg epidermis (data not shown). The temporal and spatial accumulations of Sb transcripts are consistent with functions in disc morphogenesis and bristle formation.

The Sb Protein. Sb cDNAs were sequenced and a single long open reading frame encoding a conceptual 786-residue polypeptide was identified (Fig. 3). The structure of the conceptual Sb protein has four domains that may have roles in leg and wing morphogenesis and bristle formation.

Signal/Anchor Sequence: aa 59-81. The N terminus of the Sb protein has the properties of type II transmembrane proteins (N terminus in, C terminus out) (23), which typically have long highly hydrophilic (30-40% charged residues) N



FIG. 2. Expression of the Sb gene. (a-c) Northern analysis of Sb-transcript accumulation during embryonic development (a), during larval and pupal development (b), and in mass isolated imaginal discs cultured with (+) or without (-) 20HE $(1 \ \mu g/ml)$ (c). The upper band is the Sb transcript (≈ 3.8 kb); the lower band is a loading control [rp49 (38)]. The third instar was divided into four periods for study: E, ≥ 18 ; M1, ≥ 12 ; M2, ≈ 7 ; and L, ≈ 3 hr before pupariation. (d-g) In situ hybridization. Discs from mid-third-instar larvae lack signal (d). Discs dissected 0-3 hr after pupariation (e-g) show strong reproducible patterns of Sb expression. First leg discs (e), wing discs (f), and eye-antennal discs (g) were analyzed. Note absence of signal in the precursor of the eye (arrows). ($\times 24$.)

a.	GGCACGAGCCGAATCGGAAACCGATCTG TCACGAAAAGCCCACAATACATACGCCCAC CTTAACTCAATACCAAGAGAAAAAATATTT	AACCGGTTTCGATTTCGCTGTTTGCTGTTC GACTGAAGAAAGGCAGTGAAATAAAAAAGAA TCAATTCGAGACGTCGCTATAAAGAAAACA	GCTTTTCGGTTCGCACACTTGCAGCTCCAT ATCAACCGAAACAAAAGTGAAAAACAATAA AATAAAGCTCGAAAAAAAATAAAGGCCTCA	TCTATATGGGAAAAAAACACAAAGAAGAAAT Aggccacggaaaatcacggtgcaataaaag Gtaacaattaattaagagtgtggtgcagct	118 238 358			
	ATAACAATAACAAAAAACTAGCTCACATGT	ATTTACAAAAGAGTTTTACGACTTCGCAGT	GAAATGCAAATCAATTAAACATCATATTTA	TTGTTAAAACAAAAAGAAAAAAGGAACAAC	478			
	AAGTTAAAAACGAAAAGAAATTTTCGTTAC	AATTTGATTTGCCATTACCCGAGGATTTGA	TGTGTTTTTTTGACCATGAAAACGATGGAA	AGTCCGCGACTCTAAAAAATACATATTTAA	598			
	AAATAATTAACGAGAATTCCCCAACTTGTG	CGTGCAATTATTTCGAATGAAAGCAAGAAG	AGTTCAATCGTAAATGGAAAAACTGCCAAG	TGAAAAAATGGCAAGGAGTGAAAATCAAA	718			
	TTTAAAAATCCATCAAAAAGTTTGCCGTTC	TACAAAGATAATTAAACATTTAACCGACTC	TTAAAGCAAGGCGTAAACTTTACAATCGCA	GCCAAATAGCAAATAGCCGAAACTGAACTG	838			
	AAAGACTAACATGCGCCCGGCACGTTACCG	ATGAAGCAGCCAACTTTAATCAGACCCAGA	CTCAGACACAGAAGAAGCACACCAGCGGCA	GCCACCAAGATGTGTCCCCAAAAGGCATTGG	958			
1	CENCECANCANCANCANCANCANCAN				1078			
31	I V N N D A A C S D	C S C C A A A B S B		V A L T V V N C L A	10/0			
31	ACAGCAGCAGCGGCGCTGATCACACCGCCC	GACAGCCTCGAGTCACTGGGCTCCCTGGGA	ATCCCGTCATCCTCCGCCTCCTCCGAG	GACGACGATGACATGAGTAGCGGCTTCTAC	1198			
71	TAAAALITPP	D S L E S L G S L G	IPSSSASS(E)	(D (D (D M S S G F Y				
	CGCATCCCGCACCGCCTGGAGGGCTATCCG	CAGTTGCAGCAACTGCAGCGCGGCCAGAAC	TTCAAGATCAGCCCCAAGCCATGCTCCTTT	GGCCGCGTCGAGGGCACCTGCATGTTCGTG	1318			
111	RIPHRLEGYP	QLQQLQRGQN	FKISPKP <u>CSF</u>	GRVEGTCMFV				
	TGGGAGTGCATCAAGTCCGAGGGCAAGCAC	GTGGGCATGTGCGTCGACTCCTTCATGTTC	GGCTCCTGCTGCACGCACAACTACACCGAC	AACATTGTCCTGCCCCAGACGGCCTTCTCC	1438			
151	WECIKSEGKH	VGMCVDSFMF	<u>GSCC</u> THNYTD	NIVLPQTAFS				
	TACACGAGGCCCACCAAGCCGCTCACGCTC	CGCCCGCGACCGCCGGCAGCGCCCTACAAG	CCGATGATCAGCGGCATGACCACCATCGAG	AGGCCTCATGGCGCTGGCACCCTTGTGATT	1558			
191	YTRPTKPLTL	R P R P P A A P Y K	PMISGMTTIE	R P H G A G T L V I				
	CGTCCTTCGGGTCCGCACCACCAGGGCACT	CTGGCCCGCCCGCATCCGCCGCCCTACCAG	AGCAAGCCCACCACTGCCTCGGATCTGCAT	GGCTCAGCCTCGCATCCCAGCTCCAGTTCC	1678			
231	R P S G P H H Q G T	LARPHPPYQ	S K P T T A S D L H	G S A S H P S S S S	1 7 0 0			
	AGCTCCAGCTCTAGTTCGAATCCCAATAGC	ATTTGGCATACATCGACCCAGCAGCAGCAG	CAGCAGCATCAGCAGAATCAGCAGAACCAC	TGGCAGATGACCACCGAGCCGAGCTTTATT	1/98			
2/1					1019			
311	T K D B D T C W T K	P G T V N L P M P A	R P S K P S K P T K	K P I V Y D R S P P	1910			
	CCGCCGTCGTCCGTCCCACCGTCCACATCC	ACCTCGACGACATCAACGTCGCTGATTTGG	CCGGCCCAGACGCATCCCCCGCAGCCGCAT	CGCCCCACCAGGCCGCAGCTGTCGCCGGGC	2038			
351	PPSSVPPSTS	TSTTSTSLIW	PAQTHPPQPH	R P T R P Q L S P G				
	ACATCATTAGCCGCATCCTCGTCCTCGCAT	TGGCCATCGTCAACCACCTCAACCACCTCA	TCCACCACATCCACAACCACAACAACAACA	ACCACACGTCGGACAACCACACCGACGACA	2158			
391	TSLAASSSSH	W P S S T T S T T S	STTSTTTTT	TTRRTTPTT				
	ACAACAAGGAGGACCACGACAAACAAGCCC	ACCAGGCCCTACCAGCGGCCCACCACGGCG	ACCTCCAGCTCGAGCACCTCCACCACATCG	AGCAAGACTCCCACCACGACCAGGCCGATT	2278			
431	T T R R T T T N K P	TRPYQRPTTA	TSSSSTSTTS	SKTPTTTRPI				
	AGCTCCAGCAGTTCCAGCAGCAGCGGAATC	GTCACCAGCTCCCAACGGCCAACGCAGCCT	ACGCACCGTACACCGGTTCTGGCCACCTCC	GGCATCGAGACCAACGAGATATCAGACTCC	2398			
471	S S S S S S S G I	V T S S Q R P T Q P	THRTPVLATS	GIETNEISDS	1. 1.			
	TCCATCCCCGACGCCGGAGCTCTGGGCCGT	GTGAAGACCATTTCGGCGGCTCGCAGCGAG	TGCGGAGTACCGACTTTGGCGCGGCCCGAG	ACGCGAATCGTGGGCGGTAAGAGCGCGGCC	2518			
511	SIPDAGALGR	VKTISAARSE	CGVPTLARPE	T R I V G G K S A A				
	TTCGGTCGTTGGCCCTGGCAGGTGTCGGTG	CGGCGCACCTCCTTCTTTGGATTCTCGAGC	ACCCACCGTTGCGGTGGCGCTTTGATCAAC	F N N T N T N C H C	2030			
221			T A K C G G A L I A	GGGGTGGCCAAGAAGGTGGTGGTGCTATCCCAAG	2758			
591	V D D L L I S O I B	I R V G E Y D F S H	V O E O L P Y I E R	G V A K K V V H P K	2150			
591	TATAGCTTCCTCACGTACGAATACGACCTG	GCGCTGGTCAAGCTGGAGCAGCCGCTCGAA	TTCGCGCCGCATGTCAGTCCCATTTGCCTG	CCCGAGACGGACAGCCTCCTAATTGGCATG	2878			
631	YSFLTYEYDL	ALVKLEQPLE	FAPHVSPICL	PETDSLLIGM				
	AACGCCACGGTCACCGGCTGGGGTCGTCTC	AGCGAGGGCGGCACCCTGCCCTCCGTCCTC	CAAGAGGTCTCCGTGCCGATTGTAAGTAAC	GATAATTGCAAATCGATGTTCATGCGAGCC	2998			
671	NATVTGWGRL	SEGGTLPSVL	Q E V S V P I V S N	D N C K S M F M R A				
711	GGTCGCCAGGAGTTCATTCCGGATATATTC G R Q E F I P D I F	CTGTGCGCGGGCTATGAAACGGGAGGGCAG L C A G Y E T G G Q	GACTCCTGTCAGGGCGATTCAGGAGGTCCA D S C Q G D S G G P	CTGCAGGCCAAGTCGCAGGATGGGCGCTTC L Q A K S Q D G R F	3118			
	TTTCTGGCCGGGATCATCTCCTGGGGGCATT	GGCTGTGCCGAGGCCAATTTGCCCGGAGTC	TGTACGAGAATCTCCAAGTTTACGCCGTGG	ATACTGGAGCATGTCAGATGATGATCAGAA	3238			
751	FLAGIISWGI	GCAEANLPGV	CTRISKFTPW	ILEHVR				
	GATACTGACGACGGCAGATCGACTTTTGTT	ATTTGTTTTAATTTATCAGTTGTATCTATA	AGTTTTGTGTGATTTATCTTAAAGGAATGT	TAAGCGTAGTTAGCTTCAATTTAACAATTT	3358			
	TAACATCATGCCTTGTGGGCCACTGCCACGC	CCCTATCCGCACTCACCCATTCACCTACAC	ACACACACACACAGTCAACGCCCACTAATT	TAGTCTAGCGAAATGTTAGCACCGTAAGTT	34/8			
	GTCTAACAAGGAAATCGGTCGGATTTCGAA	TGGGATTATCTATTTGTACATAGCCGAAGG	CATAGGGGGCTTGGGAGGGGTAGGGGGGGAT	TGGGAGTCATTTGCATGTCCACAAGTTGTT	3718			
	ACCABACCTCCACAGCCAATAAAACATT	CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAA		3782			
	ACGARAGE I COROROCARTARAGAT TART							
b.								
			- Clasuage site					
		531	- Cleavage Site	589				
Sh mi	WETCHCCTPHACALCHVETS		VOCKSAAFGRWRWOVSVRRTSFFC	-FSSTHREGGALTNENWTATACHEVOD	LLTS			
Hp F1	FCVDEGRLPHTQRLLEVISVCDCPRG	RFLAAICQDCGR-RKLPVDRI	VGGRDTSLGRWPWQVSLR	-YDGAHLCGGSLLSGDWVLTAAHCFPE	RHRV			
Тр	PSATMS	ALLILAL-VGAAVAFPLEDDDKI	VGGYTCPEHSVPYQVS	LNSGYHFCGGSLINDQWVVSAAHC	-YKS			

Sn	Sn AQRISATKCQEYNAAARRLHLTDTVRTFSGKQ-C-V-PSV-	PLIVGGTP	GTPTRHGLFPHMAALGWTQGSGSKDQDIKWGC GGALVSELYVLTAAHCATSGSKP		
		639	659	and a state of the	
Sb	Sb giairvgeydfshvgeglpyiergvakkvvhpkysb	FLTYEY DLALVKLEQ	PLEFAPHVSPICLP-ETDSLLIG	MNATVTGWG-RLSEGGTLPSVLQEVSV	
Hp	Hp LSRWRVFAGAVAQASPHGLQLGVQAVVYHGGYLPFRDPNS	SEENSNDIALVHLSSI	LPLTEYIQPVCLPAAGQALVDG	KICTVTGWG-NTQYYGQQAGVLQEARV	
Тр	Tp RIQVRLGEHNINVLEGDEQFINAAKIIKHPNYSS	SWTLNNDIMLIKLSS	PVKLNARVAPVALP SACAPAG	TQCLISGWGNTLSNGVNNPDLLQCVDA	
Sn	Sn pdmvrlgarqlnetsatq-qdikiliivlhpkyR	SSAYYH DIALLKLTR	RVKFSEQVRPACLW-QLPELHI-	PTVVAAGWG-RTEFLGAKSNALRQVDL	
	737				
Sb	Sb pivsndncksmfmragr-gefipdificagyetgggdscggdsggp.	LQAKSQDGRF	FLAGIISWGIGCAEANLPGVCTR	ISKFTPWILEHVR	
Hp	Hp PIISNDVC-NGADFYGN-Q-IKPKMF-CAGYPEGGIDACQGDSGGP	FVCEDSISRTPRW	RLCGIVSWGTGCALAQKPGVYTE	VSDFREWIFQAIKTHSEASGMVTQLXPVA	
Тр	Tp PVLSQADCEAAYP-GE-ITSSMICVGFLEGGKDSCQGDSGGP-	VVCNGQ	LQGIVSWGYGCALPDNPGVYTK	VCNFVGWIQDTIAANXISSVSLQSVC	
Sn	SO DAVPONTCKOTYRKERRLPRGTTEGOFCAGYIPGGRDTCDGDSGGP	IHALLPEYNCVA		LYSYLDWIEKIAF	

FIG. 3. (a) The Sb cDNA sequence and predicted 786-aa sequence. The sequence begins with the 5' end of an \approx 3.8-kb Sb cDNA and is based on sequence from both strands of this cDNA and identical sequences from portions of several other cDNAs. The 22-residue hydrophobic domain beginning at aa 59 is boxed. Charged amino acids surrounding this proposed transmembrane domain are circled. The disulfide knotted domain is underlined. The C-terminal region with homology to serine proteases is shaded. A polyadenylylation signal (nt 3735–3740) begins 15 nt before the poly(A) tail. (b) The C-terminal region of the predicted Sb protein (aa 504–786) aligned with the catalytic domains of three serine proteases of the trypsin superfamily. Numbers refer to the predicted Sb sequence. Hp, hepsin, a human liver serine protease (18, 19); Tp, a rat pancreatic trypsin (20); Sn, snake, a *Drosophila* serine protease (21). Shaded regions indicate generally conserved regions of serine-protease catalytic domains (22). Residues in the catalytic triad (His-589, Asp-639, Ser-737) are highlighted in black. Eight cysteines (in clear boxes) occupy conserved sites; six form disulfide bonds within the serine protease domain, and two span the activation site and form disulfide bonds in two-chain proteases.

termini preceding a hydrophobic transmembrane signal/ anchor sequence (22–29 aa) (24–26). The cytoplasmic sequence preceding the membrane-spanning domain characteristically has a net positive charge relative to that of the sequence immediately following the membrane-spanning domain (25). The proposed cytoplasmic domain of the Sb protein has 58 aa, including 18 charged residues (16 basic; Fig. 3a) preceding a signal/anchor sequence of 22 aa. A second in-frame methionine codon at nt 938, if used as a translational start, would produce a 763-aa protein with a 35-aa putative intracellular domain. The 20 aa immediately preceding the proposed transmembrane domain have a net charge of +5 relative to those immediately following the domain. The abundance of positively charged amino acids on



FIG. 4. Effect of exogenous protease on Sb discs. A pair of third leg discs were dissected from a Sb^{G3b}/Sb^{G3b} prepupa 6 hr after pupariation. One disc (a) was incubated in Robb's culture medium without trypsin, and the other (b) was treated for 1 min in medium containing 0.001% trypsin. Both discs were fixed and photographed. The length of the untreated disc is less than normal for this stage of development. The treated disc has unfolded and elongated to almost wild-type proportions. (×40.)

the N-terminal side of the hydrophobic sequence predicts that the N terminus of Sb is intracellular.

Serine-protease domain: aa 542-786. The C-terminal 244 aa of Sb have the properties characteristic of a trypsin-like protease that preferentially cleaves after basic amino acids (22, 27). Three invariant amino acids (His-589, Asp-639, Ser-737), correspond to those in chymotrypsinogen (His-57, Asp-102, Ser-195) and are essential for catalysis (28). Six cysteines needed to form three intramolecular disulfide bonds that stabilize the catalytic pocket also occur in conserved domains. The putative proteolytic activation site of the Sb zymogen (after Arg-542) is in a characteristic Arg-Ile-Val-Gly-Gly (RIVGG) motif. Thus, the Sb protease can potentially cleave the Sb zymogen. The Sb protein contains two additional cysteines (Cys-531 and Cys-659) corresponding to those used in "two-chain" proteases, including other arthropod proteases (21, 29-31), to attach the processed enzyme to its noncatalytic N-terminal fragment via a disulfide bond. After proteolytic activation, the proposed disulfide bond would attach the active Sb protease to its transmembrane N-terminal fragment and tether the enzyme to the surface of the cell in which it was synthesized. The tethered structure is consistent with the autonomous development of Sb and sbd bristles in somatic mosaics of mutant and wildtype cells (32). Whether Sb-sbd mutations act autonomously in appendage morphogenesis is not known.

Disulfide knotted domain: aa 138–173. The noncatalytic N-terminal region of Sb contains a cluster of 6 cysteine residues like those in the "disulfide knotted domain" (33)

identified in three other arthropod serine proteases [a Limulus (horseshoe crab) proclotting protein and the Drosophila proteases easter and snake, involved in dorsal-ventral patterning (21, 30, 31)]. The intramolecular covalent bonding configurations in this region have been established in the proclotting enzyme from horseshoe crab (31). The knot may function as a receptor for an activating protein of the proclotting enzyme. In all, Sb has 14 conserved cysteine residues common to arthropod serine proteases: 6 in the noncatalytic N terminus, 6 in the catalytic region, and 2 in possible tethering positions.

The protease domain of the Sb primary sequence is most similar to that of hepsin (Fig. 3b), a mammalian protease that is also an apparent type II transmembrane protein with a demonstrated cell surface location (18, 19). Many two-chain serine proteases, after activation, localize to cell surfaces either by forming a membrane-bound complex with a cofactor (e.g., the association of factor IX with factor VIII in blood clotting) or by binding to a receptor (e.g., binding of urokinase-type plasminogen activator to its receptor) (34, 35). For hepsin and the Sb protease, localization of the active protease to the surface of the cell in which it is synthesized is apparently an inherent property of a single gene product. The sequence similarities between hepsin and trypsin are limited to the protease domain and the region around the tethering cysteines. Hepsin, for example, lacks the disulfide knotted domain of Sb. Additional sequence comparisons are needed to clarify the complex and intriguing evolutionary relationships between Sb, hepsin, and other serine proteases.

Stem: aa 250-500. The Sb zymogen contains repeats of several amino acids (principally serine and threonine, but also glutamine, proline, and basic residues) that separate the protease domain from the disulfide knotted domain and would potentially extend the C-terminal protease domain from the cell surface. We refer to this region as the stem. Less extensive serine- and threonine-rich domains also occur in this region in the *Limulus* proclotting enzyme and in the *Drosophila* proteases easter and snake. In the *Limulus* enzyme, this region is extensively glycosylated (31).

Proteases in Imaginal Disc Morphogenesis. During the first 6 hr after pupariation, leg and wing discs undergo a complex, 20HE-dependent series of cell shape changes resulting in the elongation, unfolding, and eversion of appendages (9). Several observations predict important roles for proteases in disc morphogenesis. Exogenous trypsin and chymotrypsin accelerate 20HE-induced disc morphogenesis (4, 7), and some



FIG. 5. Effects of recessive sbd and dominant Sb mutations on bristle morphogenesis. (a-c) Light micrographs of the dorsal thoraces of pharate adults of wild type (a), sbd^{VX1}/sbd^{VX1} (b), and Sb^{63b}/Sb^{63b} (c) showing the prominent scutellar and dorsocentral bristles examined here. (d-j) Confocal micrographs (negative images) of phalloidin-labeled bristles showing the distribution of actin bundles in longitudinal (d-h) and transverse (i and j) views of wild type (d, e, and j), sbd^{201}/sbd^{201} (f and g), and Sb^{63b}/Sb^{63b} (h and i). i is from a 6-µm frozen transverse section of a Sb^{63b}/Sb^{63b} pupa; e, f, and h are Z-series projections of entire bristles. The discontinuous appearance of some of the filament bundles (d, e, g, and h) is probably an artifact of preparation. [Bar in a (for a-c) = 0.1 mm; all other bars = 20 µm.]

serine-protease inhibitors block disc morphogenesis (6). The 20HE-dependent appearance of endogenous soluble serine proteases, presumably not encoded by Sb, has been described (6). Six hours after pupariation, leg discs from Sb and sbd homozygotes are significantly shorter than wild-type leg discs (36). This developmental defect can be ameliorated within seconds by exposing explanted Sb (Fig. 4) or sbd (data not shown) leg discs to trypsin. These results confirm that disc morphogenesis depends on endogenous proteases and suggest that both Sb and sbd mutants lack Sb protease activity.

The basis of the function of protease in disc morphogenesis is not understood, but discs, like other epithelia, are encased by extracellular matrices (9). Modification of extracellular matrices by proteolysis has been proposed in general to be a key component of tissue reorganization during morphogenesis (35). We believe proteolysis would free disc cells from attachments to apical or basal extracellular matrices and facilitate the cell shape changes needed for disc morphogenesis (8). Indeed, type IV collagen is proteolytically cleaved in discs within 2 hr of exposure to 20HE (5), providing an example of a modification of the basal matrix that is associated with disc morphogenesis.

Cytoskeletal Defects in Sb and sbd Bristles. A possible function of the Sb cytoplasmic domain is suggested by the autonomous effects of Sb-sbd mutants on bristle development (Fig. 5). Bristles are long apical extensions of single epithelial cells (Fig. 5a). They contain submembranous bundles of actin filaments (10) that run the length of the shaft. Bristles grow at their tips (37) with the concomitant extension of microfilament bundles, the history of bristle morphogenesis being recorded in the structure of microfilament bundles from the base to the tip. The extension of a bristle through an overlying extracellular matrix could be facilitated by a surface-bound protease. In wild-type bristles, 15–18 uniformly distributed submembranous bundles of actin merge gradually to a tapered tip (Fig. 5 d, e, and j). Sb-sbd mutations affect filament bundle assembly in bristles. In sbd mutants bristle development begins with normal numbers and distribution of filament bundles, but the bundles ultimately become disorganized at the tip (Fig. 5 f and g), accounting for the frayed ends of sbd bristles (Fig. 5b). Severe defects at the tips of sbd bristles are first evident at 40 hr with abrupt reduction in the number of filament bundles. In contrast, in Sb mutants, bristles are defective from their inception. Sb bristles have more filament bundles than normal (10). In bristles of Sb^{63b} homozygotes, 25-30 filament bundles of variable thickness are irregularly arranged, with some bundles running through the center of the cell (Fig. 5 h and i). By 40 hr filament elongation stops abruptly, resulting in short blunt-ended bristles (Fig. 5c).

The functional nature of Sb mutants and the basis of the gain-of-function bristle phenotype remain to be determined. However, some Sb mutations (at least Sb^{1} and Sb^{63b}) shorten the Sb transcript and have insertions in the genomic region corresponding to the polypeptide sequence between the signal/anchor sequence and the proteolytic activation site (Fig. 1). These are expected to produce truncated proteins that lack the extracellular protease. So, alterations potentially affecting the extracellular structure of the Sb protein affect the intracellular organization of filament bundles. These alterations may act by inducing conformational changes in the intracellular domain of the Sb protein. This suggests that Sb has a dual function. It acts through its proteolytic extracellular domain to detach imaginal disc cells from extracellular matrices, and it carries an outside to inside signal via its intracellular domain to reorganize the cytoskeleton and facilitate cell shape changes underlying morphogenesis.

We thank K. V. Anderson, C. Craik, D. S. King, and C. Goodman for comments; P. Simpson and H. Heitzler for new *sbd* mutations; C. Bayer for experimental contributions; and G. Hammonds and M. Cummings for advice on sequence comparisons. L.F.A. was a U.S. Public Health Service predoctoral trainee and M.P. a postdoctoral trainee (GM07127 and HD07229, respectively). This work was supported in part by a U.S. Public Health Service research grant (GM19937).

- 1. Hynes, R. O. (1992) Cell 69, 11-25.
- Pollanen, J., Hedman, K., Nielsen, L. S., Dano, K. & Vaheri, A. (1988) J. Cell Biol. 106, 87–95.
- 3. Jalink, K. & Moolenaar, W. H. (1992) J. Cell Biol. 118, 411-419.
- Fekete, E., Fristrom, D., Kiss, I. & Fristrom, J. W. (1975) Wilhelm Roux Arch. Entwicklungsmech. Org. 173, 123–138.
- Fessler, L. I., Condic, M., Nelson, R., Fessler, J. & Fristrom, J. (1993) Development, in press.
- 5. Pino-Heiss, S. & Schubiger, G. (1989) Dev. Biol. 132, 282-291.
- Poodry, C. A. & Schneiderman, H. A. (1971) Dev. Biol. 168, 1-9.
- Condic, M. L., Fristrom, D. & Fristrom, J. W. (1991) Development 111, 23-33.
- 9. Fristrom, D. & Fristrom, J. W. (1993) in *The Development of Drosophila*, eds. Bate, M. & Martinez-Arias, A. (Cold Spring Harbor Lab., Plainview, NY), in press.
- 10. Overton, J. (1967) J. Morphol. 122, 367-380.
- 11. Beaton, A., Kiss, I., Fristrom, D. & Fristrom, J. W. (1988) Genetics 120, 453-464.
- 12. Appel, L. F. (1992) Ph. D. thesis (Univ. of California, Berkeley).
- 13. Maroni, G. & Stamey, S. C. (1983) Drosoph. Inf. Serv. 59, 142-143.
- 14. Eugene, O., Yund, M. A. & Fristrom, J. W. (1979) *Tissue Culture Assoc. Manual* 5, 1055-1062.
- Brutlag, D. L., Dautricourt, J. P., Maulik, S. & Pelph, J. (1990) Comput. Appl. Biosci. 6, 237-245.
- 16. Tautz, D. & Pfeifle, C. (1989) Chromosoma 98, 81-85.
- 17. Lewis, E. B. (1951) Cold Spring Harbor Symp. Quant. Biol. 16, 159-174.
- Tsuji, A., Torres-Rosado, A., Arai, T., Le Beau, M. M., Lemons, R. S., Chou, S.-H. & Kurachi, K. (1991) J. Biol. Chem. 266, 16948-16953.
- Leytus, S. P., Loeb, K. R., Hagen, F. S., Kurachi, K. & Davie, E. W. (1988) *Biochemistry* 27, 1067–1074.
- MacDonald, R. J., Stary, S. J. & Swift, G. H. (1982) J. Biol. Chem. 257, 9724-9732.
- 21. De Lotto, R. & Spierer, P. (1986) Nature (London) 323, 688-692.
- Furie, B., Bing, D. H., Feldmann, R. J., Robinson, D. J., Burnier, J. P. & Furie, B. C. (1982) J. Biol. Chem. 257, 3875-3882.
- 23. Singer, S. J. (1990) Annu. Rev. Cell Biol. 6, 247-296.
- Andrews, D. W., Young, J. C., Mirels, L. F. & Czarnota, G. J. (1992) J. Biol. Chem. 267, 7761–7769.
- Hartmann, E., Rapoport, T. A. & Lodish, H. F. (1989) Proc. Natl. Acad. Sci. USA 86, 5786–5790.
- 26. Parks, G. D. & Lamb, R. A. (1991) Cell 64, 777-787.
- Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G. & Bartunik, H. (1983) J. Mol. Biol. 164, 237-282.
- Hartley, B. S., Brown, J. R., Kauffman, D. L. & Smillie, L. B. (1965) Nature (London) 207, 1157-1159.
- 29. Nakamura, T., Morita, T. & Iwanaga, S. (1985) J. Biochem. (Tokyo) 97, 1561–1574.
- 30. Chasen, R. & Anderson, K. (1989) Cell 56, 391-400.
- Muta, T., Hasimoto, R., Miyata, T., Nishimura, H., Toh, Y. & Iwanaga, S. (1990) J. Biol. Chem. 265, 22426-22433.
- 32. Golic, K. G. (1991) Science 252, 958-961.
- 33. Smith, C. L. & DeLotto, R. (1992) Protein Sci. 1, 1225-1226.
- 34. Furie, B. & Furie, B. C. (1988) Cell 53, 505-518.
- 35. Saksela, O. & Rifkin, D. B. (1988) Annu. Rev. Cell Biol. 4, 93-126.
- Fristrom, D., Beaton, A., Kiss, I. & Fristrom, J. W. (1987) in Molecular Biology of Invertebrate Development, ed. O'Connor, J. D. (Liss, New York), pp. 235-246.
- 37. Lees, A. D. & Picken, L. E. R. (1945) Proc. R. Soc. London B 132, 396-423.
- O'Connell, P. O. & Rosbash, M. (1984) Nucleic Acids Res. 12, 5495–5513.