European Journal of Cell Biology ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

European Journal of Cell Biology

journal homepage: www.elsevier.de/ejcb

The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase

Annett Kilic [a,1], Sven Klose [b,1], Bernhard Dobberstein [a], Elisabeth Knust [b], Katja Kapp [a,*]

[a] Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH-Allianz, Universität Heidelberg, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany
[b] Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstr. 108, 01307 Dresden, Germany

ARTICLE INFO

Article history:
Received 16 November 2009
Received in revised form
25 January 2010
Accepted 1 February 2010

Keywords:
signal peptide
ER targeting
signal peptide peptidase
Crumbs
Drosophila

ABSTRACT

N-terminal signal sequences mediate nascent protein targeting to and protein insertion into the membrane of the endoplasmic reticulum. They are typically 15-30 amino acid residues long with a core hydrophobic region flanked by an N-terminal (n-) and a C-terminal region. Following cleavage by signal peptidase, some of the resulting signal peptides are further processed by signal peptide peptidase (SPP) and fragments are liberated into the cytosol. Such fragments can have independent, post-targeting functions affecting diverse cellular processes. We show that *Drosophila melanogaster* Crumbs, a transmembrane protein controlling cell polarity and morphogenesis, is synthesized with an 83 residues-long signal sequence. To our knowledge, this is currently the longest signal sequence described for an eukaryotic protein. The unusual length is caused by an extended n-region, but the extension does neither affect protein targeting nor signal sequence cleavage. The signal sequence is cleaved off and the resulting signal peptide, SP^{Crb}, is proteolytically processed by SPP, thus representing the first substrate described for the *Drosophila* enzyme. We further show that signal peptide fragments can be degraded by the proteasome. Expression of transgenes encoding tagged variants of Crumbs in *Drosophila* embryos suggests that the signal peptide is short-lived *in vivo*. Our findings support a model suggesting that besides generating fragments with post-targeting functions, SPP-mediated processing is the first step in the degradation of signal peptides.

Introduction

Signal sequences are the N-terminal extensions of nascent secretory and membrane proteins. They mediate protein targeting to and insertion into the endoplasmic reticulum (ER) membrane (Blobel and Dobberstein, 1975). Most signal sequences range between 15-30 amino acid residues in length but some extraordinary long ones comprise more than 100 residues (Hegde and Bernstein, 2006; Martoglio and Dobberstein, 1998). A common feature of signal sequences is their tripartite structure. A net positively charged N-terminal (n)-region of variable length is followed by a central, hydrophobic (h)-region of 7-15 amino acid residues, and a C-terminal (c)-region of 3-7 residues representing the consensus cleavage site for signal peptidase. Beyond that, signal sequences are highly variable in their amino acid sequence (von Heijne, 1990). Following protein insertion into the ER membrane, most signal sequences are cleaved off by signal peptidase (Paetzel et al., 2002). The high degree of variation accomplishes that individual signal sequences differently affect ER targeting (e.g. the interaction with the signal recognition particle), the efficiency of protein insertion or signal sequence cleavage (Hegde and Bernstein, 2006; Martoglio and Dobberstein, 1998).

Signal peptidase-mediated cleavage results in signal peptides that initially appear as small, transmembrane molecules (Kapp et al., 2009; Shaw et al., 1988). They can be processed by the intramembrane processing protease signal peptide peptidase (SPP), and the generated signal peptide fragments are liberated from the ER membrane. This fate has been described for the signal sequences of bovine prolactin, the human leukocyte antigen-A*0301 and some viral glycoproteins. In these cases, the signal peptide fragments play a role in processes as different as immune surveillance, cellular signaling, protein maturation or virus maturation (Weihofen and Martoglio, 2003). SPP has been identified in several species suggesting that SPP-mediated processing of signal peptides generally occurs in the animal kingdom (Golde et al., 2009). In *Drosophila* embryos, transcripts for SPP were found from germ band extension onwards, and animals with mutations in the *Spp* gene die as larvae (Casso et al., 2005). Alternatively to processing by SPP, signal peptides can accumulate as membrane-inserted peptides. For example, the 58 residues-long signal peptides derived from different arenaviral

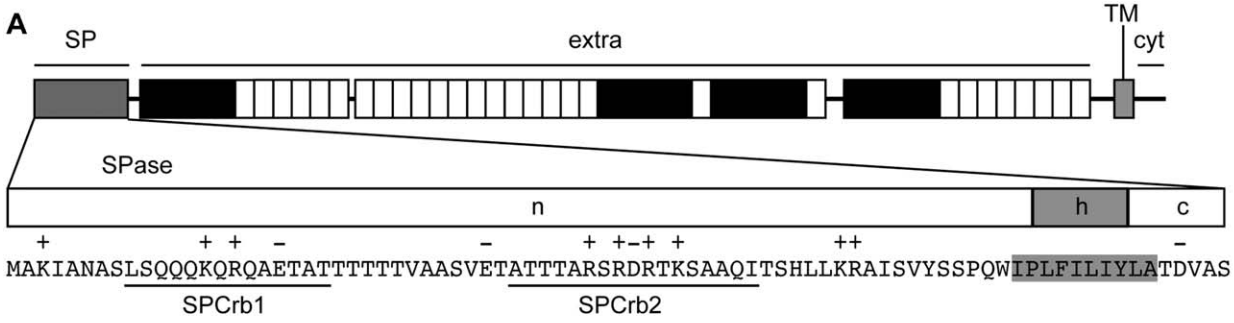
* Corresponding author. Present address: Knust Lab, Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstr. 108, 01307 Dresden, Germany. Tel.: +49 351 210 2474; fax: +49 351 210 1309.
E-mail address: kapp@mpi-cbg.de (K. Kapp).
[1] These authors have contributed equally.

0171-9335/$ - see front matter © 2010 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ejcb.2010.02.001

glycoproteins are required for viral glycoprotein processing, transport, and viral infectivity (Hegde and Bernstein, 2006; Schrempf et al., 2007). Finally, entire signal peptides can be released from the ER membrane and are transported into a different compartment, e.g. the nucleus. This fate has been described for the 98 residues-long signal peptide of the mouse mammary tumor virus Rem protein (Dultz et al., 2008). In summary, the common feature of these signal peptides with known post-targeting functions is their unusual length. Secondly, these yet analyzed signal peptides are predominately of viral origin.

In *Drosophila melanogaster*, proteins synthesized with extended signal sequences are also known, e.g. Crumbs (Tepass et al., 1990). Crumbs is described as a type I transmembrane protein comprising 29 epidermal growth factor (EGF)-like domains and four laminin A G-like domains in its extracellular region. The transmembrane region is followed by a short, highly conserved cytoplasmic region, which contains a 4.1/Ezrin/Radixin/Moesin (FERM) domain- and a C-terminal Postsynaptic density 95/Discs large/Zonula occludens-1 (PDZ) domain-binding site (Fig. 1A; Klebes and Knust, 2000). Crumbs is detected from gastrulation onwards at the apical membrane of ectodermally-derived epithelia (Tepass et al., 1990). It tethers an apical protein complex, consisting of the core components Crumbs, Stardust, *D*PATJ and *D*Lin-7, to the apical membrane. In many epithelia of the *Drosophila* embryo, the Crumbs-complex is required to maintain epithelial polarity (Bulgakova and Knust, 2009).

Using cell-free *in vitro* assays and cell-based experiments, we demonstrate that Crumbs is synthesized with a signal sequence of 83 amino acid residues. The signal sequence is co-translationally cleaved off. The resulting signal peptide is processed by *Drosophila* SPP and can be further degraded by the proteasome. Results based on transgene expression in *Drosophila* embryos suggest a short half-life *in vivo*.



B

```
D.me. -----MAKIANASLSQQQKQRQAETATTTTTTVAASVETATTTARSRDRTKSAAQIT-----SHLLKRAISVYSSPQWIPLFILIYLATDVAS 83
D.si. -----MAKIANASLSQQQKQRQAATATTTTTTVAASVETATTTARSRDRTKSAAQIT-----SHLLKRAISVYSSPQWIPLFILIYLATDVAS 83
D.se. -----MAKIANASLSQQQKQRQ--TATTTTTTVAASVETATTTARSRDRTKSAAQIT-----SHLLKRAISVYSSPQWIPLFILIYLATDVAS 81
D.ya. -----MAIIANASLPKQQKQRQ-ATTG-TTTTVAASVETATTTARSRDRTKSAAQIT-----SHLLKRAISVYSSPQWIPLFILIYLATDVAS 81
D.er. -----MAIIANASLPQQHKQRQ-ATTTTTTTTVAASVETATTTARSRDRTKSAAQIT-----SHLLKRAISVYSSPQWIPLFILIYLATDVAS 82
D.an. -----MAVIANASARQQQ---------------PAKTTTTTTTTRSRDRTKSAAQITTT---SHLLKRAIS---APQWIPLFILIYLATDVAS 67
D.ps. ----------------------------------------~QSRARTKSAAQITTS-TAQHLLKRAIS---APQWIFLFILIYLATDVAS 45
D.pe. ----------------------------------------~QSRARTKSAAQITTS-TAQHLLKRAIS---APQWIFLFILIYLATDVAS 45
D.wi. --MTTKA------------TKQQQQQ-----------------LRQAQQH-------------LLKRAIS---APQWIFLLILIYLATDVAS 45
D.vi. MAMTSAAVYASASANNAHVTQQQQQQ--------TTNQQRQQQRLSRARTKSAAQMTST--YTLLLKRANS---APQWIFLFILIYLATDVAS 80
D.mo. MAMTAAATYASATANNAYVTQQQQQQQQKQRRQRTTTNQQRQQRRLSRARTKSPAQVTSAALSTLLLKRACS---APQWIFLFILIYLATDVAS 90
D.gr. MAMTATA-YASATHVTQQQQQQQQQQQQQQQQ---TTNQQRQQQRLSRARTKNAAQLTST--LTPLLKRAIS---APQWIFLFILIYLATDVGA 84
                                                  : : :             ***** *   :**** *:*********.:
```

C

```
D.m.      MAKIANASLSQQQKQRQAETATTTTTTVAASVETATTTARSRDRTKSAAQITSHLLKRAISVYSSPQWIPLFILIYLATDVAS 83
A.a.      ------------------------------------------------------MRIRTNLFL--VLIGIASVSVLNH-TEA 25
T.c.      ---------------------------------------------------------MGFLS--SWRLGPISVFLGGGLFS 22
N.v.      ------------------------------------------------------MRAALHVVA--VLAT---LLTTST-SYG 22
H.s._Crb1 -----------------------------------------------------MALKNINYLL--IFYLSFSLLIYIKNSFC 27
H.s._Crb2 ----------------------------------------------------MALARPGTPDPQ--ALASVLLLLLWAP-ALS 28
H.s._Crb3 -----------------------------------------------------MANPGLGLLL--ALGLPFLLARWGR-AWG 26
```

Fig. 1. Schematic representation of Crumbs, its signal sequence and multi-alignments with the Crumbs signal sequences. (A) The cartoon represents *Drosophila melanogaster* Crumbs comprising a signal sequence (marked as grey box), a large extracellular region (extra) with 4 laminin G-like domains (shown as black boxes) and 29 EGF-like domains (depicted as white boxes), a transmembrane region (TM, marked in grey), and a short cytoplasmic region (cyt) of 37 amino acid residues. The signal sequence is enlarged to show its tripartite structure with the n-, h- and c-region. The amino acid sequence of the signal sequence is given at the bottom (N-terminal 83 residues of *Drosophila melanogaster* Crumbs), positively (+) and negatively (-) charged residues are indicated, and the peptides used to raise antibodies are marked (SPCrb1, SPCrb2). (B) A ClustalW multiple sequence alignment of the predicted signal sequence of *Drosophila melanogaster* (D.me., M33753) together with the predicted signal sequences of *Drosophila simulans* (D.si., XM_002104666), *Drosophila sechellia* (D.se., XM_002032451), *Drosophila yakuba* (D.ya., XM_002099219), *Drosophila erecta* (D.er., XM_001981972), *Drosophila ananassae* (D.an., XM_001967430), *Drosophila pseudoobscura pseudoobscura* (D.ps., XM_001359106), *Drosophila persimilis* (D.pe., XM_002016819), *Drosophila willistoni* (D.wi., XM_002072392), *Drosophila virilis* (D.vi., XM_002053259), *Drosophila mojavensis* (D.mo., XM_002000104), and *Drosophila grimshawi* (D.gr., XM_001989410). Identical amino acid residues are indicated by asterisks. Similarity between amino acids is shown by dots, and conserved substitutions are marked by colons, respectively. The predicted h-regions are underlined. (C) A ClustalW multiple sequence alignment of the predicted Crumbs signal sequences of *Drosophila melanogaster* (D.m., M33753) together with Crumbs from different insects including the mosquito *Aedes aegyptii* (A.a. XM_001658640), the beetle *Tribolium castaneum* (T.c., XM_965547), and the jewel wasp *Nasonia vitripennis* (N.v., XM_001604088) as well as with the three human Crumbs proteins (*Homo sapiens*, H.s.; Crb1, NM_201253; Crb2, NM_173689; Crb3, NM_139161). The predicted h-regions are underlined.

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

Material and methods

Plasmids

To generate pRK5rs_Crb$^{1-321}$, pUAST_crb^{wt} (Wodarz et al., 1995) was digested with EcoRI and the fragment was cloned into pRK5rs (Eaton et al., 1986). For pRK5rs_Crb^{1-321*}, alanine at position 2 was exchanged for methionine by site-directed mutagenesis using the primers 5′-TTC GCC CTT CCA CCA TGA TGA AAA TCG CCA ATG CGT CAC and 5′-GTG ACG CAT TGG CGA TTT TCA TCA TGG TGG AAG GGC GAA. The asterisk indicates this mutation. For the signal peptide marker, pRK5rs_SP^{Crb*}, representing the first 83 residues, a PCR fragment based on pUAST_crb^{wt} as template and the primers 5′-CGA ATT CCA CCA TGA TGA AAA TCG CCA ATG CG and 5′-ATC TCG AGT CAT GAG GCG ACA TCT GTT GC was cloned, thereby introducing a STOP codon at the C-terminus. pRK5rs_Crb$^{Myc-intra}$ and pRK5rs_Crb$^{Myc-intra*}$ were generated by PCR with pUAST_crb$^{intra-myc}$ (Wodarz et al., 1995) as template using the forward primer 5′-TAG AAT TCC ACC ATG GCT AAA ATC GCC AAT GC and the reverse primer 5′-CGC TCG AGG CAA AAT ATG TTT TTT ATT TG, and 5′-TAG AAT TCC ACC ATG ATG AAA ATC GCC AAT GC as forward primer for the mutant. Of note, we renamed the insert 'crb$^{intra-myc}$' into 'Crb$^{Myc-intra}$', referring to the localization of the Myc tag.

To generate pRK5rs_Crb$^{56-321}$, a PCR using pUAST_crb^{wt} as template and the primers 5′- CGA ATT CCA CCA TGC GCG CCA TTT CAG TTT ATT CG and 5′-CTC GGG AAT TCT CCA GGC AGG AAC C was performed and the fragment cloned. For pRK5rs_Crb$^{\Delta(1-56)-Myc-intra}$, pUAST_crb$^{intra-myc}$ and the primers 5′-CGA ATT CCA CCA TGC GCG CCA TTT CAG TTT ATT CG and 5′-CGC TCG AGG CAA AAT ATG TTT TTT ATT TG were used for PCR.

pBS_Crb$^{8xMyc-intra}$ (unpublished, S.K. & E.K.) was used to clone pRK5rs_Crb$^{8xMyc-intra*}$ by PCR using the primers 5′-TAG AAT TCC ACC ATG ATG AAA ATC GCC AAT GC and 5′-CGC TCG AGG CAA AAT ATG TTT TTT ATT TG. pRK5rs_HACrb$^{8xMyc-intra}$ was cloned by overlap-PCR using pBS_Crb$^{8xMyc-intra}$ as template and the primers 5′-GGA ATT CCC CCC CCC AAA AAA AAT A and 5′-GTG ACG CAT TGG CGA TTT TAG CAG CGT AAT CTG GAA CAT CGT ATG GGT ACA TGA TCG CTG ATT TGT TGG CTC AAG as well as 5′-CTT GAG CCA ACA AAT CAG CGA TCA TGT ACC CAT ACG ATG TTC CAG ATT ACG CTG CTA AAA TCG CCA ATG CGT CAC and 5′-CCG CTC GAG GCA AAA TAT GTT TTT TAT TTG. The mutant with a non-cleavable (nc) signal sequence, pRK5rs_HA$_{nc}$Crb$^{8xMyc-intra}$ has the exchanges A82P and S83W and was obtained by site-directed mutagenesis using the primers 5′-CTT AGC AAC AGA TGT CCC TTG GGT GGC GGT GCC GAC GGA G and 5′-CTC CGT CGG CAC CGC CAC CCA AGG GAC ATC TGT TGC TAA G. The inserts Crb$^{8xMyc-intra}$, HA-Crb$^{8xMyc-intra}$ and HA$_{nc}$Crb$^{8xMyc-intra}$ were subcloned into the pUAST vector.

To obtain the cDNA of SPP, an *Drosophila melanogaster* ovary cDNA library (Grosshans et al., 1999) was used as a template with the primers 5′-CAC AGC ATG GCG GAG GAA GTC ATC G and 5′-CTA CTT GCC CTT TTT CGA CTC C in a PCR (FlyBase: CG11840, GenBank Accession Number: NM_078720; Casso et al., 2005). Similarly, the SPP-L3 cDNA (FlyBase: CG17370, GenBank Accession Number: NM_143180) was obtained using the primers 5′-GCG AGC ATG TCG CAC GGT GGA GC and 5′-TCA GAC TTC CAG TTG TTT TGA TGG. For generation of pRK5rs_SPP-D274A, the primers 5′-GCA ATG CTG GGA CTG GGC GCC ATC GTT ATT CCG GGC ATC and 5′-GAT GCC CGG AAT AAC GAT GGC GCC CAG TCC CAG CAT TGC were applied. For introduction of a Myc-tag N-terminal of the KKGK signal, the primer 5′-TAG AAT TCC TAC TTG CCC TTT TTC AGG TCC TCC TCG GAG ATC AGC TTC TGC TCC GAC TCC TTC TTT TTG C was used. pRK5rs_SPP-L3-Myc was cloned by PCR using the primers 5′-TAG AAT TCC ACC ATG TCG CAC GGT GGA GCC and 5′-TAC TCG AGT CAT AGA TCC TCT TCA CTA ATA AGT TTT TGT TCG ACT TCC AGT TGT TTT GAT GG. Its dominant negative SPP-L3 mutant, pRK5rs_SPP-L3-D307A-Myc was generated by *in vitro* mutagenesis using the primers 5′-TCT ATG CTG GGC CTG GGT GCT GTG GTG ATG CCG GGC CTG and 5′-CAG GCC CGG CAT CAC CAC AGC ACC CAG GCC CAG CAT AGA. All plasmid inserts were confirmed by full-length sequencing (MWG Biotech) using standard primers.

SPCrb1 and SPCrb2 peptides and antibodies

Peptides representing the amino acid residues 9-22 (LSQQQKQRQAETATC) and 35-51 (ATTARSRDRTKSAAQIC) of the Crumbs signal sequence were synthesized and coupled to keyhole limpet hemocyanin (Peptide Specialty Laboratories). Two antisera termed SPCrb1 and SPCrb2 were raised in rabbits (Charles River Laboratories). For affinity purification of both sera, the synthetic peptides were reduced and coupled to a SulfoLink column (Pierce). The antibodies were eluted with 0.1 M glycine-HCl, pH 2.0 and immediately neutralized with 2 M K_2HPO_4.

In vitro translation/translocation assay

Plasmid DNA was linearized by restriction enzymes, purified and used for *in vitro* transcription with SP6 polymerase as described before (Lyko et al., 1995). Transcripts were treated with DNase (Promega) and purified using G25 columns (GE Healthcare) as suggested by the manufacturers, respectively. *In vitro* translation was done for 30 min at 30 °C in 10 µl reactions using rabbit reticulocyte lysate (Promega), Redivue Pro-mix L-[^{35}S] cell labeling mix (GE Healthcare) and 1-1.5 eq of either canine pancreas or *Drosophila melanogaster* S2 cells rough microsomes produced according to the protocol from Walter and Blobel or Lundin et al., respectively (Lundin et al., 2007; Walter and Blobel, 1983). Signal peptide peptidase inhibitor (Z-LL)$_2$-ketone (Calbiochem-Merck) was dissolved in DMSO and added as indicated. In case that no stop codon is present, the reactions were terminated by the addition of 4 mM puromycin and incubation at 20 °C for 15 min. *In vitro* reactions were precipitated by adding two volumes of saturated ammonium sulphate solution for 20 min on ice and centrifugation at 16,000 *g* for 5 min. Proteins were resuspended in dH_2O, again precipitated with two volumes ice-cold absolute ethanol, pelleted and resuspended in SDS sample buffer (75 mM Tris-HCl, pH 6.8; 12% glycerol; 4% SDS; 0.01% Serva blue G and 100 mM DTT).

To separate microsomes from the supernatant, *in vitro* reactions were layered on a 50 µl cushion (50 mM Hepes-KOH, pH 7.6; 150 mM KOAc; 2 mM $Mg(OAc)_2$; 1 mM DTT; 500 mM sucrose) and membranes were pelleted by 5 min centrifugation at 100,000*g* and 4 °C (Beckman TLA 100 rotor or Sorvall S100-AT3 rotor). For SDS-PAGE, the supernatant was precipitated with ammonium sulphate as described above while the pellet was directly resuspended in SDS sample buffer.

Cells and transient transfection

HEK293 cells (ATCC: CRL-1573^{TM}) were grown in Dulbecco's modified Eagle's medium/F12-medium containing 17.5 mM glucose, 10% fetal calf serum and 2 mM L-glutamine. Cells were controlled for the absence of mycoplasma using standard PCR. HEK293 cells were transfected by the calcium phosphate precipitation method as described previously (Dultz et al., 2008). Transient expression is controlled by the CMV promoter when using pRK5rs-based constructs. As control, empty vector was used.

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

Metabolic labeling and immunoprecipitation

42 h post-transfection, cells were depleted of methionine and cysteine for 2 h (Dulbecco's modified Eagle's medium without L-methionine and L-cysteine, containing 4.5 g/l glucose, 2 mM L-glutamine and 10% fetal calf serum dialysed against PBS to remove small molecules up to 12 kD). As indicated in the figure legends, 5 μM epoxomycin, solved in methanol, was added 14 h before labeling and throughout labeling. Similarly, the SPP inhibitor $(Z\text{-}LL)_2$- ketone, solved in DMSO, was added using different concentrations as indicated in the figures. Labeling was done with 75 μCi/ml Redivue Pro-mix L-[^{35}S] cell labeling mix (GE Healthcare) as detailed in the figure legends. Cells were lysed in 1% Triton X-100 containing lysis buffer (50 mM Hepes-NaOH, pH 7.5; 150 mM NaCl; 1% (v/v) Triton X-100; 10% (v/v) glycerol; 1,5 mM $MgCl_2$; 2 mM EGTA; 1 mM phenylmethylsulfonyl fluoride, 10 μg/ml aprotinin) for 10 min at 4 °C. Non-solubilized material was separated by centrifugation (16,000*g*, 5 min, 4 °C). For immunoprecipitation, cleared lysates were diluted 1:1 with HNTG buffer (20 mM Hepes-NaOH, pH 7.5; 150 mM NaCl; 0.1% (v/v) Triton X-100; 10% (v/v) glycerol). Protein-A sepharose beads (GE Healthcare) and about 0.5 μg antibody per reaction were added. Antibodies were either SPCrb1 and SPCrb2 (both compare above) or Myc antibody prepared from 9E10 hybridoma supernatant using standard procedures. Samples were rotated for 3 h at 4 °C. The beads were washed four times with HNTG buffer and resuspended in SDS sample buffer.

Cell lysis and Western blot analyses

42 h post-transfection cells were lysed as described above. For deglycosylation, samples were incubated with 1000 units PNGaseF or EndoH (both: New England Biolabs) for 1 h at 37 °C according to instructions provided by the manufacturers. Protein samples were mixed with SDS sample buffer. To generate whole-cell-lysates, cells were directly resuspended in SDS sample buffer. Following SDS-PAGE and semi-dry blotting on nitrocellulose, blocking and antibody incubation were done with 5% skim milk (Roth) in 1 x TBS-T (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 5 mM EDTA, pH 8.0; 0.05% Triton X-100). Primary antibodies are as described above. Goat anti-mouse antibodies and goat anti-rabbit antibodies were conjugated with peroxidase (Sigma-Aldrich). Immunodetection was done with BM chemiluminescence blotting substrate (Roche).

Tris-Tricine SDS gel electrophoresis

The samples were analyzed by Tris-Tricine SDS polyacrylamide gel electrophoresis (PAGE) as described by Schägger and von Jagow (1987) using 16.5% polyacrylamide gels and Tricine-containing cathode buffer.

Drosophila stocks

Oregon R was used as wild type. For the overexpression experiments flies carrying $GAL4^{daG32}$ were crossed to flies carrying either $UAS::Crb^{8xMyc\text{-}intra}$, $UAS::HACrb^{8xMyc\text{-}intra}$ or $UAS::HA_{nc}Crb^{8x\text{-}Myc\text{-}intra}$. In the embryo, $GAL4^{daG32}$ is ubiquitously expressed (Wodarz et al., 1995). *Df(3R)Exel6199*, which removes the *crb* locus, was recombined with *Df(3L)H99*, which removes the pro-apoptotic genes *grim, reaper, sickle, and head involution defective* (Chen et al., 1996; Grether et al., 1995; White et al., 1994). All crosses were performed at 25 °C.

Lysis of embryos

Embryos were collected from apple-juice agar plates, dechorinated, lysed by sonification in 1% Triton X-100 containing lysis buffer (compare above) and incubated for 10 min on ice. The non-solubilized material was pelleted by centrifugation at 16,000*g* for 10 min. The supernatant was recovered and treated with PNGaseF as described above.

Immunohistochemistry of embryos

Embryos were heat fixed (Muller and Wieschaus, 1996) and stored in methanol before antibody staining. The primary antibodies were used with following dilutions: rabbit anti SPCrb1, 1:100; rat anti Crb2.8 (unpublished, S.K. & E.K.), 1:400; mouse anti c-Myc 9E10 (Developmental Studies Hybridoma Bank, University Iowa), 1:100. Secondary antibodies, conjugated to Cy2, Cy3 (Jackson) and Alexa647 (Dianova) were used 1:200. Stained embryos were mounted in glycerol-propylgallate and imaged with a confocal microscope LSM 510 (Zeiss) using LSM 510 software (Zeiss). Images were processed by ImageJ, Adobe Photoshop CS3 and Adobe Illustrator CS3.

In silico analyses

The annotated Crumbs sequences were retrieved from FlyBase (http://flybase.org/blast/) and NCBI databases (http://blast.ncbi.nlm.nih.gov). The GenBank accession numbers are given in the figure legends. To predict a cleavable signal sequence and to determine the signal peptidase cleavage site, computational methods were used including SignalP3.0, SignalP2.0 (Bendtsen et al., 2004; Nielsen et al., 1999), Predisi (Hiller et al., 2004) and Phobius (Kall et al., 2007). Of note, many signal sequence prediction programs have a threshold, e.g. for 70 residues. When analyzing proteins such as Crumbs, this parameter needs to be changed. Per definition an h-region of a signal sequence starts with a first quadruplet with at least three hydrophobic residues (isoleucine, valine, leucine, phenylalanine, methionine and alanine) thereby ignoring the initiating methionine residue. The h-region ends if no further quadruplet can be found (compare: von Heijne, 1983). Multiple sequence alignments were performed with ClustalW (Thompson et al., 1994).

Results

The potential signal sequence of Crumbs is conserved within the Drosophilidae

The protein sequence of *Drosophila melanogaster* Crumbs is deduced from cDNA sequences and was confirmed by genomic clones. The cDNA sequences contain a 5′ untranslated region, and the cDNA size is in agreement with the size of transcripts as detected by Northern Blot (Tepass et al., 1990). The deduced Crumbs protein precursor (GenBank: M33753, UniProtKB/Swiss-Prot P10040, version 113) comprises 2146 residues with a hydrophobic region located 65 residues C-terminal of the translation initiation site and a transmembrane domain close to the C-terminal end. Considering signal sequence cleavage, Crumbs is described as a type I membrane protein, i.e. exposing its N-terminus to the extracellular site and the C-terminus to the cytoplasm. The signal sequence prediction program SignalP V2.0.b2 (SignalP-NN) and the program PrediSi suggest a signal sequence of 83 residues. Other programs like SignalP 3.0 or Phobius either suggest different signal sequence cleavage sites

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

with low probabilities or no signal sequence cleavage. According to the UniProtKB/Swiss-Prot annotation, the signal sequence comprises 88 residues. Thus, it is unclear if the N-terminus of *D. melanogaster* Crumbs encodes a cleavable signal sequence and if so, how long it is.

To get further support for the length of the signal sequence, we compared the *D. melanogaster* sequence with those of other *Drosophila* species. Starting with the highly conserved cytoplasmic region, we identified Crumbs of 11 *Drosophila* species by BlastP searches using the FlyBase interface. Signal sequences were predicted by the programs SignalP V2.0.b2 and PrediSi, and used for a multiple sequence alignment (Fig. 1B). The *D. melanogaster* Crumbs signal sequence is almost identical to that of *D. simulans* and has a high degree of similarity to the predicted signal sequences of *D. sechellia*, *D. yakuba* and *D. erecta*. In particular, these five signal sequences are threonine- and glutamine-rich. The five species belong to the *Drosophila melanogaster* subgroup (Lemeunier et al., 1986). For *D. pseudoobscura pseudoobscura* and *D. persimilis* only incomplete signal sequences of 45 residues were available, which are identical to each other. These two signal sequences as well as other *Drosophila* species-derived signal sequences have a lower similarity and are glutamine- but not threonine-rich compared with the sequences of the *Drosophila melanogaster* subgroup. However, the C-terminal 25 residues of the predicted signal sequences, including the 'LLKRAIS' motif, the hydrophobic region, and the c-region are highly conserved (Fig. 1B).

Next, we compared the *D. melanogaster* Crumbs signal sequences with sequences from other insects (*Aedes aegyti*, *Tribolium castaneum, and Nasonia vitripennis*) and found essentially no significant conservation (Fig. 1C). While there is only one *crb* gene in these insects, three genes have been identified in mammals, including humans. Multiple sequence alignment of the *D. melanogaster* signal sequence with that of human (Crb1, Crb2, and Crb3) also revealed no significant conservation (Fig. 1C). On average, the Non-*Drosophilidae* Crumbs signal sequences have a length of 26 residues, which is a typical size for a signal sequence. In contrast, the *Drosophilidae* Crumbs signal sequences are unusually long in comprising about 83 residues. The increase in length is predominately caused by the net positively charged n-region.

The N-terminus of Crumbs is a cleavable signal sequence

Next, we experimentally addressed the question whether the N-terminus of *D. melanogaster* Crumbs is a cleavable signal sequence. We applied a cell-free *in vitro* translation/translocation assay in which mRNA encoding Crumbs is translated in the presence of ER-derived rough microsomes (RMs). Here, the detection of newly synthesized proteins depends on the presence of methionine and cysteine residues used for radioactive labeling. The predicted Crumbs signal sequence contains only a single methionine residue defining the translation initiation followed by an alanine residue. Since this methionine residue is prone for removal by methionine aminopeptidase (Meinnel et al., 2006, and data not shown), we exchanged the alanine residue at position 2 by a methionine residue. (All constructs containing this exchange are marked with ∗ throughout the text). To test ER targeting and signal sequence cleavage, we used transcripts encoding the N-terminal 321 residues of Crumbs for *in vitro* translation/translocations assays (Fig. 2A). The synthesis of $Crb^{1\text{-}321*}$ in the absence of RMs resulted in a protein of about 35 kD, representing the precursor, $preCrb^{1\text{-}321*}$ (Fig. 2B, lane 2). In the presence of canine pancreas-derived RMs (cRMs), a protein of about 38 kD and a peptide triplet of 8-9 kD were identified (Fig. 2B, lane 3). The sizes of the precursor (lane 2) and the protein running slower than the precursor (lane 3) are inferred from the predicted molecular weights. Upon addition of the glycosylation inhibitor acceptor tripeptide, we obtained a 35 and a 24 kD protein as well as the peptide triplet (Fig. 2B, lane 4). Taken together, these results suggest that the 38 kD protein is a glycoprotein differing from the 35 kD precursor by the increase of about 12 kD due to four glycosylation events and the concurrent reduction of about 8-9 kD due to cleavage of an 83 residues-long signal sequence. Thus, this Crumbs variant is targeted to the ER and translocated across the ER membrane, the signal sequence is cleaved off and further processed while the translocated protein accumulates as glycoprotein ($gpCrb^{84\text{-}321}$). To corroborate these findings, we *in vitro* translated transcripts encoding the N-terminal 83 residues of Crumbs, i.e. the predicted signal peptide, SP^{Crb*}. This signal peptide marker was detected as a 9 kD protein (Fig. 2B, lane 5, 10) and co-migrated with the largest peptide of the peptide triplet described above. Furthermore, we applied RMs derived from *Drosophila* S2 cells (Fig. 2B, *D*RMs, lane 6-9) and obtained similar results for ER targeting and translocation.

To verify that the identified peptide triplet represents the signal peptide and its fragments, two antisera (SPCrb1 and SPCrb2) were raised against peptides within the signal sequence (Fig. 1A). Both antibodies were used for immunoprecipitation of the *in vitro* generated $Crb^{1\text{-}321*}$ and its derivatives (Fig. S1). In both cases, the precursor ($preCrb^{1\text{-}321*}$) and the three peptides of about 8-9 kD were detected (Fig. S1, lane 5, 9). The respective pre-immune sera as well as an unrelated antibody failed to detect these proteins (Fig. S1, lane 4, 8, 12). A pre-incubation of the SPCrb1 or SPCrb2 antibodies with the specific peptide but not with an unspecific peptide prevented the immunoreaction (Fig. S1, lane 6, 7, 10, 11). Taken together, both antibodies specifically detect the precursor and the peptide triplet indicating that the peptides of the triplet are derived from the Crumbs signal sequence.

To test Crumbs targeting and signal sequence cleavage in cells, we transiently expressed a truncated Crumbs variant, $Crb^{Myc\text{-}intra*}$ (Fig. 2A), that was previously used for successful rescue experiments in flies (Wodarz et al., 1995). In embryos homozygous mutant for *crumbs*, no contiguous cuticle but only cuticle crumbs are formed, resulting in the naming of this mutant. The overexpression of $Crb^{Myc\text{-}intra}$ (previously named $crb^{intra\text{-}myc}$) leads to the same degree of restoration of large parts of the cuticule as overexpression of full-length Crumbs (Wodarz et al., 1995). Metabolically labeled cells were lysed and used for immunoprecipitation. The SPCrb1 antibody precipitated two peptides of about 8-9 kD (Fig. 2C, lane 5). Since the larger one co-migrates with the signal peptide marker (Fig. 2C, lane 4), we conclude that the two peptides represent SP^{Crb*} and its fragment SPF^{Crb*}. In contrast, the SPCrb2 antibody precipitated SPF^{Crb*} and a further fragment of about 6 kD (Fig. 2C, lane 6). In a control, the Myc antibody precipitated a 19 kD and a 14 kD protein (Fig. 2C, lane 7). The 19 kD protein was also precipitated by the SPCrb1 and the SPCrb2 antibodies (Fig. 2C, lane 5, 6), indicating that this protein is the precursor, $preCrb^{Myc\text{-}intra*}$. The 14 kD protein is the mature, glycosylated protein, $gpCrb^{Myc\text{-}intra}$, which is in agreement with the predicted molecular weight of the signal sequence-cleaved glycoprotein. The glycosylation was experimentally verified by treatment with Peptide:N-glycosidase F (PNGaseF), which reduced the mass to about 11 kD (data not shown). From these results we conclude that in cells Crumbs is targeted to the ER, is inserted and that the signal sequence is cleaved off. Additionally, the resulting signal peptide is further processed into signal peptide fragments.

To confirm the length of the signal sequence, we experimentally determined the signal sequence cleavage site by Edman

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

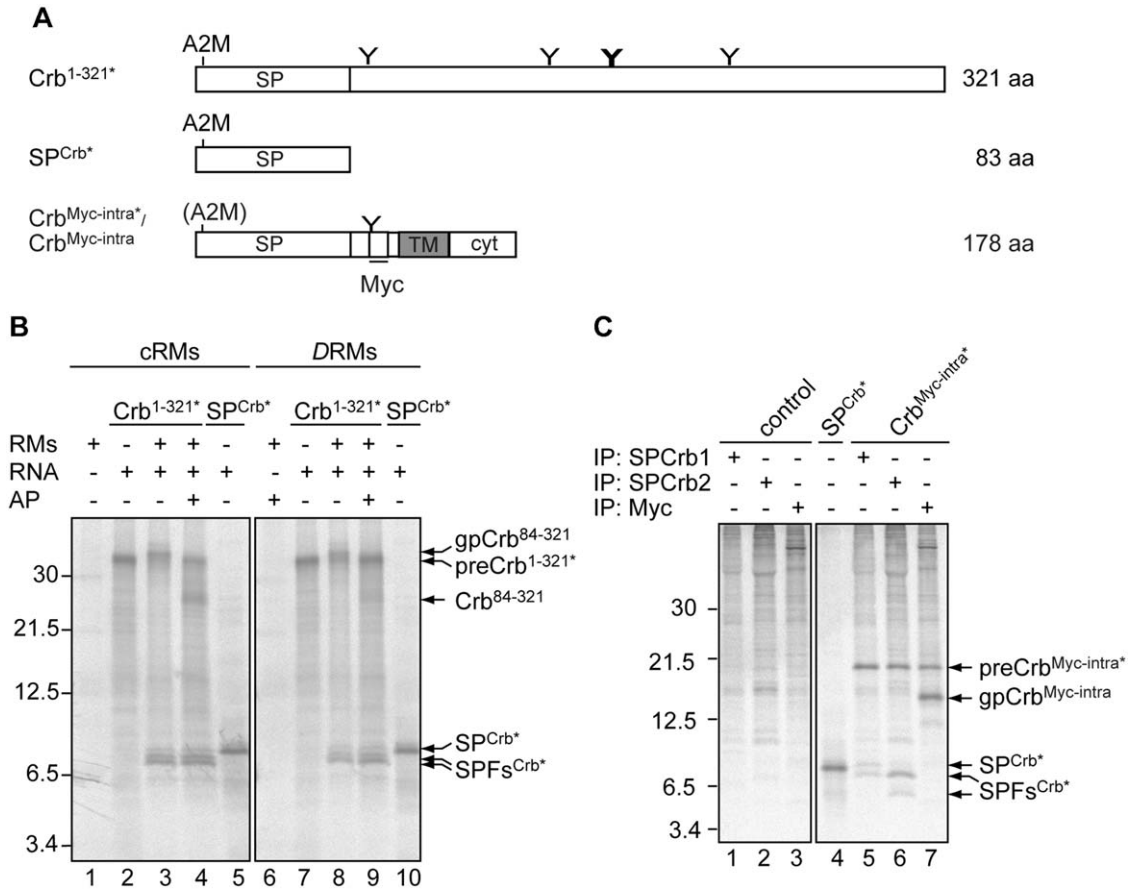


Fig. 2. Identification of the Crumbs signal peptide and its fragments by a cell-free *in vitro* assay and in cells. (A) The cartoon depicts the constructs used to analyze Crumbs targeting and signal sequence cleavage. $Crb^{1\text{-}321*}$ encodes the N-terminal 321 amino acids of Crumbs including five glycosylation sites as indicated (Y). Note that two glycosylation sites are very close to each other and are therefore not resolved (bold). To detect proteins derived from the N-terminus by radioactive labeling even when the N-terminal methionine residue is cleaved off, we exchanged the alanine residue at position 2 by a methionine residue (compare Fig. 1A). All constructs encoding this A2M mutation are marked by an asterisk (*). As marker, we used a construct encoding the N-terminal 83 residues of Crumbs followed by a stop codon. This marker (SP^{Crb*}) also includes the A2M mutation. $Crb^{Myc\text{-}intra}$/$Crb^{Myc\text{-}intra*}$ encode proteins consisting of the Crumbs signal sequence, a linker with a glycosylation site and a Myc-tag, the Crumbs transmembrane region (TM), and the Crumbs cytoplasmic region (cyt). (B) $Crb^{1\text{-}321*}$ and SP^{Crb*} were translated in the absence and presence of canine pancreas-derived rough microsomes (cRMs) or *Drosophila* S2 cell-derived rough microsomes (*D*RMs), respectively. Reactions in lane 4 and 9 were treated with acceptor tripeptide (AP) to inhibit glycosylation. The *in vitro* translation/translocation reactions were precipitated, proteins were separated by SDS-PAGE, and analyzed by autoradiography. (C) HEK293 cells transiently expressing $Crb^{Myc\text{-}intra*}$ were labeled for 30 min. Antigens were immunoprecipitated with SPCrb1, SPCrb2 or Myc antibodies, separated by SDS-PAGE and analyzed by autoradiography. *In vitro* translated SP^{Crb*} was co-separated as size-marker. Of note, in some figures, e.g. Fig. 2B and C, it seems that the size marker migrates slightly faster than the signal peptidase cleavage-derived signal peptide, but compare Fig. 3A, C. The most likely explanation is that the migration behavior of the marker is different due to the low net protein content in comparison to lanes including RMs and signal peptide fragments. However, we cannot completely rule out that the marker and the processed signal peptide are either different in length or might undergo different post-translational modifications. RMs, rough microsomes; cRMs, canine-pancreas derived rough microsomes; *D*RMs, *Drosophila* S2 cell-derived rough microsomes; pre, precursor; gp, glycoprotein; SPF, signal peptide fragment.

degradation. We expressed $Crb^{Myc\text{-}intra}$ in HEK293 cells and used cell lysates for immunoprecipitation with the Myc antibody. After separation by SDS-PAGE, we stained the gel with Coomassie. As before, we received a 19 kD and a 14 kD protein (data not shown, but see Fig. 2C). The 14 kD protein was N-terminally sequenced (data not shown) and found to start with amino acid 84 of the $Crb^{Myc\text{-}intra}$ precursor, which is consistent with an 83 residues-long signal sequence.

The extended n-region has no effect on ER targeting and signal sequence cleavage

Signal sequences can differently affect ER targeting, protein insertion and signal sequence cleavage. Therefore, we wanted to know if the extended n-region, causing the unusual length of the Crumbs signal sequence, contributes to the function and efficiency of the signal sequence. To test this, we deleted the N-terminal 56 amino acid residues of the Crumbs signal sequence. The resulting minimal signal sequence consists essentially of the hydrophobic core region flanked by a short, positively charged n-region and the c-region ($Crb^{56\text{-}321}$, Fig. S2A). Using the *in vitro* translation/translocation assay, we detected the precursor $Crb^{56\text{-}321}$ and its translocated glycoprotein $gpCrb^{84\text{-}321}$ (Fig. S2B, lanes 5-7). The potential minimal signal peptide of about 3 kD is not resolved on the gel. A densitometric analysis revealed no significant difference in the amount of translocated proteins when comparing $Crb^{1\text{-}321*}$ and $Crb^{56\text{-}321}$ (data not shown). Thus, similar to the full-length signal sequence, the minimal signal sequence accomplishes ER targeting and translocation.

Next, we tested the translocation potential of both signal sequences in cells. To this end, we replaced the wild type signal sequence of $Crb^{Myc\text{-}intra*}$ with the minimal signal sequence ($Crb^{\Delta(1\text{-}56)\text{-}Myc\text{-}intra}$, Fig. S2A). Upon transient expression and metabolic labeling, we found SP^{Crb*}, SPF^{Crb*}, and the precursor $preCrb^{Myc\text{-}intra*}$ by immunoprecipitation with the SPCrb1 antibody. No proteins were detected analyzing cells expressing $Crb^{\Delta(1\text{-}56)\text{-}Myc\text{-}intra}$, since the epitope of SPCrb1 is deleted

(Fig. S2C, lanes 1-4). Using the Myc antibody, the precursors and mature proteins of $Crb^{Myc\text{-}intra*}$ and $Crb^{\Delta(1\text{-}56)\text{-}Myc\text{-}intra}$ were found (Fig. S2C, lanes 5-7). Again, we observed no significant difference in the expression and translocation, thereby confirming the results from the *in vitro* translation/translocation experiments. We conclude that the extended n-region of the unusually long signal sequence of Crumbs does not affect the efficiency of targeting and translocation under the experimental conditions used.

SP^{Crb} is processed by Drosophila SPP

Considering that the signal peptide was detected as peptide triplet, we next studied the proteolytic processing of the signal peptide. Since signal peptides can be substrates for the intramembrane cleaving aspartyl protease signal peptide peptidase (SPP), the SPP-specific inhibitor $(Z\text{-}LL)_2$-ketone (Weihofen et al., 2002) was tested for its ability to prevent proteolysis of the Crumbs signal peptide. In the presence of $(Z\text{-}LL)_2$-ketone, the largest peptide of the peptide triplet accumulated, whereas adding just the organic solvent did not influence the processing of SP^{Crb*} (data not shown, but see Fig. 3A, lanes 5-8). Usually, SPP processing is accompanied by the liberation of N-terminal fragments to the cytosol. To determine the localization of the full-length Crumbs signal peptide and its fragments, the microsomal membranes of the *in vitro* translation/translocation reactions were pelleted by centrifugation through a sucrose cushion. The pellet and the supernatant, which represent the membranes and the cytosol, respectively, were analyzed independently. SP^{Crb*} was only detected in the pellet fraction, while the fragments were found in both fractions (Fig. 3A, lane 3-8). This result suggests that SP^{Crb*} is processed by SPP *in vitro* and that fragments can be liberated into the cytosol.

In order to analyze if SP^{Crb} processing also occurs in a cellular context, we took advantage of the observation that $(Z\text{-}LL)_2$-ketone is a membrane-permeable compound (Bland et al., 2003). HEK293 cells transiently expressing $Crb^{Myc\text{-}intra*}$ were incubated with or without the inhibitor, metabolically labeled, and lysed. Without the inhibitor, the $Crb^{Myc\text{-}intra*}$ precursor, the signal peptide SP^{Crb*}

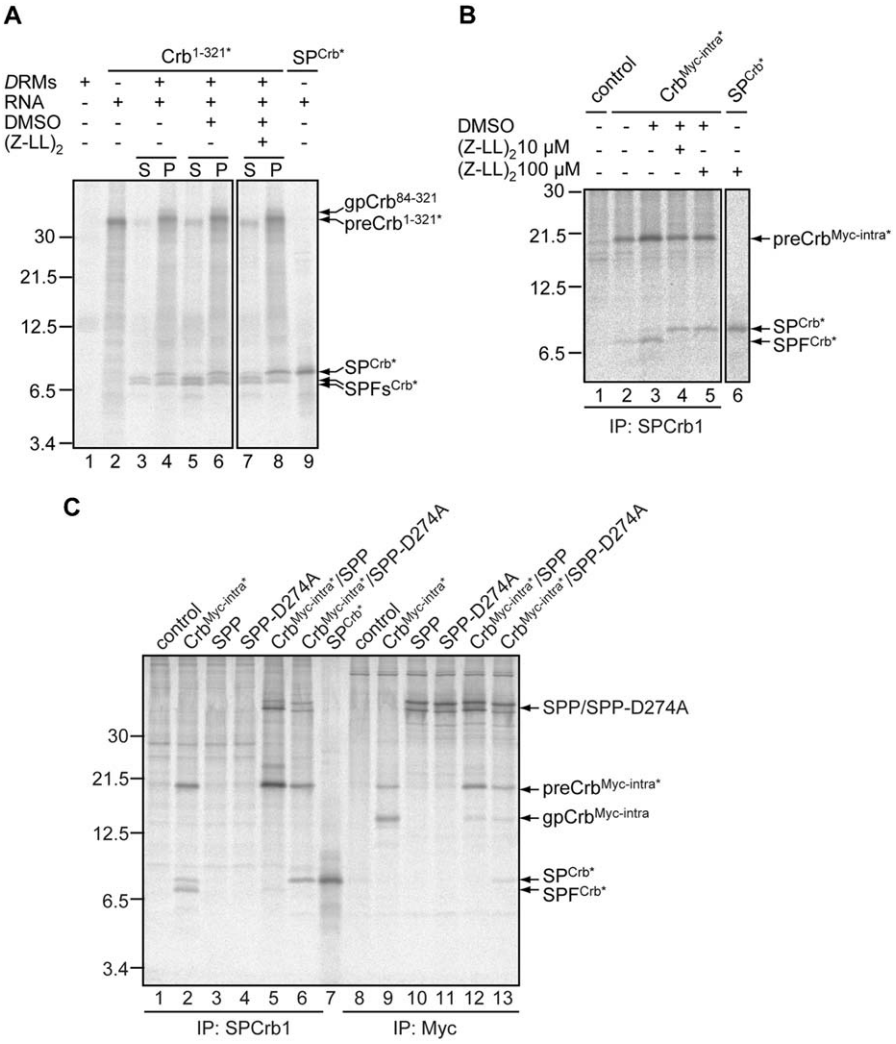


Fig. 3. SPP-mediated proteolytic processing of the Crumbs signal peptide. (A) $Crb^{1\text{-}321*}$ was translated in the presence of *D*RMs and 10 µM $(Z\text{-}LL)_2$- ketone. The microsomal membranes were separated from the cytosol by centrifugation through a sucrose cushion. The pellet (P) and the supernatant (S) were analyzed as described above (Fig. 2A). Of note, the 6 kD fragment observed in lane 5 corresponds to the antigen identified by the SPCrb2 as shown in Fig. 2C. (B) HEK293 cells expressing $Crb^{Myc\text{-}intra*}$ were incubated for 14 hours with different concentrations of $(Z\text{-}LL)_2$- ketone (as indicated), metabolically labeled for 30 min and lysed. Antigens were immunoprecipitated with the SPCrb1 antibody, separated by SDS-PAGE and visualized by phosphorimaging. (C) HEK293 cells expressing $Crb^{Myc\text{-}intra*}$ and either SPP or SPP-D274A were metabolically labeled for 30 min. Both SPP constructs include a Myc-tag, which is inserted N-terminally of the ER retrieval signal (four C-terminal residues encoding KKGK). Cell lysates were used for immunoprecipitation with the SPCrb1 or with the Myc antibody, respectively, and analyzed as described above. As a size marker, *in vitro* translated SP^{Crb*} was applied on each gel.

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

and the 8 kD signal peptide fragment were immunoprecipitated by the SPCrb1 antibody (Fig. 3B, lane 3). In the presence of $(Z\text{-}LL)_2$-ketone, only the precursor and SP^{Crb*} were detected (Fig. 3B, lane 4, 5), indicating that SPP cleavage also occurs in cells.

To test whether the previously described *Drosophila* SPP (Casso et al., 2005) is able to process SP^{Crb}, we cloned the corresponding cDNA. SPP family members share two aspartic acid residues at their catalytically active sites, one of which is aspartic acid 274 in *Drosophila* SPP. To receive a catalytically inactive mutant (Golde et al., 2009; Weihofen et al., 2002), we replaced this residue by alanine. Furthermore, we included a Myc-tag in both cDNAs to detect the proteins, here referred to as SPP and SPP-D274A. Upon expression of $Crb^{Myc\text{-}intra*}$ in HEK293 cells, the SPCrb1 antibody detected the 19 kD precursor, $preCrb^{Myc\text{-}intra*}$, as well as the 9 kD SP^{Crb*} and its 8 kD fragment SPF^{Crb*} (Fig. 3C, lane 2). Upon co-overexpression of SPP, the precursor and small amounts of the signal peptide fragment were detected (Fig. 3C, lane 5), suggesting an efficient processing of SP^{Crb*}. Of note, only small amounts of the signal peptide fragment are detectable since upon overexpression of SPP the $preCrb^{Myc\text{-}intra*}$ precursor accumulates resulting in the reduced production of SP^{Crb*} and its fragment. A second reason might be that due to efficient SPP processing of the available SP^{Crb*}, the fragments are also efficiently degraded as shown later. In contrast, co-overexpression of $Crb^{Myc\text{-}intra*}$ and SPP-D274A resulted in the accumulation and stabilization of SP^{Crb*} (Fig. 3C, lane 6), which co-migrated with the *in vitro* synthesized signal peptide marker (Fig. 3C, lane 7). As control for the expression of $Crb^{Myc\text{-}intra*}$, SPP, and SPP-D274A, we analyzed cell lysates by immunoprecipitation with the Myc antibody. We detected the precursor $preCrb^{Myc\text{-}intra*}$ and its mature protein, $gpCrb^{Myc\text{-}intra}$, (compare Fig. 2C). Secondly, we identified SPP and its mutant as doublet of a 43 kD and a 49 kD protein, respectively. The two forms correspond to differently glycosylated SPP proteins, as known for human SPP (Nyborg et al., 2004; and data not shown). In the presence of $Crb^{Myc\text{-}intra*}$, SPP and SPP-D274A were also co-immunoprecipated by the SPCrb1 antibody (Fig. 3C, lane 5, 6). Similar co-immunoprecipitation results have been described upon overexpression of SPP and SPP-like proteases in other systems (Fluhrer et al., 2006) and coincide with the accumulation of the substrate precursor proteins. In summary, these results demonstrate that SP^{Crb*} can be processed by *Drosophila* SPP.

Besides SPP, a second member of the SPP family of intramembrane cleaving proteases, SPP-L3, is predicted in the *Drosophila* genome (Casso et al., 2005). Consequently, we tested whether *D. melanogaster* SPP-L3 can also process SP^{Crb}. We cloned SPP-L3 and the corresponding catalytically inactive mutant SPP-L3-D307A, and co-overexpressed both with $Crb^{Myc\text{-}intra*}$ in HEK293 cells (Fig. S3). Since we found that the SPCrb1 antibody extensively cross-reacts with SPP-L3 (data not shown), the SPCrb2 antibody was applied for this analysis. Neither co-overexpression of SPP-L3 nor SPP-L3-D307A with $Crb^{Myc\text{-}intra*}$ affected the amounts or processing of SP^{Crb*}. As described for SPP and depending on the presence of $Crb^{Myc\text{-}intra*}$, SPP-L3 and SPP-L3-D307A were co-immunoprecipitated with the SPCrb2 antibody. By immunoprecipitation with the Myc antibody, the expression of $Crb^{Myc\text{-}intra*}$, SPP-L3, and SPP-L3-D307A was confirmed (Fig. S3, lanes 9-13). Taken together, these results demonstrate that SP^{Crb} is a substrate for SPP, but not a substrate for SPP-L3.

SP^{Crb*} has a short half-life and is degraded by the proteasome

To see whether SP^{Crb} or its fragments accumulate in cells, we addressed their half-lifes. We pulse-labeled HEK293 expressing $Crb^{Myc\text{-}intra*}$ for 30 min and chased the cells for 15, 30 and 45 min. The amounts of SP^{Crb*} and its fragments were reduced by time but small amounts of SP^{Crb*} were still detectable after 45 min (Fig. 4A, lane 5). As control, the amounts of the mature protein and the precursor were not affected (Fig. 4A, lanes 8-11). Thus, SP^{Crb*} is most likely degraded by time. Although it is unlikely, we cannot rule out that SP^{Crb*} by time is tightly associated with a further molecule thereby preventing its immunoprecipitation and detection. In a next step, we tried to identify SP^{Crb} by Western Blot analysis, i.e. under steady state conditions and failed using cells overexpressing $Crb^{Myc\text{-}intra}$ (Fig. 4B, lane 2). In contrast, upon co-expression of SPP-D274A, SP^{Crb} accumulates and is detected by Western Blot (Fig. 4B, lane 6). Of note, we detected the wild type signal peptide in this experiment and not the A2M mutant as required for radioactive labeling.

A potential event following SPP-mediated processing of a signal peptide is the proteasomal degradation (Bland et al., 2003). To test whether this applies for SP^{Crb}, we added the proteasome inhibitor epoxomycin in a pulse-chase experiment in which $Crb^{Myc\text{-}intra*}$ was overexpressed in HEK293 cells. Under this condition, SP^{Crb*} and its fragments were still detected after 120 min, in contrast to 30 min without the proteasome inhibitor (Fig. 4C). In summary, we conclude that the signal peptide fragments are degraded by the proteasome.

No endogenous Crumbs signal peptide is detected in the Drosophila embryo but SPCrb1 detects its epitope when signal sequence cleavage is prevented

To test whether the signal peptide can be detected in wild type *Drosophila* tissues, we performed immunofluorescence microscopy, using the two antibodies generated against the signal sequence. The SPCrb2 antibody did not result in any staining of tissues known to express *crb* (embryonic epithelia, eye imaginal discs, adult eyes). In contrast, the SPCrb1 antibody detected an epitope at the leading edge during dorsal closure (data not shown). Secondly, the antibody stained punctate structures in cells of the amnioserosa throughout embryonic development (compare Fig. 6A, B). This staining is unspecific, since it is also found in embryos lacking the *crumbs* locus (data not shown). Since in these embryos the amnioserosa undergoes apoptosis at early developmental stages, we used embryos homozygous mutant for the deficiency *Df(3R)Exel6199*, which lacks the *crumbs* locus, and the deficiency *Df(3L)H99* to delay apoptosis. The specificity of the staining at the leading edge could not be addressed, since the respective cells of *Df(3R)Exel6199*, *Df(3L)H99* homozygous mutant embryos are still severely affected.

Next, we addressed the question whether the SPCrb1 antibody can detect SP^{Crb} at all when its antigen is expressed in a *Drosophila* tissue. Therefore, we overexpressed a transgene that encodes a Crumbs variant with a non-cleavable (nc) signal sequence named $HA_{nc}Crb^{8xMyc\text{-}intra}$, which should result in the accumulation of the precursor. To facilitate detection, we fused an HA-tag to the very N-terminus and an 8x Myc-tag N-terminal to the transmembrane domain. As controls a non-mutated form, $HACrb^{8xMyc\text{-}intra}$ as well as a non-HA-tagged form, $Crb^{8xMyc\text{-}intra*}$, were used (Fig. 5A). First, we tested signal sequence cleavage and protein glycosylation of the respective proteins by *in vitro* translation/translocation assays. $Crb^{8xMyc\text{-}intra*}$ and $HACrb^{8xMyc\text{-}intra}$ have a cleavable signal sequence that is further processed by SPP, whereas the signal sequence of $HA_{nc}Crb^{8xMyc\text{-}intra}$ was not cleaved (Fig. 5B, lanes 2, 3, 5, 6 for $Crb^{8xMyc\text{-}intra*}$, lanes 8, 9, 11, 12 for $HACrb^{8xMyc\text{-}intra}$, and lanes 13, 14, 16, 17 for $HA_{nc}Crb^{8xMyc\text{-}intra}$). Furthermore, all three proteins were translocated since the presence of acceptor tripeptide interferes with glycosylation (Fig. 5B, lanes 4, 10, 15 in comparison to lanes 3, 9, 14). Next, we analyzed the processing of

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

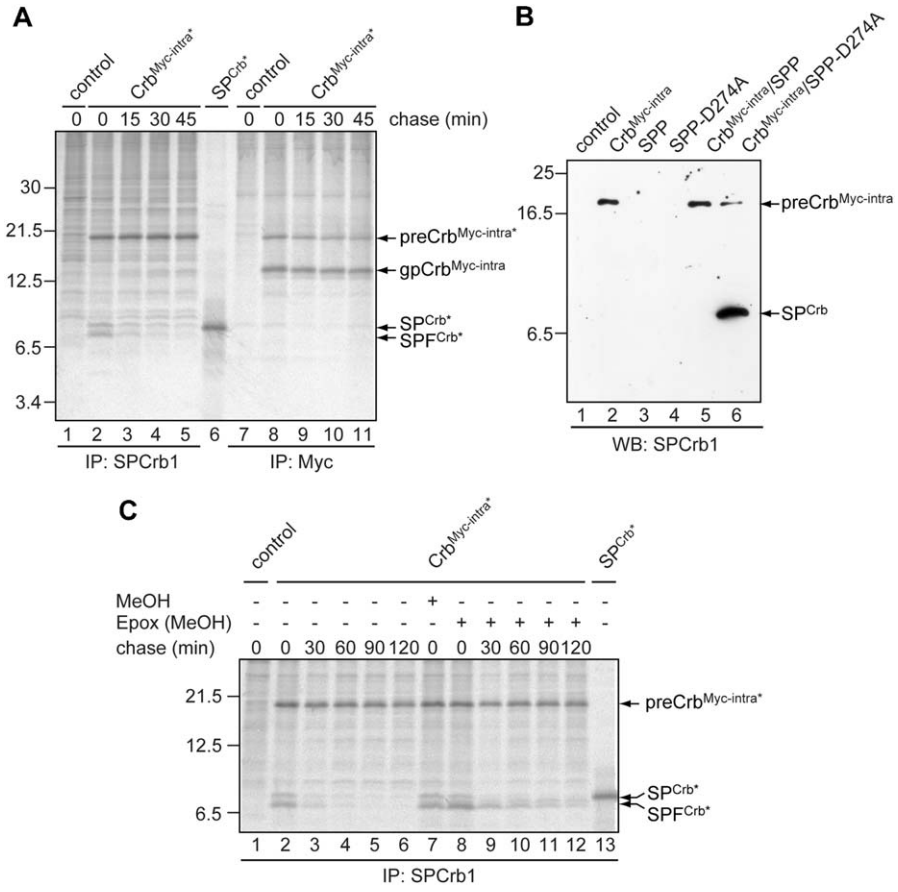


Fig. 4. The half-life and fate of the Crumbs signal peptide and its fragments. (A) HEK293 cells expressing $Crb^{Myc\text{-}intra*}$ were metabolically labeled for 30 min. At the indicated time points, proteins were immunoprecipitated using the SPCrb1 antibody or the Myc antibody, respectively. Antigens were separated by SDS-PAGE and analyzed by autoradiography. (B) HEK293 cells expressing $Crb^{Myc\text{-}intra}$ alone or together with SPP or SPP-D274A were lysed, proteins were separated by SDS-PAGE and then analyzed by Western blotting. (C) HEK293 cells expressing $Crb^{Myc\text{-}intra*}$ were treated with 5 μM epoxomycin as indicated. Cells were metabolically labeled for 30 min and chased as indicated. Cell lysates were used for immunoprecipitation with the SPCrb1 antibody, analyzed by SDS-PAGE and autoradiography. As a size marker, *in vitro* translated SP^{Crb*} was co-migrated when samples were analyzed by autoradiography.

these proteins in HEK293 cells. While signal sequence cleavage was observed in $Crb^{8xMyc\text{-}intra*}$ and $HACrb^{8xMyc\text{-}intra}$, it did not take place in the signal sequence cleavage mutant. In the presence of co-overexpressed SPP-D274A, which inhibits signal peptide processing, we also identified SP^{Crb} and the HA-tagged signal peptide (data not shown).

To address the question whether these proteins remain associated with the ER membrane or are reliably transported along the secretory pathway, we analyzed their glycosylation status in HEK293 cells and in *Drosophila* embryos expressing any of these transgenes. For this purpose, cell and tissue lysates were treated with endoglycosidase H (EndoH) and PNGaseF to determine the glycosylation pattern (Fig. 5C, D). We observed different sensitivities of $Crb^{8xMyc\text{-}intra}$ and $HACrb^{8xMyc\text{-}intra}$ towards the two enzymes, while $HA_{nc}Crb^{8xMyc\text{-}intra}$ did not exhibit any difference (Fig. 5C). $Crb^{8xMyc\text{-}intra*}$ and $HACrb^{8xMyc\text{-}intra}$ obtain EndoH-resistant, complex-type oligosaccharides, suggesting that both proteins reach at least the medial-Golgi apparatus. In contrast, $HA_{nc}Crb^{8xMyc\text{-}intra}$ did not receive complex-type oligosaccharides and thus presumably remains in the ER membrane or cis-Golgi apparatus. In *Drosophila* embryo lysates, we found that all three proteins are sensitive for PNGaseF treatment (Fig. 5D), suggesting that at least each protein is efficiently inserted into the ER membrane.

In order to determine the behavior of the transgene-encoded proteins ($Crb^{8xMyc\text{-}intra}$, $HACrb^{8xMyc\text{-}intra}$ and $HA_{nc}Crb^{8xMyc\text{-}intra}$) in *Drosophila* embryos, we analyzed the localizations of the mature proteins and the respective signal sequences. Upon overexpression of $Crb^{8xMyc\text{-}intra}$ and $HACrb^{8xMyc\text{-}intra}$, the Myc staining was associated with the plasma membrane as well as with the ER (data not shown). The non-cleavable $HA_{nc}Crb^{8xMyc\text{-}intra}$ was mostly retained in the ER, with some staining in the cytosol. Co-localization of the Myc epitope and the Crumbs signal sequence in the ER and the cytoplasm led to the conclusion, that, unlike $Crb^{8xMyc\text{-}intra}$ and $HACrb^{8xMyc\text{-}intra}$, the non-cleavable protein was not transported to the plasma membrane (Fig. 6C-C″).

From these results we conclude that the SPCrb1 antibody is able to detect its epitope in *Drosophila,* but only upon overexpression and when part of the precursor protein. Considering that in the cell-based assay SP^{Crb*} is degraded by the proteasome, these data suggest that the endogenous signal peptide is short-lived.

Discussion

D. melanogaster Crumbs is synthesized with an unusually long, cleavable signal sequence of 83 residues

Starting our investigation, it was unclear whether the N-terminus of Crumbs is a cleavable signal sequence, since experimental data were not available and several prediction programs did not suggest signal sequence cleavage. Assuming that signal sequence cleavage takes place, Crumbs will be a type I membrane protein with a long, extracellular N-terminus, including EGF-like domains and laminin A

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

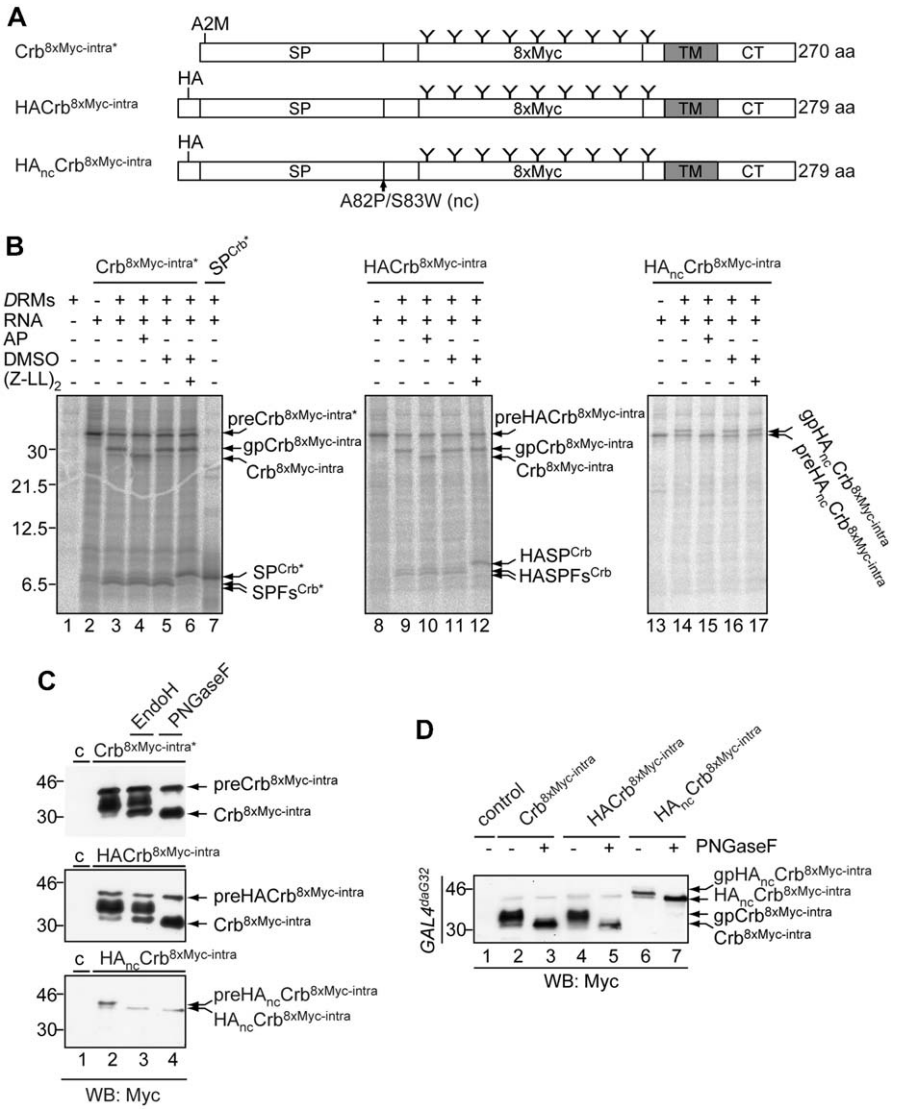


Fig. 5. Biochemical analysis of a signal sequence cleavage site mutant. (A) The cartoon shows the construct $Crb^{8xMyc\text{-}intra*}$, $HACrb^{8xMyc\text{-}intra}$ and $HA_{nc}Crb^{8xMyc\text{-}intra}$. In $Crb^{8xMyc\text{-}intra*}$ the single Myc-tag of $Crb^{Myc\text{-}intra*}$ is replaced by eight Myc-tags to improve its detection. The asterisk again indicates the A2M mutation required to detect radioactively labeled signal peptides. $HACrb^{8xMyc\text{-}intra}$ is the N-terminally HA-tagged analogue to $Crb^{8xMyc\text{-}intra}$. $HA_{nc}Crb^{8xMyc\text{-}intra}$ has the two mutations A82P and S83W, thereby destroying the signal sequence cleavage site. (B) Transcripts were raised from the constructs described in (A) and applied for *in vitro* translation/translocation assays (compare Figs. 2B and 3A). As a size marker, *in vitro* translated SP^{Crb*} was used. (C) HEK293 cells expressing $Crb^{8xMyc\text{-}intra*}$, $HACrb^{8xMyc\text{-}intra}$ and $HA_{nc}Crb^{8xMyc\text{-}intra}$ were lysed in the presence of 1% Triton X-100 to subsequently treat the lysates with EndoH and PNGaseF as indicated. As control (c) cells transfected with empty vector were used. (D) *Drosophila* embryos ubiquitously overexpressing various transgene-encoded proteins (as indicated) using $GAL4^{daG32}$ were lysed. The lysates were treated with PNGaseF as indicated, separated by SDS-PAGE and analyzed by Western blotting.

G-domain like repeats. Following a classical transmembrane region, the C-terminus is cytosolic. Alternatively, if the signal sequence is not cleaved, Crumbs will be a double-pass membrane protein with both termini facing the cytosol. In this scenario, a short, about 65 residues-long, cytosolic N-terminus is followed by a hydrophobic region that spans the membrane. Again, the EGF-like domains are in the extracellular space separated by the transmembrane region from the cytosolic C-terminus. Indeed, the association of the cytosolic C-terminus with the PDZ-domain containing proteins Stardust, *D*PATJ and *D*Lin7 (Bachmann et al., 2008; Bachmann et al., 2001; Klebes and Knust, 2000) and the described glycosylation (Koles et al., 2007) are in agreement with both scenarios.

Our data clearly confirm ER targeting and insertion of the Crumbs precursor, signal sequence cleavage, and glycosylation of the mature protein. Furthermore, we identified the cleaved signal peptide, termed SP^{Crb}. Importantly, signal sequence cleavage not only takes place in the presence of canine pancreas-derived rough microsomes but also when *Drosophila* S2 cell-derived rough microsomes, i.e. from the same species, are applied. Furthermore, no co-staining of *Drosophila* tissues using an antibody against the signal sequence in combination with an antibody against the mature (endogenous) protein was obtained, arguing that the signal sequence is efficiently cleaved off in flies. Similarly, the overexpression of a transgene with a cleavable signal sequence as determined by *in vitro* assays did not result in a co-staining applying both antibodies. By contrast, we obtained co-staining upon overexpression of a precursor protein with a signal sequence cleavage site mutation. Thus, we conclude that Crumbs is a type I membrane protein with a cleavable signal sequence and suggest that signal sequence cleavage occurs in the animal tissues.

The signal sequence of *D. melanogaster* Crumbs comprises 83 residues. We deduce this from the following findings. (i) A marker peptide comprising the first 83 residues of Crumbs was found to co-migrate with the *in vitro*- or *in cellulo*-synthesized, cleaved signal sequence using high-resolution gels. (ii) The signal sequence cleavage site was confirmed by N-terminal protein

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

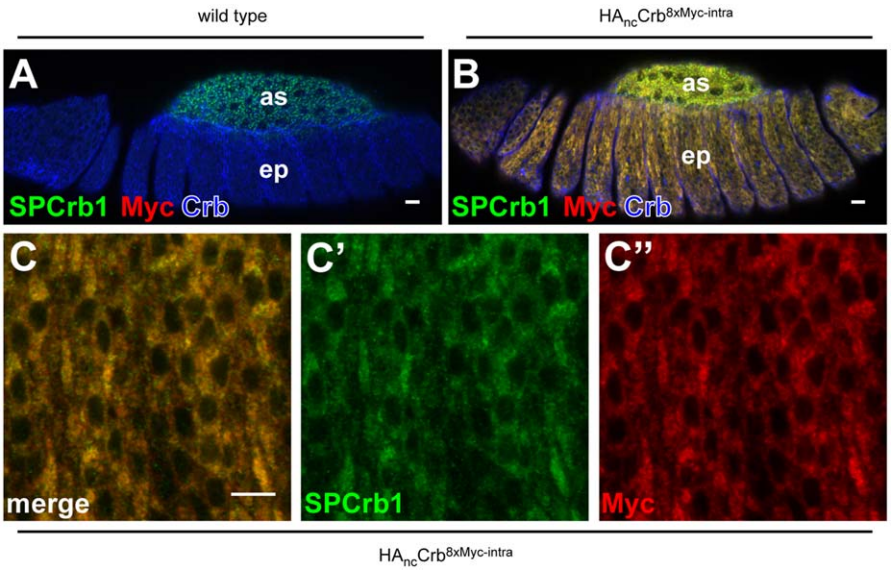


Fig. 6. The SPCrb1 antibody can detect the epitope of SP^{Crb} in *Drosophila*. (A) Lateral view of a wild type *Drosophila* embryo (stage 13). The amnioserosa (as) covers the dorsal part of the embryo, between the epidermis (ep). Crumbs (blue, detection of endogenous full-length Crumbs) is expressed in both tissues and localizes to the plasma membrane. The SPCrb1 antibody detects punctate structures in the cells of the amnioserosa (green), which are not SP^{Crb}. There is no signal for the anti-Myc antibody detectable. (B) Lateral view of a *Drosophila* embryo (stage 14) overexpressing the non-cleavable HA$_{nc}$Crb$^{8xMyc-intra}$ ubiquitously using GAL4^{daG32}, stained with anti-Myc (red), SPCrb1 antibody (green) and anti-Crb (blue). The SPCrb1 antibody detects a signal, which co-localizes with the Myc-tagged HA$_{nc}$Crb$^{8xMyc-intra}$ protein in the cytoplasm and ER. Of note, the punctate staining within the amnioserosa is not changed in comparison to A. (C-C″) as in B with higher magnification of the epidermis. In A-C anterior is left. Scale bars represent 10 μm (A, B) and 5 μm (C-C″).

sequencing of a Crumbs-derived protein, comprising the signal sequence and the following 13 residues. This suggests that also in the endogenous protein this cleavage site is used. (iii) The mutation of the signal cleavage site prevented the accumulation of SP^{Crb}. (iv) The length and amino acid sequence of the signal sequence is highly conserved within the *Drosophilidae* Crumbs proteins.

Comprising 83 residues, the Crumbs signal sequence is unusually long, with the increase in length caused by an extended n-region. Typically, signal sequences are in the range of 15-30 residues but some longer ones are also known (Kapp et al., 2009). The neuropilin-like product of the *murine discoidin, CUB and LCCL domain containing 2- (Dcbld2)* gene, has a signal sequence of 63 residues and was previously described as the longest secretory protein signal sequence among eukaryotes (Kobuke et al., 2001). In *D. melanogaster*, besides Crumbs, the proteins Serrate (Fleming et al., 1990; Thomas et al., 1991) and DE-Cadherin (Oda et al., 1994) are currently annotated in the UniProt KB (The UniProt Consortium, 2009) to have long signal sequences of 79 and 69 residues, respectively, but to our knowledge no experimental evidence for their cleavage is available. Until now, predominately viral glycoproteins were shown to have extremely long, cleavable signal sequences, e.g. the arenaviral glycoprotein C-derived signal peptides with 58 residues (Hegde and Bernstein, 2006) and the MMTV Env/Rem signal sequence with 98 residues (Dultz et al., 2008). The Crumbs signal sequence comprising 83 residues is currently the longest signal sequence from an eukaryotic type I membrane protein, further promoting the finding that not only viral proteins have long signal sequences.

SP^{Crb} is specifically processed by SPP

SP^{Crb} is processed by SPP and different signal peptide fragments are generated. In particular, SP^{Crb} is processed by SPP of canine pancreas-derived rough microsomes, by *Drosophila* SPP of *Drosophila* S2 cells-derived rough microsomes, or after overexpression of *Drosophila* SPP in mammalian cells. Processing can be inhibited using either (Z-LL$_2$)-ketone, a compound known to specifically inhibit SPP and SPP-like proteases (Golde et al., 2009), or by overexpression of *Drosophila* SPP-D274A, a mutant variant of an active-site aspartyl residue. Importantly, not only the mutant SP^{Crb*}, carrying an amino acid exchange required for radioactive labeling and autoradiography, but also the wild type SP^{Crb} are detected in the presence of SPP-D274A. In contrast, *Drosophila* SPP-L3 does not process SP^{Crb}, suggesting substrate specificity for SPP.

SP^{Crb} fulfils the two prerequisites for SPP processing as previously described: a helix-breaking motif in the h-region and the preceding cleavage by signal peptidase (Lemberg and Martoglio, 2002). In detail, SP^{Crb} has a 10 residues-long h-region including a proline residue at position 2. Signal sequence cleavage is a prerequisite for SPP processing since a mutant of the signal sequence cleavage site was not processed by SPP. Beyond that, processing results in two fragments of about 8 kD. The presence of two different SPP-processing products has also been observed for TNFα processed by SPP-L2b (Fluhrer et al., 2006) and for the amyloid precursor protein or Notch processed by the SPP-related presenilin (reviewed in Golde et al., 2009), suggesting a common feature of aspartyl protease-mediated intramembrane cleavage processes.

Previous data suggested that zebrafish SPP and SPP-L3 may have overlapping functions (Krawitz et al., 2005). Furthermore, human SPP-L3 was shown to process a bovine prolactin signal peptide based substrate (Nyborg et al., 2006), and recombinant *Drosophila* SPP and SPP-L3 were shown to process the bovine prolactin signal peptide using an *in vitro* assay with chemically-synthesized signal peptides (Narayanan et al., 2007). In contrast, substrate specificity for the three human SPP-like proteases (SPP-L2a, -L2b, -L2c) in comparison to SPP and SPP-L3 is assumed, since the few substrates identified so far are typically exclusively processed by either SPP/SPP-L3 or members of the SPP-L2 family (Golde et al., 2009). In *D. melanogaster*, currently only two SPP proteins are described, SPP and SPP-L3 (Casso et al., 2005). Thus, with SP^{Crb} we have identified the first and specific substrate for *Drosophila* SPP, which is not a

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

substrate for *Drosophila* SPP-L3. Clearly, the identification of further SPP substrates in combination with analyses of *spp* mutants is required to understand substrate specificity.

Towards the function of the Crumbs signal sequence and the signal peptide

By multiple sequence alignments we found that the signal sequence length of Crumbs proteins, especially the extended n-region, is highly conserved within the *Drosophilidae* but not beyond, even not in other insects. This exclusive sequence conservation suggests that a potential function is attributed to *Drosophilidae* Crumbs signal sequences. In principle, the sequence variations in signal sequences can affect either ER targeting, insertion and signal sequence cleavage or a post-targeting function of the signal peptide, or both (Hegde and Bernstein, 2006; Martoglio and Dobberstein, 1998). We did not find that the extended n-region of the Crumbs signal sequence affects targeting, insertion or signal sequence cleavage, at least not under the conditions used. However, it is still possible that targeting in a specific fly tissue is affected.

Post-targeting functions of several signal peptides or SPP processing-derived fragments have been identified. One example is the upregulation of TGFα expression (eosinophil cationic protein; Chang et al., 2007; Wu and Chang, 2004). Signal peptide fragments released into the ER lumen have been shown to be presented as a cytotoxic T-cell epitope (preprocalcitonin; El Hage et al., 2008), while fragments released to the cytosol have been shown to bind calmodulin (prolactin; Martoglio et al., 1997) or are presented as antigen to report the production of the corresponding proteins (different HLA proteins; Bland et al., 2003; Lemberg et al., 2001). For the latter, a series of events of precursor protein insertion into the ER membrane, signal peptidase cleavage, SPP-mediated processing, liberation of an N-terminal fragment to the cytosol, proteasome-mediated processing and TAP-dependent ER import is conceivable and in part formerly proven. Our data clearly demonstrate that the Crumbs signal peptide fragments can be degraded by the proteasome. This supports the model suggesting that SPP processing generates fragments, which are either rapidly degraded by the proteasome or that have functions on their own. Rapid degradation is presumably the most prominent scenario for SP^{Crb} in most fly tissues expressing Crumbs, since we did not identify the signal peptide or its fragments in fly tissues. However, the detection of either SP^{Crb} or its fragments at the leading edge argues that they may accumulate and may have a tissue- or stage-specific post-targeting function. Future experiments need to address the specificity of this staining.

Considering a post-targeting function of SP^{Crb}, its amino acid sequence and composition, especially of the n-region, are of particular interest. Besides having positively charged residues, the extended n-region is glutamine- and threonine-rich. While these stretches are conserved within the *Drosophila melanogaster* subgroup of the *Drosophilidae*, only the glutamine stretch is conserved beyond this subgroup. Homotypic stretches are more abundant in *Drosophila* proteins than in human proteins, and most of these fly proteins are essential developmental proteins (Karlin and Burge, 1996). *Drosophila* now provides an ideal experimental system to study the functional importance of the signal peptide *in vivo*.

Acknowledgements

We are especially grateful to J. Großhans and C. Wenzl for substantial support and several tools. We thank H. Heid (DKFZ Heidelberg) for Edman degradation analysis and S. Özüyaman for the pUAST construct. Many thanks to Maria Leidenberger and Klaus Meese for technical assistance. This work was funded by the Deutsche Forschungsgemeinschaft by the SFB 638/A2 (to B.D.) and the GRK 1188 (B.D. and K.K.), and by the Max Planck Society (E.K.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejcb.2010.02.001.

References

Bachmann, A., Schneider, M., Theilenberg, E., Grawe, F., Knust, E., 2001. *Drosophila* Stardust is a partner of Crumbs in the control of epithelial cell polarity. Nature 414, 638–643.

Bachmann, A., Grawe, F., Johnson, K., Knust, E., 2008. *Drosophila* Lin-7 is a component of the Crumbs complex in epithelia and photoreceptor cells and prevents light-induced retinal degeneration. Eur. J. Cell Biol. 87, 123–136.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795.

Bland, F.A., Lemberg, M.K., McMichael, A.J., Martoglio, B., Braud, V.M., 2003. Requirement of the proteasome for the trimming of signal peptide-derived epitopes presented by the nonclassical major histocompatibility complex class I molecule HLA-E. J. Biol. Chem. 278, 33747–33752.

Blobel, G., Dobberstein, B., 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J. Cell Biol. 67, 835–851.

Bulgakova, N.A., Knust, E., 2009. The Crumbs complex: from epithelial-cell polarity to retinal degeneration. J. Cell Sci. 122, 2587–2596.

Casso, D.J., Tanda, S., Biehs, B., Martoglio, B., Kornberg, T.B., 2005. *Drosophila* signal peptide peptidase is an essential protease for larval development. Genetics 170, 139–148.

Chang, H.T., Kao, Y.L., Wu, C.M., Fan, T.C., Lai, Y.K., Huang, K.L., Chang, Y.S., Tsai, J.J., Chang, M.D., 2007. Signal peptide of eosinophil cationic protein upregulates transforming growth factor-alpha expression in human cells. J. Cell Biochem. 100, 1266–1275.

Chen, P., Nordstrom, W., Gish, B., Abrams, J.M., 1996. grim, a novel cell death gene in *Drosophila*. Genes Dev. 10, 1773–1782.

Dultz, E., Hildenbeutel, M., Martoglio, B., Hochman, J., Dobberstein, B., Kapp, K., 2008. The signal peptide of the mouse mammary tumor virus Rem protein is released from the ER membrane and accumulates in nucleoli. J. Biol. Chem.

Eaton, D.L., Wood, W.I., Eaton, D., Hass, P.E., Hollingshead, P., Wion, K., Mather, J., Lawn, R.M., Vehar, G.A., Gorman, C., 1986. Construction and characterization of an active factor VIII variant lacking the central one-third of the molecule. Biochemistry 25, 8343–8347.

El Hage, F., Stroobant, V., Vergnon, I., Baurain, J.F., Echchakir, H., Lazar, V., Chouaib, S., Coulie, P.G., Mami-Chouaib, F., 2008. Preprocalcitonin signal peptide generates a cytotoxic T lymphocyte-defined tumor epitope processed by a proteasome-independent pathway. Proc. Natl. Acad. Sci. USA 105, 10119–10124.

Fleming, R.J., Scottgale, T.N., Diederich, R.J., Artavanis-Tsakonas, S., 1990. The gene Serrate encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in *Drosophila melanogaster*. Genes Dev. 4, 2188–2201.

Fluhrer, R., Grammer, G., Israel, L., Condron, M.M., Haffner, C., Friedmann, E., Bohland, C., Imhof, A., Martoglio, B., Teplow, D.B., et al., 2006. A gamma-secretase-like intramembrane cleavage of TNFalpha by the GxGD aspartyl protease SPPL2b. Nat. Cell Biol. 8, 894–896.

Golde, T.E., Wolfe, M.S., Greenbaum, D.C., 2009. Signal peptide peptidases: a family of intramembrane-cleaving proteases that cleave type 2 transmembrane proteins. Semin. Cell Dev. Biol. 20, 225–230.

Grether, M.E., Abrams, J.M., Agapite, J., White, K., Steller, H., 1995. The head involution defective gene of *Drosophila melanogaster* functions in programmed cell death. Genes Dev. 9, 1694–1708.

Grosshans, J., Schnorrer, F., Nusslein-Volhard, C., 1999. Oligomerisation of Tube and Pelle leads to nuclear localisation of dorsal. Mech. Dev. 81, 127–138.

Hegde, R.S., Bernstein, H.D., 2006. The surprising complexity of signal sequences. Trends Biochem. Sci. 31, 563–571.

Hiller, K., Grote, A., Scheer, M., Munch, R., Jahn, D., 2004. PrediSi: prediction of signal peptides and their cleavage positions. Nucleic Acids Res. 32, W375–W379.

Kall, L., Krogh, A., Sonnhammer, E.L., 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res. 35, W429–W432.

Kapp, K., Schrempf, S., Lemberg, M.K., Dobberstein, B., 2009. Post-targeting functions of signal peptides. In: Zimmermann, R. (Ed.), Protein Transport into the Endoplasmic Reticulum. Landes Bioscience, Austin.

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001

Karlin, S., Burge, C., 1996. Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. Proc. Natl. Acad. Sci. USA 93, 1560–1565.

Klebes, A., Knust, E., 2000. A conserved motif in Crumbs is required for E-cadherin localisation and zonula adherens formation in *Drosophila*. Curr. Biol. 10, 76–85.

Kobuke, K., Furukawa, Y., Sugai, M., Tanigaki, K., Ohashi, N., Matsumori, A., Sasayama, S., Honjo, T., Tashiro, K., 2001. ESDN, a novel neuropilin-like membrane protein cloned from vascular cells with the longest secretory signal sequence among eukaryotes, is up-regulated after vascular injury. J. Biol. Chem. 276, 34105–34114.

Koles, K., Lim, J.M., Aoki, K., Porterfield, M., Tiemeyer, M., Wells, L., Panin, V., 2007. Identification of N-glycosylated proteins from the central nervous system of *Drosophila melanogaster*. Glycobiology 17, 1388–1403.

Krawitz, P., Haffner, C., Fluhrer, R., Steiner, H., Schmid, B., Haass, C., 2005. Differential localization and identification of a critical aspartate suggest non-redundant proteolytic functions of the presenilin homologues SPPL2b and SPPL3. J. Biol. Chem. 280, 39515–39523.

Lemberg, M.K., Bland, F.A., Weihofen, A., Braud, V.M., Martoglio, B., 2001. Intramembrane proteolysis of signal peptides: an essential step in the generation of HLA-E epitopes. J. Immunol. 167, 6441–6446.

Lemberg, M.K., Martoglio, B., 2002. Requirements for signal peptide peptidase-catalyzed intramembrane proteolysis. Mol. Cell 10, 735–744.

Lemeunier, F., David, J.R., Tsacas, L., Ashburner, M., 1986. The *Melanogaster* Species Subgroup. Academic Press, London, New York.

Lundin, C., Kall, L., Kreher, S.A., Kapp, K., Sonnhammer, E.L., Carlson, J.R., Heijne, G.V., Nilsson, I., 2007. Membrane topology of the *Drosophila* OR83b odorant receptor. FEBS Lett. 581, 5601–5604.

Lyko, F., Martoglio, B., Jungnickel, B., Rapoport, T.A., Dobberstein, B., 1995. Signal sequence processing in rough microsomes. J. Biol. Chem. 270, 19873–19878.

Martoglio, B., Graf, R., Dobberstein, B., 1997. Signal peptide fragments of preprolactin and HIV-1 p-gp160 interact with calmodulin. EMBO J. 16, 6636–6645.

Martoglio, B., Dobberstein, B., 1998. Signal sequences: more than just greasy peptides. Trends Cell Biol. 8, 410–415.

Meinnel, T., Serero, A., Giglione, C., 2006. Impact of the N-terminal amino acid on targeted protein degradation. Biol. Chem. 387, 839–851.

Muller, H.A., Wieschaus, E., 1996. Armadillo, bazooka, and stardust are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in *Drosophila*. J. Cell Biol. 134, 149–163.

Narayanan, S., Sato, T., Wolfe, M.S., 2007. A C-terminal region of signal peptide peptidase defines a functional domain for intramembrane aspartic protease catalysis. J. Biol. Chem. 282, 20172–20179.

Nielsen, H., Brunak, S., von Heijne, G., 1999. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein Eng. 12, 3–9.

Nyborg, A.C., Kornilova, A.Y., Jansen, K., Ladd, T.B., Wolfe, M.S., Golde, T.E., 2004. Signal peptide peptidase forms a homodimer that is labeled by an active site-directed gamma-secretase inhibitor. J. Biol. Chem. 279, 15153–15160.

Nyborg, A.C., Ladd, T.B., Jansen, K., Kukar, T., Golde, T.E., 2006. Intramembrane proteolytic cleavage by human signal peptide peptidase like 3 and malaria signal peptide peptidase. Faseb J. 20, 1671–1679.

Oda, H., Uemura, T., Harada, Y., Iwai, Y., Takeichi, M., 1994. A *Drosophila* homolog of cadherin associated with armadillo and essential for embryonic cell–cell adhesion. Dev. Biol. 165, 716–726.

Paetzel, M., Karla, A., Strynadka, N.C., Dalbey, R.E., 2002. Signal peptidases. Chem. Rev. 102, 4549–4580.

Schagger, H., von Jagow, G., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368–379.

Schrempf, S., Froeschke, M., Giroglou, T., von Laer, D., Dobberstein, B., 2007. Signal peptide requirements for LCMV glycoprotein C maturation and virus infectivity. J. Virol. 81, 12515–12524.

Shaw, A.S., Rottier, P.J., Rose, J.K., 1988. Evidence for the loop model of signal-sequence insertion into the endoplasmic reticulum. Proc. Natl. Acad. Sci. USA 85, 7592–7596.

Tepass, U., Theres, C., Knust, E., 1990. Crumbs encodes an EGF-like protein expressed on apical membranes of *Drosophila* epithelial cells and required for organization of epithelia. Cell 61, 787–799.

The UniProt Consortium, 2009. The Universal Protein Resource (UniProt) 2009. Nucleic Acids Res. 37, D169–D174.

Thomas, U., Speicher, S.A., Knust, E., 1991. The *Drosophila* gene Serrate encodes an EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs. Development 111, 749–761.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

von Heijne, G., 1983. Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. 133, 17–21.

von Heijne, G., 1990. The signal peptide. J. Membr. Biol. 115, 195–201.

Walter, P., Blobel, G., 1983. Preparation of microsomal membranes for cotranslational protein translocation. Methods Enzymol. 96, 84–93.

Weihofen, A., Binns, K., Lemberg, M.K., Ashman, K., Martoglio, B., 2002. Identification of signal peptide peptidase, a presenilin-type aspartic protease. Science 296, 2215–2218.

Weihofen, A., Martoglio, B., 2003. Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides. Trends Cell Biol. 13, 71–78.

White, K., Grether, M.E., Abrams, J.M., Young, L., Farrell, K., Steller, H., 1994. Genetic control of programmed cell death in *Drosophila*. Science 264, 677–683.

Wodarz, A., Hinz, U., Engelbert, M., Knust, E., 1995. Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of *Drosophila*. Cell 82, 67–76.

Wu, C.M., Chang, M.D., 2004. Signal peptide of eosinophil cationic protein is toxic to cells lacking signal peptide peptidase. Biochem. Biophys. Res. Commun. 322, 585–592.

Please cite this article as: Kilic, A., et al., The *Drosophila* Crumbs signal peptide is unusually long and is a substrate for signal peptide peptidase. Eur. J. Cell Biol. (2010), doi:10.1016/j.ejcb.2010.02.001