12. B. Sharpe, M. Beresford, R. Bowen, J. Mitchard, A. D. Chalmers, *Stem Cell Rev.* **9**, 721–730 (2013).
13. J. Trepel, M. Mollapour, G. Giaccone, L. Neckers, *Nat. Rev. Cancer* **10**, 537–549 (2010).
14. T. Gutschner, M. Hämmerle, S. Diederichs, *J. Mol. Med. (Berlin)* **91**, 791–801 (2013).
15. C. F. Schaefer *et al.*, *Nucleic Acids Res.* **37** (Database), D674–D679 (2009).
16. B. Gottlieb, L. K. Beitel, A. Nadarajah, M. Paliouras, M. Trifiro, *Hum. Mutat.* **33**, 887–894 (2012).
17. C. Lu, J. Luo, *Transl. Androl. Urol.* **2**, 178–186 (2013).
18. E. S. Antonarakis *et al.*, *N. Engl. J. Med.* **371**, 1028–1038 (2014).
19. M. D. Balbas *et al.*, *eLife* **2**, e00499 (2013).
20. J. D. Joseph *et al.*, *Cancer Discov* **3**, 1020–1029 (2013).
21. S. Carreira *et al.*, *Sci. Transl. Med.* **6**, 254ra125 (2014).
22. J. Zhang, J. L. Manley, *Cancer Discov* **3**, 1228–1237 (2013).
23. L. Kurlender *et al.*, *Biochim. Biophys. Acta* **1755**, 1–14 (2005).
24. H. I. Scher *et al.*, *N. Engl. J. Med.* **367**, 1187–1197 (2012).
25. M. Katoh, M. Katoh, *Clin. Cancer Res.* **13**, 4042–4045 (2007).
26. S. Takahashi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4938–4943 (2011).
27. D. Zheng *et al.*, *Mol. Cancer Res.* **11**, 482–493 (2013).
28. T. S. Gujral *et al.*, *Cell* **159**, 844–856 (2014).
29. V. K. Arora *et al.*, *Cell* **155**, 1309–1322 (2013).
30. N. Rath, M. F. Olson, *EMBO Rep.* **13**, 900–908 (2012).
31. F. Tang *et al.*, *Nat. Methods* **6**, 377–382 (2009).
32. H. Hieronymus *et al.*, *Cancer Cell* **10**, 321–330 (2006).

ACKNOWLEDGMENTS

We thank C. Sawyers for helpful discussions; A. McGovern, E. Stadtmueller, and B. Abebe for clinical trial support; and L. Libby and L. Nieman for technical assistance. This work was supported by grants from the Prostate Cancer Foundation (D.A.H., S.M., M.T., M.R.S., and R.J.L.), Charles Evans Foundation (D.A.H.), Department of Defense (D.T.M., R.J.L., and D.T.T.), Stand Up to Cancer (D.A.H., M.T., S.M., and L.V.S.), Howard Hughes Medical Institute (D.A.H.), National Institute of Biomedical Imaging and Bioengineering (NIBIB), NIH, EB008047 (M.T.), NCI 2RO1CA129933 (D.A.H.), National Cancer Institute, NCI, Federal Share Program and Income (S.M. and D.T.M.), Affymetrix, Inc. (D.T.T., K.A., and N.D.), Mazzone Program–Dana-Farber Harvard Cancer Center (D.T.M.), Burroughs Wellcome Fund (D.T.T.), and the Massachusetts General Hospital–Johnson & Johnson Center for Excellence in CTC Technologies (D.A.H., M.T., and S.M.). D.T.T. is a paid consultant for Affymetrix, Inc.; R.J.L. is a paid consultant for Janssen LLC. The Massachusetts General Hospital has filed for patent protection for the CTC-iChip technology. RNA-sequencing data have been deposited in GEO under accession number GSE67980.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6254/1351/suppl/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S7
References (*33*–*37*)

10 March 2015; accepted 3 August 2015
10.1126/science.aab0917

SMALL PEPTIDES

Pri sORF peptides induce selective proteasome-mediated protein processing

J. Zanet,[1,2]* E. Benrabah,[1,2]* T. Li,[3] A. Pélissier-Monier,[1,2] H. Chanut-Delalande,[1,2] B. Ronsin,[1,2] H. J. Bellen,[3,4] F. Payre,[1,2]† S. Plaza[1,2]†

A wide variety of RNAs encode small open-reading-frame (smORF/sORF) peptides, but their functions are largely unknown. Here, we show that *Drosophila polished-rice* (*pri*) sORF peptides trigger proteasome-mediated protein processing, converting the Shavenbaby (Svb) transcription repressor into a shorter activator. A genome-wide RNA interference screen identifies an E2-E3 ubiquitin-conjugating complex, UbcD6-Ubr3, which targets Svb to the proteasome in a *pri*-dependent manner. Upon interaction with Ubr3, Pri peptides promote the binding of Ubr3 to Svb. Ubr3 can then ubiquitinate the Svb N terminus, which is degraded by the proteasome. The C-terminal domains protect Svb from complete degradation and ensure appropriate processing. Our data show that Pri peptides control selectivity of Ubr3 binding, which suggests that the family of sORF peptides may contain an extended repertoire of protein regulators.

Eukaryotic genomes encode many noncoding RNAs (ncRNAs) that lack the classical hallmarks of protein-coding genes. However, both ncRNAs and mRNAs often contain small open reading frames (sORFs), and there is growing evidence that they can produce peptides, from yeast (*1*) to plants (*2*, *3*) or humans (*4*, *5*). The *polished rice* or *tarsal-less* (*pri*) RNA contains four sORFs that encode highly related 11– to 32–amino acid peptides, required for embryonic development across insect species (*6*–*8*). In flies, *pri* is essential for the differentiation of epidermal outgrowths called trichomes (*7*, *8*). Trichome development is governed by the Shavenbaby (Svb) transcription factor (*9*–*11*); however, only in the presence of *pri* can Svb turn on the program of trichome development, i.e., activate expression of cellular effectors (*12*, *13*). Indeed, the Svb protein is translated as a large repressor, *pri* then induces truncation of its N-terminal region, which leads to a shorter activator (*12*). Thereby, *pri* defines the developmental timing of epidermal differentiation, in a direct response to systemic ecdysone hormonal signaling (*14*). Although we now have a clear framework for the developmental functions of *pri*, how these small peptides can trigger Svb processing is unknown.

To identify factors required for Svb processing in response to *pri*, we performed a genome-wide RNA interference (RNAi) screen in a cell line co-expressing green fluorescent protein (GFP)–tagged Svb and *pri* (Fig. 1A). We set up an automated assay quantifying Svb processing for each of the *Drosophila* genes, with an inhibitory score reflecting the proportion of cells unable to cleave off the Svb N terminus (see the supplementary materials). *pri* RNAi displayed the highest score, which validated our approach to identifying molecular players in Svb processing. Methods used to evaluate results from genome-wide screening all converged on a key role for the proteasome. For instance, COMPLEAT, a bioinformatic frame-

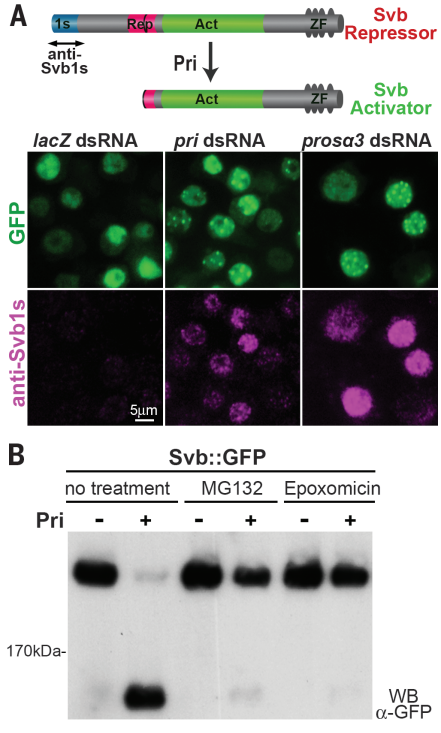

Fig. 1. Pri-dependent processing of Svb requires proteasome activity. (**A**) Drawing of Svb processing (antibody against Svb1s recognizes the repressor-specific N-terminal region) and snapshots from the screen illustrating the effect of double-stranded RNA against *lacZ* (negative control), *pri*, and *proteasomeα3* subunit (*prosα3*) on Svb::GFP processing. Cells were stained for Svb1s (purple) and GFP (green). (**B**) Western blot analysis of cells that express Svb::GFP, with or without *pri* and proteasome inhibitors (MG132, epoxomicin).

[1]Centre de Biologie du Développement, Université de Toulouse III—Paul Sabatier, Bâtiment 4R3, 118 route de Narbonne, F-31062 Toulouse, France. [2]CNRS, UMR5547, Centre de Biologie du Développement, F-31062 Toulouse, France. [3]Program in Developmental Biology, Baylor College of Medicine, Houston, TX 77030, USA. [4]Department of Molecular and Human Genetics, Howard Hughes Medical Institute, Neurological Research Institute, Baylor College of Medicine, Houston, TX 77030, USA.
*These authors contributed equally to this work. **†Corresponding author. E-mail: francois.payre@univ-tlse3.fr (F.P.); serge.plaza@univ-tlse3.f (S.P.)**

work based on protein complex analysis (*15*), identified the proteasome in 66 out of the 71 top predictions (fig. S1A and table S1). A survey of individual proteasome subunits indicated that both the 20*S* catalytic core and the 19*S* regulatory particles are required for Svb processing (fig. S1B and table S2). Chemical proteasome inhibitors independently confirmed this conclusion, because they also prevented *pri*-induced Svb processing (Fig. 1B). These data thus provide compelling evidence that Svb processing results from a *pri*-dependent proteolysis by the proteasome.

To investigate how *pri* regulates proteolysis of Svb, we first identified the protein region(s) in Svb that are involved in *pri*-dependent processing. Systematic deletions demonstrated the importance of the Svb N terminus for *pri* response and restricted the minimal motif to the N-terminal 31 amino acids (fig. S2, A and B). Deletion of this motif within an otherwise full-length protein (Δ31) made Svb refractory to *pri* (fig. S2, A and B). Conversely, the Svb N terminus when fused to GFP (1s::GFP) was sufficient to transform this protein into a *pri* target and to make GFP sensitive to *pri*. Unlike Svb, however, 1s::GFP was completely degraded by the proteasome upon *pri* expression (Fig. 2A and fig. S2, C and D).

Recent studies have shown that structural features of proteins influence their degradation by the proteasome (*16*): Whereas unstructured substrates, such as intrinsically disordered regions, favor degradation (*17*), tightly folded domains can resist proteasome progression (*18*). Analysis of Svb sequences predicted intrinsically disordered features (Fig. 2A and fig. S3A) throughout its N-terminal moiety, which is degraded. By contrast, the proteasome-resistant C-terminal moiety comprises two folded regions: the transcriptional activation and zinc finger domains. Within the transcriptional activation region, we found that amino acids 532 to 701 protected Svb from complete degradation. Indeed, the C-terminally truncated mutants of 1 to 701 amino acids (and longer) were still processed, whereas mutants shortened by 1 to 532 amino acids (and shorter) were fully degraded (fig. S3B). We tested whether other folded domains would also protect Svb from complete degradation and found that attaching zinc fingers to short Svb mutants—otherwise degraded upon *pri* expression—was sufficient to restore processing (Fig. 2A). Likewise, the DNA binding domain of Gal4 protected against degradation (fig. S3, B and C), which indicated that even a heterologous protein domain with strong structure can protect Svb from full degradation in response to *pri*. Hence, distinct regions of Svb mediate its processing by the proteasome: the 31 N-terminal residues act as a *pri*-dependent degradation signal, or degron, and C-terminal domains act as stabilizing features that prevent complete degradation.

Proteins are targeted to the proteasome by the covalent attachment of ubiquitin to Lys residues (*19*). The Svb N terminus is highly conserved from insects to human (Fig. 2B); it comprises two invariant Lys residues (K3 and K8) and a third one at a less constrained position (K28 in *Drosophila*). We found that individual Lys substitutions had only a weak effect or no effect, whereas simultaneous mutation of all three Lys (3Kmut) abolished Svb processing (Fig. 2B). Furthermore, we detected strong *pri*-dependent ubiquitination of Svb when the proteasome was inhibited (Fig. 2C). By contrast, this was no longer seen in the 3Kmut variant, which demonstrated the key role of these three Lys in ubiquitin-dependent Svb processing.

Ubiquitin conjugation requires three enzymes (E1, E2, and E3); specificity is generally conferred by the E3 ubiquitin ligases that recognize and bind to substrates (*19*). A prominent hit from our RNAi screen was *Ubr3* (7 hits out of the top 15), which encodes an E3. Ranking all *Drosophila* ubiquitin enzymes by their inhibitory score confirmed that Ubr3 was the major E3 required for Svb processing and identified UbcD6 (Rad6) as its associated E2 (figs. S4 and S5 and table S3), consistent with evidence that human Ubr3 also forms a complex with UbcD6 (*20*). Like many

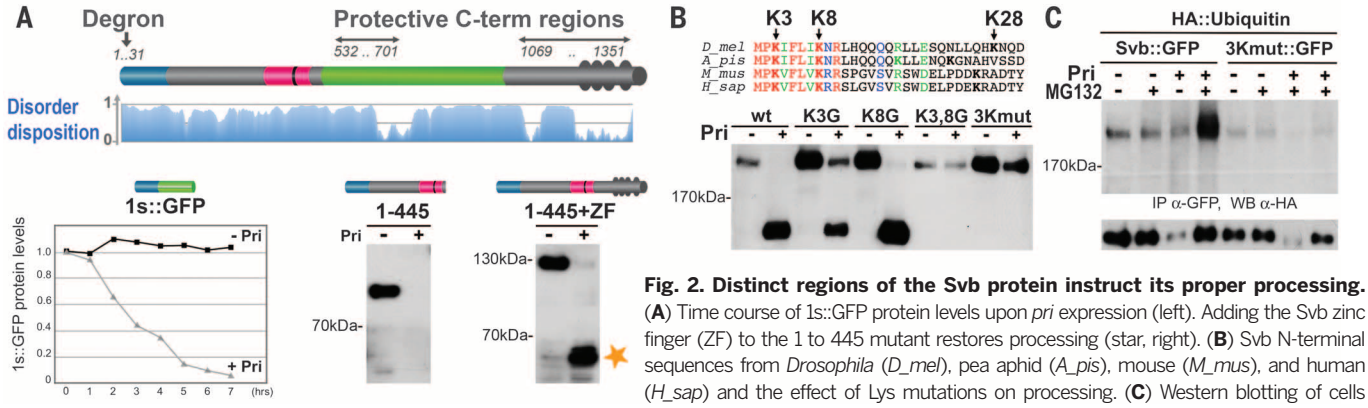

Fig. 2. Distinct regions of the Svb protein instruct its proper processing. (**A**) Time course of 1s::GFP protein levels upon *pri* expression (left). Adding the Svb zinc finger (ZF) to the 1 to 445 mutant restores processing (star, right). (**B**) Svb N-terminal sequences from *Drosophila* (*D_mel*), pea aphid (*A_pis*), mouse (*M_mus*), and human (*H_sap*) and the effect of Lys mutations on processing. (**C**) Western blotting of cells coexpressing hemagglutinin (HA)-tagged ubiquitin, Svb::GFP, or the 3Kmut variant, with or without *pri* and MG132.

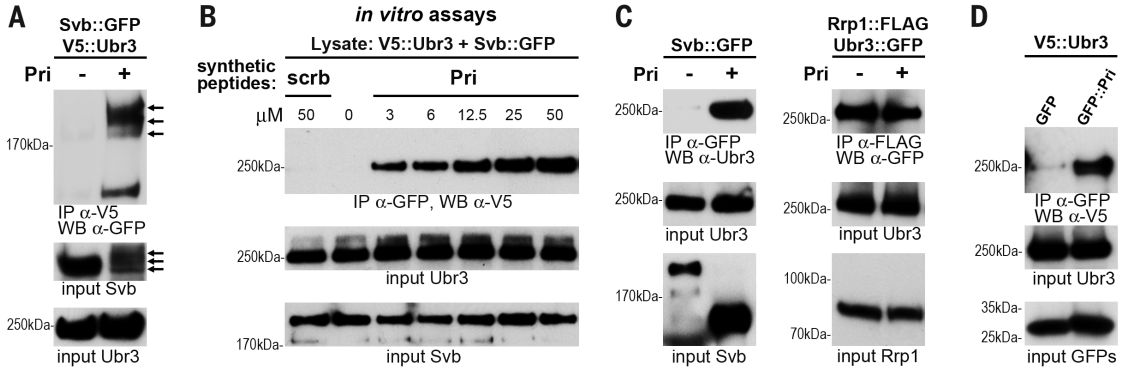

Fig. 3. Ubr3 mediates Svb ubiquitination in the response to Pri peptides. (**A**) Coimmunoprecipitation of Ubr3::V5 with Svb::GFP in the presence of MG132. Arrows indicate ubiquitinated Svb. (**B**) Protein extracts from cells coexpressing V5::Ubr3 and Svb::GFP, in the absence of *pri*, were incubated in vitro with a synthetic Pri peptide (MAAYLDPTGQY) (*31*) and immunoprecipitated using antibody against GFP. A scrambled peptide (MKTYPGALYDA, scrb) was used for control. (**C**) Coimmunoprecipitation of Ubr3 with Svb::GFP or with Rrp1::FLAG in the presence of MG132. (**D**) GFP and GFP::Pri were immunopurified and incubated with V5::Ubr3 protein extracts. Bound fractions were analyzed by antibody against V5 and Western blotting.

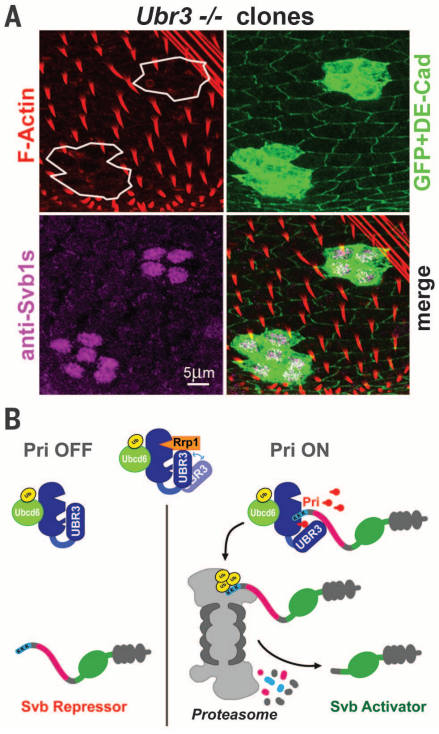

Fig. 4. *Ubr3* is required in vivo for Svb processing. (**A**) Clones of *Ubr3*-null epidermal cells in mosaic pupae (GFP-positive, green cytoplasm) do not form trichomes (F-actin, red) and retain unprocessed Svb repressor (Svb1s-specific antibody, purple). Cells contours are revealed by DE-cadherin (green). (**B**) Model of Svb processing in response to *pri*. After the binding of Pri peptides, Ubr3 becomes able to bind the Svb N terminus and, together with UbcD6, ubiquitinates three Lys residues. N-terminal unstructured regions of Svb are degraded by the proteasome, whereas C-terminal folded regions (green and gray ovals) protect from complete degradation and allow release of the truncated Svb activator. In contrast, Ubr3 binds to other substrates (e.g., Rrp1 and DIAP1), with or without Pri peptides.

proteasome factors, Ubr3 has a broad subcellular distribution in cytoplasm and nuclei, whereas Svb and UbcD6 are nuclear proteins (fig. S6). Svb processing still occurred normally when nuclear export was impaired (figs. S6 and S7 and table S4), which indicated that the proteolytic activation of Svb takes place within the nucleus.

Several additional lines of evidence support the conclusion that Ubr3 mediates the function of *pri* for Svb ubiquitination. First, Ubr3 coimmunoprecipitated with Svb in a *pri*-dependent manner and ubiquitinated Svb was found in a complex with Ubr3 upon proteasome inhibition (Fig. 3A). Second, the N terminus of Svb was sufficient for Ubr3 binding in response to *pri* (fig. S8). Note that a functional N-terminal degron in Svb was required for its interaction with Ubr3, because the ubiquitin-resistant 3Kmut variant no longer bound Ubr3. Third, in protein extracts from cells that do not express *pri*, addition of synthetic Pri peptide was sufficient to promote Ubr3-Svb interaction in vitro, in a dose-dependent manner (Fig. 3B). By contrast, a peptide of the same composition but in a "scrambled" sequence lacked activity.

Although critical for the binding of Ubr3 to the Svb N terminus, Pri peptides are, however, not indispensable for Ubr3 activity. We found that *pri* did not influence the binding of Ubr3 to Ape1 (Rrp1) (Fig. 3C), a factor involved in DNA repair and regulated by Ubr3-dependent proteasome degradation (*21*). Also, the interaction of Ubr3 with DIAP1, which inhibits apoptosis (*22*, *23*), occurred with or without *pri* (fig. S9). Moreover, we found that Pri peptides interacted with Ubr3, even in the absence of Svb (Fig. 3D and fig. S8). Finally, the isolated UBR-box of Ubr3 no longer required Pri peptides to bind Svb (fig. S10), which suggested that other Ubr3 motifs prevent Svb interaction in the absence of *pri*. We therefore conclude that Pri peptides directly regulate the selectivity of Ubr3 for binding to the Svb N terminus and, thereby, trigger Svb ubiquitination and processing by the proteasome.

We recently isolated a *Ubr3* loss-of-function allele (*24*) and assayed its phenotype in the differentiation of epidermal cells. As observed for *pri* mutants, embryos lacking *Ubr3* were unable to differentiate trichomes and to process Svb (fig. S11). Moreover, inactivation of either *UbcD6* or *Ubr3* prevented formation of adult trichomes in mosaic animals (Fig. 4A and fig. S12). When compared with their wild-type neighbors, *Ubr3*-null cells accumulated the repressor form of Svb, which demonstrated Ubr3's essential role for Svb processing in vivo.

Taken together, our data show that Pri peptides control the binding of the Ubr3 ubiquitin ligase to Svb and activate its processing by the proteasome (see Fig. 4B). In the absence of Pri, Ubr3 nonetheless recognizes other substrates (*21*–*23*), which shows that a main role for Pri peptides is to modify the binding selectivity of Ubr3. This could potentially be achieved through a conformational change in Ubr3 protein, as proposed for Ubr1 (*25*), that unmasked the recognition site for Svb upon Pri peptide binding to Ubr3.

Although recent work has uncovered thousands of novel sORF peptides (*1*–*5*), only a handful of their molecular targets have yet been identified. sORF peptides have recently been found to bind and regulate the Ca^{2+} uptake SERCA protein (*26*, *27*), the heterotrimeric guanine nucleotide–binding protein–coupled signaling APJ (Apelin) (*28*), and the DNA repair protein Ku (*29*). Protein-protein interactions often involve small protein regions, and artificial peptides that mimic these binding surfaces have been proven to be potent modulators of protein complexes (*30*). We propose that sORF-encoded peptides provide an unexplored reservoir of protein-binding interfaces, well suited to regulate the activity of a wide range of cellular factors.

REFERENCES AND NOTES

1. J. E. Smith *et al.*, *Cell Reports* **7**, 1858–1866 (2014).
2. K. Hanada *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2395–2400 (2013).
3. D. Lauressergues *et al.*, *Nature* **520**, 90–93 (2015).
4. N. T. Ingolia *et al.*, *Cell Reports* **8**, 1365–1379 (2014).
5. J. Ma *et al.*, *J. Proteome Res.* **13**, 1757–1765 (2014).
6. J. Savard, H. Marques-Souza, M. Aranda, D. Tautz, *Cell* **126**, 559–569 (2006).
7. T. Kondo *et al.*, *Nat. Cell Biol.* **9**, 660–665 (2007).
8. M. I. Galindo, J. I. Pueyo, S. Fouix, S. A. Bishop, J. P. Couso, *PLOS Biol.* **5**, e106 (2007).
9. F. Payre, A. Vincent, S. Carreno, *Nature* **400**, 271–275 (1999).
10. D. L. Stern, N. Frankel, *Philos. Trans. R. Soc. London B Biol. Sci.* **368**, 20130028 (2013).
11. J. Crocker *et al.*, *Cell* **160**, 191–203 (2015).
12. T. Kondo *et al.*, *Science* **329**, 336–339 (2010).
13. D. Menoret *et al.*, *Genome Biol.* **14**, R86 (2013).
14. H. Chanut-Delalande *et al.*, *Nat. Cell Biol.* **16**, 1035–1044 (2014).
15. A. Vinayagam *et al.*, *Sci. Signal.* **6**, rs5 (2013).
16. S. Fishbain *et al.*, *Nat. Struct. Mol. Biol.* **22**, 214–221 (2015).
17. S. Prakash, L. Tian, K. S. Ratliff, R. E. Lehotzky, A. Matouschek, *Nat. Struct. Mol. Biol.* **11**, 830–837 (2004).
18. L. Tian, R. A. Holmgren, A. Matouschek, *Nat. Struct. Mol. Biol.* **12**, 1045–1053 (2005).
19. A. Varshavsky, *Annu. Rev. Biochem.* **81**, 167–176 (2012).
20. T. Tasaki *et al.*, *J. Biol. Chem.* **282**, 18510–18520 (2007).
21. C. Meisenberg *et al.*, *Nucleic Acids Res.* **40**, 701–711 (2012).
22. M. Ditzel *et al.*, *Mol. Cell* **32**, 540–553 (2008).
23. Q. Huang *et al.*, *Cell Death Differ.* **21**, 1961–1970 (2014).
24. S. Yamamoto *et al.*, *Cell* **159**, 200–214 (2014).
25. G. C. Turner, F. Du, A. Varshavsky, *Nature* **405**, 579–583 (2000).
26. E. G. Magny *et al.*, *Science* **341**, 1116–1120 (2013).
27. D. M. Anderson *et al.*, *Cell* **160**, 595–606 (2015).
28. A. Pauli *et al.*, *Science* **343**, 1248636 (2014).
29. S. A. Slavoff, J. Heo, B. A. Budnik, L. A. Hanakahi, A. Saghatelian, *J. Biol. Chem.* **289**, 10950–10957 (2014).
30. A. Patgiri, K. K. Yadav, P. S. Arora, D. Bar-Sagi, *Nat. Chem. Biol.* **7**, 585–587 (2011).
31. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

ACKNOWLEDGMENTS

We thank N. Perrimon, S. Mohr, O. Coux, P. Meier, N. Tapon, P. Demange, L. Twyffels, the *Drosophila* Screening Center, C. Polesello, M. Soulard, Y. Latapie, and P. Valenti for invaluable support. This work was supported by Agence Nationale de la Recherche (smORFpep and ChronoNet), Association pour la Recherche sur le Cancer, Fondation Recherche et Innovation Thérapeutique en Cancérologie, and Howard Hughes Medical Institute. The authors declare no conflict of interest.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6254/1356/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S4
References (*32*–*40*)

14 May 2015; accepted 10 August 2015
10.1126/science.aac5677

Editor's Summary

Small peptide regulates protein activity

Coding and noncoding RNAs can produce peptides from small open reading frames (smORFs), with a variety of mostly unknown functions. Using a genome-wide screen, Zanet *et al.* show that Polished rice (Pri) smORF peptides control fruit fly development by binding to an E3 ubiquitin ligase. This changes the ligase's selectivity and triggers proteasome-dependent maturation of the developmental transcription factor Shavenbaby. Other smORF peptides may act by a similar mechanism to regulate protein activity.

Science, this issue p. 1356