

**Table 2 | Correlations controlled for cognitive and performance factors**

Measure (task)	$r_p^2$	$t$ d.f. = 60	$P$
With TEMA-2			
Symbolic maths (TEMA-2)	0.146	3.205	0.002
Intelligence (WASI-full)	0.013	0.887	0.379
Task demands (RAN-colour)	0.004	0.492	0.625
With WJ-Rcalc			
Symbolic maths (WJ-Rcalc)	0.155	3.325	0.003
Intelligence (WASI-full)	0.070	2.124	0.038
Task demands (RAN-colour)	0.017	1.023	0.310

ANS acuity ( $w$ ) measured in ninth grade retroactively correlated with third grade symbolic maths achievement and other measures.  $r_p^2$  values represent the proportion of the variance in ANS acuity accounted for by the listed variable when controlling for the two remaining variables in each analysis (TEMA-2 or WJ-Rcalc).  $t$  values represent the distance, measured in units of standard error, between the obtained correlation and the null hypothesis of no correlation.  $P$  values represent the probability of obtaining the observed correlation in a sample of data by random chance when there is truly no relation in the population.

can be retrospectively predicted by a subject's ANS acuity in young adulthood, as measured by the simple task of determining which of two quickly flashed arrays has more dots, even with extensive controls for other cognitive and performance factors.

Our results are consistent with at least two interpretations. Given that it is functional in infancy<sup>7</sup>, long before the onset of symbolic mathematics instruction, the ANS may have a causal role in determining individual maths achievement. Indeed, neuropsychological evidence suggests that the ANS is activated during symbolic mathematical reasoning across the lifespan<sup>13</sup>; therefore individual differences in ANS acuity might give rise to individual differences in maths ability.

**Table 3 | Correlations controlled for all available factors**

Measure (task)	$r_p^2$	$t$ d.f. = 40	$P$
With TEMA-2			
Symbolic maths (TEMA-2)	0.167	2.831	0.007
Intelligence (WASI-full)	0.005	0.472	0.640
Task demands (RAN-colour)	0.023	0.981	0.332
Verbal IQ (WASI-verbal)	0.005	0.459	0.649
Performance IQ (WASI-performance)	0.006	0.482	0.632
Executive functions (CNT-B3)	0.021	0.918	0.364
Visual working memory (MemPuzl)	0.067	1.694	0.098
Visual segmentation (DTVPfg)	0.009	0.599	0.552
Object perception (DTVPfc)	0.001	0.172	0.864
Visual reasoning (DTVPvc)	0.025	1.004	0.321
Spatial reasoning (DTVPps)	0.012	0.701	0.488
Visual motor integration (VMI)	0.035	1.213	0.232
Word knowledge (WJ-RIwid)	0.012	0.706	0.484
Reading (WJ-Rwa)	0.001	0.225	0.823
Rapid lexical access (RAN-letter)	0.049	1.435	0.149
Rapid lexical access (RAN-number)	0.012	0.685	0.497
Gender	0.028	1.069	0.291
With WJ-Rcalc			
Symbolic maths (WJ-Rcalc)	0.200	3.149	0.003
Intelligence (WASI-full)	0.013	0.736	0.466
Task demands (RAN-colour)	0.004	0.391	0.698
Verbal IQ (WASI-verbal)	0.009	0.605	0.549
Performance IQ (WASI-perf)	0.013	0.727	0.472
Executive functions (CNT-B3)	0.035	1.208	0.234
Visual working memory (MemPuzl)	0.084	1.916	0.062
Visual segmentation (DTVPfg)	0.032	1.148	0.254
Object perception (DTVPfc)	0.001	0.201	0.842
Visual reasoning (DTVPvc)	0.008	0.578	0.566
Spatial reasoning (DTVPps)	0.018	0.869	0.390
Visual motor integration (VMI)	0.013	0.725	0.473
Word knowledge (WJ-RIwid)	0.014	0.757	0.454
Reading (WJ-Rwa)	0.000	0.014	0.988
Rapid lexical access (RAN-letter)	0.014	0.757	0.454
Rapid lexical access (RAN-number)	0.000	0.037	0.970
Gender	0.012	0.684	0.498

ANS acuity ( $w$ ) measured in ninth grade retroactively correlated with third grade symbolic maths achievement and other measures.  $r_p^2$  values represent the proportion of the variance in ANS acuity accounted for by the listed variable when controlling for all other variables in the list.  $t$  values represent the distance, measured in units of standard error, between the obtained correlation and the null hypothesis of no correlation.  $P$  values represent the probability of obtaining the observed correlation in a sample of data by random chance when there is truly no relation in the population.

Alternatively, individual differences in the quantity or quality of engagement in formal mathematics might increase ANS acuity. This latter possibility is hinted at by cross-cultural differences in Weber fractions, with maths-educated adults having better ANS acuity than adults from indigenous cultures lacking maths education<sup>5,14</sup>. These causal relationships, possible tertiary factors and the trainability of ANS acuity<sup>26</sup> remain to be explored. Further evidence will add to the present results, which suggest that our ability to reason over symbolic numbers is deeply entwined with an evolutionarily ancient system for numerical approximation.

## METHODS SUMMARY

At age 14 (that is, ninth grade), ANS acuity was assessed for 64 subjects (see Methods). The percentage correct on the ANS task was modelled for each individual subject as  $1 - \text{error rate}$ , where error rate is defined as:

$$\frac{1}{2} \operatorname{erfc} \left( \frac{n_1 - n_2}{\sqrt{2w} \sqrt{n_1^2 + n_2^2}} \right)$$

where  $\operatorname{erfc}(x)$  is the complementary error function related to the integration of the normalized Gaussian distribution. This model fits percentage correct as a function of the Gaussian approximate number representations for the two sets displayed on a trial ( $n_1$  and  $n_2$ , that is, blue dots and yellow dots) with a single free parameter, the Weber fraction ( $w$ ; see Supplementary Information)<sup>5</sup>. Correlations presented were between this estimate of ANS acuity ( $w$ ), measured at age 14, and scores on standardized cognitive and performance measures, from kindergarten to sixth grade.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 10 February; accepted 4 July 2008.

Published online 7 September 2008.

- Carey, S. Bootstrapping and the origin of concepts. *Daedalus* **133**, 59–68 (2004).
- Feigenson, L., Dehaene, S. & Spelke, E. S. Core systems of number. *Trends Cogn. Sci.* **8**, 307–314 (2004).
- Barth, H., Kanwisher, N. & Spelke, E. S. The construction of large number representations in adults. *Cognition* **86**, 201–221 (2003).
- Cordes, S., Gelman, R. & Gallistel, C. R. Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychon. Bull. Rev.* **8**, 698–707 (2001).
- Pica, P., Lemer, C., Izard, V. & Dehaene, S. Exact and approximate arithmetic in an Amazonian indigene group. *Science* **306**, 499–503 (2004).
- Gordon, P. Numerical cognition without words: Evidence from Amazonia. *Science* **306**, 496–499 (2004).
- Xu, F. & Spelke, E. S. Large number discrimination in 6-month-old infants. *Cognition* **74**, B1–B11 (2000).
- Nieder, A. & Miller, E. A. Parieto-frontal network for visual numerical information in the monkey. *Proc. Natl Acad. Sci. USA* **101**, 7457–7462 (2004).
- Meck, W. H. & Church, R. M. A mode control model of counting and timing processes. *J. Exp. Psychol. Anim. Behav. Process.* **9**, 320–334 (1983).
- Brannon, E. M. & Terrace, H. S. Ordering of the numerosities 1–9 by monkeys. *Science* **282**, 746–749 (1998).
- Dehaene, S., Dehaene-Lambertz, G. & Cohen, L. Abstract representations of numbers in the animal and human brain. *Trends Neurosci.* **21**, 355–361 (1998).
- Gallistel, C. R. *The Organization of Learning* (MIT Press, 1990).
- Dehaene, S. *The Number Sense: How the Mind Creates Mathematics*. (Oxford Univ. Press, 1997).
- Halberda, J. & Feigenson, L. Developmental change in the acuity of the “Number Sense”: The approximate number system in 3-, 4-, 5-, 6-year-olds and adults. *Dev. Psychol.* (in the press).
- Lachance, J. A. & Mazzocco, M. M. M. A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learn. Individ. Differ.* **16**, 195–216 (2006).
- Mandler, G. & Shebo, B. J. Subitizing: an analysis of its component processes. *J. Exp. Psychol. Gen.* **111**, 1–22 (1982).
- Ginsburg, H. & Baroody, A. *Test of Early Mathematics Ability* 2nd edn (Austin, 1990).
- Woodcock, R. W. & Johnson, M. B. *Woodcock-Johnson Psycho-Educational Battery-Revised* (DLM Teaching Resources, Allen, 1989).
- Denckla, M. B. & Rudel, R. G. Rapid automatized naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia* **14**, 471–479 (1976).
- Wechsler, D. *Wechsler Abbreviated Scale of Intelligence (WASI)* (Psychological Corporation, 1999).
- Fias, W. & Fischer, M. H. in *Handbook of Mathematical Cognition* (ed. Campbell, J.) 43–54 (Psychology Press, 2005).
- Mazzocco, M. M. M. & Myers, G. F. Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Ann. Dyslexia* **53**, 218–253 (2003).

23. Kurdek, L. A. & Sinclair, R. J. Predicting reading and mathematics achievement in fourth-grade children from kindergarten readiness scores. *J. Educ. Psychol.* **93**, 451–455 (2001).
24. Espy, K. A. *et al.* The contribution of executive functions to emergent mathematic skills in preschool children. *Dev. Neuropsychol.* **26**, 465–486 (2004).
25. Mazzocco, M. M. M. & Kover, S. T. A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychol.* **13**, 18–45 (2007).
26. Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L. & Dehaene, S. An open trial assessment of “The Number Race”, an adaptive computer game for remediation of dyscalculia. *Behav. Brain Funct.* **2**, 19 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** Longitudinal data collection was supported by a National Institutes of Health grant RO1 HD 034061 to M.M. We thank the Baltimore County Public School District and the children and parents who participated in this research, D. Naimen and R. Goodman for help with modelling, and L. Stapleton for help with statistical analyses.

**Author Contributions** J.H., M.M. and L.F. conceived the experiment; J.H. designed the numerical discrimination procedure; M.M. provided longitudinal data and oversaw data collection; J.H. performed the modelling and data analysis; J.H., L.F. and M.M. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to J.H. ([halberda@jhu.edu](mailto:halberda@jhu.edu)).

## METHODS

**Subjects.** Sixty-four 14-yr-old children participated (32 male; mean age 14 yr 10 months, ranging from 14 yr 3 months to 15 yr 9 months). They were from lower-middle to upper-middle economic backgrounds and were enrolled in normal (that is, non-learning-disabled) public classrooms in a single suburban school district outside of Baltimore, Maryland, USA. Throughout the longitudinal portion of the study, subjects were tested yearly in one, two or three 1-h sessions. The numerical discrimination task was performed during a laboratory visit during the tenth year of the longitudinal study (that is, ninth grade). Further details are given in the Supplementary Information. A total of 80 14-yr-olds were tested during the tenth year assessment but 9 were removed from the final sample owing to high variability in their performance on the ANS acuity assessment (see Modelling and analysis section), and 7 were removed because of missing data for some portion of the standardized tests in grades kindergarden to sixth grade.

**Numerical discrimination.** Subjects completed this task twice at an interval of approximately 60 min, as the first and last sub-tasks in a larger test battery from the longitudinal study. Each run of the task lasted 5 min. Subjects viewed dot arrays on a computer screen and judged whether there were more blue or more yellow dots. For each trial, pressing the space bar initiated a 250 ms blank-screen delay followed by a 200 ms appearance of an array of intermixed blue and yellow dots. After the array had disappeared, subjects had an unlimited amount of time to indicate their response by pressing a colour-coded keyboard button and saying the name of the more numerous colour aloud. Reaction time averaged approximately 1,100 ms across subjects. Subjects were told that, if they wished to change their choice, they could correct an erroneous key-press response by reporting the intended response to the experimenter, who noted it on a score sheet. Self corrections were reported by only 6 of the 64 subjects and accounted for only 7 out of the 5,120 total trials recorded in the study. The number of dots in each set in the array ranged from 5 to 16. Whether the yellow or blue set was larger was randomized. Each trial was drawn from one of four ratio bins in which the ratio of the smaller to the larger set was 1:2, 3:4, 5:6 or 7:8. For each of two runs of the experiment, subjects received 10 practice trials randomly selected from these ratios followed by 40 randomly ordered test trials (10 trials per ratio). Half of the trials in each ratio were 'dot-size controlled': the size of the average blue dot was equal to the size of the average yellow dot. On these trials, the set with more dots necessarily also had a larger total area on screen. The other half of trials were 'area controlled': the total number of blue pixels equalled the total number of yellow pixels such that the total cumulative area of the two sets was identical. The set with more dots thereby had smaller dots on average. Because the two sets were spatially overlapping and each dot was randomly placed in a shared display window, area-controlled trials also controlled for other continuous variables associated with number such as total dot density, inter-dot distance and the total envelope size of each set. Preliminary analyses showed similar results for dot-size-controlled and area-controlled trials and these data were

therefore combined for each subject. On both dot-size-controlled and area-controlled trials, individual dot size varied randomly by up to  $\pm 35\%$  of the set average to discourage the use of individual dot size as a proxy for number. The diameter of a typical dot subtended approximately 1 degree of visual angle from a viewing distance of 50 cm.

**Modelling and analysis.** Previous investigations have modelled numerical representations either as having linearly increasing means and linearly increasing standard deviation<sup>27</sup>, or as having logarithmically compressed means with constant standard deviation<sup>8</sup>. Both of these formats capture the performance pattern that is characteristic of the ANS (error that increases linearly with target numerosity). We used a classical psychophysics model that relies on a linear format of the ANS (although a logarithmic model makes the same predictions for our simple numerical acuity task), which provides a psychologically plausible model of performance in numerical discrimination<sup>5</sup>. Percentage correct was modelled as a function of increasing ratio (larger set/smaller set, or  $n_2/n_1$ ). The numerosity for the blue set and yellow set were represented as Gaussian random variables (that is,  $X_2$  and  $X_1$ ) with means  $n_2$  and  $n_1$  and standard deviations equal to the Weber fraction  $w \times n$ . Subtracting the Gaussian for the smaller set from the larger set returned a new Gaussian with a mean of  $n_2 - n_1$  and a standard deviation of  $w\sqrt{n_1^2 + n_2^2}$  (simply the difference of two Gaussian random variables). Percentage correct was then equal to  $1 - \text{error rate}$ , in which error rate is defined as the area under the tail of the resulting Gaussian curve, computed as:

$$\frac{1}{2} \operatorname{erfc} \left( \frac{n_1 - n_2}{\sqrt{2} w \sqrt{n_1^2 + n_2^2}} \right)$$

This model fits percentage correct on the numerical discrimination task as a function of the Gaussian approximate number representations for the two sets (that is, blue and yellow dots) with a single free parameter, the Weber fraction ( $w$ ). An individual subject's Weber fraction ( $w$ ) describes the standard deviations for the Gaussian representations of the ANS, thereby describing the amount of overlap between any two Gaussian representations, and thereby predicting percentage correct for any numerical discrimination. Using this model, the best-fit value for the Weber fraction ( $w$ ) was determined by a program implementing the Levenberg–Marquardt algorithm for nonlinear least-squares fit on the average percentage correct in each ratio bin for each subject. The model attempts to determine the best-fit value for  $w$  in 50 iterations, each iteration being an attempt to reduce the sum of squared error. The model did not settle on a value for 9 of the original 80 subjects we tested, owing to high variability in the accuracy of their responses. These subjects were removed from the analysis.

27. Brannon, E. M., Wusthoff, C. J., Gallistel, C. R. & Gibbon, J. Numerical subtraction in the pigeon: evidence for a linear subjective number scale. *Psychol. Sci.* 12, 238–243 (2001).

# UNC-6/netrin and its receptor UNC-5 locally exclude presynaptic components from dendrites

Vivian Y. Poon<sup>1</sup>, Matthew P. Klassen<sup>1</sup> & Kang Shen<sup>1,2</sup>

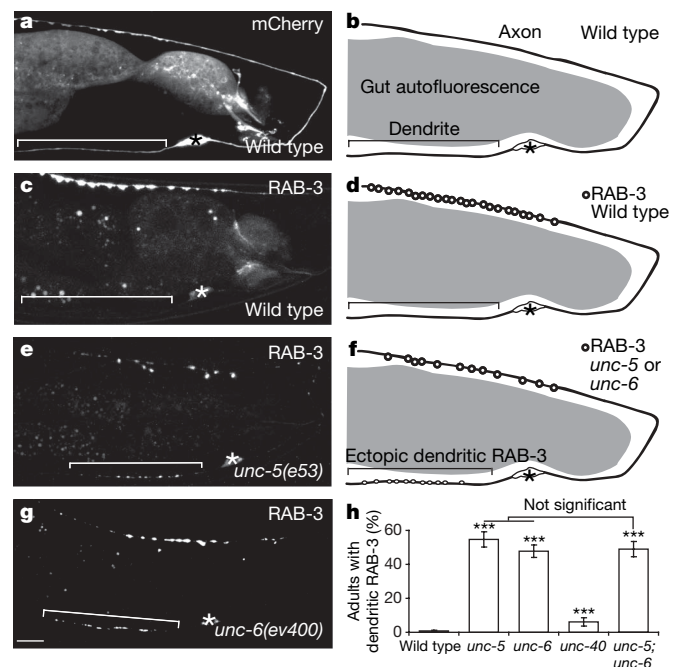
Polarity is an essential feature of many cell types, including neurons that receive information from local inputs within their dendrites and propagate nerve impulses to distant targets through a single axon. It is generally believed that intrinsic structural differences between axons and dendrites dictate the polarized localization of axonal and dendritic proteins<sup>1</sup>. However, whether extracellular cues also instruct this process *in vivo* has not been explored. Here we show that the axon guidance cue UNC-6/netrin and its receptor UNC-5 act throughout development to exclude synaptic vesicle and active zone proteins from the dendrite of the *Caenorhabditis elegans* motor neuron DA9, which is proximal to a source of UNC-6/netrin. In *unc-6/netrin* and *unc-5* loss-of-function mutants, presynaptic components mislocalize to the DA9 dendrite. In addition, ectopically expressed UNC-6/netrin, acting through UNC-5, is sufficient to exclude endogenous synapses from adjacent subcellular domains within the DA9 axon. Furthermore, this anti-synaptogenic activity is interchangeable with that of LIN-44/Wnt despite being transduced through different receptors, suggesting that extracellular cues such as netrin and Wnts not only guide axon navigation but also regulate the polarized accumulation of presynaptic components through local exclusion.

The *C. elegans* motor neuron DA9 elaborates a molecularly and functionally distinct axon and dendrite<sup>2</sup> (Fig. 1a, b). In wild-type animals, presynaptic components are excluded from the dendrite and accumulate in a stereotyped and discrete domain within the DA9 dorsal axon<sup>3</sup> (Fig. 1c, d). These presynaptic components include synaptic vesicle proteins such as RAB-3, SNB-1/synaptobrevin and SNG-1/synaptogyrin (Supplementary Fig. 1), the L-type voltage-gated calcium channel  $\beta$ -subunit CCB-1, and the active zone protein SYD-2/ $\alpha$ -liprin (Supplementary Fig. 2). In exploring whether extracellular cues instruct this polarized localization, we found that these presynaptic proteins mislocalize to the DA9 dendrite in *unc-6/netrin* (*ev400*) and *unc-5(e53)* null mutants (Fig. 1e–h, and Supplementary Figs 1 and 2). This mislocalization defect is not enhanced in *unc-5;unc-6/netrin* double mutants, suggesting that UNC-5 and UNC-6/netrin function in the same pathway (Fig. 1h). A null mutation in the other principal UNC-6/netrin receptor, UNC-40, results in a minor mislocalization defect (Fig. 1h). We further observed that this mislocalization is partly suppressed by a mutation in the presynaptic assembly gene, *syd-2/liprin- $\alpha$*  (Supplementary Fig. 3), suggesting that SYD-2/ $\alpha$ -liprin promotes the accumulation of GFP::RAB-3 in the DA9 dendrite. In addition to the mislocalization defect, the average number of GFP::RAB-3 puncta in the dorsal axon of DA9 is reduced in *unc-5* mutants compared with wild-type animals (Fig. 1e–g and Supplementary Fig. 4).

Netrins are evolutionarily conserved axon guidance molecules present in worms<sup>4</sup>, flies<sup>5</sup> and mammals<sup>6</sup>. The activity of these secreted

molecules is transmitted through two distinct cell-surface receptors: UNC-5 repels axons<sup>7</sup> whereas UNC-40/DCC/Frazzled<sup>8–10</sup> attracts axons to a source of UNC-6/netrin. Similar to mammals, UNC-6/netrin in *C. elegans* is expressed in many classes of ventral cells and its expression persists into adulthood<sup>4</sup>. The UNC-5 receptor is expressed in DA motor neurons, as indicated by antibody staining, transgene expression<sup>11</sup> and microarray analysis<sup>12</sup>. We confirmed that UNC-5 is expressed in DA9 with a transgenic line expressing dsRed driven by the *unc-5* promoter (Supplementary Fig. 5a–c).

To distinguish if UNC-5 is required for localization of presynaptic components in DA9 itself or in other cells like the postsynaptic partners of DA9 (VD/DD neurons and dorsal body-wall muscles), we performed cell-autonomous rescue experiments with various promoters. The *mig-13* promoter is expressed only in DA9 within the tail region at the early larval L1 stage<sup>13</sup> and a *mig-13::unc-5* transgene



**Figure 1** | GFP::RAB-3 is mislocalized to the dendrite in *unc-5* and *unc-6/netrin* mutants. **a–d**, Micrographs and diagrams of representative wild-type adults expressing cytoplasmic mCherry (**a, b**) or GFP::RAB-3 (**c, d**). **e–g**, Micrographs and diagram of representative *unc-5* and *unc-6/netrin* mutant adults expressing GFP::RAB-3. Signal in the middle of the worm is gut autofluorescence. Anterior, left; dorsal, top; brackets, dendrites; asterisks, cell bodies. Scale bar, 10  $\mu$ m. **h**, Penetration of dendritic GFP::RAB-3 in adults with no DA9 guidance defects. Error bars, standard error of proportion;  $n > 100$ ; \*\*\* $P < 0.0001$  (versus wild-type animals),  $\chi^2$  test.

<sup>1</sup>Neuroscience Program, Stanford University School of Medicine, 300 Pasteur Drive, California 94305, USA. <sup>2</sup>Department of Biology, Howard Hughes Medical Institute, Stanford University, 385 Serra Mall, California 94305, USA.

robustly rescues the mislocalization defect in *unc-5* mutant L1 and adult animals, suggesting that UNC-5 acts cell-autonomously in DA9 to exclude presynaptic components from the dendrite. Furthermore, we did not observe any rescue when UNC-5 was expressed in the postsynaptic partners of DA9 using the *unc-25* and *unc-129m* promoters<sup>14,15</sup> (Fig. 2a and Supplementary Fig. 6). To substantiate these observations, we created *unc-5* mutant animals expressing a rescuing *unc-5::unc-5* transgene together with a cytoplasmic DA9 marker in a mosaic pattern. In two independent transgenic lines, we observed a strong correlation between the expression of *unc-5* in DA9 and rescue of the mislocalization defect, consistent with a cell-autonomous function for UNC-5 in DA9 (Fig. 2b). Using an UNC-5::YFP fusion construct expressed specifically in DA9, we observed a higher level of UNC-5::YFP in the dendrite and ventral axon than in the dorsal axon (Supplementary Fig. 7).

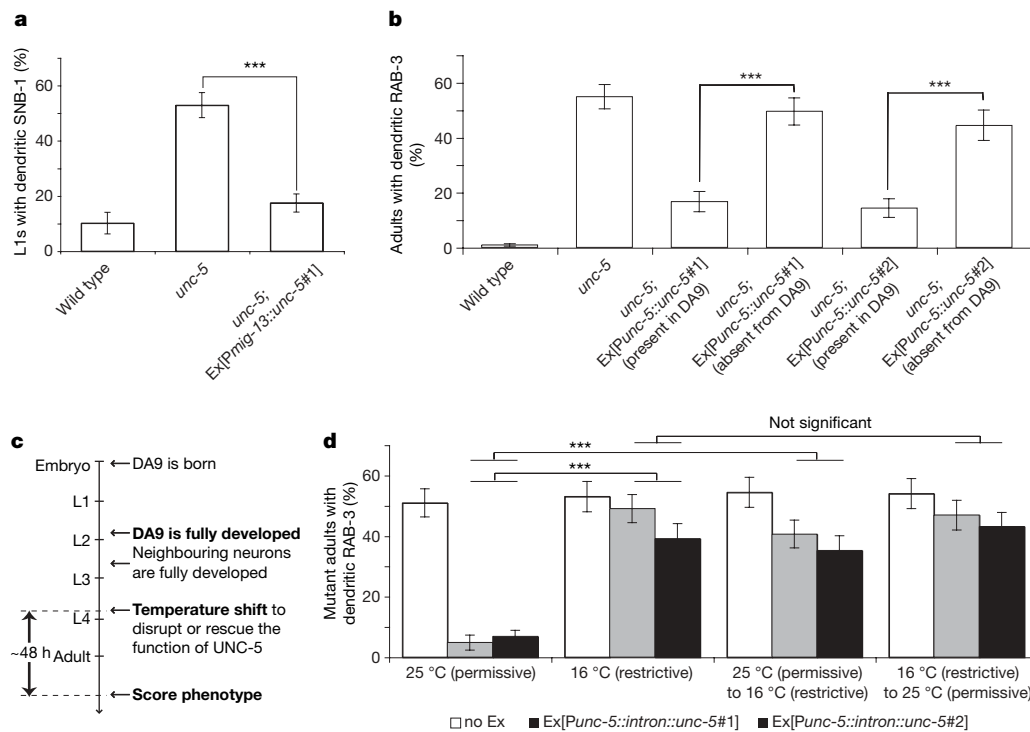
The disrupted distribution of presynaptic components in *unc-5* and *unc-6/netrin* mutants is observable at the early L1 larval stage, when the DA9 dendrite begins to form, and persists throughout the life of the animal (Supplementary Figs 1 and 2a–f). To differentiate whether the mislocalization of presynaptic components is a consequence of an early axodendritic polarity or guidance defect or a later developmental defect, we used a modified version of a silencing intron cassette (M. Chalfie, personal communication) to regulate *unc-5* temporally. We observed that culturing *unc-5* mutant animals expressing the *unc-5::intron::unc-5* transgene at 25 °C throughout development resulted in a significant rescue of the mislocalization defect in *unc-5* mutants, whereas culturing them at 16 °C led to no rescue, suggesting that the transgene produces functional UNC-5 at 25 °C but not at 16 °C (Fig. 2d and Supplementary Fig. 8a). A shift to the restrictive temperature at the L4 larval stage, after DA9 and surrounding neurons are fully developed, results in a mislocalization defect in the transgenic mutant animals that is comparable to *unc-5* mutants. Conversely, a shift to the permissive temperature is insufficient to rescue the mislocalization defect, suggesting that the defect

is irreversible. Therefore, compromising the activity of *unc-5* in mature DA9 neurons leads to a mislocalization defect, suggesting a novel function for UNC-5 in maintaining the polarized localization of GFP::RAB-3 independent of early polarization and guidance.

To elucidate further whether this novel function of UNC-5 can be separated from its previously known role in axon guidance, we examined axon guidance in the DD and VD neurons using the same temperature-shift experimental paradigm. We found that a shift to the restrictive temperature at the L4 larval stage did not cause further errors in axon guidance, suggesting that UNC-5 is only required during the early outgrowth phase to guide axons, and is not required later to maintain axon trajectory (Supplementary Fig. 8b).

The two distinct roles of UNC-5 in axon guidance and GFP::RAB-3 localization is further supported by the following observations. First, approximately half of *unc-5* or *unc-6/netrin* mutant animals have no detectable DA9 guidance defects, yet the mislocalization defect is still observed (Supplementary Figs 9 and 10a). Second, *unc-5* and *unc-6/netrin* mutant animals with defective DA9 guidance do not display more penetrant mislocalization defects (Supplementary Fig. 10a). Third, UNC-129/TGF- $\beta$ , like UNC-5 and UNC-6/netrin, is important for the dorsal guidance of DA neurons<sup>15</sup>, and we observed that approximately half of *unc-129/Tgf- $\beta$*  mutant animals have severe DA9 guidance defects but none exhibit dendritic GFP::RAB-3 (Supplementary Fig. 10b, c).

What cell biological processes do UNC-5 and UNC-6/netrin affect in localizing presynaptic components? It is possible that they are important for the establishment and maintenance of dendritic fate. Hence, we examined the localization of four dendritically localized proteins: CAM-1/ROR<sup>16</sup>, a receptor tyrosine kinase that localizes somatodendritically in hippocampal cultures<sup>17</sup>; UNC-9/innexin, a structural component of invertebrate gap junctions<sup>18</sup>; F35D2.3/fibrillin<sup>16</sup>; and DYS-1/dystrophin<sup>16</sup>. In wild-type animals, CAM-1, UNC-9 and DYS-1 localize to the DA9 dendrite, cell body and ventral axon, whereas F35D2.3/fibrillin localizes exclusively to the dendrite. These



**Figure 2 | UNC-5 acts cell autonomously in DA9 and is required throughout development.** **a**, Expression of UNC-5 in DA9 rescues the mislocalization defect in *unc-5* mutant L1 animals ( $n > 100$ ). **b**, Expression of UNC-5 in other cells does not rescue the mislocalization defect ( $n > 80$ ). **c**, Experimental timeline. **d**, The *unc-5::intron::unc-5* transgene rescues the

mislocalization defect of *unc-5;mec-8* mutant adults at 25 °C, not 16 °C. The mislocalization defect occurs when UNC-5 is inactivated early or late in development. Error bars, standard error of proportion;  $n > 100$ ; \*\*\* $p < 0.0001$  (within each transgenic line),  $\chi^2$  test.

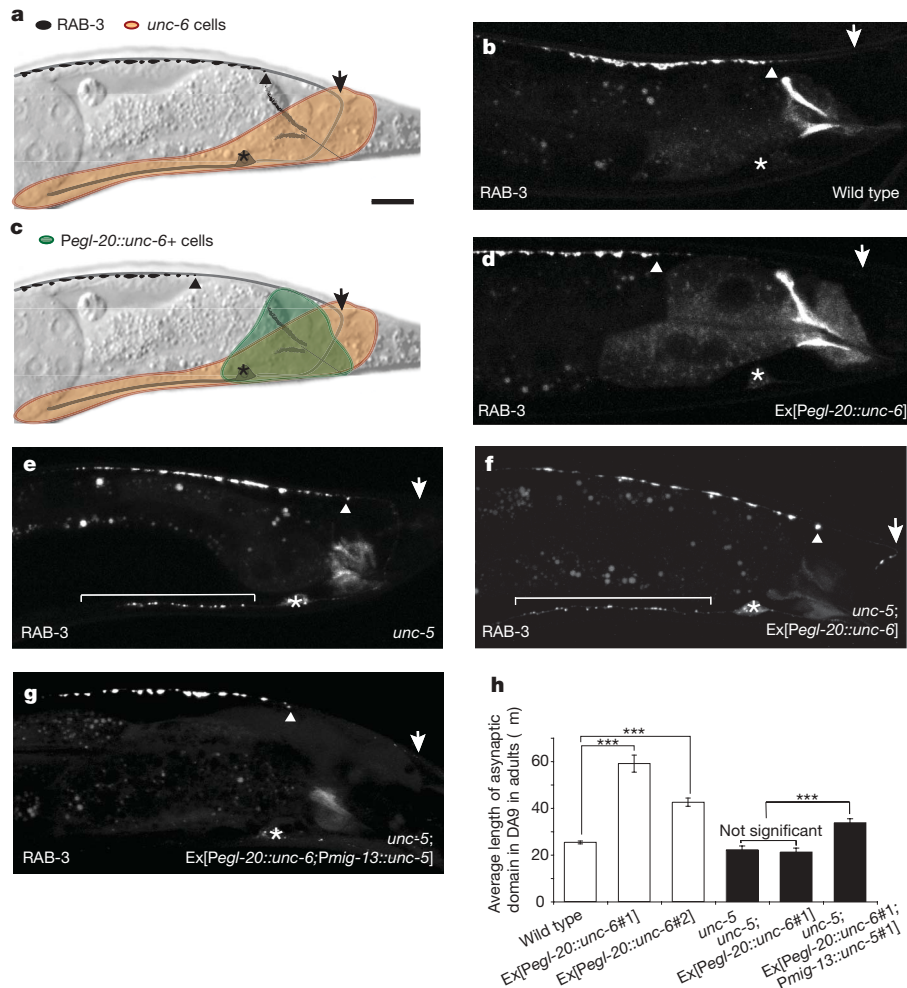
localization patterns are unaffected in *unc-5* and *unc-6/netrin* mutants, suggesting that many aspects of axodendritic polarization are maintained in these mutants (Supplementary Fig. 11). These results are consistent with the late temporal requirement of UNC-5 for proper localization of presynaptic components.

An alternative possibility is that UNC-5 and UNC-6/netrin regulate accumulation of presynaptic components in the dorsal axon, and that reduced axonal accumulation may indirectly cause the mislocalization defect. However, we did not observe any significant difference when we compared the average number of axonal GFP::RAB-3 puncta between *unc-5* mutants with and without dendritic GFP::RAB-3 (Supplementary Fig. 3). We conclude that this model is unlikely to be true.

To test directly whether UNC-6/netrin provides instructive information for the localization of presynaptic components, we generated a posterior to anterior gradient of UNC-6/netrin near DA9 using the *egl-20* promoter<sup>19</sup> where the posterior segment of the DA9 dorsal axon is exposed to an abnormally high level of UNC-6/netrin. This ectopic expression of UNC-6/netrin causes DA9 guidance defects in a small fraction of animals ( $12.2 \pm 3.2\%$ ). When we examined animals with normal DA9 guidance, we observed that the *egl-20::unc-6/netrin* transgene dramatically displaces GFP::RAB-3 anteriorly, creating an enlarged asynaptic zone in the posterior segment of the DA9 dorsal axon compared with wild-type animals (Fig. 3a–d, h). As the transgene does not affect the dorsal

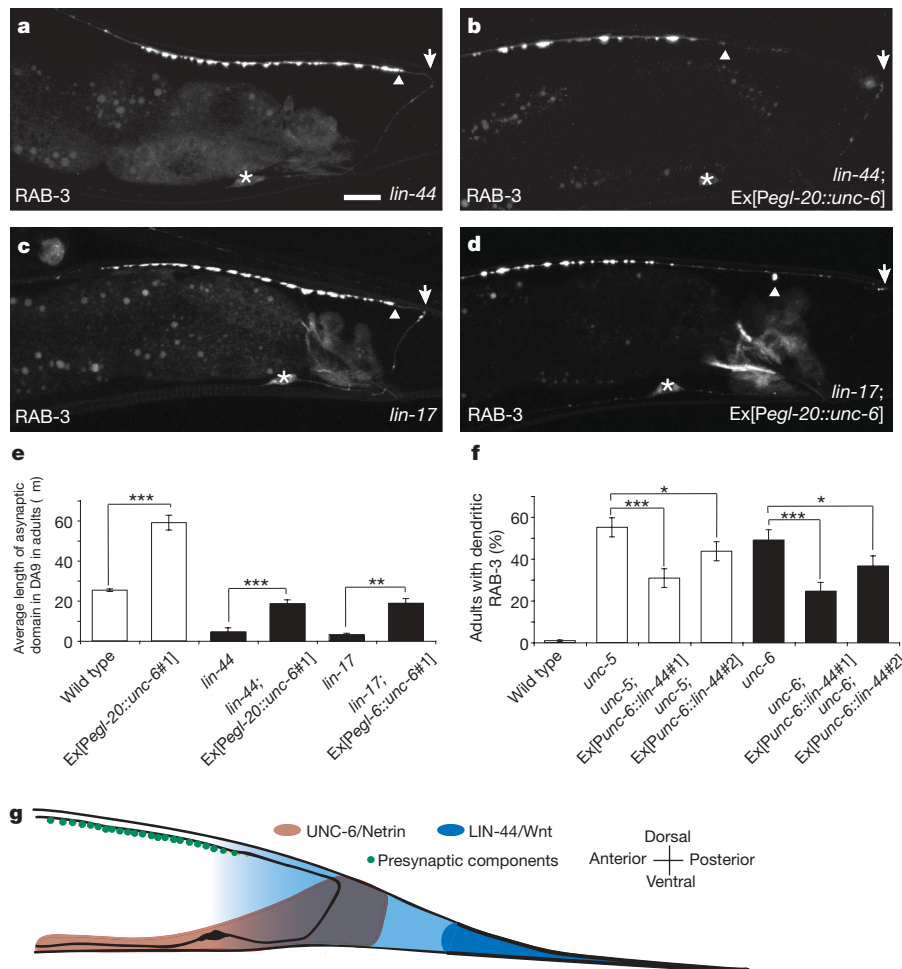
axon length of DA9 (Supplementary Fig. 12), it is unlikely that altered axonal outgrowth of DA9 leads to the enlarged asynaptic zone. We further observed that *unc-5* mutants expressing the *egl-20::unc-6/netrin* transgene do not have an enlarged asynaptic domain (Fig. 3e, f). In addition, the enlarged asynaptic domain is partly restored in *unc-5* mutants expressing the *mig-13::unc-5* transgene, demonstrating that UNC-5 acts cell-autonomously in DA9 to mediate ectopic UNC-6/netrin-induced exclusion of presynaptic components (Fig. 3g, h).

The striking similarity between the displacement of presynaptic components induced by the *egl-20::unc-6/netrin* and *egl-20::lin-44/wnt* transgenes<sup>3</sup> suggests that UNC-6/netrin and LIN-44/Wnt can both exclude synapses. We observed that the asynaptic domains in *lin-44/wnt(n1792);Ex[Pegl-20::unc-6/netrin]* or *lin-17/fz(n671);Ex[Pegl-20::unc-6/netrin]* mutants are significantly larger than those in *lin-44/wnt* or *lin-17/fz* mutants alone, suggesting that ectopically expressed UNC-6/netrin is sufficient to rescue the mislocalization defect in *lin-44/wnt* and *lin-17/fz* mutants (Fig. 4a–e). We further observed a significant reduction in penetrance of the mislocalization defects in *unc-5* and *unc-6/netrin* mutants expressing *lin-44/wnt* under the control of the *unc-6* promoter (Fig. 4f). If UNC-6/netrin and LIN-44/Wnt have similar functions, one might expect that *unc-6/netrin;lin-44/wnt* double mutants would have more severe mislocalization defects. However, these double mutants exhibit a fully penetrant guidance defect such that the DA9 axon turns posteriorly (data not shown), precluding analysis of presynaptic



**Figure 3 | UNC-6/netrin is sufficient to exclude GFP::RAB-3 locally and acts through UNC-5.** **a**, Endogenous UNC-6/netrin-expressing cells (beige). **c**, Ectopically expressed UNC-6/netrin (green). **b, d**, Representative wild-type L4 animal expressing GFP::RAB-3 in the absence (**b**) or presence (**d**) of ectopic UNC-6/netrin. **e–g**, Representative *unc-5* mutant adults expressing

GFP::RAB-3 with (**f**) or without (**e**) ectopic UNC-6/netrin or, alternatively, ectopic UNC-6/netrin with the *mig-13::unc-5* transgene (**g**). Arrows, posterior end of dorsal axon; arrowheads, posterior border of presynaptic domain; brackets, dendrites; asterisks, cell bodies. Scale bar, 10  $\mu\text{m}$ . **h**, Error bars, s.e.m. ( $n > 50$ ); \*\*\* $P < 0.0001$ ,  $t$ -test.



**Figure 4 | UNC-6/netrin and LIN-44/Wnt function interchangeably.** **a–d**, Representative *lin-44/wnt* or *lin-17/fz* mutant adults expressing GFP::RAB-3 with (**b, d**) or without (**a, c**) ectopic UNC-6/netrin. Arrows, posterior end of dorsal axon; arrowheads, posterior border of presynaptic domain; brackets, dendrites; asterisks, cell bodies. Scale bar, 10 μm. **e**, Error

bars, s.e.m. ( $n > 50$ ). \*\*\* $P < 0.0001$ ,  $t$ -test. **f**, Error bars, standard error of proportion ( $n > 100$ ); \*\*\* $P < 0.0001$ ; \*\* $P < 0.005$ ; \* $P < 0.05$ ;  $\chi^2$  test. **g**, Model for the roles of UNC-6/netrin and LIN-44/Wnt in subcellular patterning of presynaptic specializations in DA9.

localization. Collectively, these results suggest that UNC-6/netrin and LIN-44/Wnt play parallel roles in specifying the discrete presynaptic domain of DA9 by excluding presynaptic components from inappropriate compartments (Fig. 4g). Ventrally secreted UNC-6/netrin excludes presynaptic components from the DA9 dendrite, whereas LIN-44/Wnt secreted by the tail hypodermal cells performs a similar function in the posterior segment of the DA9 dorsal axon.

Here we demonstrate a novel role for UNC-6/netrin in providing spatial information for the exclusion of presynaptic components throughout development. Interestingly, UNC-6/netrin was recently shown to promote presynaptic formation in the amphid interneuron AIY in *C. elegans*<sup>20</sup>. These opposing effects of UNC-6/netrin on presynaptic formation might be explained by the different receptors used: UNC-5 in DA9 and UNC-40/DCC in AIY. The roles of these receptors in synaptic polarization parallel their contrasting functions in axon guidance, with UNC-5 functioning in repulsion and UNC-40/DCC in attraction<sup>7–10</sup>.

In addition to its well-characterized function in axon guidance, UNC-6/netrin was recently implicated in the initial polarization of the *C. elegans* hermaphrodite-specific neuron (HSN) neuronal cell body<sup>21</sup>. However, it is unclear whether UNC-6/netrin is required for later stages of neuronal polarity. Our findings suggest that UNC-6/netrin and UNC-5 activity coordinate two temporally distinct functions in DA9: axons are first guided to the appropriate locations, and presynaptic components are later localized in a polarized manner.

Consistent with this hypothesis, netrin is expressed in the adult mammalian nervous system long after axon guidance is complete<sup>22</sup>.

The conventional view of synapse formation is that contact between synaptic partners triggers assembly of the pre- and postsynaptic apparatus through the interaction of adhesion molecules like neuroligin/neurexin, SynCAM and EphrinB/EphB receptor<sup>23</sup> across the synaptic cleft. However, extracellular cues such as members of the Wnt<sup>24,25</sup>, fibroblast growth factor<sup>26</sup> and bone morphogen protein<sup>27</sup> families can also promote synapse formation. Our studies in DA9 suggest that negative regulators also pattern synaptogenesis by inhibiting the accumulation of presynaptic components in inappropriate subcellular domains. The UNC-6/netrin gradient is high ventrally and low dorsally<sup>4</sup>, encompassing the dendrite and ventral axon of DA9. The LIN-44/Wnt gradient is high posteriorly and low anteriorly<sup>28</sup>, effectively reaching the ventral axon, commissure and posterior region of the DA9 dorsal axon. Signalling through independent receptors, both UNC-6/netrin and LIN-44/Wnt, excludes presynaptic components, setting negative constraints on presynaptic formation in DA9 and forcing synapses to form in a discrete domain within the DA9 dorsal axon (Fig. 4g). Thus, inhibitory factors play essential roles in patterning the subcellular distribution of synapses.

#### METHODS SUMMARY

**Strains and genetics.** Worms were raised on OP50 *Escherichia coli*-seeded nematode growth medium (NGM) plates at 22 °C. The following mutant strains were

obtained through the Caenorhabditis Genetics Center: CB271 *unc-40(e271)I*, MT1306 *lin-17(n671)I*, MT4051 *lin-44(n1792)I*, NW987 *unc-129(ev554)IV*, MT464 *unc-5(e53)IV*; *dpy-11(e224)V*; *lon-2(e678)X*, NW434 *unc-6(ev400)X*, CB620 *lin-18(e620)X* and CZ900 *syd-2(ju37)X*. TU218 *mec-8(u218)I* was provided by M. Chalfie. N2 Bristol was used as the wild-type reference strain.

**Cloning and constructs.** Expression clones were made in a derivative of pPD49.26 (A. Fire), the pSM vector (S. McCarroll and C. I. Bargmann, personal communication). The plasmids and strains *wyIs85(Pitr-1 pB::gfp::rab-3)*, *wyIs92(Pmig-13::snb-1::yfp)*, *wyEx403(Pitr-1 pB::cam-1::yfp)*, *wyEx771(Pmig-13::gfp::ccb-1; Pmig-13::mcherry::rab-3)* and *wyIs75(Punc-47::rfp; Pexp-1::gfp)* (G. Maro and K.S., unpublished observations) were generated as previously described<sup>3</sup>. The following plasmids and strains were generated using standard techniques<sup>29</sup>: *wyEx1902(Pitr-1 pB::mcherry)*, *wyEx2055(Pitr-1 pB::gfp::syd-2; Pitr-1 pB::mcherry::rab-3)*, *wyIs109(Pmig-13::cfp::rab-3; Pmig-13::snb-1::yfp; Pmig-13::sng-1::mcherry)*, *wyEx1054(Pitr-1 pB::unc-9::yfp)*, *wyEx2396(Pitr-1 pB::F35D2.3/fibrillin::yfp)*, *wyEx2430(Pitr-1 pB::dys-1::yfp)*, *wyEx1311(Pmig-13::unc-5#1)*, *wyEx1485(Pmig-13::unc-5#2)*, *wyEx1498(Punc-5::unc-5::sl2::dsred)*, *wyEx1228(Punc-5::unc-5)*, *wyEx2419(Pmig-13::mcherry; Punc-5::unc-5#1)*, *wyEx2418(Pmig-13::mcherry; Punc-5::unc-5#2)*, *wyEx1277(Pitr-1 pB::unc-5::yfp)*, *wyEx2306(Punc-5::intron::unc-5#1)*, *wyEx2308(Punc-5::intron::unc-5#2)*, *wyEx1904(Pegl-20::unc-6#1)*, *wyEx1916(Pegl-20::unc-6#2)*, *wyEx2093(Punc-6::lin-44#1)* and *wyEx2094(Punc-6::lin-44#2)*. We used the co-injection markers *Podr-1::GFP* or *dsRED* injected at 20 ng  $\mu\text{l}^{-1}$  or *Pttx-3::CFP* injected at 50 ng  $\mu\text{l}^{-1}$ ; transgenes were introduced into *C. elegans* as previously described<sup>29</sup>.

**Fluorescence microscopy and confocal imaging.** Images of fluorescently tagged fusion proteins were captured in live *C. elegans* using a Plan-Apochromat 63 $\times$ /1.4 objective on a Zeiss LSM510 confocal microscope. Worms were immobilized using 10 mM levamisole (Sigma). Only animals with no detectable DA9 guidance defects were used for imaging and quantification, unless otherwise specified.

**Quantification of average synaptic domain or dorsal axon length.** Asynaptic domain length was quantified as previously described<sup>3</sup>. Animals in the mid-L4 larval stage expressing a DA9-specific cytoplasmic marker (*wyEx1902*) were imaged using a Zeiss Axiophot with AxioCam/Axiovision digital imaging. As the DA9 dorsal axon extends slightly anterior to the vulva, the distance between the end of the axon and the vulva, a landmark for the mid-body, was measured.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 29 May; accepted 29 July 2008.

Published online 7 September 2008; corrected 2 October 2008 (details online).

- Horton, A. C. & Ehlers, M. D. Neuronal polarity and trafficking. *Neuron* **40**, 277–295 (2003).
- White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the ventral nerve cord of *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B* **275**, 327–348 (1976).
- Klassen, M. P. & Shen, K. Wnt signaling positions neuromuscular connectivity by inhibiting synapse formation in *C. elegans*. *Cell* **130**, 704–716 (2007).
- Wadsworth, W. G., Bhatt, H. & Hedgecock, E. M. Neuroglia and pioneer neurons express UNC-6 to provide global and local netrin cues for guiding migrations in *C. elegans*. *Neuron* **16**, 35–46 (1996).
- Mitchell, K. J. *et al.* Genetic analysis of Netrin genes in *Drosophila*: Netrins guide CNS commissural axons and peripheral motor axons. *Neuron* **17**, 203–215 (1996).
- Serafini, T. *et al.* Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system. *Cell* **87**, 1001–1014 (1996).
- Leung-Hagsteeijn, C. *et al.* UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in *C. elegans*. *Cell* **71**, 289–299 (1992).
- Chan, S. S. *et al.* UNC-40, a *C. elegans* homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues. *Cell* **87**, 187–195 (1996).
- Keino-Masu, K. *et al.* Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell* **87**, 175–185 (1996).

- Kolodziej, P. A. *et al.* *frazzled* encodes a *Drosophila* member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. *Cell* **87**, 197–204 (1996).
- Su, M. *et al.* Regulation of the UNC-5 netrin receptor initiates the first reorientation of migrating distal tip cells in *Caenorhabditis elegans*. *Development* **127**, 585–594 (2000).
- Fox, R. M. *et al.* A gene expression fingerprint of *C. elegans* embryonic motor neurons. *BMC Genomics* **6**, 42 (2005).
- Sym, M., Robinson, N. & Kenyon, C. MIG-13 positions migrating cells along the anteroposterior body axis of *C. elegans*. *Cell* **98**, 25–36 (1999).
- Jin, Y., Jorgensen, E., Hartwig, E. & Horvitz, H. R. The *Caenorhabditis elegans* gene *unc-25* encodes glutamic acid decarboxylase and is required for synaptic transmission but not synaptic development. *J. Neurosci.* **19**, 539–548 (1999).
- Colavita, A., Krishna, S., Zheng, H., Padgett, R. W. & Culotti, J. G. Pioneer axon guidance by UNC-129, a *C. elegans* TGF- $\beta$ . *Science (New York, N. Y.)* **281**, 706–709 (1998).
- Sieburth, D. *et al.* Systematic analysis of genes required for synapse structure and function. *Nature* **436**, 510–517 (2005).
- Paganoni, S. & Ferreira, A. Expression and subcellular localization of Ror tyrosine kinase receptors are developmentally regulated in cultured hippocampal neurons. *J. Neurosci. Res.* **73**, 429–440 (2003).
- Phelan, P. Innexins: members of an evolutionarily conserved family of gap-junction proteins. *Biochim. Biophys. Acta* **1711**, 225–245 (2005).
- Whangbo, J. & Kenyon, C. A Wnt signaling system that specifies two patterns of cell migration in *C. elegans*. *Mol. Cell* **4**, 851–858 (1999).
- Colon-Ramos, D. A., Margeta, M. A. & Shen, K. Glia promote local synaptogenesis through UNC-6 (netrin) signaling in *C. elegans*. *Science* **318**, 103–106 (2007).
- Adler, C. E., Fetter, R. D. & Bargmann, C. I. UNC-6/Netrin induces neuronal asymmetry and defines the site of axon formation. *Nature Neurosci.* **9**, 511–518 (2006).
- Manitt, C., Wang, D., Kennedy, T. E. & Howland, D. R. Positioned to inhibit: netrin-1 and netrin receptor expression after spinal cord injury. *J. Neurosci. Res.* **84**, 1808–1820 (2006).
- McAllister, A. K. Dynamic aspects of CNS synapse formation. *Annu. Rev. Neurosci.* **30**, 425–450 (2007).
- Hall, A. C., Lucas, F. R. & Salinas, P. C. Axonal remodeling and synaptic differentiation in the cerebellum is regulated by WNT-7a signaling. *Cell* **100**, 525–535 (2000).
- Packard, M. *et al.* The *Drosophila* Wnt, wingless, provides an essential signal for pre- and postsynaptic differentiation. *Cell* **111**, 319–330 (2002).
- Umemori, H., Linhoff, M. W., Ornitz, D. M. & Sanes, J. R. FGF22 and its close relatives are presynaptic organizing molecules in the mammalian brain. *Cell* **118**, 257–270 (2004).
- McCabe, B. D. *et al.* The BMP homolog Gbb provides a retrograde signal that regulates synaptic growth at the *Drosophila* neuromuscular junction. *Neuron* **39**, 241–254 (2003).
- Herman, M. A., Vassilieva, L. L., Horvitz, H. R., Shaw, J. E. & Herman, R. K. The *C. elegans* gene *lin-44*, which controls the polarity of certain asymmetric cell divisions, encodes a Wnt protein and acts cell nonautonomously. *Cell* **83**, 101–110 (1995).
- Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol.* **48**, 451–482 (1995).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by the W. M. Keck Foundation, the McKnight Endowment Fund, the Searle Scholar Award and the Howard Hughes Medical Institute. We thank the International Caenorhabditis Genetic Center and M. Chalfie for strains. We further thank M. Chalfie for sharing unpublished results on the silencing system. We also thank C. Gao and Y. Fu for technical assistance, and T. Clandinin, C. Bargmann, S. McConnell, and members of the Shen laboratory for comments on the manuscript.

**Author Contributions** V.Y.P., M.P.K. and K.S. designed the experiments; V.Y.P. performed the experiments and data analysis; V.Y.P., M.P.K. and K.S. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to K.S. ([kangshen@stanford.edu](mailto:kangshen@stanford.edu)).



## METHODS

**Temperature shift experiments.** Animals were either cultured at 16 °C or 25 °C for multiple generations before being shifted to a different temperature. Experimental animals in the L3 and early L4 larval stages were placed at 16 °C for three days or 25 °C for two days before the phenotype was analysed. Scoring was performed in gravid adults with normal DA9 guidance for all experiments. *unc-5; mec-8* double mutants were analysed for experiments in Fig. 2d. It was later discovered that the temperature-dependent regulation of Ex[P*unc-5::intron::unc-5*] was independent of the *mec-8* mutation (Supplementary Fig. 8a).

**Constructs and transgenic worms.** *wyIs109*: a *XmaI*–*NheI* PCR fragment containing *cfp* was subcloned into *Pttx-3::rab-3* pSM obtained from *Pttx-3::mcherry::rab-3* (ref. 20) to make *Pttx-3::cfp::rab-3*. A *SphI*–*Ascl* fragment containing *Pmig-13* (ref. 3) was subcloned into *cfp::rab-3* pSM derived from *Pttx-3::cfp::rab-3* described above to make *Pmig-13::cfp::rab-3*. A *KpnI*–*Apal* fragment containing mCherry obtained from *Pmig-13::lin-17::mcherry*<sup>3</sup> was subcloned into *Pmig-13::gateway* pSM from *Pmig-13::gateway::yfp*<sup>3</sup> to make *Pmig-13::gateway::mcherry*. The *sng-1* entry clone was obtained from the ORFeome project (<http://worfdb.dfc.harvard.edu/>) and cloned into the destination vector *Pmig-13::gateway::mcherry* using the gateway strategy with LR clonase (Invitrogen) to make *Pmig-13::sng-1::mcherry*. *Pmig-13::cfp::rab-3*, *Pmig-13::sng-1::mcherry* and *Pmig-13::snb-1::yfp*<sup>3</sup> were injected with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  and integrated into chromosome V using trimethylpsoralen/ultraviolet mutagenesis. *cfp* primers: 5'-TCCCCCGGGATGAGTAAAGGAGAAGAAGTTTTCAC and 3'-CTAGCTAGCTTTGTATAGTTCATCCATGCCATG.

*wyEx2055*: a *NheI*–*KpnI* fragment containing the *syd-2* genomic sequence from *Pgcy-8::mCherry::syd-2* was subcloned into *Pitr-1 pB::gfp* pSM to make *Pitr-1 pB::gfp::syd-2*; a *FseI*–*Ascl* PCR fragment containing *Pitr-1 pB* was subcloned into *mcherry::rab-3* pSM from *Pglr-3::mcherry::rab-3*. *Pitr-1 pB::gfp::syd-2* was injected at 0.5 ng  $\mu\text{l}^{-1}$  with *Pitr-1 pB::mcherry::rab-3* at 10 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into N2 animals. *Pitr-1 pB* primers: 5'-GAAAGGGGCGCCATCTATTCAGAGTTCGTTCCCGAGC and 3'-CTTCCGGCGCGCCC AATTCTGTGCTTCCACCACCAC.

*wyEx1902*: a *SphI*–*Ascl* PCR fragment containing *Pitr-1 pB* was subcloned into *mcherry::gateway* from *Pmig-13::mcherry::gateway*. *Pitr-1 pB::mcherry::gateway* was injected at 5 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into N2 animals. *Pitr-1 pB* primers: 5'- (ref. 3) and 3'-GAAAGGGGCGCGCCAATTCTGTGCTTC CACCAC.

*wyEx1311*, *wyEx1485*: a *KpnI*–*Apal* fragment containing the *unc-5* 3' untranslated region (UTR) was subcloned into *Pmig-13::snb-1* pSM and a *Ascl*–*KpnI* PCR fragment containing the *unc-5* genomic sequence was subcloned into *Pmig-13::unc-5* 3' UTR pSM. The *Pmig-13::unc-5* plasmid was injected at 2 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into *unc-5; wyIs85* mutants. The two arrays are separate lines obtained in independent injections. Higher levels of the plasmid were toxic and there would be no germline transmission, whereas lower levels did not rescue the defect. *unc-5* primers: 5'-GAAAGGGGCGCGCCCATG GACGAAATCACAATCACAACAACAAC and 3'-GAAGGGTACCAGTGGGG ACACAATTTGTGAAAAGCTG; *unc-5* 3' UTR primers: 5'-GAAAGGGGT ACCGTCGAATTTTGTGACAAAACAACACTAG and 3'-GAAAGGGGGCC CCGGTCTTTCTGCATAGAAAATCGC.

*wyEx1498*: an *Ascl*–*KpnI* PCR fragment containing the *unc-5* genomic sequence and a *SphI*–*Ascl* PCR fragment containing the *Punc-5* were subcloned

into *sl2 dsred* pSM<sup>20</sup> to make *Punc-5::unc-5::sl2::dsred*. This plasmid was injected at 10 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into *unc-5; lin-18; wyIs85* mutants.

*wyEx1228*, *wyEx2418*, *wyEx2419*: a *SphI*–*Ascl* PCR fragment containing *Punc-5* was subcloned into *unc-5::unc-5* 3' UTR from *Pmig-13::unc-5::unc-5* 3' UTR. *Punc-5::unc-5* was injected at 10 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into *unc-5; wyIs85* mutants (*wyEx1228*). *Punc-5::unc-5* was injected at 8 ng  $\mu\text{l}^{-1}$  with *Pmig-13::mcherry::gateway*<sup>3</sup> at 8 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into *unc-5; wyIs85* mutants. *wyEx2418* and *wyEx2419* were two arrays obtained that had animals where UNC-5 was absent in DA9. Most of the other lines obtained had no animals where UNC-5 was absent in DA9. *Punc-5* primers: 5'-GAAAGGGCATGCTGAGCTTTTCCAACTAGAGAGCTTC and 3'-GAAAG GGGCGCCTACTGGAATAGAAATTATGATTAGTGACAACTTG.

*wyEx1277*: an *EcoRI*–*Apal* PCR fragment containing the *unc-5* 3' UTR was subcloned into *Pmig-13::snb-1::yfp* pSM<sup>3</sup> to make *Pmig-13::snb-1::yfp::unc-5* 3' UTR. An *Ascl*–*KpnI* PCR fragment containing the *unc-5* genomic sequence (similar to *wyEx1311*) and a *SphI*–*Ascl* PCR fragment containing *Pitr-1 pB* (similar to *wyEx2055*) were subcloned into *unc-5* 3' UTR pSM derived from *Pmig-13::snb-1::yfp::unc-5* 3' UTR. *Pitr-1 pB::unc-5::yfp* was injected at 80 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into N2 animals. *unc-5* 3' UTR primers: 5'-GAAAGGGGAATTCGTCGAATTTTGTGACAAAACAACAACACTAG and 3'-GAAAGGGGGCCCGGTCTTCTGCATAGAAAATCGC.

*wyEx1904*, *wyEx1916*: a *SphI*–*Ascl* PCR fragment containing *Pegl-20* was subcloned into *unc-6* pSM derived from *Punc-6::unc-6* (ref. 20). The *Pegl-20::unc-6* plasmid was injected at 20 ng  $\mu\text{l}^{-1}$  with *Pttx-3::cfp* at 50 ng  $\mu\text{l}^{-1}$  into N2 animals. The two arrays were separate lines obtained from one injection. *Pegl-20* primers: 5'-GAAAGGGCATGCAAGTTCCCTTTTATTTTTGAAGTCA TCC and 3'-GAAAGGGGCGCGCCTATTCTGAAATTGAGATGTTTGA AATTC.

*wyEx2093*, *wyEx2094*: an *Ascl*–*KpnI* PCR fragment containing the *lin-44* complementary DNA was subcloned into *Punc-6* pSM derived from *Punc-6::unc-6::mcherry*<sup>20</sup>. The *Punc-6::lin-44* plasmid was injected at 20 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into *unc-5; wyIs85* mutants. The two arrays were separate lines obtained in the same injection. *lin-44* primers: 5'-GAAAGGGGCG CGCCATCGAGCAGCTCTTTTGTATTTC and 3'-GAAAGGGGTACCTTAA AAAATTAGGCTTTTTCGGCGGTG.

*wyEx2306*, *wyEx2308*: an *Ascl* PCR fragment obtained from pAC13, a gift from M. Chalfie, containing *mec-2 intron9* was subcloned into *Punc-5::unc-5* (similar to *wyEx1494*). This plasmid was injected at 10 ng  $\mu\text{l}^{-1}$  with *Podr-1::gfp* at 20 ng  $\mu\text{l}^{-1}$  into *unc-5; mec-8; wyIs85* mutants. The two arrays were separate lines obtained from one injection. *unc-5; wyIs85; wyEx2308* mutant animals were obtained by crossing *unc-5; mec-8; wyIs85; wyEx2308* animals with N2 animals. *mec-2 intron9* primers: 5'-GAAAGGGGCGCGCCACCACCGCTAAAGTG TAAGTTTTC and 3'-GAAAGGGGCGCGCCGACGGTGGCTCCTCACTGAA AAC.

*wyEx1054*, *wyEx2396*, *wyEx2430*: the *unc-9*, *F35D2.3* and *dys-1* entry clones were obtained from the ORFeome project (<http://worfdb.dfc.harvard.edu/>) and cloned into the destination vector *Pitr-1 pB::gateway::yfp*<sup>3</sup> using the gateway strategy with LR clonase (Invitrogen) to make *Pitr-1 pB::unc-9::yfp*, *Pitr-1 pB::F35D2.3::yfp* and *Pitr-1 pB::dys-1::yfp*. These plasmids were then injected at 10 ng  $\mu\text{l}^{-1}$  or 40 ng  $\mu\text{l}^{-1}$  (for *dys-1*) with *Podr-1::dsRED* or *Podr-1::GFP* at 20 ng  $\mu\text{l}^{-1}$  into N2 worms.

## LETTERS

# STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling

Hiroki Ishikawa<sup>1</sup> & Glen N. Barber<sup>1</sup>

The cellular innate immune system is essential for recognizing pathogen infection and for establishing effective host defence. But critical molecular determinants responsible for facilitating an appropriate immune response—following infection with DNA and RNA viruses, for example—remain to be identified. Here we report the identification, following expression cloning, of a molecule (STING; stimulator of interferon genes) that appears essential for effective innate immune signalling processes. It comprises five putative transmembrane regions, predominantly resides in the endoplasmic reticulum and is able to activate both NF- $\kappa$ B and IRF3 transcription pathways to induce expression of type I interferon (IFN- $\alpha$  and IFN- $\beta$ ) and exert a potent anti-viral state following expression. In contrast, loss of STING rendered murine embryonic fibroblasts extremely susceptible to negative-stranded virus infection, including vesicular stomatitis virus. Further, STING ablation abrogated the ability of intracellular B-form DNA, as well as members of the herpesvirus family, to induce IFN- $\beta$ , but did not significantly affect the Toll-like receptor (TLR) pathway. Yeast two-hybrid and co-immunoprecipitation studies indicated that STING interacts with RIG-I and with SSR2 (also known as TRAP $\beta$ ), which is a member of the translocon-associated protein (TRAP) complex required for protein translocation across the endoplasmic reticulum membrane following translation<sup>1,2</sup>. Ablation by RNA interference of both TRAP $\beta$  and translocon adaptor SEC61 $\beta$  was subsequently found to inhibit STING's ability to stimulate expression of IFN- $\beta$ . Thus, as well as identifying a regulator of innate immune signalling, our results imply a potential role for the translocon in innate signalling pathways activated by select viruses as well as intracellular DNA.

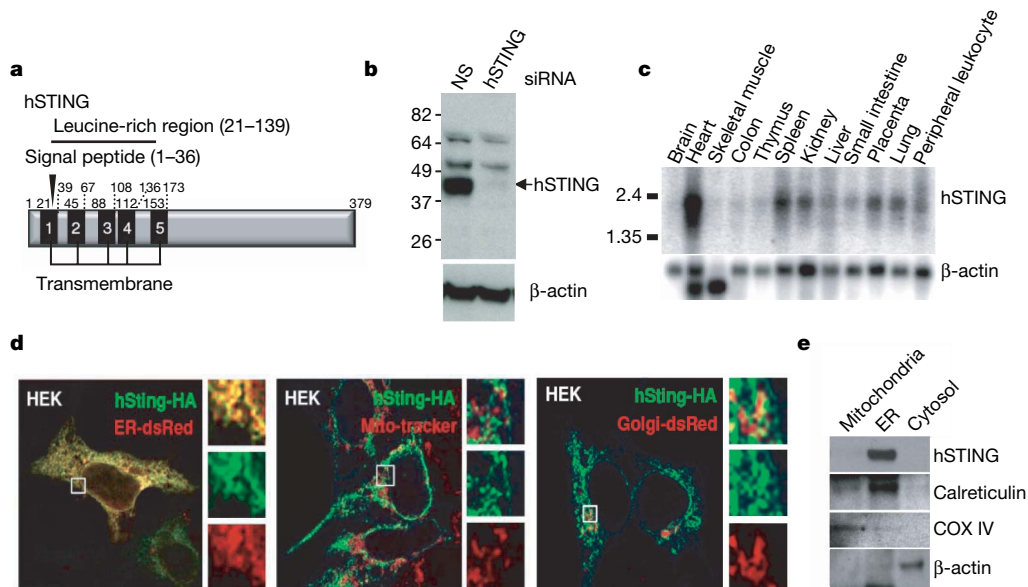
Cellular host defence responses to pathogen invasion principally involve the detection of pathogen associated molecular patterns (PAMPs), such as viral nucleic acid or bacterial cell wall components (including lipopolysaccharide or flagellar proteins), that results in the induction of anti-pathogen genes<sup>3–9</sup>. For example, viral RNA can be detected by membrane bound TLRs present in the endoplasmic reticulum (ER) and/or endosomes (for example, TLR3 and TLR7/8) or by TLR-independent intracellular DExD/H box RNA helicases, referred to as retinoic acid inducible gene 1 (RIG-I) or melanoma differentiation associated antigen 5 (MDA5, also referred to as IFIH1 and helicard)<sup>3–10</sup>. Pathogen DNA can be recognized by TLR9 present in plasmacytoid dendritic cells, although it is now apparent that important TLR-independent pathways also exist to recognize DNA in alternative tissue, the mechanisms of action of which remain to be determined<sup>3–10</sup>. These events culminate in the activation of downstream signalling events, leading to the transcription of NF- $\kappa$ B and IRF3/7-dependent genes, including type I IFN.

To further determine the mechanisms of innate immune signalling, we employed an expression screening system in which approximately

5,500 human and 9,000 murine full length complementary DNAs were individually transfected into 293T cells harbouring a luciferase gene under control of the IFN- $\beta$  promoter (IFN- $\beta$ -Luc). The top five hits whose overexpression lead to the significant induction of IFN- $\beta$ -Luc were found to be IPS-1 (also referred to as VISA/CARDIF/MAVS) (Supplementary Fig. 1)<sup>11–14</sup>. However, we also isolated a previously uncharacterized molecule (gi:38093659/NP\_938023/2610307O08RIK) which we referred to as STING that harboured five predicted transmembrane motifs (in humans) and existed as a 379 amino acid protein in human cells and 378 amino acids in murine cells (Fig. 1a and Supplementary Fig. 1). A putative signal cleavage motif was found to exist at position 1–36 and a leucine rich region was apparent between amino acids 21 and 139 (Fig. 1a). The predicted molecular weight of human STING was 42,192 Da, which approximately corresponded to its observed molecular weight in human 293 cells following immunoblot analysis using a rabbit antiserum raised to a STING peptide (Fig. 1b). RNAi studies confirmed that the observed 42 kDa band was indeed STING (Fig. 1b). STING was found to be ubiquitously expressed in a variety of tissues, as determined by northern analysis, and was found to predominantly reside in the ER region of the cell as determined by confocal microscopy and fractionation analysis (Fig. 1c–e).

Overexpression of STING in 293T cells was subsequently confirmed to robustly induce the expression of the IFN promoter (IFN- $\beta$ -Luc) by up to 400-fold, but not a control TK promoter driving luciferase (pRL-TK), interferon regulatory factor 3 (IRF3) responsive promoters (PRD-III-I-Luc) up to 1,000-fold, an NF- $\kappa$ B responsive promoter (NF- $\kappa$ B-Luc) 12-fold, and interferon-inducible promoters (interferon sensitive response element, ISRE-Luc) up to 800-fold (Fig. 2a–d). STING did not activate control promoters driving luciferase reporters such as those derived from the RB, p53 or E2F genes (Supplementary Fig. 2). Increased dimerization of IRF3 was also observed in STING expressing 293T cells, confirming that STING regulates the induction of type I IFN at locations upstream of IRF3 activation, dimerization and translocation (Supplementary Fig. 2)<sup>15</sup>. Endogenous *Irfn* mRNA and IFN- $\beta$  protein was induced significantly in murine embryonic fibroblasts (MEFs) transiently transfected with STING (Fig. 2e, f). DNA microarray analysis of STING expression in 293T cells further emphasized STING's ability to induce primary innate immune response genes (Fig. 2g). Accordingly, MEF cells expressing STING or IPS-1 were significantly resistant to vesicular stomatitis virus (VSV) infection (Fig. 2h, i). Subsequent analysis indicated that STING function was ablated in the absence of the  $\kappa$ B kinase family member TBK-1, confirming that STING's activity involved activation of IRF3 and was indeed upstream of this kinase (Fig. 2j)<sup>15</sup>. Finally, we observed that STING did not exert robust activity in the absence of FADD, which has also been shown to be important for efficient innate immune signalling processes (Supplementary Fig. 2)<sup>16</sup>.

<sup>1</sup>Department of Medicine and Sylvester Comprehensive Cancer Center, University of Miami School of Medicine, Miami, Florida 33136, USA.



**Figure 1 | STING is an ER protein.** **a**, Schematic of human STING (hSTING) indicating transmembrane and leucine rich regions. **b**, Immunoblot analysis of STING in HEK 293 cells treated with RNAi to STING (hSTING) or control RNAi (NS). **c**, Northern blot analysis of human STING and control  $\beta$ -actin. **d**, Confocal analysis of HEK 293 cells (HEK) transfected with human

STING tagged at the carboxyl end with HA. Transfected cells were also analysed using ER-dsRed, Mito Tracker Red or Golgi-dsRed. **e**, Fractionation experiments confirm that STING resides in the ER. Control antibodies indicate accuracy of fractionation (Calreticulin, ER; COX IV, mitochondria;  $\beta$ -actin, cytosol).

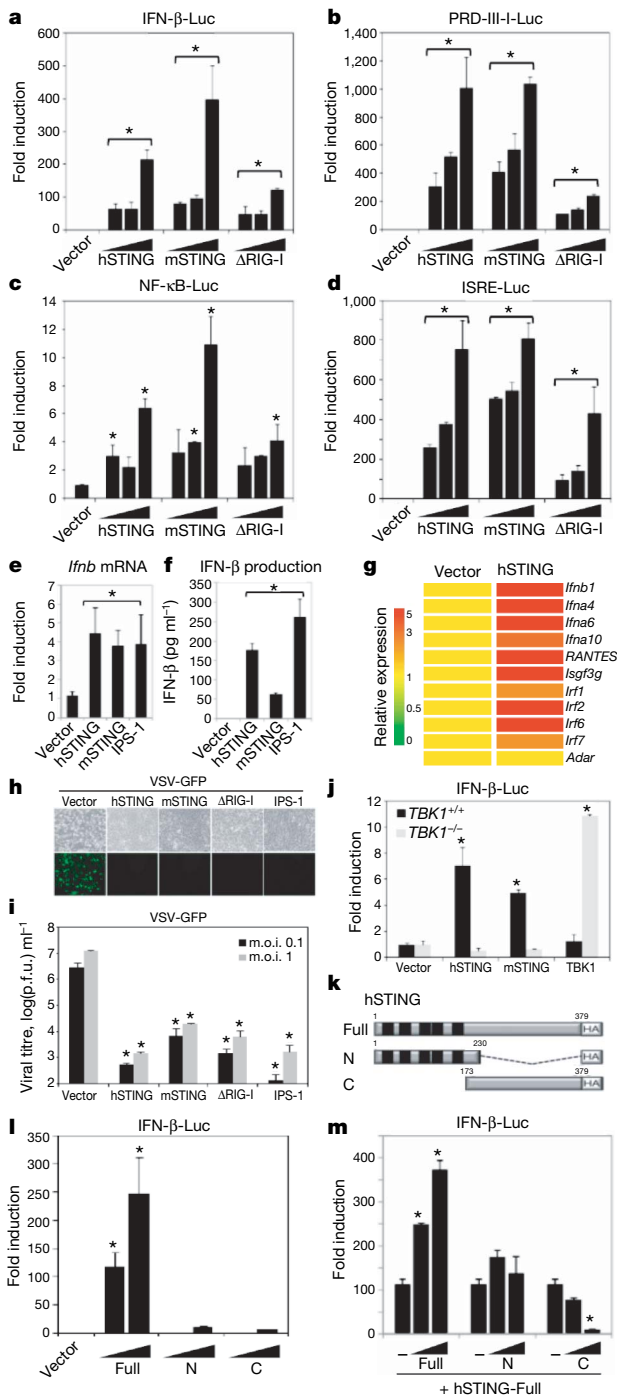
We also established that while STING that had been tagged with HA at the carboxyl region retained activity, this activity was lost when STING was tagged at the amino terminus or C terminus with GFP (data not shown). Subsequent analysis indicated that the N-terminal region of STING containing the five putative transmembrane regions (amino acids 1–230) or just the carboxyl region of STING (amino acids 173–379) did not exhibit significant ability, alone, to induce the IFN- $\beta$ -Luc promoter (Fig. 2k, l). Thus, full length, intact STING is required for efficient function. However, we further observed that the carboxyl region of STING exerted a dominant-negative inhibitory effect and could impede the ability of full-length STING to stimulate IFN- $\beta$ -Luc (Fig. 2m). Collectively, these data indicate that expression of STING activates the innate immune response, including type I IFN, leading to the induction of an antiviral state.

To further analyse STING function, we used an RNAi approach to ablate STING in a number of cell types. Our data indicated that knockdown of STING in HEK 293 cells modestly reduced the ability of the negative-stranded rhabdovirus VSV-GFP to induce IFN- $\beta$ , presumably because these viruses are only weak activators of IFN- $\beta$  (Supplementary Fig. 3). However, such cells were rendered extremely susceptible to virus infection and replication. To confirm a requirement for STING in the regulation of type I IFN induction and in host defence, we generated STING deficient (*Sting*<sup>-/-</sup>) mice by targeted homologous recombination in embryonic stem (ES) cells (Supplementary Fig. 4). *Sting*<sup>-/-</sup> animals were born at the Mendelian ratio, and developed and bred normally. Accordingly, MEFs from wild-type and *Sting*<sup>-/-</sup> animals were infected with VSV-GFP at varying multiplicity of infection (m.o.i., 0.01–1) for up to 36 h post infection. This study confirmed that more progeny virus was produced in MEFs lacking STING compared to controls (by a factor of 10<sup>2</sup>; 24–36 h) (Fig. 3a–d). These data were verified using VSV expressing a luciferase reporter gene and VSV- $\Delta$ M, which exhibits a defect in the viral matrix protein normally responsible for inhibiting cellular messenger RNA export from the nucleus<sup>17</sup> (Supplementary Fig. 4 and Fig. 3e). Reconstitution of STING to *Sting*<sup>-/-</sup> MEFs rescued the susceptibility to VSV infection (Supplementary Fig. 4). Similar analysis also indicated that loss of STING also reduced the ability of Sendai virus to induce IFN- $\beta$

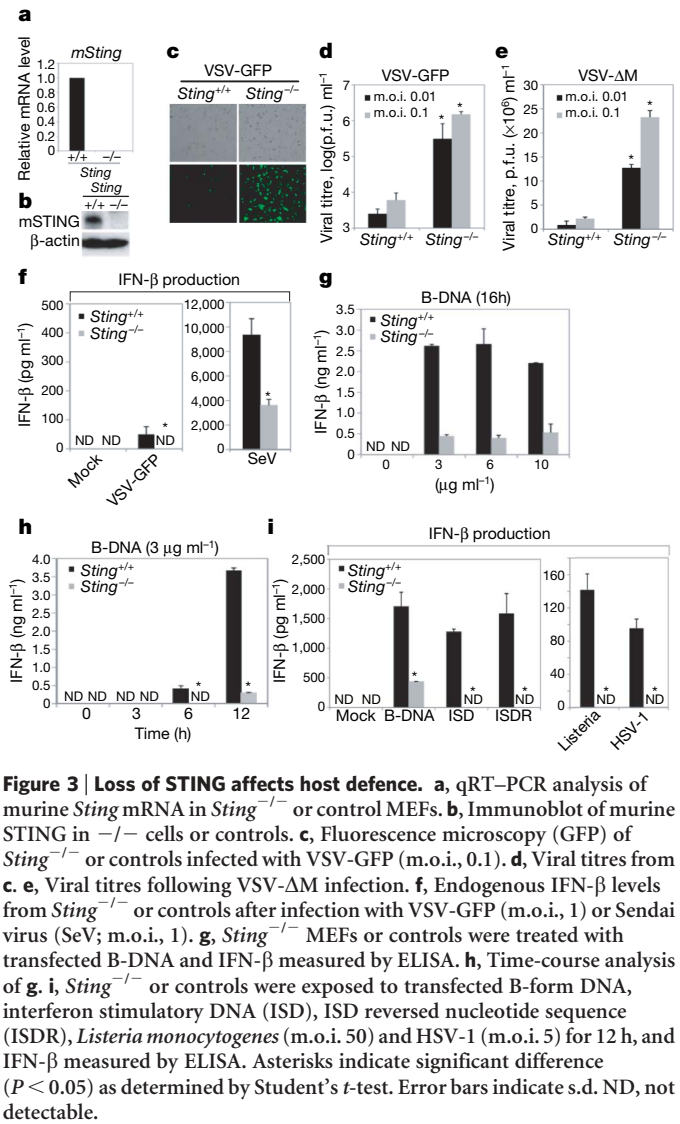
(Fig. 3f). In contrast, we did not observe a strong requirement for STING to mediate the ability of transfected poly I:C to induce IFN- $\beta$  induction, which is largely governed by the intracellular RIG-I homologue MDA5 (ref. 18) (Supplementary Fig. 5). We also observed that the positive-stranded encephalomyocarditis virus (EMCV), a member of the picornavirus family, did not effectively induce IFN- $\beta$  or replicate differently, regardless of the presence of STING (Supplementary Fig. 5).

Thus we conclude that STING may play a more predominant role in facilitating RIG-I mediated innate signalling, rather than MDA5. Interestingly, we did notice a significant defect (>5-fold) in the ability of transfected B-form DNA (poly dA-dT) to induce IFN- $\beta$  in MEFs lacking STING compared to controls (Fig. 3g, h). More strikingly, the non CpG containing interferon stimulatory DNA (ISD) was completely unable to induce IFN- $\beta$  in *Sting*<sup>-/-</sup> MEFs, as were the DNA virus herpes simplex virus 1 and bacteria *Listeria monocytogenes* (Fig. 3i). TLR9 is considered to govern CpG DNA-mediated induction of IFN- $\beta$ , but is not active in MEFs<sup>19,20</sup>. Thus, it is plausible that STING may function in TLR9-independent, DNA-mediated induction of type I IFN. This effect was similarly observed in murine STING-lacking bone marrow derived macrophages (BMDM) or bone marrow derived dendritic cells (GM-DC) cultured using granulocyte-macrophage colony stimulating factor (GM-CSF) (Supplementary Fig. 6). However, no significant difference was observed in the ability of exogenous poly I:C or lipopolysaccharide (LPS) to induce IFN- $\beta$ , when comparing *Sting*<sup>-/-</sup> BMDMs or GM-DCs to controls, events which depend on TLR3 and TLR4, respectively. Whereas loss of STING rendered MEFs highly susceptible to VSV-GFP, less susceptibility was observed following VSV-GFP infection of *Sting*<sup>-/-</sup> GM-DCs or BMDMs, indicating that STING may be more important in facilitating negative-stranded virus-mediated innate signalling in fibroblasts. Collectively, our data indicate that loss of STING leads to a defect in RIG-I mediated type I IFN induction but does not affect the TLR pathway. In addition, we report that STING functions in the pathway used by intracellular B-form DNA to induce IFN- $\beta$ .

To further examine the mechanisms of STING's function in innate signalling, we attempted to determine if STING interacted with RIG-I

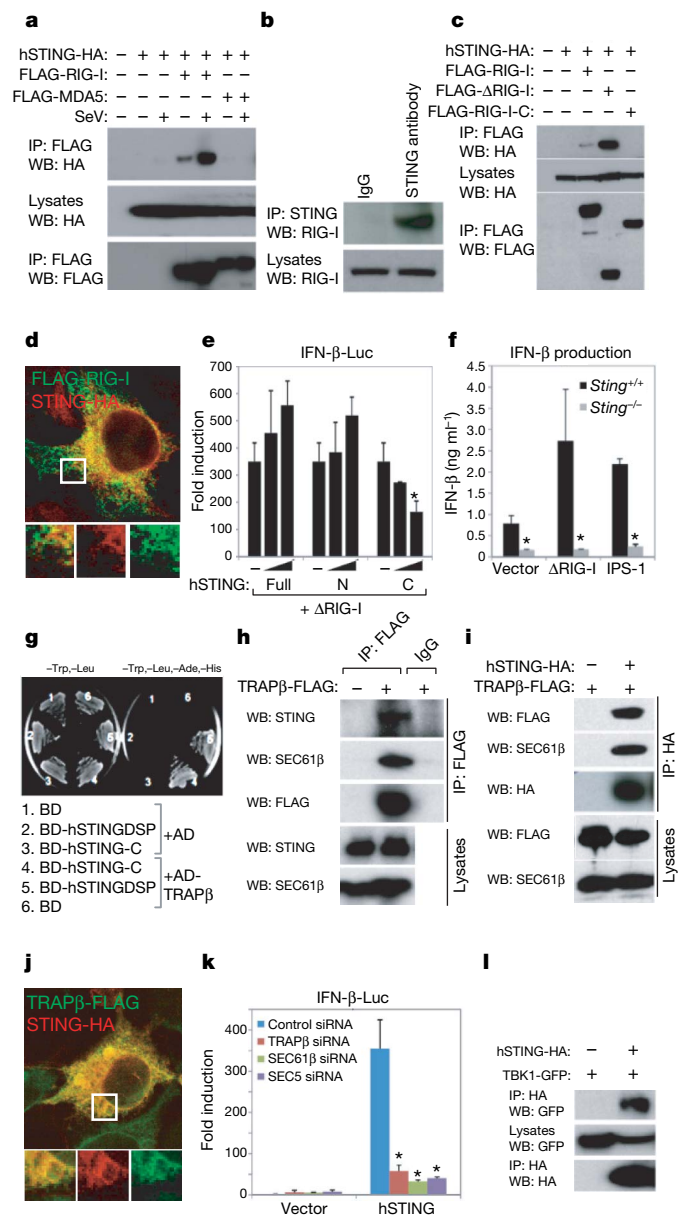


**Figure 2 | STING facilitates IFN induction.** **a**, 293T cells were transfected with an IFN- $\beta$ -Luc (p110-Luc) plasmid and increasing amounts of human STING (hSTING), murine STING (mSTING) or control  $\Delta$ RIG-I. **b–d**, 293T cells transfected as in **a** with either PRD-III-I-Luc (**b**), NF- $\kappa$ B-Luc (**c**) or ISRE-Luc (**d**) reporter plasmids were analysed similarly. **e**, MEFs were transfected as in **a**, and *Ifnb* mRNA was analysed by qRT-PCR. **f**, Medium from transfected MEFs was analysed for IFN- $\beta$  protein by ELISA. **g**, Microarray analysis of 293T cells transfected with hSTING. **h**, MEFs transfected with STING or  $\Delta$ RIG-I or IPS-1 are resistant to VSV-GFP infection (m.o.i., 1). Top panel is light microscopy, lower panel is fluorescent microscopy. **i**, Plaque assay from **h**. p.f.u., plaque-forming units. **j**, TBK-1 deficient MEFs do not facilitate STING signalling. **k**, Schematic of hSTING variants. **l**, 293T cells were transfected as in **a** with hSTING or variants, and luciferase measured. **m**, 293T cells were transfected with STING and increasing amounts of full length STING (hSTING-Full), hSTING-N or hSTING-C with luciferase plasmids as in **a**. Asterisks indicate significant difference ( $P < 0.05$ ) as determined by Student's *t*-test. Error bars indicate s.d.



**Figure 3 | Loss of STING affects host defence.** **a**, qRT-PCR analysis of murine *Sting* mRNA in *Sting*<sup>-/-</sup> or control MEFs. **b**, Immunoblot of murine STING in *Sting*<sup>-/-</sup> cells or controls. **c**, Fluorescence microscopy (GFP) of *Sting*<sup>-/-</sup> or controls infected with VSV-GFP (m.o.i., 0.1). **d**, Viral titres from **c**. **e**, Viral titres following VSV- $\Delta$ M infection. **f**, Endogenous IFN- $\beta$  levels from *Sting*<sup>-/-</sup> or controls after infection with VSV-GFP (m.o.i., 1) or Sendai virus (SeV; m.o.i., 1). **g**, *Sting*<sup>-/-</sup> MEFs or controls were treated with transfected B-DNA and IFN- $\beta$  measured by ELISA. **h**, Time-course analysis of **g**. **i**, *Sting*<sup>-/-</sup> or controls were exposed to transfected B-form DNA, interferon stimulatory DNA (ISD), ISD reversed nucleotide sequence (ISDR), *Listeria monocytogenes* (m.o.i. 50) and HSV-1 (m.o.i. 5) for 12 h, and IFN- $\beta$  measured by ELISA. Asterisks indicate significant difference ( $P < 0.05$ ) as determined by Student's *t*-test. Error bars indicate s.d. ND, not detectable.

and MDA5, which are putative innate immune signalling receptors for negative- or positive-stranded viral RNA, respectively<sup>3,4,10</sup>. Co-immunoprecipitation experiments in 293T cells principally indicated that FLAG-tagged RIG-I but not MDA5 could associate with HA-tagged STING in co-transfection experiments (Fig. 4a). The binding of RIG-I to STING was augmented upon infection of 293T cells with Sendai virus (Fig. 4a). We confirmed in normal human umbilical vein endothelial cells (HUVECs) that endogenous RIG-I could associate directly, or indirectly as a complex, with endogenous STING (Fig. 4b). This is in agreement with previous data indicating that STING seems to preferentially modulate the RIG-I, rather than the MDA5, regulated pathway (Fig. 3d and Supplementary Fig. 5). We also determined that the CARD domains of RIG-I (amino acids 1–284) were preferentially able to associate with HA-STING in transfected 293T cells (Fig. 4c). RIG-I was subsequently shown to colocalize with STING in the co-transfected 293T cells (Fig. 4d). It was similarly observed that the carboxyl region of STING could inhibit the function of RIG-I, again indicating that this region of STING can exhibit a dominant-inhibitory effect (Fig. 4e). We also noted association of the RIG-I downstream adaptor IPS-1 with STING although, similarly to the situation with RIG-I, it is not yet clear whether IPS-1 directly interacts with STING or exists as a complex with RIG-I/STING (Supplementary Fig. 7). Accordingly, we observed that STING was able to induce the expression of an IFN- $\beta$  driven luciferase construct in MEFs lacking RIG-I or IPS-1, probably confirming that that STING functions downstream of these latter molecules



**Figure 4 | STING associates with the translocon.** **a**, 293T cells were co-transfected with HA-STING, FLAG-RIG-I or MDA5 and infected with Sendai virus (SeV; m.o.i. 1). Lysates were immunoprecipitated (IP) and immunoblotted (IB) using antibodies to HA or FLAG. **b**, Endogenous human STING associates with RIG-I in HUVECs. **c**,  $\Delta$ RIG-I (amino acids 1–284) and not RIG-I-C (amino acids 218–925) associate with STING in co-transfected 293T cells. **d**, Confocal image of 293T cells co-transfected with tagged STING and RIG-I. **e**, 293T cells were co-transfected with control vector (–) or increasing amounts of full-length, amino (amino acids 1–230) or carboxyl (amino acids 173–379) STING and  $\Delta$ RIG-I, and IFN- $\beta$ -Luc was measured. **f**, Control or *Sting*<sup>–/–</sup> MEFs were transfected with  $\Delta$ RIG-I (amino acids 1–284) or IPS-1 and IFN- $\beta$  was measured by ELISA. **g**, GAL4 binding domain (BD) fused to the carboxyl region of hSTING, BD-hSTING-C, interacts with Ssr2/TRAP $\beta$  fused to the GAL4 activation domain (AD-hTRAP $\beta$ ) in yeast-two hybrid screening (BD-hSTING-GASP, amino acids 36–369; BD-hSTING-C, amino acids 173–379). **h**, HEK 293 cells were transfected with FLAG-tagged TRAP $\beta$  and endogenous STING or SEC61 $\beta$  measured by immunoblot. **i**, STING and TRAP $\beta$  were co-transfected HEK 293 cells and analysis carried out as in **h**. **j**, Co-localization of STING and TRAP $\beta$  in 293T cells. **k**, RNAi to TRAP $\beta$ , SEC61 $\beta$  or SEC5 in HEK 293 cells ablates STING signalling. **l**, HA-STING associates with GFP-TBK-1 in co-transfected HEK 293 cells. Asterisks indicate significant difference ( $P < 0.05$ ) as determined by Student's *t*-test. Error bars indicate s.d.

(Supplementary Fig. 8). The ability of  $\Delta$ RIG-I or IPS-1 to induce IFN- $\beta$  appeared diminished in *Sting*<sup>–/–</sup> MEFs (Fig. 4f). However, there was also a marked reduction in the induction of IFN- $\beta$  by all transfected plasmids (including vector alone) in the absence of STING compared to control MEFs, probably because endogenous DNA innate signalling pathways are defective (Fig. 3g–i). These data indicate that STING may be an important downstream adaptor molecule that facilitates RIG-I and perhaps IPS-1 function.

To gain further insight into the molecular mechanisms of the action of STING, we screened an IFN-induced, human fibroblast yeast two-hybrid cDNA library using STING (amino acids 173–379) as a bait and repeatedly isolated SSR2/TRAP $\beta$ , a member of the TRAP complex comprising four subunits ( $\alpha$ – $\Delta$ ) that facilitates translocation of proteins into the ER following translation (Fig. 4g)<sup>1,2</sup>. The TRAP complex is known to associate with the translocon, comprising three subunits, SEC61 $\alpha$ , SEC61 $\beta$  and SEC61 $\gamma$ . Given this information, we confirmed that TRAP $\beta$  can indeed associate with endogenous STING in HEK 293 cells following co-immunoprecipitation experiments (Fig. 4h). Using this approach, we confirmed that TRAP $\beta$  also co-immunoprecipitated with endogenous SEC61 $\beta$  (Fig. 4h). We next verified that STING could also associate not only with TRAP $\beta$  but also SEC61 $\beta$ , probably as a complex (Fig. 4i). STING was also observed to colocalize with TRAP $\beta$  in the ER region of the cell (Fig. 4j). Indeed, loss of TRAP $\beta$  or SEC61 $\beta$  reduced STING's ability to induce an IFN- $\beta$  promoter driving luciferase (Fig. 4k; Supplementary Fig. 9). Taken together, these data indicate that the STING may be involved in translocon function, and that the translocon may be able to influence the induction of type I IFN.

Thus, STING is predominantly an ER resident protein that may link RIG-I and DNA-mediated intracellular innate signalling to the translocon. We speculate that RIG-I may detect translating viral RNAs at the intersection of ribosome/ER translocon association and require STING to exert effective function<sup>2</sup>. Alternatively, STING may participate in mediating ER stress response pathways, but this remains to be verified. Although it is not clear how signalling from the translocon to IRF3/NF- $\kappa$ B occurs, it has recently been established that the translocon physically associates with the exocyst—the octameric Sec6–Sec8 complex that also associates with the ER and tethers secretory vesicles to membranes, and facilitates protein synthesis and secretion<sup>21,22</sup>. Recently, the exocyst complex was found to recruit and activate TBK1 and play a role in type I IFN- $\beta$  induction<sup>23</sup>. Our preliminary analysis indicates that STING also co-immunoprecipitates with TBK1, and that RNAi ablation of Sec5 also rendered cells defective in the STING function (Fig. 4k, l). Thus STING may facilitate the detection of intracellular viral RNA species as well as B-form DNA, indicating convergence of these intracellular PAMP recognition pathways.

## METHODS SUMMARY

**Plasmid constructs.** Human STING (hSTING), murine STING (mSTING), hSTING-N (amino acids 1–230) and hSTING-C (amino acids 173–379) sequences were amplified by PCR and were cloned into pcDNA3 plasmids to generate C-terminally HA-tagged expression constructs. Expression plasmids encoding Flag-tagged RIG-I,  $\Delta$ RIG-I (amino acids 1–284), IPS-1 and TBK-1 were described previously<sup>24</sup>. GFP-tagged RIG-I and TBK-1 were generated by cloning into pAcGFP1-C1 (Clontech). Other plasmids were obtained as follows: p110-Luc and PRD-III-I-Luc (T. Maniatis), IFN- $\beta$ -Luc and ISRE-Luc (J. Hiscott), NF $\kappa$ B-Luc (Stratagene), ER-dsRED and Golgi-dsRED (Clontech).

**Antibodies.** Rabbit polyclonal antibody against a synthetic peptide corresponding to residues 324–340 of human STING was obtained from EvoQuest Custom Antibody Services (Invitrogen). Other antibodies were obtained from following sources: SEC61 $\beta$  (Upstate),  $\beta$ -Actin, HA, FLAG (Sigma), COX IV, Calreticulin, IRF3 and RIG-I (Abcam).

**Confocal microscopy.** ER-dsRED and Golgi-dsRED (Clontech) were used for ER and Golgi marker, respectively. For mitochondria staining, living cells were incubated with 300 nM of Mito Tracker Red (Invitrogen) for 45 min at 37 °C.

**RNA interference.** Chemically synthesized 21-nucleotide siRNA duplexes were obtained from Dharmacon. RNA oligonucleotides used for human and murine STING were as follows: hSTING, GCAUCAAGGAUCGGGUUU; mSTING, CCAACAGCGUCUACGAGA.

**Generation of the *Sting* knockout mice.** The linearized targeting vector was electroporated into E14.1 ES cells originated from 129SvEv strain, followed by the selection in G418. One positive clone was injected into C57BL/6J blastocysts and *Sting*<sup>-/-</sup> mice generated on a C57BL/6J background.

**Yeast two-hybrid analyses.** To screen for interacting partners of STING, a yeast two-hybrid approach was used, adopting a novel IFN-induced, telomerase immortalized human fibroblast cDNA library, generated by our laboratory. hSTING-C (amino acids 173–379) sequence was amplified by PCR and was cloned into pGBK-T7 (Clontech) to generate a bait plasmid pGBK-T7-hSTING-C. 100 µg of the IFN-induced library's plasmid DNA was used to perform a library-scale transformation of yeast strain Y187.

**Statistics.** Students *t*-test was used to analyse data.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 7 July; accepted 6 August 2008.**

**Published online 24 August 2008.**

- Hartmann, E. *et al.* A tetrameric complex of membrane proteins in the endoplasmic reticulum. *Eur. J. Biochem.* **214**, 375–381 (1993).
- Menetret, J. F. *et al.* Architecture of the ribosome-channel complex derived from native membranes. *J. Mol. Biol.* **348**, 445–457 (2005).
- Takeuchi, O. & Akira, S. Recognition of viruses by innate immunity. *Immunol. Rev.* **220**, 214–224 (2007).
- Beutler, B. *et al.* Genetic analysis of resistance to viral infection. *Nature Rev. Immunol.* **7**, 753–766 (2007).
- Takahashi, K. *et al.* Nonspecific RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses. *Mol. Cell* **29**, 428–440 (2008).
- Pichlmair, A. *et al.* RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* **314**, 997–1001 (2006).
- Hornung, V. *et al.* 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**, 994–997 (2006).
- Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunol.* **5**, 730–737 (2004).
- Loo, Y. M. *et al.* Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. *J. Virol.* **82**, 335–345 (2008).
- Onomoto, K., Yoneyama, M. & Fujita, T. Regulation of antiviral innate immune responses by RIG-I family of RNA helicases. *Curr. Top. Microbiol. Immunol.* **316**, 193–205 (2007).
- Kawai, T. *et al.* IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nature Immunol.* **6**, 981–988 (2005).
- Meylan, E. *et al.* Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* **437**, 1167–1172 (2005).
- Seth, R. B. *et al.* Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF- $\kappa$ B and IRF 3. *Cell* **122**, 669–682 (2005).
- Xu, L. G. *et al.* VISA is an adapter protein required for virus-triggered IFN- $\beta$  signaling. *Mol. Cell* **19**, 727–740 (2005).
- McWhirter, S. M. *et al.* IFN-regulatory factor 3-dependent gene expression is defective in Tbk1-deficient mouse embryonic fibroblasts. *Proc. Natl Acad. Sci. USA* **101**, 233–238 (2004).
- Balachandran, S., Thomas, E. & Barber, G. N. A FADD-dependent innate immune mechanism in mammalian cells. *Nature* **432**, 401–405 (2004).
- Faria, P. A. *et al.* VSV disrupts the Rae1/mrnp41 mRNA nuclear export pathway. *Mol. Cell* **17**, 93–102 (2005).
- Kato, H. *et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**, 101–105 (2006).
- Ishii, K. J. *et al.* A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nature Immunol.* **7**, 40–48 (2006).
- Stetson, D. B. & Medzhitov, R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* **24**, 93–103 (2006).
- Guo, W. & Novick, P. The exocyst meets the translocon: A regulatory circuit for secretion and protein synthesis? *Trends Cell Biol.* **14**, 61–63 (2004).
- Lipschutz, J. H., Lingappa, V. R. & Mostov, K. E. The exocyst affects protein synthesis by acting on the translocation machinery of the endoplasmic reticulum. *J. Biol. Chem.* **278**, 20954–20960 (2003).
- Chien, Y. *et al.* RalB GTPase-mediated activation of the I $\kappa$ B kinase TBK1 couples innate immune signaling to tumor cell survival. *Cell* **127**, 157–170 (2006).
- Balachandran, S. *et al.* Fas-associated death domain-containing protein-mediated antiviral innate immune signaling involves the regulation of Irf7. *J. Immunol.* **178**, 2429–2439 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank T. Venkataraman, J. Hyun, T. Sato and M. Conkright for technical assistance, M. Gale for RIG-I and IPS-1 lacking MEFs, Y. C. Weh for TBK-1 lacking MEFs, and S. Nagata, T. Maniatis, J. Hiscott and N. Reich for plasmid constructs.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to G.N.B. ([gbarber@med.miami.edu](mailto:gbarber@med.miami.edu)).

## METHODS

**Cells, viruses and reagents.** 293T and HEK293 cells were obtained from the ATCC and were maintained in DMEM medium supplemented with 10% FBS. Human umbilical vein endothelial cells (HUVECs) were obtained from Lonza, and were maintained in endothelial growth medium, EGM-2 (Lonza). *Fadd*<sup>+/+</sup>, *Fadd*<sup>-/-</sup>, *Tbk-1*<sup>+/+</sup> and *Tbk-1*<sup>-/-</sup> MEFs were provided by W.-C. Yeh<sup>16</sup>. VSV-GFP was used in infections and titred as described<sup>16</sup>. VSV-ΔM was constructed as described<sup>17</sup>. Murine IFN-β ELISA Kit was obtained from PBL. EMCV was purchased from ATCC. Synthetic ds B-DNA (poly dA-dT) and poly I:C were purchased from GE Healthcare. Interferon stimulatory DNA (ISD) and reverse sequence of ISD were described previously<sup>20</sup>. For stimulation of cells, B-DNAs or poly I:C were mixed with Lipofectamine 2000 (Invitrogen) at a ratio of 1:1 (vol/wt), and then added to cells at a final concentration of 3 μg ml<sup>-1</sup>.

**Reporter analysis.** 293T cells or MEFs seeded on 24-well plates were transiently transfected with 50 ng of the luciferase reporter plasmid together with a total of 250 ng of various expression plasmids or empty control plasmids. As an internal control, 10 ng of pRL-TK was transfected simultaneously. Then, 36 h later, the luciferase activity in the total cell lysates was measured.

**Northern blot.** Human multiple tissue RNA blots (Clontech) were hybridized with a <sup>32</sup>P-labelled full-length human STING probe.

**Real-time PCR.** Total RNA was isolated from cells using the RNeasy RNA extraction kit (Qiagen) and cDNA synthesis was performed using 1 μg of total RNA (Roche). Fluorescence real-time PCR analysis was performed using a LightCycler 2.0 instrument (Roche Molecular Biochemicals) and TaqMan Gene Expression Assays (Applied Biosystems). Relative amounts of mRNA were normalized to the 18S ribosomal RNA levels in each sample.

**DNA microarray analysis.** Total RNA was extracted from 293T cells transfected with STING expressing vectors. Preparation of cDNA and microarray analysis was performed at the W.M. Keck Foundation Biotechnology Research Laboratory DNA microarray facility at Yale University. The Human Genome U133 Plus 2.0 Array (Affymetrix) was used. Data analysis was performed with GeneSpring software (Silicon Genetics)<sup>16</sup>.

**Mitochondria and ER fraction isolation.** Mitochondria and ER membranes were purified on discontinuous sucrose gradients as previously described<sup>25</sup>. Briefly, HUVEC cells in MTE buffer (0.27 M mannitol, 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.4) were lysed by sonication. Lysed cells were centrifuged at 700g for 10 min to remove nuclei and cellular debris. Mitochondria were obtained by centrifugation at 15,000g for 10 min, and post mitochondrial supernatant was used for purification of ER fractions. Mitochondria pellet was resuspended in MTE buffer, and was layered on discontinuous sucrose gradients consisting of 1.0 M and 1.7 M sucrose and banded by centrifugation at 40,000g for 22 min. Mitochondria fraction was collected and pelleted by centrifugation at 15,000g for 10 min. Purified mitochondria were resuspended in PBS and used for western blot analysis. To isolate ER fractions, postmitochondrial supernatant described above was layered on discontinuous sucrose gradients consisting of 1.3 M, 1.5 M and 2.0 M sucrose, and banded by centrifugation at 100,000g for 70 min. The ER fraction at the interface between the supernatant and the 1.3 M sucrose was collected, and pelleted by centrifugation at 100,000g for 45 min. The ER membranes were resuspended in PBS and were used for western blot analysis.

**RNA interference.** 93T cells were plated on 24-well plates at 5 × 10<sup>4</sup> cells per well and transfected with 10 pmol of siRNA duplex per well using Lipofectamine RNAiMAX (Invitrogen). MEFs were transfected by using an Amaxa nucleofactor apparatus (program A-023) and Amaxa MEF nucleofactor kit 1 according to the manufacturer's recommendations. At 72 h after transfection, cells were used for further experiments.

**Primers.** The following primers were used for cloning: hSTING forward, 5'-CCCAAGCTTGGCCGCCACCATGCCACTCCAGCCTGC-3'; hSTING reverse,

5'-CCCAAGCTTGGCCGCCACCATGCCACTCCACCTGCATCC-3'; mSTING forward, CCGCTCGAGAGAGAAATCCGTGCGGAGAG; mSTING reverse, CCGC TCGAGGATGAGGTCA.

The sequences of each siRNA oligonucleotide used in this study are follows: hSTING siRNA, 5'-GCAUCAAGGAUCGGGUUU-3'; mSTING siRNA, 5'-CCAACAGCGUCUACGAGA-3'; SSR2 siRNA, 5'-UCUCAAGGCUUGU AUUU-3', 5'-GUGGAACUAUCUGAUGAU-3', 5'-CAUCUACAAUGUUG GCUC-3', 5'-AAACGAAGAAGAACUGAU-3'; hSEC61B siRNA, 5'-CAGUA UUGGUUAUGAGUC-3'; 5'-GUUCGUAGAUUCAGUUAC-3', 5'-GCUCA AAGUUGGCCUGU-3', 5'-CUGUAAGCUUGCGUUUU-3'; SEC5 siRNA, 5'-CGUCACACCUUCCUAAAU-3', 5'-CGGCAUCUCUCCAAAUGA-3', 5'-CAACAGGUGUCAGAAACU-3', 5'-GAAAGGCGGUCUCAGUAC.

**Co-immunoprecipitation, native PAGE and immunoblot analysis.** Cells were seeded on 100-mm dishes at 1 × 10<sup>6</sup> cells per dish. Cells were transfected with a total of 10 μg of empty plasmid or various expression plasmids using Lipofectamine 2000. At 36 h after transfection, cells were lysed in M-PER buffer (Pierce) or 1% digitonin (Calbiochem) buffer (20 mM Tris-HCl, 150 mM NaCl, 1% digitonin) containing protease inhibitors (Roche). Lysates were incubated with HA affinity matrix (Covance) or FLAG affinity beads (Sigma) overnight. The beads were washed three times by TBS containing 0.05% Tween-20 and immunoprecipitates were eluted with non-reducing sample buffer by boiling for 5 min. For endogenous STING immunoprecipitation, HUVECs were lysed in 1% digitonin lysis buffer. Cleared supernatants were incubated with 10 μg of anti-STING antibody, followed by incubation with immobilized protein G (Pierce). The beads were washed four times by 1% digitonin lysis buffer and immunoprecipitates were eluted with SDS sample buffer by boiling for 5 min. Native PAGE and immunoblotting was carried out as previously described<sup>24</sup>.

**Generation of the *Sting* knockout mice.** The linearized targeting vector was electroporated into E14.1 ES cells originated from 129SvEv strain, followed by the selection in G418. Targeted clones were screened by PCR. From 90 clones, 1 positive clone was identified. This ES clone was subjected to the generation of chimaera mice by injection using C57BL/6J blastocysts as the host. The resulting male chimaeras were further mated with C57BL/6J female mice for germline transmission. The heterozygous mice (F<sub>1</sub> mice) were interbred to obtain wild-type, heterozygous and homozygous littermates (F<sub>2</sub>). The genotypes of the mice were determined by PCR. Animals were bred at the University of Miami School of Medicine Transgenic Core Facility. Mice were allowed to freely access to food and water and housed at an ambient temperature of 23 °C and at a 12 h light/dark cycle. Animal care and handling was performed as per IACUC guidelines.

**IFN induced library construction.** To screen for interacting partners of STING, a yeast two hybrid approach was adopted using a novel IFN-induced, telomerase immortalized human fibroblast cDNA library. To construct this library, cells were treated overnight with 1,000 units each of human IFN-α and -β to induce interferon dependent genes. Poly(A)<sup>+</sup> RNA was extracted and cDNA synthesis was carried out using BD Powerscript RT (Clontech). The cDNA was then amplified by PCR, digested with *Sfi*I and ligated to the yeast prey vector pGADT7-RecAB. The ligated mixture was transformed into *Escherichia coli* strain DH10B. The number of independent clones in the unamplified library was estimated to be 3.2 × 10<sup>6</sup>, with an average size of 1.52 kb and inserts ranging in size from 0.8 to 3.0 kb. 15 independent colonies were isolated and DNA extracted to check for size by PCR.

25. Mavinakere, M. S. *et al.* Processing of human cytomegalovirus UL37 mutant glycoproteins in the endoplasmic reticulum lumen prior to mitochondrial importation. *J. Virol.* **80**, 6771–6783 (2006).

# Modelling Myc inhibition as a cancer therapy

Laura Soucek<sup>1</sup>, Jonathan Whitfield<sup>1</sup>, Carla P. Martins<sup>1</sup>, Andrew J. Finch<sup>1</sup>, Daniel J. Murphy<sup>1</sup>, Nicole M. Sodir<sup>1</sup>, Anthony N. Karnezis<sup>1</sup>, Lamorna Brown Swigart<sup>1</sup>, Sergio Nasi<sup>2</sup> & Gerard I. Evan<sup>1</sup>

**Myc is a pleiotropic basic helix–loop–helix leucine zipper transcription factor that coordinates expression of the diverse intracellular and extracellular programs that together are necessary for growth and expansion of somatic cells<sup>1</sup>. In principle, this makes inhibition of Myc an attractive pharmacological approach for treating diverse types of cancer. However, enthusiasm has been muted by lack of direct evidence that Myc inhibition would be therapeutically efficacious, concerns that it would induce serious side effects by inhibiting proliferation of normal tissues, and practical difficulties in designing Myc inhibitory drugs. We have modelled genetically both the therapeutic impact and the side effects of systemic Myc inhibition in a preclinical mouse model of Ras-induced lung adenocarcinoma by reversible, systemic expression of a dominant-interfering Myc mutant. We show that Myc inhibition triggers rapid regression of incipient and established lung tumours, defining an unexpected role for endogenous Myc function in the maintenance of Ras-dependent tumours *in vivo*. Systemic Myc inhibition also exerts profound effects on normal regenerating tissues. However, these effects are well tolerated over extended periods and rapidly and completely reversible. Our data demonstrate the feasibility of targeting Myc, a common downstream conduit for many oncogenic signals, as an effective, efficient and tumour-specific cancer therapy.**

Myc is deregulated and overexpressed in most cancer cells, where it hijacks the diverse intracellular and extracellular regenerative programs that drive normal cell expansion. Consistent with this, de-activation of Myc in established, Myc-induced transgenic tumours triggers proliferative arrest and re-differentiation of tumour cells, and collapse of the tumour microenvironment and vasculature, usually resulting in rapid tumour regression<sup>2–7</sup>. Although this indicates that Myc might be a good therapeutic target, there are caveats. First, Myc exerts its biological influence through protein–protein and protein–DNA interactions that have proven difficult to disrupt with small molecules. Second, aberrant Myc expression in most human cancers is usually not due to mutation in the *Myc* gene itself but a consequence of its induction by ‘upstream’ oncogenic signals. The therapeutic utility of inhibiting Myc when its aberrant expression is a consequence, not a cause, of oncogenesis is unclear. Finally, Myc is essential for proliferation and stem cell compartment maintenance of regenerative adult tissues such as the gastrointestinal tract, skin and bone marrow. Hence, blocking Myc function systemically might trigger devastating and irreversible side effects. Together, such concerns greatly undermine the credibility of Myc inhibition as an anti-cancer strategy.

Myc-dependent transactivation requires heterodimerization with its bHLHZip partner protein Max<sup>8–10</sup>. Dimerization with Max is also essential for Myc proliferative and oncogenic functions<sup>8,11</sup> and its inhibition has been shown to have potential therapeutic value<sup>12,13</sup>. To model both the therapeutic impact and side effects of Myc inhibition *in vivo*, we constructed a mouse in which endogenous Myc

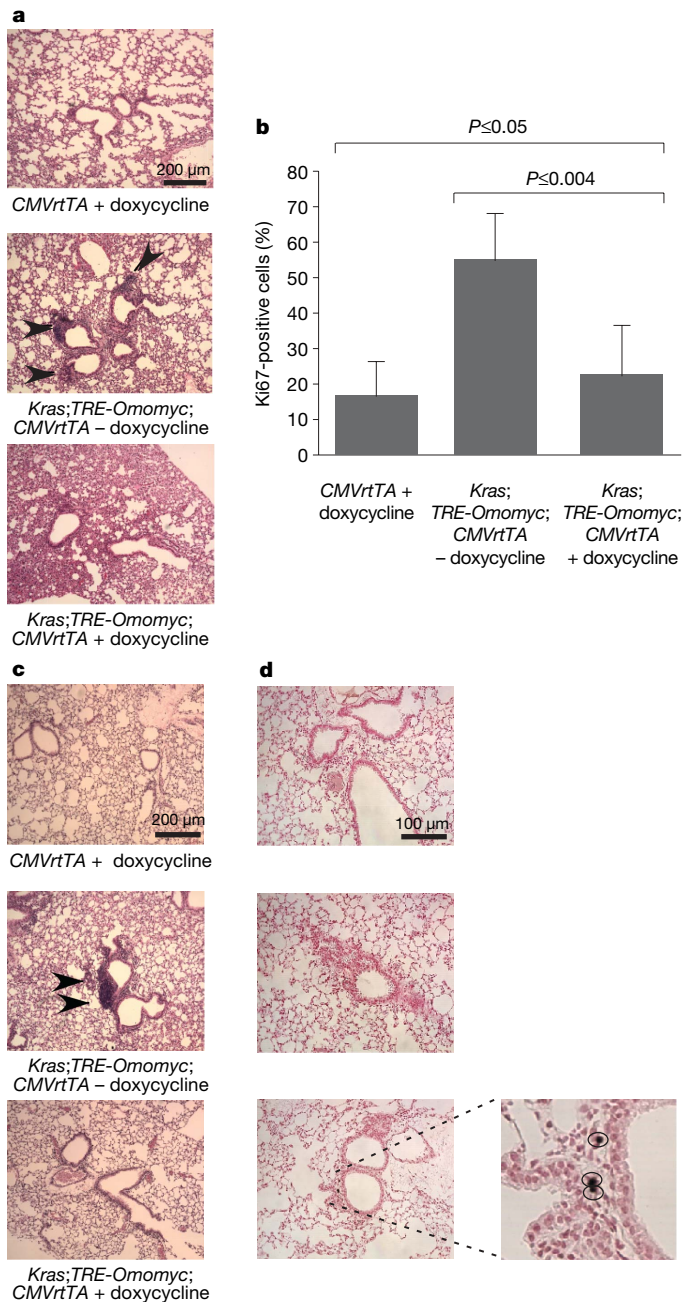
function may be systemically and reversibly inhibited in tissues of adult animals through inducible expression of the dominant interfering Myc bHLHZip dimerization domain mutant Omomyc. Omomyc has four designed amino acid substitutions that facilitate homodimerization with all three oncogenic Myc proteins (c-Myc, N-Myc and L-Myc) but result in little, if any, interaction with Mad proteins. Omomyc dimerization denies Max access to Myc<sup>14</sup> and, because Myc–Omomyc heterodimers cannot bind Myc–Max E-box consensus recognition elements, it efficiently blocks Myc-dependent transcriptional activation<sup>15</sup>. Omomyc expression reverses Myc-induced transformation *in vitro*<sup>15</sup> and Myc-driven tumorigenesis *in vivo*<sup>16</sup>.

We conditionally expressed Omomyc using the highly promiscuous cytomegalovirus (CMV) early promoter, which drives relatively high levels of expression in multiple tissue types<sup>17–19</sup>. The *Omomyc* coding sequence was placed downstream of a tetracycline-responsive promoter element (*TRE*) and mice harbouring the *TRE-Omomyc* transgene were then crossed into a background expressing the rtTA transactivator from the CMV early promoter<sup>20</sup>. Omomyc expression was then induced in *TRE-Omomyc;CMVrtTA* double transgenic animals by administration of doxycycline in their drinking water. Real-time polymerase chain reaction with reverse transcription (RT–PCR) confirmed that *Omomyc* expression was detectable in all tested tissues of *TRE-Omomyc;CMVrtTA* double transgenic mice on doxycycline treatment, and was absent from both *TRE-Omomyc* and *CMVrtTA* single transgenic mice, as well as *TRE-Omomyc;CMVrtTA* double transgenic mice not treated with doxycycline (data not shown). Despite some variation between different tissues, steady-state *Omomyc* mRNA levels were demonstrably higher than those of endogenous *c-myc* in most tissues (intestine, kidney, pancreas, heart and lung) and comparable in the rest (for example, spleen and skin) (Supplementary Fig. 1).

To ascertain the therapeutic potential of inhibiting endogenous Myc function in cancer, we chose a well-characterized and validated lung cancer mouse model in which tumorigenesis is initiated by oncogenic activation of endogenous *Kras*<sup>21,22</sup>, a common mutation in human non-small-cell lung cancers. *Kras* is activated in this *LSL-Kras<sup>G12D</sup>* model by inhalation of adenovirus expressing Cre recombinase, which sporadically excises a transcriptional stop element in bronchioalveolar duct junction (BADJ) epithelial cells, triggering expression of oncogenic *Kras<sup>G12D</sup>* driven from the endogenous *Kras* promoter<sup>22</sup>. Tumorigenesis involves development of hyperplasia/adenoma 4–6 weeks after *Kras* activation followed by sporadic progression of some lesions to adenocarcinoma. The resulting mouse lung tumours closely resemble their human counterparts<sup>22,23</sup>. To assess the contribution of endogenous Myc to the initiation, early and late phases of *Kras*-induced lung tumorigenesis, we crossed switchable *TRE-Omomyc;CMVrtTA* mice into the heterozygous *LSL-Kras<sup>G12D</sup>* lung cancer model. To ascertain whether inhibition of Myc prevents initiation of *Kras<sup>G12D</sup>*-induced lung tumours,

<sup>1</sup>Department of Pathology and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94143-0875, USA. <sup>2</sup>Istituto di Biologia e Patologia Molecolari, C.N.R., University La Sapienza, 00185 Rome, Italy.

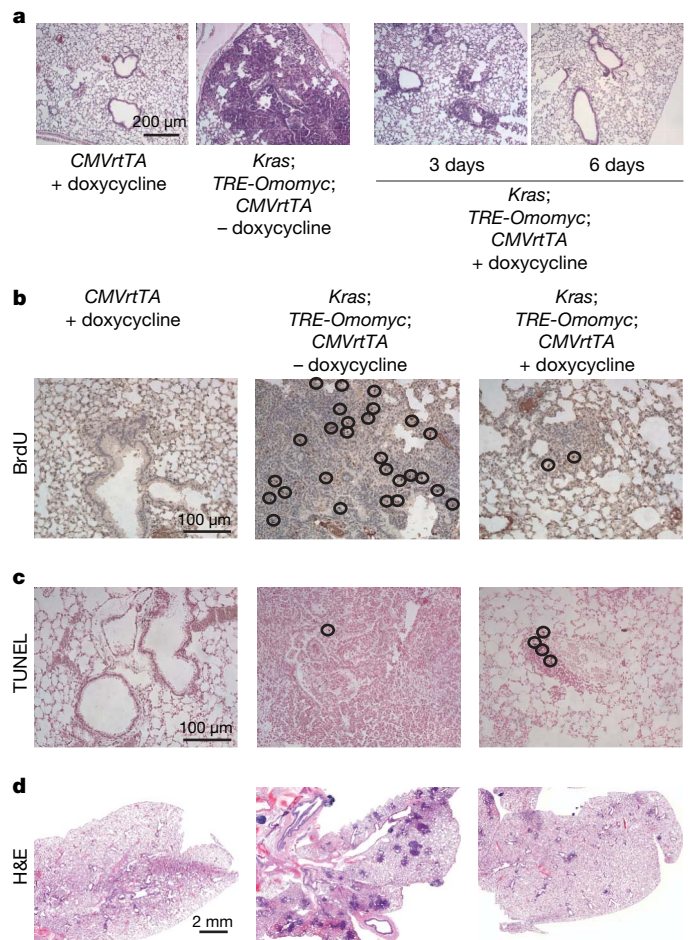




**Figure 1 | Endogenous Myc function is required for formation and maintenance of early-stage Kras-induced lung hyperplasias/adenomas.** **a**, Representative haematoxylin-and-eosin-stained sections of lungs from CMVrtTA single transgenic mice (CMVrtTA + doxycycline), untreated Kras; TRE-Omomyc; CMVrtTA triple transgenic mice (Kras; TRE-Omomyc; CMVrtTA - doxycycline) and doxycycline-treated LSL-Kras; TRE-Omomyc; CMVrtTA triple transgenic mice (Kras; TRE-Omomyc; CMVrtTA + doxycycline), 4 weeks after infection with adenoviral Cre. Hyperplastic lesions are indicated by black arrowheads. **b**, Graphical representation of total BADJ cells scored as proliferating (Ki67-positive). Error bars represent standard deviation derived from approximately 100 BADJs per mouse. At least three mice were used per series. **c**, Myc inhibition triggers regression of early-stage lung adenomas. Haematoxylin and eosin staining of lungs from mice treated or not with doxycycline for 1 week, starting 6 weeks after Cre-recombinase-expressing adenovirus infection, is shown. A small adenoma is indicated by black arrowheads. **d**, TUNEL staining reveals positive cells in Omomyc and Ras co-expressing samples (Kras; TRE-Omomyc; CMVrtTA + doxycycline), but not in untreated or single transgenic controls (Kras; TRE-Omomyc; CMVrtTA - doxycycline and CMVrtTA + doxycycline). Higher magnification of positive cells is shown in the insert on the right.

680

$5 \times 10^7$  plaque-forming units (p.f.u.) of Cre-recombinase-expressing adenovirus was administered intranasally to LSL-Kras; TRE-Omomyc; CMVrtTA mice and Omomyc induced after 24 h and thereafter maintained for 4 weeks. Mice were then killed and lung morphology, cell proliferation and apoptosis assessed immunohistochemically. For simplicity, only representative controls are shown: doxycycline-treated CMVrtTA mice exemplify all strains lacking an activatable Kras<sup>G12D</sup> allele, whereas untreated LSL-Kras; TRE-Omomyc; CMVrtTA mice represent all strains of mice with activated Kras<sup>G12D</sup> but no Omomyc. All mice expressing Kras<sup>G12D</sup> without Omomyc (Fig. 1a, LSL-Kras; TRE-Omomyc; CMVrtTA - doxycycline) developed multiple hyperplasias and atypical



**Figure 2 | Myc inhibition elicits regression of established lung tumours.** **a**, Haematoxylin-and-eosin-stained lungs from control CMVrtTA single transgenic mice (CMVrtTA + doxycycline), Kras<sup>G12D</sup>-expressing mice (Kras; TRE-Omomyc; CMVrtTA - doxycycline) and 3- or 6-day-treated Kras<sup>G12D</sup> and Omomyc co-expressing mice (Kras; TRE-Omomyc; CMVrtTA + doxycycline) 18 weeks after infection with adenovirus expressing Cre recombinase. An example of a frank tumour is shown from a mouse expressing Kras<sup>G12D</sup> only (Kras; TRE-Omomyc; CMVrtTA - doxycycline). **b**, BrdU staining shows a considerable reduction in BrdU-positive cells in Kras<sup>G12D</sup> and Omomyc co-expressing mice (Kras; TRE-Omomyc; CMVrtTA + doxycycline) compared with tissues from mice expressing Kras<sup>G12D</sup> only (Kras; TRE-Omomyc; CMVrtTA - doxycycline). **c**, TUNEL staining indicates the presence of apoptotic cells in Omomyc and Kras<sup>G12D</sup> co-expressing sections (Kras; TRE-Omomyc; CMVrtTA + doxycycline), but not in untreated or single transgenic control tissues (Kras; TRE-Omomyc; CMVrtTA - doxycycline and CMVrtTA + doxycycline). **d**, Haematoxylin and eosin staining of lung tissue from mice treated for 4 weeks with doxycycline shows clearance of tumour lesions as a consequence of Omomyc expression (compare Kras; TRE-Omomyc; CMVrtTA + doxycycline with Kras; TRE-Omomyc; CMVrtTA - doxycycline).

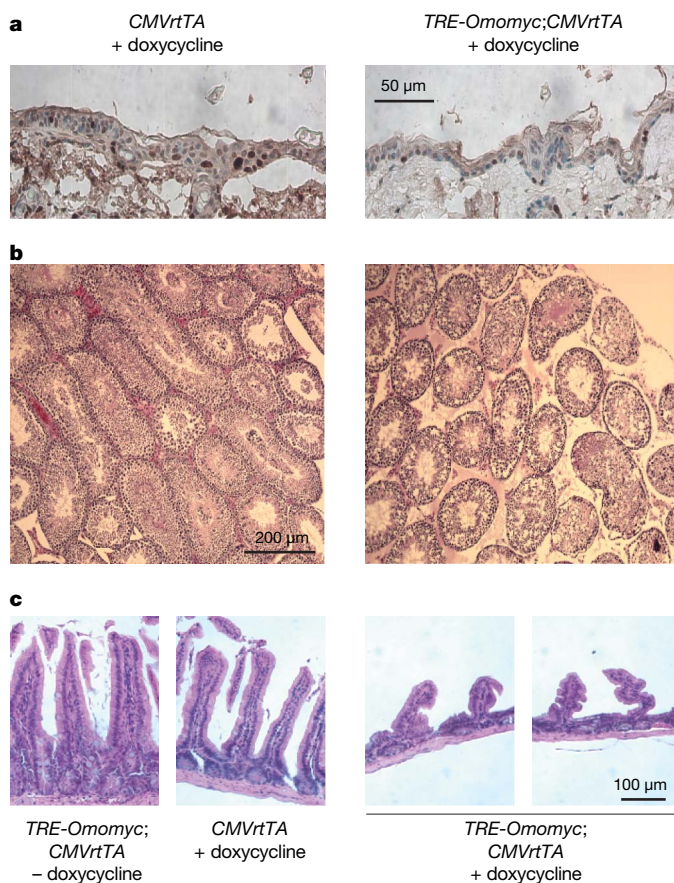
adenomatous hyperplasias (present in  $33 \pm 5.2\%$  of total scored BADJs). By contrast, lesions were absent from all mice expressing *Kras*<sup>G12D</sup> together with Omomyc (*LSL-Kras;TRE-Omomyc;CMVrtTA* + doxycycline). Overall, Omomyc expression significantly suppressed the fraction of cells proliferating (Ki67-positive) at the BADJs (Fig. 1b), consistent with the notion that endogenous Myc is required for proliferation of *Kras*<sup>G12D</sup>-induced early-stage tumour cells. No apoptotic (TdT-mediated dUTP nick end labelling (TUNEL)-positive) cells were present in the BADJs of lungs from mice in any of the test or control groups (data not shown).

To assess the impact of blocking endogenous Myc on maintenance of early-stage Ras-induced lung adenomas, *Kras*<sup>G12D</sup> was activated in *LSL-Kras;TRE-Omomyc;CMVrtTA* mice and nascent tumours allowed to develop for 6 weeks. Representative mice from cohorts of all strains were then killed and examined histologically. All mice harbouring activated *Kras*<sup>G12D</sup> exhibited multiple bronchiolar hyperplasias and adenomatous hyperplasias (present in  $48.7 \pm 6\%$  of total BADJs). Doxycycline was then administered to one-half of the remaining mice for 1 week. Myc inhibition elicited marked shrinkage of lesions and reduction in their multiplicity (identifiable lesions were observed only in  $1.7 \pm 0.8\%$  of BADJs) (Fig. 1c). Consistent with the involution of adenomas, we noted significant apoptosis in the residual lesions of mice expressing Omomyc ( $5.7 \pm 2.5\%$  of total cells). Apoptosis was negligible in control *LSL-Kras;TRE-Omomyc;CMVrtTA* mice in which Omomyc expression had not been induced (0.001%; Fig. 1d). As in the above prevention study, Myc inhibition profoundly reduced the overall numbers of proliferating cells at the BADJs (data not shown).

The induction of apoptosis by Myc inhibition in early-stage *Kras*-induced lung adenomas suggests that, in addition to proliferation, endogenous Myc is required to maintain survival of early-stage *Kras*<sup>G12D</sup>-induced lung tumours. To ascertain whether the same holds true for 'advanced' *Kras*-induced tumours, *Kras*<sup>G12D</sup> was sporadically activated in lungs of *LSL-Kras;TRE-Omomyc;CMVrtTA* mice and tumours allowed to evolve for 18 weeks. At this time, all representative animals exhibited many neoplastic lesions (present in  $89 \pm 6.2\%$  of total BADJs), with some large and highly vascularized adenocarcinomas, as described<sup>22</sup>. Endogenous Myc function was then inhibited in one-half of each cohort and groups of three mice each killed 3, 6 and 28 days later. After only 3 days of sustained Myc inhibition, histological analysis of lungs revealed marked shrinkage of tumours (Fig. 2a). By 6 days we noted a significant reduction in the proliferation of bronchioalveolar epithelial cells ( $1.7 \pm 0.7\%$  5-bromodeoxyuridine (BrdU)-positive cells in BADJs of mice co-expressing Omomyc and *Kras*<sup>G12D</sup> versus  $4.8 \pm 2.2\%$  BrdU-positive cells in mice expressing activated *Kras*<sup>G12D</sup> alone, and  $0.9 \pm 0.3\%$  BrdU-positive cells in *Kras*<sup>G12D</sup>-negative mice; Fig. 2b) accompanied by a significant increase in expression of the senescence marker SA- $\beta$ -galactosidase (data not shown). In addition, we noted significant apoptosis in regressing tumours of Omomyc-expressing mice ( $5.8 \pm 1.3\%$  TUNEL-positive cells versus  $\sim 0.001\%$  in Omomyc negative controls; Fig. 2c). By 28 days of sustained Myc blockade, lungs of animals appeared to be free of obvious tumours with only scattered foci of atypical cells at the BADJs (Fig. 2d). Moreover, during this extended period of Omomyc expression we did not detect any occasional tumours that failed to regress and/or continued to grow, even though each animal developed multiple (hundreds) discrete tumours. Thus, endogenous Myc function is continuously required for proliferation and survival of incipient, early and advanced *Kras*<sup>G12D</sup>-driven lung tumours *in vivo*. We do not yet know if Myc inhibition eradicates all the tumours. However, experiments are underway to monitor for relapse and, should some or all re-grow, for the susceptibility of resurgent tumours to further rounds of Myc inhibition.

Myc has an essential role in the proliferation of all normal cells, raising the possibility that systemic Myc inhibition would trigger severe side effects, especially in continuously proliferating tissues.

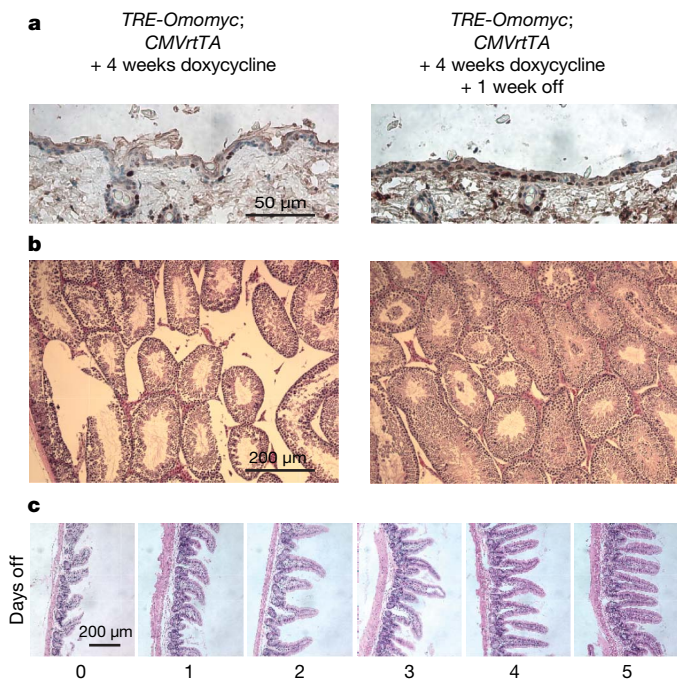
Therefore, to explore the impact of systemic Myc inhibition on normal tissues, Omomyc-inducible double transgenic (*TRE-Omomyc;CMVrtTA*) mice were treated continuously with doxycycline for 4 weeks and monitored throughout. Over this extended period, general health and activity of doxycycline-treated *TRE-Omomyc;CMVrtTA* mice appeared indistinguishable from that of non-Omomyc-expressing controls (Omomyc-negative *TRE-Omomyc* and *CMVrtTA* single transgenic mice), with no significant changes in body weight, general activity or blood chemistry (Supplementary Fig. 2 and Supplementary Table 3). After 4 weeks, animals were killed and organs examined. We observed no discernible histological changes in any adult organs with low proliferative indices (for example, pancreas, kidney, liver, heart or lung) (data not shown). However, tissues exhibiting rapid turnover were significantly affected. The basal layer of skin epidermis exhibited marked reduction of proliferating cells ( $12.6 \pm 3.7\%$  of total cells in Omomyc-expressing skin versus  $28.5 \pm 5.5\%$  of basal layer cells in controls) and thinning of the epidermis (Fig. 3a). Furthermore, hair re-growth after shaving was completely inhibited (Supplementary Fig. 3). Nonetheless, Myc inhibition induced no apoptosis, aberrant differentiation or loss of tissue integrity and mice exhibited no signs of skin tearing, ulceration or evident discomfort, even after 8 weeks of sustained Omomyc expression. In testis, Myc inhibition induced



**Figure 3 | Inhibition of endogenous Myc suppresses proliferation in skin, testis and GI tract. a**, Anti-Ki67-stained sections of epidermis from 4-week doxycycline-treated *TRE-Omomyc;CMVrtTA* mice and *CMVrtTA* single transgenic controls. **b**, Anti-Ki67-stained sections of testis from 4-week doxycycline-treated *TRE-Omomyc;CMVrtTA* mice show reduced proliferation in seminiferous tubules of doxycycline-treated *TRE-Omomyc;CMVrtTA* mice compared with Omomyc-negative controls. **c**, Haematoxylin-and-eosin-stained small intestine sections show blunted villi in *TRE-Omomyc;CMVrtTA* mice treated for 4 weeks with doxycycline (*TRE-Omomyc;CMVrtTA* + doxycycline) compared with intestine from untreated *TRE-Omomyc;CMVrtTA* and doxycycline-treated *CMVrtTA* single transgenic (*CMVrtTA* + doxycycline) controls.

dramatic atrophy marked by loss of spermatogonia and spermatocytes (Fig. 3b) and some inhibition of cell proliferation ( $32.7 \pm 1.8\%$  of total luminal cells in cycle in Omomyc-expressing mice versus  $42.3\% \pm 3.7\%$  Ki67-positive cells in control mice). We also observed a significant reduction in proliferation in the intestinal crypts (not shown) and striking attrition of villi in the small intestine (Fig. 3c). As with skin, Myc inhibition induced no increase in apoptosis or perturbation of differentiation, and integrity of the epithelium was maintained. Remarkably, all animals retained normal weight, hydration and blood chemistry, confirming maintenance of adequate intestinal absorption and barrier function against bacterial incursion (Supplementary Fig. 2 and Supplementary Table 1).

Despite its widespread activity in most adult tissues, the CMV early promoter functions only weakly in many haematopoietic lineages<sup>20</sup>, precluding its use in modelling the collateral impact of Myc inhibition on bone marrow and lymphoid function. We therefore made use of a complementary model in which doxycycline-dependent expression of Omomyc is driven from the  $\beta$ -actin promoter, which is ubiquitously active in haematopoietic lineages and the stem cell compartments of most tissues, including skin and bone marrow<sup>24–27</sup>. Induction of Omomyc in *TRE-Omomyc;  $\beta$ -actin-rtTA* mice recapitulated the Omomyc phenotypes in skin, gut and testis (Supplementary Fig. 4 and data not shown) and induced profound inhibition of proliferation in bone marrow (50% reduction of BrdU incorporation after 1 week) plus rapid onset of anaemia and leucopenia (Supplementary Table 2). However, by 2 weeks of sustained Myc inhibition, blood counts had returned close to normal levels, aside from mild polycythaemia (Supplementary Table 2). This recovery coincided with onset of extramedullary haematopoiesis in the spleen,



**Figure 4 | The degenerative phenotypes induced by systemic Myc inhibition are rapidly and completely reversible on restoration of Myc function.** Myc transactivation function was blocked for 4 weeks by sustained Omomyc expression. Doxycycline was then withdrawn for 1 week. **a**, Ki67 staining indicates rapid recovery of cell proliferation in skin. **b**, Rapid recovery of spermatogenesis in seminiferous tubules. Haematoxylin-and-eosin-stained sections from Omomyc-expressing testis (left panel) and the same tissue 1 week after discontinuing Omomyc expression (right panel) are shown. **c**, Rapid recovery of intestinal villus architecture in *TRE-Omomyc;CMVrtTA* mice after restoration of Myc function. After 4 weeks of sustained Omomyc expression, doxycycline treatment was discontinued and cohorts of mice killed each day for 5 days after doxycycline withdrawal. Representative haematoxylin-and-eosin-stained sections are shown for each time point.

primarily in the erythroid and megakaryocytic lineages (Supplementary Fig. 4). Throughout, all animals showed no overt signs of ill health or distress.

Myc is crucial for the maintenance of stem cell compartments and the balance between self-renewal and differentiation in multiple tissues, including the skin and intestinal crypts<sup>28,29</sup>. This raised the issue of whether even transient Myc inhibition might trigger irreversible changes in sensitive tissues, undercutting the therapeutic practicability of Myc inhibition therapy. Therefore, to assess the reversibility of Omomyc-induced tissue attrition, *TRE-Omomyc;CMVrtTA* mice were treated with doxycycline for 4 weeks. Doxycycline was then withdrawn to restore endogenous Myc function, and the status of affected tissues was followed over time. Within only 1 week of doxycycline withdrawal, cell proliferation had returned to pre-Omomyc levels in the basal layer of skin epithelium (Fig. 4a), hair re-growth re-initiated, and thereafter mice revealed no discernible deficits in any aspect of skin biology (data not shown). Likewise, testis rapidly recovered its full complement of spermatogenic cells including mature sperm (Fig. 4b). In the small intestine, villus length was restored within only 4–5 days (Fig. 4c). Long-term (~1 yr) observation of such mice has revealed no discernible pathology in any tissue. Thus, although systemic suppression of Myc activity has profound effects on proliferating somatic tissues, these effects are well tolerated, without any discernible negative impact on animal well-being and, moreover, they are completely reversible.

Why Myc inhibition should trigger viable arrest in normal tissues but apoptosis in tumours is unclear. Perhaps sustained flux through endogenous Myc is needed to support the increased metabolic and biosynthetic demands made by oncogenic Ras signalling. Endogenous Myc may, like its oncogenic counterpart, also be essential for maintenance of the tumour microenvironment<sup>30</sup>. Whatever the underlying mechanism, our studies suggest that pharmacological inhibition of Myc offers both specificity and efficacy in the treatment of neoplastic disease.

## METHODS SUMMARY

**Transgene construction and generation of mice.** *Omomyc* cDNA was cloned into the pTRE2 plasmid (Clontech). The *TRE-Omomyc* transgene was microinjected into the male pronucleus of day-1-fertilized (*CBA*  $\times$  *C57BL/6*) F<sub>1</sub> embryos, which were transferred into day-1-plugged pseudopregnant foster mice. *CMVrtTA* mice were obtained from the Jackson Laboratory (Tg(*rtTA*hCMV)4Bjd/J). *LSL-Kras<sup>G12D</sup>* mice were gifts from T. Jacks and the  *$\beta$ -actin-rtTA+* mice from S. Artandi. Excision of the STOP element was triggered by infection with a Cre-recombinase-expressing adenovirus ( $5 \times 10^7$  p.f.u.)<sup>22</sup>. Doxycycline ( $2 \text{ mg ml}^{-1}$  plus 5% sucrose) was added to drinking water. All mice were treated in accordance with protocols approved by the Institutional Animal Care and Use Committee at UCSF (IACUC approval number AN076148).

**Histology and immunohistochemistry.** Tissues were stained with haematoxylin and eosin. For antigen retrieval, sections were boiled for 1 min in 0.01 M citrate buffer (pH 6.0). Primary antibody (anti-Ki67 (Neomarkers)) was applied in blocking buffer for 2 h. Incorporated BrdU and apoptotic cells were identified with the BrdU Detection kit II (Roche) and Apoptag kit (Chemicon). HRP-conjugated secondary antibodies (Dako and Molecular Probes) were applied for 30 min and visualized with DAB (Vector Laboratories). Cell proliferation was quantified by scoring Ki67-positive cells as a proportion of total cells at the BADJs in lung tissue (approximately 100 BADJs per mouse), as a percentage of total basal cells in skin (at least 600 total cells per mouse), and as a proportion of total cells in each seminiferous tubule in testis (approximately 80 tubules per section). At least three mice were used for every experimental condition and genotype (Supplementary Table 3).

**Determination of mRNA levels using real-time PCR.** Real-time quantitative RT-PCR (TaqMan) was used to quantify mRNA levels. Total RNA was isolated using Trizol reagent (Gibco). Total RNA ( $1 \mu\text{g}$ ) was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad). The *Gus* gene was used as an internal amplification control.

Received 1 March; accepted 11 July 2008.

Published online 17 August 2008.

- Oster, S. K., Ho, C. S., Soucie, E. L. & Penn, L. Z. The myc oncogene: Marvelously Complex. *Adv. Cancer Res.* **84**, 81–154 (2002).

2. Arvanitis, C. & Felsher, D. W. Conditionally MYC: insights from novel transgenic models. *Cancer Lett.* **226**, 95–99 (2005).
3. Felsher, D. W. & Bishop, J. M. Reversible tumorigenesis by MYC in hematopoietic lineages. *Mol. Cell* **4**, 199–207 (1999).
4. Flores, I. *et al.* Defining the temporal requirements for Myc in the progression and maintenance of skin neoplasia. *Oncogene* **23**, 5923–5930 (2004).
5. Jain, M. *et al.* Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science* **297**, 102–104 (2002).
6. Pelengaris, S., Khan, M. & Evan, G. I. Suppression of Myc-induced apoptosis in beta cells exposes multiple oncogenic properties of Myc and triggers carcinogenic progression. *Cell* **109**, 321–334 (2002).
7. Pelengaris, S. *et al.* Reversible activation of c-Myc in skin: induction of a complex neoplastic phenotype by a single oncogenic lesion. *Mol. Cell* **3**, 565–577 (1999).
8. Amati, B., Littlewood, T. D., Evan, G. I. & Land, H. The c-Myc protein induces cell cycle progression and apoptosis through dimerization with Max. *EMBO J.* **12**, 5083–5087 (1993).
9. Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B. & Burley, S. K. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature* **363**, 38–45 (1993).
10. Nair, S. K. & Burley, S. K. Structural aspects of interactions within the Myc/Max/Mad network. *Curr. Top. Microbiol. Immunol.* **302**, 123–143 (2006).
11. Amati, B. *et al.* Oncogenic activity of the c-Myc protein requires dimerisation with Max. *Cell* **72**, 233–245 (1993).
12. Chen, J. *et al.* Effects of the MYC oncogene antagonist, MAD, on proliferation, cell cycling and the malignant phenotype of human brain tumour cells. *Nature Med.* **1**, 638–643 (1995).
13. Prochownik, E. V. c-Myc as a therapeutic target in cancer. *Expert Rev. Anticancer Ther.* **4**, 289–302 (2004).
14. Soucek, L. *et al.* Design and properties of a Myc derivative that efficiently homodimerizes. *Oncogene* **17**, 2463–2472 (1998).
15. Soucek, L. *et al.* Omomyc, a potential Myc dominant negative, enhances Myc-induced apoptosis. *Cancer Res.* **62**, 3507–3510 (2002).
16. Soucek, L., Nasi, S. & Evan, G. I. Omomyc expression in skin prevents Myc-induced papillomatosis. *Cell Death Differ.* **11**, 1038–1045 (2004).
17. Baskar, J. F. *et al.* The enhancer domain of the human cytomegalovirus major immediate-early promoter determines cell type-specific expression in transgenic mice. *J. Virol.* **70**, 3207–3214 (1996).
18. Furth, P. A. *et al.* The variability in activity of the universally expressed human cytomegalovirus immediate early gene 1 enhancer/promoter in transgenic mice. *Nucleic Acids Res.* **19**, 6205–6208 (1991).
19. Kothary, R. *et al.* Unusual cell specific expression of a major human cytomegalovirus immediate early gene promoter-lacZ hybrid gene in transgenic mouse embryos. *Mech. Dev.* **35**, 25–31 (1991).
20. Zhan, Y., Brady, J. L., Johnston, A. M. & Lew, A. M. Predominant transgene expression in exocrine pancreas directed by the CMV promoter. *DNA Cell Biol.* **19**, 639–645 (2000).
21. Jackson, E. L. *et al.* The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Res.* **65**, 10280–10288 (2005).
22. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
23. Sweet-Cordero, A. *et al.* An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. *Nature Genet.* **37**, 48–55 (2005).
24. Okabe, M. *et al.* 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* **407**, 313–319 (1997).
25. Sarin, K. Y. *et al.* Conditional telomerase induction causes proliferation of hair follicle stem cells. *Nature* **436**, 1048–1052 (2005).
26. Sawamura, D. *et al.* Promoter/enhancer cassettes for keratinocyte gene therapy. *J. Invest. Dermatol.* **112**, 828–830 (1999).
27. Wright, D. E. *et al.* Cyclophosphamide/granulocyte colony-stimulating factor causes selective mobilization of bone marrow hematopoietic stem cells into the blood after M phase of the cell cycle. *Blood* **97**, 2278–2285 (2001).
28. Korinek, V. *et al.* Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature Genet.* **19**, 379–383 (1998).
29. Murphy, M. J., Wilson, A. & Trumpp, A. More than just proliferation: Myc function in stem cells. *Trends Cell Biol.* **15**, 128–137 (2005).
30. Evan, G. I. Can't kick that oncogene habit. *Cancer Cell* **10**, 345–347 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank T. Jacks and S. Artandi for their gifts of *LSL-Kras<sup>G12D</sup>* and *β-actin-rtTA* + mice, respectively. We thank F. Rostker for technical assistance and Y. Yaron and L. Johnson for advice on the *LSL-Kras<sup>G12D</sup>* model and adenovirus inhalation. We thank our laboratory colleagues for their comments and feedback. This study was supported by grant 2R01 CA98018 from the National Cancer Institute (to G.I.E.). S.N. acknowledges support from AIRC, ASI, CNR, MIUR FIRB and FIRS. C.P.M. is a Leukemia and Lymphoma Society Fellow. J.W. acknowledges support from Human Frontier Science Program. This paper is dedicated to the memory of Judah Folkman.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to G.I.E. ([gevan@cc.ucsf.edu](mailto:gevan@cc.ucsf.edu)).

## LETTERS

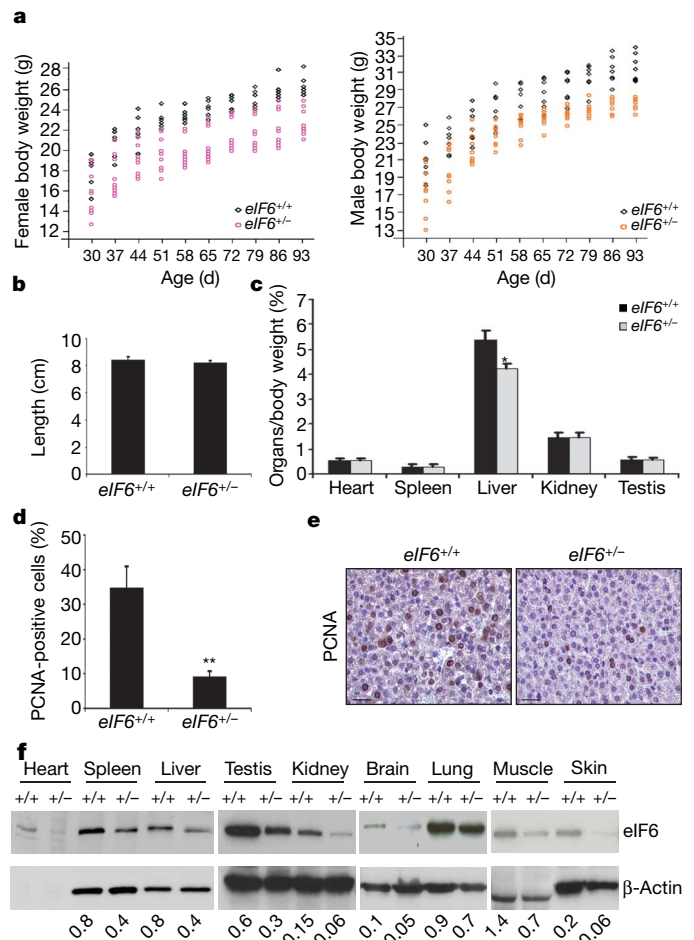
# Eukaryotic initiation factor 6 is rate-limiting in translation, growth and transformation

Valentina Gandin<sup>1</sup>, Annarita Miluzio<sup>1</sup>, Anna Maria Barbieri<sup>1</sup>, Anne Beugnet<sup>1</sup>, Hiroaki Kiyokawa<sup>2\*</sup>, Pier Carlo Marchisio<sup>1,3\*</sup> & Stefano Biffo<sup>1,4</sup>

Cell growth and proliferation require coordinated ribosomal biogenesis and translation. Eukaryotic initiation factors (eIFs) control translation at the rate-limiting step of initiation<sup>1,2</sup>. So far, only two eIFs connect extracellular stimuli to global translation rates<sup>3</sup>: eIF4E acts in the eIF4F complex and regulates binding of capped messenger RNA to 40S subunits, downstream of growth factors<sup>4,5</sup>, and eIF2 controls loading of the ternary complex on the 40S subunit and is inhibited on stress stimuli<sup>6,7</sup>. No eIFs have been found to link extracellular stimuli to the activity of the large 60S ribosomal subunit. eIF6 binds 60S ribosomes precluding ribosome joining *in vitro*<sup>8–10</sup>. However, studies in yeasts showed that eIF6 is required for ribosome biogenesis rather than translation<sup>11–14</sup>. Here we show that mammalian eIF6 is required for efficient initiation of translation, *in vivo*. *eIF6* null embryos are lethal at preimplantation. Heterozygous mice have 50% reduction of eIF6 levels in all tissues, and show reduced mass of hepatic and adipose tissues due to a lower number of cells and to impaired G1/S cell cycle progression. *eIF6*<sup>+/-</sup> cells retain sufficient nucleolar eIF6 and normal ribosome biogenesis. The liver of *eIF6*<sup>+/-</sup> mice displays an increase of 80S in polysomal profiles, indicating a defect in initiation of translation. Consistently, isolated hepatocytes have impaired insulin-stimulated translation. Heterozygous mouse embryonic fibroblasts recapitulate the organism phenotype and have normal ribosome biogenesis, reduced insulin-stimulated translation, and delayed G1/S phase progression. Furthermore, *eIF6*<sup>+/-</sup> cells are resistant to oncogene-induced transformation. Thus, eIF6 is the first eIF associated with the large 60S subunit that regulates translation in response to extracellular signals.

The *eIF6* gene was deleted by homologous recombination using embryonic stem cell technology (Supplementary Fig. 1). The portion of the gene containing the first two exons and the first two introns was substituted by a cassette containing the neomycin resistance gene. The presence of the neomycin resistance cassette did not affect expression of wild-type *eIF6* and of adjacent genes (Supplementary Fig. 2). Germ-line transmission was achieved, and intercrossing of *eIF6* heterozygous mice did not produce any *eIF6*<sup>-/-</sup> offspring. We did not observe *eIF6*<sup>-/-</sup> mice as early as 3.5 days post coitum, indicating that eIF6 is essential for proper development at the stage of preimplantation (Supplementary Table 1). The lethality of *eIF6*<sup>-/-</sup> embryos is consistent with the early expression of the protein in the blastocysts (Supplementary Fig. 1d).

Heterozygous *eIF6*<sup>+/-</sup> mice were viable and indistinguishable from wild-type counterparts up to 30 days after birth. At three months of age, heterozygous mice, independently from gender and genetic background, weighed less than their wild-type littermates (Fig. 1a). The head–anus length of *eIF6*<sup>+/-</sup> and wild-type mice was



**Figure 1 | eIF6 reduction leads to diminished body weight and affects liver growth at the proliferation level.** Age-matched mice were weighed and the growth of organs was followed after birth. **a**, Body weight of females and males recorded from 30 to 93 days after birth. **b**, Length of mice at 93 days, expressed as mean  $\pm$  s.d. ( $n = 8$ ). **c**, The relative weight of indicated organs was calculated as the ratio between organ weight and body weight. Values represent mean  $\pm$  s.d. ( $n = 8$ ). **d**, Percentage of PCNA-positive cells in 9-day-old livers. Values represent mean  $\pm$  s.d. ( $n = 4$ ); **e**, corresponding representative immunohistochemical staining. Scale bar, 100  $\mu$ m. **f**, Representative result of western blot analysis of eIF6 expression in organs from *eIF6*<sup>+/+</sup> and *eIF6*<sup>+/-</sup> adult mice; bottom, densitometric analysis normalized to actin levels. \* $P \leq 0.01$ , \*\* $P \leq 0.05$ .

<sup>1</sup>Molecular Histology and Cell Growth Laboratory, San Raffaele Science Institute, Via Olgettina 58, 20132 Milan, Italy. <sup>2</sup>Northwestern University, Feinberg School of Medicine, 303 East Avenue Chicago, Chicago, Illinois 60611, USA. <sup>3</sup>Università Vita-Salute, Medical School, Via Olgettina 58, 20132 Milan, Italy. <sup>4</sup>DISAV, University of Eastern Piedmont, Via Bellini 8G, 15100 Alessandria, Italy.

\*These authors contributed equally to this work.

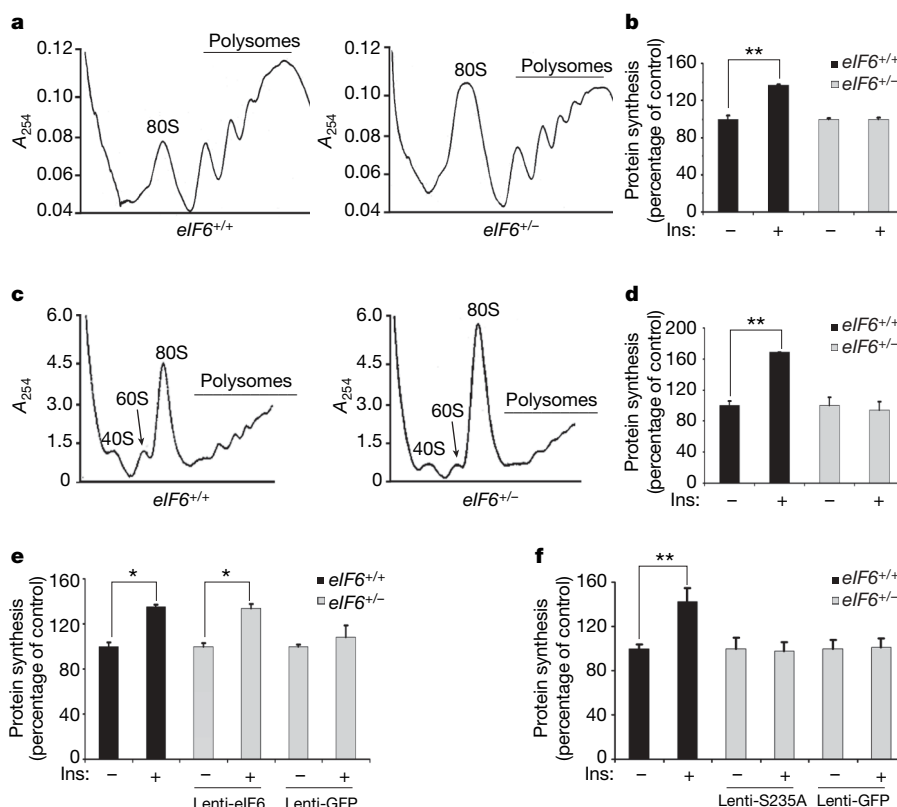
identical (Fig. 1b), suggesting that the reduction of body mass in  $eIF6^{+/-}$  mice could be due to the smaller size of specific organs. Biometric measures of  $eIF6^{+/-}$  and wild-type organs at autopsy revealed that heart and brain were equal to those in wild type, whereas all other organs from  $eIF6^{+/-}$  mice showed a slight trend to mass reduction (not shown). When the weight of the organs was normalized to the body weight of the corresponding animal, for most organs the differences disappeared, whereas the liver (Fig. 1c) and the adipose tissue (Supplementary Fig. 3) exhibited a reduction in their weight/body ratio. A reduction in liver mass could be due to a reduction in cell number, or to a reduction in cell size. Histological staining showed that the liver of  $eIF6^{+/-}$  mice was normal (Supplementary Fig. 4b, d), and further morphometric analysis excluded differences in cell size (Supplementary Fig. 4a, c). In contrast, we observed fewer cycling cells in the liver of  $eIF6^{+/-}$  mice (Fig. 1d, e), suggesting that its reduced mass was due to a reduction in cell number. We performed a similar analysis on adipocytes derived from mesenchymal stem cells: in summary, we found that the reduction of adipose tissue was due to a decreased proliferation of pre-adipocytes, whereas the lipid content per cell was constant (Supplementary Fig. 3). We conclude that the reduction in liver and fat tissue is probably due to a cell autonomous defect marked by reduced proliferative capability.

Quantitative western blotting analysis of eIF6 levels showed that the levels of eIF6 in all heterozygous organs were reduced to half that of the wild type (Fig. 1f; Supplementary Fig. 3c for fat tissue). In spite of reduction of eIF6 levels, we did not detect obvious phenotypes, *in vivo*, in other organs, including the skin (Supplementary Fig. 5), where we reported binding of eIF6 (p27BBP) to  $\beta_4$  integrin<sup>15</sup>. The reduction of eIF6 in all organs from heterozygous mice suggests that under specific circumstances eIF6 haploinsufficiency may become rate-limiting.

We performed polysomal profiles of livers from  $eIF6^{+/-}$  and  $eIF6^{+/+}$  littermates. Approximately 30% of the heterozygous/wild-type couples analysed did not present detectable differences in polysomal profiles. In the remaining 70% heterozygous/wild-type couples, in  $eIF6^{+/-}$  mice polysomal profiles showed an increased 80S peak, accompanied by decreased polysomes (Fig. 2a); this suggests that in mammalian liver  $eIF6^{+/-}$  hepatocytes have a defect in initiation of translation rather than in ribosome biogenesis. We tested this hypothesis using insulin to stimulate basal translation on primary hepatocytes *ex vivo*<sup>16</sup>. We found that insulin administration led to an increase of methionine incorporation in  $eIF6^{+/+}$  hepatocytes, but had little or no effect in  $eIF6^{+/-}$  hepatocytes (Fig. 2b).

We explored further primary mouse embryonic fibroblasts (MEFs) from  $eIF6^{+/-}$  mice. Remarkably, primary MEFs from  $eIF6^{+/-}$  mice showed a phenotype identical to liver hepatocytes with an accumulation of the 80S peak (Fig. 2c) and a reduction in insulin-stimulated translation (Fig. 2d), demonstrating that eIF6 depletion alters translation in different cell types. We previously reported that phorbol 12-myristate 13-acetate (PMA) stimulated methionine incorporation in conditions of eIF6 overexpression<sup>10</sup>. We tested whether PMA-stimulated methionine incorporation was reduced in  $eIF6^{+/-}$  cells, and found that it was (Supplementary Fig. 6). To understand whether basal translation was affected in conditions of eIF6 haploinsufficiency, we pooled data from different experiments. In brief, basal translation is slightly affected in primary hepatocytes, but not in MEFs, consistent with the observation that, *in vivo*, only liver development is affected by eIF6 haploinsufficiency (Supplementary Fig. 7).

We decided to test whether the reduction in methionine incorporation observed in conditions of eIF6 haploinsufficiency could be



**Figure 2 | eIF6 reduction results in 80S complex accumulation and a blunted translational response to insulin.** **a**, Polysomal profiles from livers showing the different amount of 80S ribosomal complex and polysomes between  $eIF6^{+/+}$  and  $eIF6^{+/-}$  *in vivo*. **b**, <sup>35</sup>S-methionine labelling experiments in isolated primary hepatocytes stimulated with 100 nM insulin (Ins). The basal rate of translation is indicated as 100%. Increase of translation is not detected in  $eIF6^{+/-}$  hepatocytes on insulin stimulation.

**c**, Polysomal profile indicates 80S accumulation in heterozygous MEFs; **d**, representative <sup>35</sup>S-methionine labelling of MEFs shows a blunted response to insulin in  $eIF6^{+/-}$ . **e**, **f**, Lentiviral-induced re-expression of wild-type  $eIF6$  in MEFs results in *de novo* insulin stimulation of methionine incorporation (**e**), whereas lentiviral-induced re-expression of  $eIF6^{Ser235Ala}$  does not (**f**). All data represent mean  $\pm$  s.d. of three independent experiments. \* $P \leq 0.01$ , \*\* $P \leq 0.05$ .

due to an indirect inactivation of the mTOR (also known as Frap1) pathway, which regulates translation at the level of initiation. We found that in MEFs, the activation of the PI(3)K–mTOR pathway, measured through the phosphorylation of Akt, rpS6 and 4E-BP1 (also known as Eif4ebp1), was normal (Supplementary Fig. 8). The levels of insulin receptor were also normal (Supplementary Fig. 9).

MEFs reconstituted with eIF6 to wild-type levels of eIF6 (Supplementary Fig. 10) regained insulin responsiveness (Fig. 2e), ruling out indirect effects due to chronic eIF6 depletion. To establish whether translation impairment might be due to the anti-association activity of eIF6, we exploited a Ser235Ala mutant, unable to be released from 60S, *in vitro*, after phosphorylation as we previously described<sup>10</sup>. Heterozygous MEFs reconstituted with eIF6(Ser235Ala) mutant (Supplementary Fig. 10) were not stimulated by insulin and behaved as *eIF6*<sup>+/-</sup> cells (Fig. 2f).

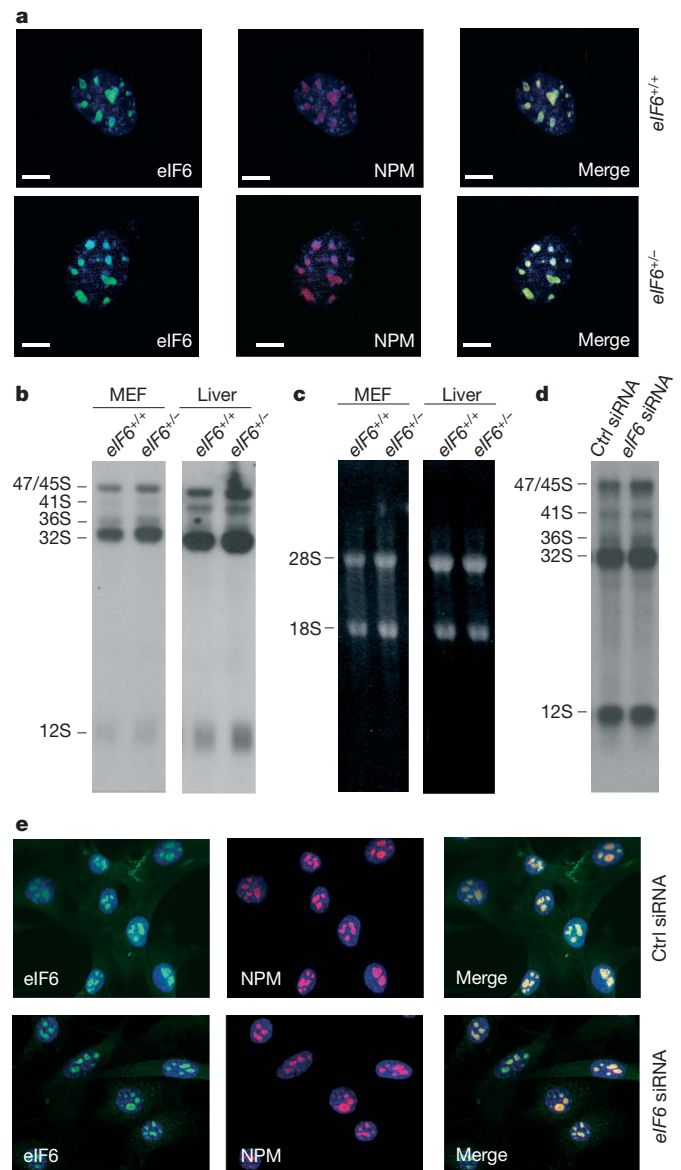
We evaluated in detail whether eIF6 haploinsufficiency can lead to deficits in ribosome biogenesis. In brief, the ratio of free 40S to 60S in *eIF6*<sup>+/-</sup> MEFs was equal to that in *eIF6*<sup>+/+</sup> MEFs, both in the cytoplasm and in the nucleus (Supplementary Fig. 11). The fact that *eIF6*<sup>+/-</sup> MEFs show altered translation and normal ribosome biogenesis can be explained by specific reduction of eIF6 in the cytoplasm, but not in nucleoli. Notably, the reduction of eIF6 that we observed in the cytoplasm of *eIF6*<sup>+/-</sup> cells was more conspicuous than the reduction in the nucleus (Supplementary Fig. 12a). In addition, the (apparent) amount of nucleolar-associated eIF6 was identical in wild-type and heterozygous cells (Fig. 3a).

To investigate whether subtle alterations in ribosome biogenesis could be observed in conditions of haploinsufficiency, we analysed by northern blotting the integrity of the ribosomal RNA processing pathway (Fig. 3b) in MEFs and liver. Our data show that the steady-state level of the rRNA precursors 47/45S, 32S and mature 28S is identical in wild-type and *eIF6*<sup>+/-</sup> cells (Fig. 3b, c). MEFs from early passages (P2/P3) rapidly grow in culture, allowing also for quantitative analysis of rRNA biogenesis<sup>17</sup> by pulse-chase with radiolabelled uridine. We found that *eIF6*<sup>+/-</sup> MEFs process rRNA with normal kinetics (Supplementary Fig. 12c). We also analysed by quantitative polymerase chain reaction with reverse transcription (Q-RT-PCR) the levels of *eIF6* messenger RNA, and two ribosomal proteins rpL5 and rpS19. Our data showed that primary *eIF6*<sup>+/-</sup> MEFs have 50% reduction of *eIF6* mRNA, but no other ribosome-associated proteins, both at the level of mRNA and of protein (Supplementary Fig. 12d, e). Finally, we asked whether a further reduction of eIF6 could lead to a defect in rRNA biogenesis. Eighty per cent depletion of eIF6 by RNA interference in NIH3T3 cells (Supplementary Fig. 13) showed normal rRNA synthesis (Fig. 3d) and detectable eIF6 in the nucleolus (Fig. 3e). Taken together, these data suggest that the reduction of cytoplasmic eIF6 leads to an increase in the steady-state levels of 80S, accompanied by an impairment of the capability of cells to upregulate protein synthesis on insulin stimulation, but a small amount of eIF6 is sufficient for rRNA biogenesis—in other words, eIF6 is an initiation factor *in vivo*.

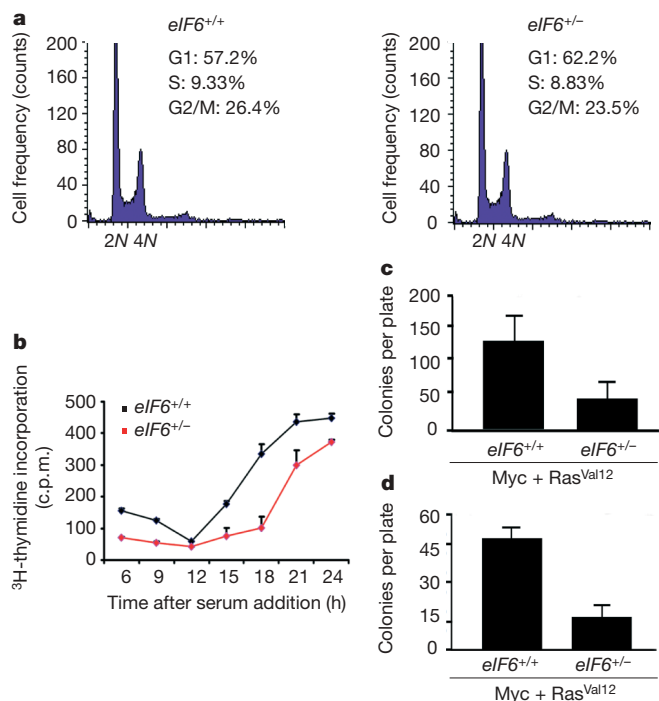
We analysed the consequences of eIF6 depletion on cell cycle progression in normal and highly proliferative cells. In early passages, *eIF6*<sup>+/-</sup> MEFs proliferated similarly to wild type, and had equal replicative senescence (Supplementary Fig. 14a, b), rate of apoptosis (Supplementary Fig. 14c, d) and cell size (Supplementary Fig. 14e). In these conditions, cell cycle distribution by fluorescence-activated cell sorting (FACS) analysis of exponentially growing MEFs derived from eIF6 wild type and matched heterozygous did not show gross differences (Fig. 4a). Next, we analysed cell cycle progression in cells synchronized by serum starving and stimulated by serum refeeding. Data obtained by <sup>3</sup>H-thymidine labelling show that *eIF6*<sup>+/-</sup> MEFs enter S-phase 18 h after stimulation (Fig. 4b), compared to the 12 h required by wild-type cells. These data show that *eIF6* haploinsufficiency affects translation and S-phase entry under specific conditions. We looked for additional models in which eIF6 haploinsufficiency could cause a phenotype. eIF6 is overexpressed

in cancer cells<sup>18,19</sup>. We decided to test whether *eIF6* haploinsufficient MEFs were efficiently transformed by infection with retroviral vectors encoding dominant-negative (DN)p53 and H-ras<sup>V12</sup>, or Myc and H-ras<sup>V12</sup>. *eIF6*<sup>+/-</sup> MEFs formed fewer transformed colonies in soft agar when infected with oncogene-carrying retroviruses compared to wild type (Fig. 4c, d). We had similar results in the focus formation assay (Supplementary Fig. 15a, b). The large difference in the transformation efficiency was not due to differences in retroviral infection and integration (Supplementary Fig. 15c).

Whether eIF6 is a true initiation factor has been unclear for a long time. The effect of eIF6, *in vitro*, ranges from no effects on translation to mild stimulation at low doses accompanied by repression at high doses<sup>8–10</sup>. The difference in the results and in the systems used did not allow a final conclusion. Tif6 (yeast homologue of mammalian eIF6)



**Figure 3 | eIF6 reduction in the cytoplasmic compartment does not affect 60S ribosomal biogenesis.** **a**, eIF6 distribution analysed in the nucleolus by anti-eIF6 immunostaining of primary MEFs. Nucleolar staining with anti-nucleophosmin (NPM) merged with anti-eIF6 (Merge) shows nucleolar eIF6 in *eIF6*<sup>+/-</sup> MEFs compared to wild-type cells. Scale bar, 3  $\mu$ m. **b, c**, Northern blot analysis of steady-state levels of 25S precursors in heterozygotes. MEFs and livers and corresponding ethidium bromide staining (**c**) showing equal steady-state levels of precursor rRNAs. **d, e**, Small interfering RNA (siRNA) of *eIF6* (80% reduction, see Supplementary Fig. 13) in NIH3T3 abolishes neither rRNA processing (**d**) nor nucleolar staining (**e**).



**Figure 4 | eIF6 reduction impairs G1/S progression in synchronized cells and causes reduced transformation.** **a**, FACS analysis of asynchronous MEFs. No significant differences in cell cycle distribution are observed between *eIF6*<sup>+/+</sup> and *eIF6*<sup>+/-</sup> MEFs. **b**, <sup>3</sup>H-thymidine labelling of synchronized cells shows a delay in S-phase entry. **c–d**, Soft agar assay of asynchronous MEFs transformed with DNp53 plus H-ras<sup>V12</sup> ( $n = 6$ ) and with Myc plus H-ras<sup>V12</sup> ( $n = 6$ ). Number of colonies expressed as mean  $\pm$  s.d. indicates that cellular potential to undergo transformation is impaired in *eIF6*<sup>+/-</sup> MEFs.

depletion led to the loss of 60S biogenesis and rRNA processing defects<sup>11–14</sup>. In yeasts, most of Tif6 seems to be localized in the nucleus<sup>20,21</sup> and in pre-ribosomal particles<sup>22</sup>, thus raising questions regarding the physiological relevance of the regulated eIF6 anti-association activity. The results we obtained in *eIF6* knockout mice demonstrate that eIF6 may be rate-limiting for translation, *in vivo*, where it may behave as a stimulatory translation initiation factor downstream of extracellular signalling *in vivo*. From this perspective, eIF6 is the first 60S associated initiation factor able to modulate translation downstream of extracellular signalling *in vivo*. These data do not contradict a role of eIF6 in ribosome biogenesis, but suggest that ribosome biogenesis requires only sequestration of a (minor) nucleolar pool of eIF6.

It remains to be understood whether eIF6 regulates global translation or specific classes of mRNA. It was recently proposed that eIF6 is the mediator regulating microRNA-mediated repression<sup>23</sup> in *Caenorhabditis elegans* and HeLa cells although not in *Drosophila*<sup>24</sup>. In conditions of *eIF6* heterozygosity we did not observe differences in expression of microRNA targets (Supplementary Fig. 16). Further studies are required.

eIF6 is consistently overexpressed in a variety of cancers including colon<sup>18</sup> carcinomas and leukaemias<sup>19</sup>. In the latter context, eIF6 upregulation was interpreted as a by-product of transformation, namely increased growth demanded of ribosomes and translation. Our data support a model in which the levels of eIF6 may directly affect tumorigenesis. In any case, our observations strengthen the concept that translational control is crucial for malignancy, as it was recently proposed for eIF4E, downstream of the mTOR pathway<sup>5</sup>.

## METHODS SUMMARY

All analyses were performed at least three times on different genetic backgrounds. For the generation of knockout mice, we deleted exons 1 and 2.

Hepatocytes and primary MEFs were prepared from littermates and kept in culture for 3 h and 2–3 weeks, respectively. Infection and rescue by eIF6-lentivirus were as detailed in Methods and Supplementary Information.

Analysis of rRNA biogenesis was performed by sucrose gradient centrifugation of dissociated ribosomes, and by pulse-chase rRNA labelling with <sup>3</sup>H-uridine and northern blotting (Supplementary Information). Analysis of translation was performed by sucrose gradient centrifugation and metabolic labelling with <sup>35</sup>S-methionine<sup>25</sup>.

Blastocysts and MEFs were fixed in 2% paraformaldehyde. Blastocysts were then stained by whole-mount immunofluorescence. Cells were attached on coverslips and stained by indirect immunofluorescence.

Primary fibroblasts were infected with retroviruses carrying Myc plus oncogenic H-ras<sup>V12</sup> or DNp53 plus oncogenic H-ras<sup>V12</sup>. Foci were counted 2–3 weeks after the primary infection<sup>26</sup>.

SDS-PAGE and western blotting analyses were performed on total extracts in RIPA buffer (10 mM Tris-HCl, pH 7.2, 1% sodium deoxycholate, 1% Triton X-100, 0.1% SDS, 150 mM NaCl, 1 mM EDTA, pH 8.0) as detailed elsewhere<sup>17</sup>. Nuclear–cytoplasmic fractionation was performed by ultracentrifugation on a sucrose cushion as described in Supplementary Methods. The purity of the fractions was checked by antibodies against Gapdh (cytoplasm) or histone H2B (also known as Hist2h2be, nucleus). We loaded equivalent amounts of cells.

Cell number and viability were analysed as previously described<sup>27</sup>. S-phase entry was analysed by the <sup>3</sup>H-thymidine assay on primary embryo fibroblasts at early passages.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 26 October 2007; accepted 15 July 2008.

Published online 10 September 2008.

- Kapp, L. D. & Lorsch, J. R. The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem.* **73**, 657–704 (2004).
- Gebauer, F. & Hentze, M. W. Molecular mechanisms of translational control. *Nature Rev. Mol. Cell Biol.* **5**, 827–835 (2004).
- Proud, C. G. Signalling to translation: how signal transduction pathways control the protein synthetic machinery. *Biochem. J.* **403**, 217–234 (2007).
- Sonenberg, N. & Pause, A. Signal transduction. Protein synthesis and oncogenesis meet again. *Science* **314**, 428–429 (2006).
- Mamane, Y., Petroulakis, E., LeBacquer, O. & Sonenberg, N. mTOR, translation initiation and cancer. *Oncogene* **25**, 6416–6422 (2006).
- Holcik, M. & Sonenberg, N. Translational control in stress and apoptosis. *Nature Rev. Mol. Cell Biol.* **6**, 318–327 (2005).
- Wek, R. C., Jiang, H. Y. & Anthony, T. G. Coping with stress: eIF2 kinases and translational control. *Biochem. Soc. Trans.* **34**, 7–11 (2006).
- Russell, D. W. & Spremulli, L. L. Purification and characterization of a ribosome dissociation factor (eukaryotic initiation factor 6) from wheat germ. *J. Biol. Chem.* **254**, 8796–8800 (1979).
- Valenzuela, D. M., Chaudhuri, A. & Maitra, U. Eukaryotic ribosomal subunit anti-association activity of calf liver is contained in a single polypeptide chain protein of Mr = 25,500 (eukaryotic initiation factor 6). *J. Biol. Chem.* **257**, 7712–7719 (1982).
- Ceci, M. *et al.* Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579–584 (2003).
- Sanvito, F. *et al.* The  $\beta$ 4 integrin interactor p27(BBP/eIF6) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly. *J. Cell Biol.* **144**, 823–837 (1999).
- Si, K. & Maitra, U. The *Saccharomyces cerevisiae* homologue of mammalian translation initiation factor 6 does not function as a translation initiation factor. *Mol. Cell Biol.* **19**, 1416–1426 (1999).
- Wood, L. C., Ashby, M. N., Grunfeld, C. & Feingold, K. R. Cloning of murine translation initiation factor 6 and functional analysis of the homologous sequence YPR016c in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **274**, 11653–11659 (1999).
- Basu, U., Si, K., Warner, J. R. & Maitra, U. The *Saccharomyces cerevisiae* TIF6 gene encoding translation initiation factor 6 is required for 60S ribosomal subunit biogenesis. *Mol. Cell Biol.* **21**, 1453–1462 (2001).
- Biffo, S. *et al.* Isolation of a novel  $\beta$ 4 integrin-binding protein (p27(BBP)) highly expressed in epithelial cells. *J. Biol. Chem.* **272**, 30314–30321 (1997).
- Clark, R. L. & Hansen, R. J. Insulin stimulates synthesis of soluble proteins in isolated rat hepatocytes. *Biochem. J.* **190**, 615–619 (1980).
- Strezoska, Z., Pestov, D. G. & Lau, L. F. Bop1 is a mouse WD40 repeat nucleolar protein involved in 28S and 5.8S rRNA processing and 60S ribosome biogenesis. *Mol. Cell Biol.* **20**, 5516–5528 (2000).
- Sanvito, F. *et al.* Expression of a highly conserved protein, p27BBP, during the progression of human colorectal cancer. *Cancer Res.* **60**, 510–516 (2000).
- Harris, M. N. *et al.* Comparative proteomic analysis of all-trans-retinoic acid treatment reveals systematic posttranscriptional control mechanisms in acute promyelocytic leukemia. *Blood* **104**, 1314–1323 (2004).



20. Senger, B. *et al.* The nucleolar Tif6p and Efl1p are required for a late cytoplasmic step of ribosome synthesis. *Mol. Cell* **8**, 1363–1373 (2001).
21. Menne, T. F. *et al.* The Shwachman–Bodian–Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nature Genet.* **39**, 486–495 (2007).
22. Volta, V. *et al.* Sen34p depletion blocks tRNA splicing *in vivo* and delays rRNA processing. *Biochem. Biophys. Res. Commun.* **337**, 89–94 (2005).
23. Chendrimada, T. P. *et al.* MicroRNA silencing through RISC recruitment of eIF6. *Nature* **447**, 823–828 (2007).
24. Eulalio, A., Huntzinger, E. & Izaurralde, E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nature Struct. Mol. Biol.* **15**, 346–353 (2008).
25. Gorrini, C. *et al.* Fibronectin controls cap-dependent translation through  $\beta 1$  integrin and eukaryotic initiation factors 4 and 2 coordinated pathways. *Proc. Natl Acad. Sci. USA* **102**, 9200–9205 (2005).
26. Zou, X. *et al.* Cdk4 disruption renders primary mouse cells resistant to oncogenic transformation, leading to Arf/p53-independent senescence. *Genes Dev.* **16**, 2923–2934 (2002).
27. Yang, Y. L. *et al.* BubR1 deficiency results in enhanced activation of MEK and ERKs upon microtubule stresses. *Cell Prolif.* **40**, 397–410 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by grants AIRC (S.B., P.C.M.), TELETHON GGB05043, CARIPL0 0578 (S.B.) and NIH-RO1 (H.K.). A.B. is supported by grant AICR 05-360. The manuscript has been improved thanks to suggestions from N. Offenhaeuser and A. Boletta. We are indebted to P. G. Pelicci for anti-NPM antibodies, H. Hirai for preliminary soft agar assays, S. Modina for blastocyst preparation, M. Vidali for hepatocytes preparation, M. Malosio for insulin receptor antibodies, S. Gregori for FACS analysis, D. Bartel for reporter constructs, G. Manfioletti for HMGA2 antibodies and F. Loreni for rpS19 antibodies. We acknowledge L. Magri for preliminary experiments, and V. Volta and S. Grosso for suggestions.

**Author Contributions** V.G., A.M., A.M.B., H.K. and S.B. planned the experiments; V.G., A.M., A.M.B., A.B. and S.B. performed the experiments; all authors analysed the data; and V.G. and S.B. wrote the paper. All authors discussed the results and contributed to the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to S.B. ([stefano.biffo@hsr.it](mailto:stefano.biffo@hsr.it)).

## METHODS

**Antibodies and reagents.** The following antibodies were used: rabbit polyclonal against eIF6 (ref. 15), caspase 3 (Cell Signaling), PCNA (Abcam) and H2B (Santa Cruz), and mouse monoclonal against  $\beta$ -actin (Sigma), nucleophosmin<sup>28</sup> and Gapdh (Chemicon International). Nuclear staining was performed with Hoechst 3342 (Molecular Probes). Human insulin was from Sigma.

**Generation of eIF6 null mice.** *eIF6* gene retrieval was obtained by screening a mouse genomic DNA library (SJ129 strain). The gene was cloned into a pPNT plasmid. The 5' long arm spanned to the exon 1 and the 3' short arm included exon 3 and exon 4. The construct was electroporated in SJ129 embryonic stem cells by the Core Facility for Conditional Mutagenesis (CFCM). eIF6 mutant cells obtained by homologous recombination were identified by Southern blot and PCR analysis (Supplementary Methods). Recombined embryonic stem cells were used to produce chimaeric animals by both injection and aggregation techniques. Chimaeras were then mated to SJ129 mice for germline transmission. Backcross was performed to the CL57Bl6 background. Genotyping analysis was performed by PCR on tail genomic DNA (Supplementary Methods). Controls for the proper expression of upstream and downstream genes to *eIF6* (*Igfb4bp*) were performed by RT-PCR as detailed in Supplementary Methods.

All experiments involving animals were performed in accordance with Italian national regulations and experimental protocols reviewed by local Institutional Animal Care and Use Committees (IACUC n. 574).

**Blastocyst preparation and immunofluorescence.** Blastocysts from hyperovulated, mated C57Bl6 mice were obtained by flushing the uterine tubes with M2 medium (Sigma). Whole-mount immunofluorescence was performed as detailed previously<sup>29</sup>. Images were obtained with fluorescent microscopy (Zeiss Axiophot) or laser confocal microscopy (Ultra View, Perkin Elmer). Image slides were taken with a digital camera (Sensicam) using Image-Pro plus software.

Immunofluorescence on MEFs was performed as described previously<sup>11</sup>, after fixation for 10 min at 25 °C in 2% buffered paraformaldehyde, 3% sucrose in PBS, followed by permeabilization in 20 mM Hepes, pH 7.4, 0.5% Triton-X100.

**Polysomal profiles.** Polysomal profiles were performed as described previously<sup>10</sup>. In brief, after measuring  $A_{254}$  of the extracts prepared in 30 mM Tris-HCl, pH 7.5, 100 mM NaCl, 30 mM MgCl<sub>2</sub>, 0.1% NP-40 and 10  $\mu\text{g ml}^{-1}$  cycloheximide, the equivalent of 5 units of absorbance were layered on a 15–50% sucrose gradient and centrifuged at 4 °C in a SW41Ti Beckman rotor for 3 h 30 min at 39,000 r.p.m., corresponding to 190,000g. Absorbance at 254 nm was recorded by BioLogic LP software (BioRad). The preparation of individual ribosomal subunits and polysomes was obtained as detailed in Supplementary Methods. In brief, ribosomal subunits were dissociated by 20 mM EDTA after isolating them at 39,000 r.p.m. (190,000g) for 16 h on a 30% sucrose cushion then separated on a 10–30% sucrose gradient at 36,000 r.p.m. (160,000g) and analysed as above.

Analysis of rRNA metabolism in primary embryo fibroblasts was performed as detailed in Supplementary Methods.

**Primary cell cultures.** Hepatocytes and primary MEFs were prepared from littermates and kept in culture for 3 h and 2–3 weeks, respectively. All the analyses were performed at least three times on different genetic backgrounds. Infection and rescue by eIF6-lentivirus was as detailed in Supplementary Information.

Primary embryo fibroblasts (MEFs) were prepared from E13.5 embryos and cultured as described<sup>30</sup>. In brief, embryos were dissociated by 0.03% trypsin in 0.2% EDTA at 37 °C for 10 min and then treated with DNaseI (New England Biolabs). After filtering fibroblasts with a 70  $\mu\text{m}$  cell strainer, they were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS), 100  $\mu\text{g ml}^{-1}$  streptomycin, 100 U  $\text{ml}^{-1}$  penicillin and 2 mM glutamine. MEFs were cultured at 9% CO<sub>2</sub> at 37 °C and experiments were performed at first passages (P1–P4). Hepatocytes were isolated as detailed in Supplementary Methods.

**Protein synthesis measurement by <sup>35</sup>S-methionine labelling.** Primary embryo fibroblasts (MEFs) and primary hepatocytes were used for analysis of translational rate. Cells were seeded at sub-confluency in 6-well plates, and insulin stimulation was performed at 100 nM for 30 min. Metabolic labelling was performed as described previously<sup>25</sup>. All experiments were done in triplicate. Statistical *P*-values calculated by two-tailed *t*-test were also indicated: two asterisks for *P*-values less than 0.05 and one asterisk for *P*-values less than 0.01.

**Immunohistochemistry and histological staining.** Immunohistochemical and histological analysis were performed on paraffin-embedded sections prepared as detailed in Supplementary Methods.

**Cell cycle analysis.** S-phase entry was analysed by <sup>3</sup>H-thymidine assay on primary embryo fibroblasts at early passages. MEFs were seeded at 60% confluency in 6-well plate and starved for 48 h in DMEM plus 0.1% FBS. MEFs were then

stimulated with 10% serum and labelled with 1  $\mu\text{Ci ml}^{-1}$  <sup>3</sup>H-thymidine for 24 h. 6, 9, 12, 15, 18, 21 and 24 h after serum addition, <sup>3</sup>H-thymidine was added to the medium and cells were lysed after 1 h as described in Supplementary Methods. Half of the radioactive sample obtained (0.5 ml) was solubilized in scintillation fluid and counted with a  $\beta$ -counter. Data derived from the mean of three independent experiments and counts per minute were normalized to the total amount of DNA.

**Retroviral infection and soft agar assay.** The Phoenix ecotropic virus packaging cells were obtained from the American Tissue Culture Collection (ATCC) and maintained in DMEM supplemented with 10% FBS, 1% glutamine and antibiotics. The pBabe-hygro vector for expression of H-ras<sup>V12</sup>, dominant-negative p53 mutant (DNp53) and Myc was described previously<sup>30</sup>. The soft agar formation assay and the focus formation assay were performed as described previously<sup>30</sup>.

**Southern blot and PCR analysis.** Southern blot hybridization was used to identify the targeted allele in the embryonic stem cells using standard protocols. Specifically, targeted clones were identified by using a 5' external probe, which detects the change of a 16.3 kb wild-type EcoRV fragment to a novel 13 kb EcoRV fragment, and then were confirmed with a 3' external probe, which detects the change of a 13.3 kb wild-type HindIII fragment into a novel 9.2 kb HindIII fragment.

Genotyping of the offspring mice was performed using a three-primers PCR strategy to simultaneously amplify target sequences from both the wild-type and the mutant alleles. PCR was performed by AmpliTaq Gold (Roche) according to the manufacturer's protocol. A shared primer number 3 (5'-GTG AGTCTTGGCTTCATGTG-3') was designed downstream of the deleted region. This primer can pair with wild-type allele-specific primer number 2 (5'-CTATGTGGCCTTGGTCCAC-3') to amplify a 320 bp PCR product or with the targeted allele-specific primer number 1 (5'-GCAGCGATCGCCTTC TATC-3') in the *neo* cassette to amplify a 650 bp fragment from the mutant allele. PCR products were resolved on 2% agarose gel.

**mRNA extraction and real-time RT-PCR.** Total RNA was extracted with TRIzol reagent (Invitrogen) from MEFs and from tested organs of *eIF6*<sup>+/+</sup> and *eIF6*<sup>+/-</sup> mice. After treatment of total RNA with RQ1 RNase-free DNase (Promega), reverse transcription was performed with SuperScript™ First-Strand (Invitrogen) according to the manufacturer's instructions. 100 ng of reverse-transcribed complementary DNA was amplified with the appropriate primers and detected with probes specific for *Igfb4bp*, *Mmp24*, *Gdf5*, *Rpl5* and *RpS19* cDNA (TaqMan gene expression assay Mm00550245\_m1, Mm00487721\_m1, Mm00433564\_m1, Mm00847026\_g1 and Mm00452264\_m1, respectively, from Applied Biosystems) in an ABI PRISM 7900HT Sequence Detection System. *Gapdh* was used as internal control gene expression (TaqMan assay reagent 4352932E). The data were expressed as the percentage of relative quantity of target genes. Results are represented as means of three independent experiments.

**Immunohistochemistry and histological staining.** Tissue samples were excised from *eIF6*<sup>+/+</sup> and *eIF6*<sup>+/-</sup> mice and immediately fixed in 10% formaldehyde in 0.1 M phosphate buffer (PBS) at pH 7.4. Serial 3- $\mu\text{m}$  sections of paraffin-embedded liver, skin and fat were stained with haematoxylin and eosin (H&E) for morphological analysis. Liver sections were subjected to immunohistochemical staining for proliferating cell nuclear antigen (PCNA) using the Vectastain Elite ABC kit (Vector). In brief, slides were passed in graded alcohol solutions, rinsed in distilled water and immersed in citrate buffer (0.1 M citric acid plus 0.1 M sodium citrate). Samples were boiled in a microwave oven three times for 5 min, washed in PBS and soaked in 3% hydrogen peroxide for 30 min to block endogenous peroxidase activity. After blocking in normal serum, slides were incubated with anti-PCNA (1:2,000) for 16 h at 4 °C. Sections were incubated with a biotinylated secondary antibody for 1 h at room temperature and then with a preformed avidin-biotinylated peroxidase complex. Signals were detected using diaminobenzidine tetrahydrochloride (DAB, BioGENEX) and sections were counterstained with haematoxylin. Stained samples were observed in a Zeiss Axiophot microscope and pictures were acquired with ScionImage Software.

**siRNA transfection of NIH 3T3 cells.** Transient transfection was performed on NIH 3T3 cells using Lipofectamin 2000 (Invitrogen), according to manufacturer's instructions.

For siRNA experiments, siGENOME duplex targeting ITGB4BP was used (NM\_002212; sense sequence, 5'-GAGCUUCGUUCGAGAACAUU-3'; anti-sense sequence, 5'-PUUGUUCUGAAGCAGCUCUU-3'; Dharmacon). SiCONTROL NON-TARGETING siRNA POOL (Dharmacon) was used as negative control.

Cells were transfected at 30–40% confluency using 100 pmol siRNA in Opti-MEM medium (Invitrogen); after 3 h at 37 °C the medium was changed, and gene silencing was examined by western blot analysis 72 h after transfection.

**Flow cytometry and measurement of cell size.** Unfixed *eIF6*<sup>+/+</sup> and *eIF6*<sup>+/-</sup> MEFs were pelleted and resuspended in PBS containing 5 ng  $\mu\text{l}^{-1}$  propidium

iodide (Sigma). Cell size (FSC) was analysed using FACS Canto II (Becton Dickinson).

**Cell fractionation.** Cells were detached by trypsinization and collected by centrifugation. The pellet was resuspended in ice-cold hypotonic wash buffer (10 mM Tris HCl, pH 7.4, 10 mM KCl, 2 mM MgCl<sub>2</sub>) and then the cytoplasmic fraction was extracted in hypotonic lysis buffer (10 mM Tris HCl, pH 7.4, 10 mM KCl, 2 mM MgCl<sub>2</sub>, 0.05% Triton X-100, 1 mM EGTA, 1 mM DTT) for 20 min on ice. Nuclei were collected by centrifugation at 700g for 15 min and resuspended in hypotonic wash buffer before loading on 1.6 M sucrose cushion for nuclear fraction purification. The nuclear pellet was obtained by ultracentrifugation (TL100, Beckman) at 50,000 r.p.m. for 1 h and then lysed in RIPA buffer.

**Lentiviral infection.** 293T cells (ATCC) were infected with packaging plasmid VSV-G, PMDLg/pRRE, pREV and transfer vector pCCL-PPT-hPGK-pre previously described<sup>31</sup>. Full-length eIF6 and GFP were cloned in transfer vector. MEFs were infected with eIF6-pCCL and GFP-pCCL at early passage (P<sub>1</sub>) and used for <sup>35</sup>S-methionine labelling as described. Multiplicity of infection (m.o.i.) was empirically determined to achieve twofold eIF6 expression by triplicate infections.

*In vitro* mutagenesis was performed on the wild-type eIF6 lentiviral transfer plasmid by the QuickChange Kit (Stratagene) using the same oligonucleotides as in ref. 32.

**Ear mesenchymal stem cell culture and adipogenesis.** Mesenchymal stem cells from outer ears of *eIF6*<sup>+/+</sup> and *eIF6*<sup>+/-</sup> mice were collected and differentiated into the adipogenic lineage as previously described<sup>33</sup>.

Differentiated cells were fixed in 10% buffered formaline and stained with Oil Red O for 10 min. The dye retained by the cells was eluted with isopropanol and quantified by measuring absorbance at 500 nm.

**Polysomal profile from liver.** Tissue was gently homogenized in 50 mM Tris HCl, pH 7.8, 240 mM KCl, 10 mM MgSO<sub>4</sub>, 5 mM DTT, 250 mM sucrose, 2% Triton X-100, 90 µg ml<sup>-1</sup> cycloheximide and 30 U ml<sup>-1</sup> RNasin using a glass douncer. After clarification, 100 µg ml<sup>-1</sup> heparin was added to the supernatant. 5 OD<sub>254</sub> of rRNA was loaded on a 15–55% sucrose gradient dissolved in 25 mM Tris HCl, pH 7.4, 25 mM NaCl, 5 mM MgCl<sub>2</sub> and 1 mM DTT and run at 39,000 r.p.m. (190,000g) for 3 h 30 min with SW41Ti swing rotor (Beckman Coulter). The polysomal profile was detected as previously described.

**RNA blot analysis.** Total RNA was isolated from MEF cells or from mouse liver using TRI REAGENT (Sigma) following the manufacturer's protocol. RNA (5 µg) was separated on 1% agarose-formaldehyde gel and analysed by northern blot hybridization using standard techniques.

The oligonucleotide ITS2-2 complementary to nucleotides 239–271 of the ITS2 region was used as probe for hybridization (5'-ACTGGTGAGGCA GCGGTCCGGGAGGCGCCGACG-3'), as described previously<sup>34</sup>. The probe was 5'-labelled using [<sup>32</sup>P] ATP and T4 polynucleotide kinase.

**Preparation of 40S and 60S ribosomal subunits.** Dissociation of 40S and 60S ribosomal subunits extracted from cytoplasmic and nuclear compartments were obtained as described previously<sup>20</sup>. 40S and 60S subunits were dissociated with 20 mM EDTA. The equivalent of 20 units of A<sub>254</sub> of total rRNA were loaded on a 30% sucrose cushion dissolved in 10 mM Tris HCl, pH 7.5, 10 mM KCl, 1 mM MgCl<sub>2</sub> and 1 mM DTT. Ribosomal RNA was collected by ultracentrifugation at 36,000 r.p.m. for 16 h, washed twice in cold 1 × PBS and resuspended in 30 mM Tris HCl, pH 7.5, 10 mM KCl, 1 mM MgCl<sub>2</sub>, 20 mM EDTA and 1 mM DTT. 10 OD of dissociated 40S and 60S subunits were separated through a 10–30%

sucrose gradient, ultracentrifuged at 36,000 r.p.m. for 3 h 30 min and analysed by Biological LP System (BioRad) as described previously.

**Hepatocytes isolation.** One-month-old mice were anaesthetized by intraperitoneal injection of 200 µl chloralium hydrate (8% w/v). Liver was perfused *in situ* through the portal vein with two different perfusion media: T1 solution pH 7.4 (0.9% NaCl, 0.05% KCl, 0.2% HEPES, 0.08 mg ml<sup>-1</sup> EGTA) and T2 solution pH 7.4 (0.6% NaCl, 0.05% KCl, 1.2% HEPES, 0.07% CaCl<sub>2</sub>, 3 g ml<sup>-1</sup> collagenase type I). Flow rate was 5 ml min<sup>-1</sup>. After collagenase digestion, hepatocytes were filtered through a 70 µm cell strainer and passed on a 37.5% PERCOLL cushion (Amersham). Pellet of viable cells was resuspended in DMEM and recovered at 5% CO<sub>2</sub>, 37 °C for 2 h before starting the experimental procedure.

**Analysis of ribosomal biogenesis.** Processing of rRNA was analysed by a pulse-chase experiment with <sup>3</sup>H-uridine (Amersham) with a modified method. Exponentially growing MEFs were labelled with 3 µCi ml<sup>-1</sup> <sup>3</sup>H-uridine and chased for the indicated period. Total RNA was extracted by TRIzol (Sigma) according to the manufacturer's protocol. The same counts per minute were loaded on a 1% formaldehyde-agarose gel and then transferred onto a nylon membrane HYBOND-N (Amersham) for 16 h. RNA was crosslinked by ultraviolet irradiation and nylon membrane was treated with EN3HANCE spray (Perkin Elmer) before autoradiography.

**S-phase entry analysis by the <sup>3</sup>H-thymidine assay.** The assay was performed at early passages (P1–P2), and <sup>3</sup>H-thymidine uptake was detected. MEFs were seeded at 60% confluency in 6-well plate and starved for 48 h in low serum. Synchronized cells were stimulated with 10% serum, and 1 µCi ml<sup>-1</sup> of <sup>3</sup>H-thymidine (Amersham) was added for 1 h at 6, 9, 12, 15, 18, 21 and 24 h. After 1 h, cells were incubated for 15 min with 5% trichloroacetic acid to precipitate DNA onto the dishes, and after two washes in 70% cold-ethanol macromolecules were solubilized in 1 ml solubilization buffer (0.1 M NaOH, 2% Na<sub>2</sub>CO<sub>3</sub>, 1% SDS) for 1 h at 60 °C. The mixture was neutralized with 36% of neutralizing solution (1 M HCl) and half of each radioactive sample was solubilized in scintillation fluid and counted with a β-counter. Data were derived from the mean of three independent experiments and normalized to the total amount of DNA.

28. Colombo, E., Marine, J. C., Danovi, D., Falini, B. & Pelicci, P. G. Nucleophosmin regulates the stability and transcriptional activity of p53. *Nature Cell Biol.* **4**, 529–533 (2002).
29. Betts, D. H., Barcroft, L. C. & Watson, A. J. Na/K-ATPase-mediated 86Rb<sup>+</sup> uptake and asymmetrical trophectoderm localization of alpha1 and alpha3 Na/K-ATPase isoforms during bovine preattachment development. *Dev. Biol.* **197**, 77–92 (1998).
30. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593–602 (1997).
31. De Palma, M. & Naldini, L. Transduction of a gene expression cassette using advanced generation lentiviral vectors. *Methods Enzymol.* **346**, 514–529 (2002).
32. Ceci, M. *et al.* Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579–584 (2003).
33. Rim, J. S., Mynatt, R. L. & Gawronska-Kozak, B. Mesenchymal stem cells from the outer ear: a novel adult stem cell model system for the study of adipogenesis. *FASEB J.* **19**, 1205–1207 (2005).
34. Strezoska, Z., Pestov, D. G. & Lau, L. F. Functional inactivation of the mouse nucleolar protein Bop1 inhibits multiple steps in pre-rRNA processing and blocks cell cycle progression. *J. Biol. Chem.* **277**, 29617–29625 (2002).

# CDK targets Sae2 to control DNA-end resection and homologous recombination

Pablo Huertas<sup>1</sup>, Felipe Cortés-Ledesma<sup>2</sup>, Alessandro A. Sartori<sup>1†</sup>, Andrés Aguilera<sup>2</sup> & Stephen P. Jackson<sup>1</sup>

DNA double-strand breaks (DSBs) are repaired by two principal mechanisms: non-homologous end-joining (NHEJ) and homologous recombination (HR)<sup>1</sup>. HR is the most accurate DSB repair mechanism but is generally restricted to the S and G2 phases of the cell cycle, when DNA has been replicated and a sister chromatid is available as a repair template<sup>2–5</sup>. By contrast, NHEJ operates throughout the cell cycle but assumes most importance in G1 (refs 4, 6). The choice between repair pathways is governed by cyclin-dependent protein kinases (CDKs)<sup>2,3,5,7</sup>, with a major site of control being at the level of DSB resection, an event that is necessary for HR but not NHEJ, and which takes place most effectively in S and G2 (refs 2, 5). Here we establish that cell-cycle control of DSB resection in *Saccharomyces cerevisiae* results from the phosphorylation by CDK of an evolutionarily conserved motif in the Sae2 protein. We show that mutating Ser267 of Sae2 to a non-phosphorylatable residue causes phenotypes comparable to those of a *sae2Δ* null mutant, including hypersensitivity to camptothecin, defective sporulation, reduced hairpin-induced recombination, severely impaired DNA-end processing and faulty assembly and disassembly of HR factors. Furthermore, a Sae2 mutation that mimics constitutive Ser267 phosphorylation complements these phenotypes and overcomes the necessity of CDK activity for DSB resection. The Sae2 mutations also cause cell-cycle-stage specific hypersensitivity to DNA damage and affect the balance between HR and NHEJ. These findings therefore provide a mechanistic basis for cell-cycle control of DSB repair and highlight the importance of regulating DSB resection.

To initiate HR, one strand of the broken DNA duplex is resected in the 5'→3' direction, generating single-stranded DNA (ssDNA) that can anneal with a homologous DNA duplex<sup>8</sup>. In *S. cerevisiae*, effective resection and HR require sustained Cdc28/Clb (Cdk1/cyclin B) kinase activity<sup>2,3,5</sup>, although the CDK targets mediating this control are still unknown. One potential target is Sae2, a protein first identified as being required for meiotic recombination. Sae2 controls the initiation of DNA-end resection in meiotic and mitotic cells<sup>9–12</sup> and was recently shown to be a DNA endonuclease<sup>13</sup>. Previous work has shown that Sae2 is targeted by the Mec1 and Tel1 kinases in response to DNA damage, generating forms of Sae2 with decreased mobility in SDS-polyacrylamide gels<sup>14</sup>. Such alterations in Sae2 gel-mobility also occurred in unperturbed cycling cells, specifically in S and G2, indicating that Sae2 might be a Cdc28 target (data not shown, and Supplementary Fig. 2a). In accord with this idea, the amount of slower-migrating Sae2 was diminished when Cdc28 was inactivated in G2-synchronized cultures by galactose-driven expression of the Cdc28/Clb repressor, Sic1 (ref. 15) (Fig. 1a).

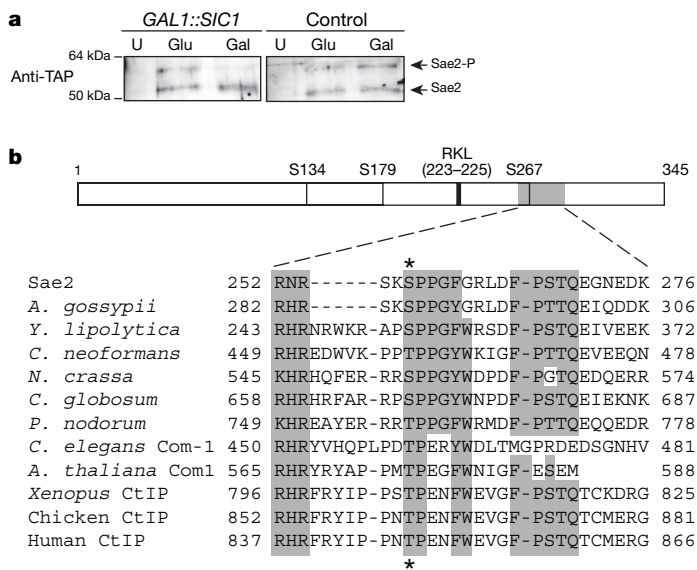
Sae2 contains three potential CDK phosphorylation sites (Fig. 1b and Supplementary Fig. 1); two of these—Ser267 and Ser134—received the highest scores for predicted phosphorylation sites in the protein (Supplementary Table 1). Ser267 maps to the Sae2 region

most highly conserved with its non-yeast orthologues (Fig. 1b), which include human CtIP, *Caenorhabditis elegans* Com1 and *Arabidopsis thaliana* Com1 (refs 16–18). To address the possible function(s) of Ser267 and other potential target sites for CDK in Sae2, we generated yeast strains in which each site was individually mutated to a non-phosphorylatable alanine residue. The *sae2-S267A* mutant showed strong hypersensitivity towards the topoisomerase I inhibitor camptothecin (Fig. 1c and Supplementary Fig. 2b; Supplementary Fig. 2c shows that this mutant is nearly as sensitive as the *sae2Δ* strain). By contrast, *sae2-S134A*, *sae2-S179A* and *sae2-S134A,S179A* cells did not show detectable hypersensitivity to camptothecin, and combining *sae2-S267A* with these other mutations showed no synergistic effect (Supplementary Fig. 2b). When we mutated Ser267 to glutamic acid to mimic constitutive phosphorylation, the resulting strain displayed no detectable hypersensitivity to camptothecin (Fig. 1c and Supplementary Fig. 2c). Together with the fact that mutation of Ser267 did not alter Sae2 protein expression (Fig. 1d), these data suggested that Ser267 phosphorylation is required for Sae2 function. Furthermore, as the major cytotoxic lesions for camptothecin are DSBs arising when replication forks encounter trapped topoisomerase I–DNA complexes<sup>19</sup>, these results suggested that phosphorylation of Sae2 on Ser267 is important for responses to DSBs generated during S phase. Indeed, the *sae2*-null and *sae2-S267A* strains, but not the *sae2-S267E* strain, showed hypersensitivity to methyl methanesulphonate (MMS), which also yields DSBs in S phase (Supplementary Fig. 2d). Consistent with analogous residues controlling the activity of Sae2-related proteins in other species, human U2OS cells downregulated for endogenous CtIP and expressing short interfering RNA (siRNA)-resistant GFP–CtIP–T847A were as sensitive to camptothecin as control cells expressing GFP (green fluorescent protein) alone, whereas cells expressing a phospho-mimicking CtIP derivative (GFP–CtIP–T847E) showed higher resistance to camptothecin (Fig. 1e; for expression levels and downregulation see Supplementary Fig. 2e).

Efficient phosphorylation of CDK substrates *in vivo* often requires the binding of cyclin to an Arg-X-Leu (RXL) motif in the target<sup>20</sup>. Such a motif is present upstream of Ser267 in Sae2 (Fig. 1b and Supplementary Fig. 1), and mutating this motif (*Sae2-R223A,L225A*) caused hypersensitivity to camptothecin as severe as that of *sae2-S267A* or *sae2Δ* cells (Fig. 1c) even though the mutated proteins were expressed at normal levels (Fig. 1d). Furthermore, the camptothecin hypersensitivity caused by the *Sae2-R223A,L225A* mutation was largely suppressed when the protein also contained the phospho-mimicking Sae2 S267E mutation (Fig. 1c). Collectively, these findings strongly suggested that Sae2 function requires its modification by CDK–cyclin complexes.

To examine the phosphorylation of Sae2 on Ser267 directly, we raised a phosphospecific antibody against this site ( $\gamma$ S267). Western

<sup>1</sup>The Wellcome Trust and Cancer Research UK Gurdon Institute, and Department of Zoology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK. <sup>2</sup>Centro Andaluz de Biología Molecular y Medicina Regenerativa CABIMER, Universidad de Sevilla-CSIC, Avenida Américo Vespucio s/n, 41092 Sevilla, Spain. †Present address: Institute of Molecular Cancer Research, University of Zurich-Irchel, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.



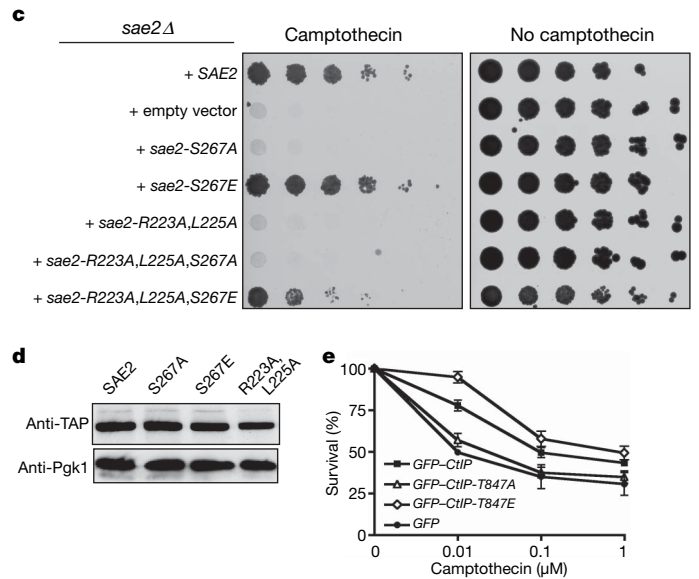
**Figure 1 | Ser 267 mutation impairs Sae2 function.** **a**, Left: TAP-tagged Sae2 was purified from cells expressing galactose-inducible *SIC1* (Gal) or not expressing *SIC1* (Glu). U, control untagged strain. Right: as above, but the strain lacked galactose-inducible *SIC1*. **b**, Sae2 diagram and homology to orthologues (see Methods for full alignment). *S. cerevisiae*, *Saccharomyces cerevisiae*; *A. gossypii*, *Ashbya gossypii*; *Y. lipolytica*, *Yarrowia lipolytica*; *C. neoformans*, *Cryptococcus neoformans*; *N. crassa*, *Neurospora crassa*; *C. globosum*, *Chaetomium globosum*; *P. nodorum*, *Phaeosphaeria nodorum*;

immunoblotting revealed that this antibody specifically detected immunoprecipitated wild-type Sae2 but not the Sae2-S267A or Sae2-R223A,L225A proteins (Fig. 2a). Furthermore, the antibody detected the slower-migrating form of Sae2 that was present at elevated levels in G2-synchronized cultures (Fig. 2b), indicating that phosphorylation of Sae2 on Ser 267 is subject to cell-cycle control. Notably, Sae2 immunoprecipitation recovered cyclins Clb3 and Clb2 from extracts prepared from G2-synchronized cells but not from G1 cells (Fig. 2b). Also consistent with Sae2's being a direct target of CDK, incubation of purified glutathione S-transferase-fused Sae2 protein—but not the S267A mutant—with recombinant CDK–cyclin complexes and ATP produced Ser 267 phosphorylation (Fig. 2c).

Further analyses suggested that all aspects of Sae2 function require phosphorylation at Ser 267. Thus, like *sae2Δ* cells, *sae2-S267A* cells were severely compromised in hairpin-induced recombination<sup>21</sup>, whereas *sae2-S267E* cells behaved similarly to the wild type (Fig. 2d). Furthermore, whereas homozygous diploid *sae2-S267E* cells produced viable spores at levels similar to those produced by the wild-type strain, like *sae2Δ* cells<sup>10,12</sup>, the *sae2-S267A* homozygous mutant strain showed a severe sporulation defect and almost no spore viability (Figs 2e and 2f). A defect in spore viability due to mutations in Ser 267 or Pro 268 has also been reported recently<sup>18</sup>.

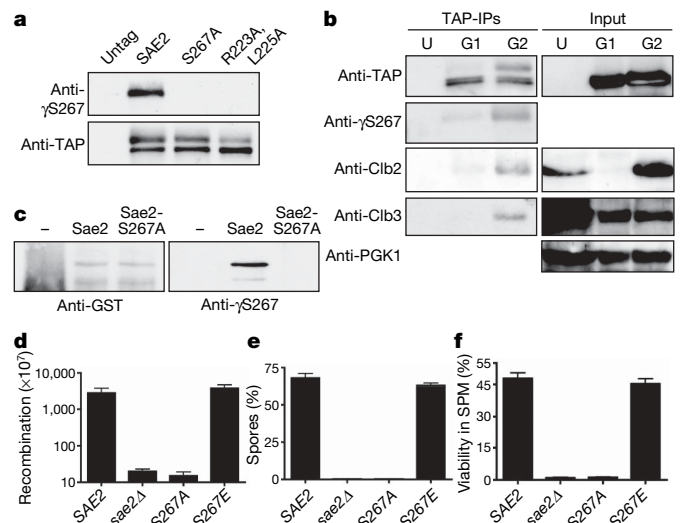
Sae2 regulates resection of chromosomal DSBs formed by the HO endonuclease<sup>9</sup>. To determine whether Sae2 Ser 267 controls this function, we generated an irreparable HO-induced DSB in the *MAT* locus and analysed resulting samples with a neutral dot-blot approach that detected only resected DNA<sup>22</sup>. For this we used three probes: one directly adjacent to the HO cleavage site, one 5 kilobases (kb) downstream and one at the *LEU2* locus, about 100 kb distal from the HO site, as a negative control. Significantly, *sae2Δ* cells and cells bearing the *sae2-S267A* mutation were impaired in resection close to the HO site, and this impairment was even more pronounced when assayed 5 kb away (Fig. 3a). By contrast, the *sae2-S267E* mutant carried out resection almost as efficiently as the wild-type strain (Fig. 3a and Supplementary Fig. 3a).

To test whether Cdc28/Cdk1-mediated Sae2 phosphorylation promotes DSB resection, we assessed resection in a strain expressing a

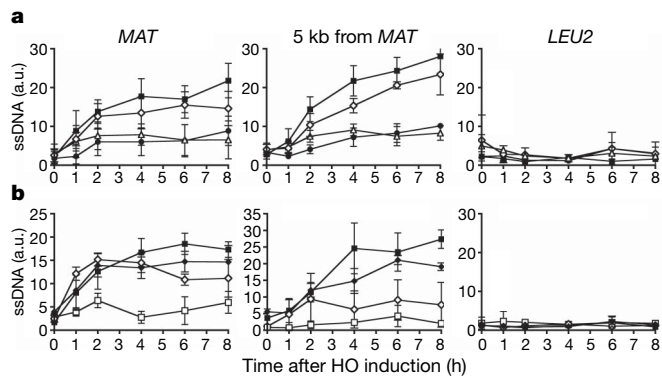


*C. elegans*, *Caenorhabditis elegans*; *A. thaliana*, *Arabidopsis thaliana*; *Xenopus*, *Xenopus laevis*; chicken, *Gallus gallus*; human, *Homo sapiens*. **c**, Fivefold serial dilutions of *sae2Δ* cultures containing the indicated SAE2 variants plated on medium lacking or containing camptothecin (5 μg ml<sup>-1</sup>). **d**, Extracts of cells harbouring TAP-tagged Sae2 variants were western immunoblotted as indicated. **e**, Survival of U2OS cells expressing GFP–CtIP fusions to 1 h treatments with the indicated doses of camptothecin. Error bars indicate s.d. (*n* = 2).

*Cdc28* derivative (*cdc28-as1*) that can be specifically inhibited by the ATP analogue 1NM-PP1 (ref. 23). Whereas inhibition of Cdk1 markedly curtailed end resection in a strain expressing wild-type Sae2, it had little effect in the *sae2-S267E* mutant when resection was measured close to the HO site (Fig. 3b and Supplementary Fig. 3b). Nevertheless, the inhibitor still had some effect on the *sae2-S267E* strain when resection was assessed at the 5 kb distal site (Fig. 3b and



**Figure 2 | Sae2 is phosphorylated by Cdc28 on Ser 267.** **a**, TAP-tagged Sae2 derivatives were immunoprecipitated and detected as indicated. **b**, TAP-tagged Sae2 was purified from G1 or G2 arrested cultures. U, G2 arrested untagged control cells. Immunoprecipitated samples and inputs (5%) were immunoblotted as indicated. **c**, Glutathione S-transferase (GST)-fused Sae2 and Sae2-S267A were purified, incubated with recombinant Cdk2/Cyclin A and ATP, resolved by 10% SDS–PAGE and immunoblotted as indicated. **d**, Recombination frequencies of strains in a hairpin-containing recombination system<sup>22</sup>. **e**, Spores after 24 h in sporulation medium. **f**, Spore viability 24 h after the addition of sporulation medium (SPM)<sup>10</sup>. Error bars in **d–f** represent s.d. (*n* = 2).



**Figure 3** | DNA-end resection is controlled by Sae2. **a**, Resection-mediated ssDNA formation at an HO DSB in wild-type *SAE2* (filled squares), *sae2-S267A* (open triangle), *sae2-S267E* (open diamonds) or empty vector (solid circles) at indicated times after HO induction at the *MAT* locus (left), 5 kb downstream of *MAT* (centre) or *LEU2* locus (right). **b**, Wild-type *SAE2* (squares) or *sae2-S267E* (diamonds) strains containing *Cdc28-as1* were grown as in **a** but in the presence of dimethylsulphoxide (filled symbols) or 1NM-PP1 (open symbols). Results are shown as means  $\pm$  s.d. ( $n = 5$ ).

Supplementary Fig. 3b). We therefore conclude that phosphorylation of Sae2 on Ser267 by Cdc28/Cdk1 is required for effective DSB resection but that additional Cdk1 target sites are required for resection to take place optimally. One candidate for such an additional CDK target is Rad9, which is phosphorylated by Cdc28 and was recently shown to affect DSB resection<sup>24</sup>.

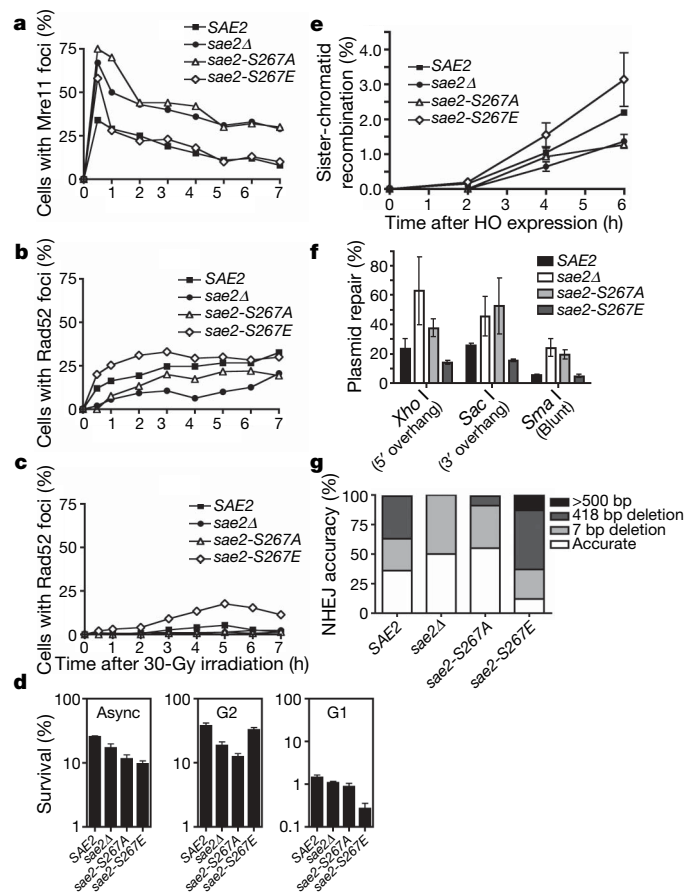
*S. cerevisiae* Mre11 is recruited quickly to DSB sites, then replaced by the HR protein Rad52 as ssDNA is formed in S and G2 cells<sup>25</sup>. We found that, like *SAE2* deletion<sup>25</sup>, the *sae2-S267A* mutation caused Mre11 foci to persist longer than in wild-type cells (Fig. 4a) and delayed Rad52 focus formation in S/G2 after X-ray treatment (Fig. 4b). In contrast, the *sae2-S267E* mutant displayed Mre11 focus disassembly kinetics similar to the wild-type strain (Fig. 4a), and in fact reproducibly formed Rad52 foci faster in S and G2 than the wild-type strain (Fig. 4b). As expected, essentially no Rad52 foci were detected in wild-type G1 cells, because DSBs are not efficiently resected. By contrast—and unlike the *sae2-S267A* mutant—the *sae2-S267E* mutant formed Rad52 foci in G1 at later time points (Fig. 4c), confirming that DNA ends are processed to some degree in this mutant even in the absence of active CDK (Fig. 3b).

To address the impact of aberrant DSB processing on DSB repair, we irradiated asynchronous, G2-arrested and G1-arrested cell cultures, then kept cells in the same state (asynchronous, G2 or G1) for 6 h to allow DNA repair. Cells were subsequently plated and colony formation was used to determine survival. When we analysed asynchronous cultures, the *sae2Δ* strain showed moderate hypersensitivity to radiation (as shown previously<sup>26</sup>), and similar, although slightly more pronounced, hypersensitivities were displayed by the *sae2-S267A* and *sae2-S267E* strains (Fig. 4d). Moreover, the two Ser267 mutations had markedly different effects in G1 and G2 (Fig. 4d). Thus, whereas the *sae2-S267A* strain was more sensitive to radiation than the control strain in G2, little or no G2 hypersensitivity was shown by the *sae2-S267E* strain. In contrast, *sae2-S267E* cells showed marked hypersensitivity to radiation in G1, whereas the *sae2*-null and *sae2-S267A* cells did not.

The above data suggested that the *sae2-S267A* and *sae2*-null strains, but not the *sae2-S267E* strain, are defective in HR because of impaired DSB resection in G2, and also suggested that the hypersensitivity of the *sae2-S267E* mutant to radiation in G1 reflects aberrant DSB resection, thus impairing NHEJ and/or triggering futile attempts to carry out HR in the absence of a sister chromatid. To test these ideas, we determined HR and NHEJ efficiencies in various *sae2* mutant backgrounds. When we used an assay in which HR intermediates were monitored by Southern blot analysis<sup>27</sup>, the *sae2*-null and *sae2-S267A* mutants showed delayed HR, whereas the *sae2-S267E* mutant showed slightly accelerated recombination (Fig. 4e and Supplementary Fig. 4).

Furthermore, by measuring NHEJ with an *in vivo* plasmid-recircularization assay<sup>28</sup>, we found that cells lacking Sae2 or bearing the *S267A* mutation had enhanced ( $P < 0.05$  compared with wild type) NHEJ efficiencies, regardless of whether the DSB contained a 5' overhang, a 3' overhang or a blunt end (Fig. 4f). In addition, and in agreement with an increased propensity for DSB resection that would impair NHEJ, *sae2-S267E* mutant cells showed a decrease in NHEJ efficiency when overhang substrates were used ( $P < 0.05$ ; Fig. 4f). By retrieving repaired plasmids from independent clones and sequencing them, we found that in wild-type cells most repair took place accurately or by micro-homology-mediated end-joining involving pairs of 4-bp repeats separated by 7 or 418 bp to create small or moderate deletions (Fig. 4g). In agreement with less efficient resection taking place in the *sae2*-null and *sae2-S267A* mutant strains, such cells favoured accurate repair or repair involving small (7-bp) deletions ( $P < 10^{-8}$ ; Fig. 4g; similar data were reported for *sae2Δ* cells<sup>29</sup>). In contrast, *sae2-S267E* mutant cells showed little accurate NHEJ and, instead, most repair products contained larger deletions of up to 2 kb ( $P < 10^{-4}$ ; Fig. 4g).

Thus, Cdc28/Cdk1-mediated Sae2 phosphorylation modulates the balance between NHEJ and HR during the cell cycle. These results lend strong support to models in which the commitment to DSB resection is highly regulated to ensure that the cell engages the most



**Figure 4** | Sae2 mutations affect Mre11 and Rad52 dynamics, and DSB repair. **a, b**, Percentages of S/G2 cells containing Mre11 (**a**) or Rad52 (**b**) foci. **c**, Percentage of G1 cells containing a Rad52 focus. **d**, Survival of *sae2Δ* mutants containing wild-type *SAE2*, *sae2-S267A*, *sae2-S267E* or empty vector grown asynchronously (Async) or arrested in G1 or G2, after irradiation with 300 Gy. Error bars represent s.d. ( $n = 2$ ). **e**, Sister-chromatid recombination measured as described previously<sup>27</sup>. Standard deviations of two independent experiments are shown (see Supplementary Fig. 4 for details and representative blot). **f**, Plasmid cleaved by *Xho*I, *Sac*I or *Sma*I was transformed into strains and NHEJ efficiency was measured. Means and s.d. of three independent experiments are shown. **g**, Classes of plasmid rejoining products from 50 independent clones of each strain transformed with the *Xho*I-cut plasmid.

appropriate DNA repair pathway, thereby optimizing genome stability. As Sae2 has endonuclease activity<sup>13</sup>, we favour a model in which Sae2, possibly in cooperation with the Mre11–Rad50–Xrs2 (MRX) complex, facilitates resection in S/G2 by mediating an endonucleolytic cleavage close to the DNA break, thus generating a clean end that can serve as an efficient substrate for nucleases such as MRX and Exo1. Sae2 activity might be particularly important to initiate resection at DSBs that contain covalently bound proteins that would otherwise resist exonuclease action; indeed, this would explain why deletion of *SAE2* causes defective removal of Spo11–DNA adducts during meiosis and marked hypersensitivity to camptothecin. Sae2 might also initiate resection at radiation-induced DSBs that are resistant to exonucleases because they bear protein–DNA crosslinks or complex damage to bases at their termini. By contrast, at sites of clean DSBs, *SAE2* deletion would only slow down resection and ensuing HR, thus explaining why *sae2* mutants are not as sensitive to radiation as other HR mutants<sup>26</sup>. Finally, we note that the motif encompassing Ser 267 of Sae2 is highly conserved in Sae2 counterparts in higher eukaryotes, and that mutation of the analogous Thr 847 site in human CtIP to Ala (but not to Glu) yields hypersensitivity to camptothecin. This suggests that analogous CDK-control mechanisms for DSB resection operate in many other organisms. One exception to this, however, is likely to be provided by *Schizosaccharomyces pombe*, whose Sae2/CtIP homologue, Ctp1, lacks a CDK site analogous to Ser 267 of Sae2. In this case, it seems that, rather than controlling Ctp1 phosphorylation, the CDK machinery instead regulates the protein expression of Ctp1 (ref. 30). Nevertheless, although some species-specific variations undoubtedly exist, we speculate that Sae2/CtIP/Com1/Ctp1 proteins will turn out to have ubiquitous functions in facilitating DSB resection in S and G2 and modulating the choice of DSB repair pathway in eukaryotic cells.

## METHODS SUMMARY

A *sae2Δ* strain in W303 background<sup>9,14</sup> was transformed with plasmids harbouring the indicated *SAE2* mutant and used in all experiments except those listed below. For Figs 1a and 2b, a Sae2-TAP strain (Open Biosystems) was used. In Fig. 2e, f, a strain harbouring the indicated Sae2 mutant at its chromosomal locus in the SK1 background was used. For Fig. 3, we deleted *SAE2* in a strain harbouring the *cdc28as1* allele in the JKM179 background<sup>5</sup>. The W5573-15D strain was used in Fig. 4a–c (ref. 25). A *sae2*-deleted OIS-15 strain was used in Fig. 4e (ref. 27). Yeasts were grown with standard procedures. When indicated, cells were arrested in G1 with  $\alpha$ -factor and in G2 with nocodazole. DNA resection assays<sup>22</sup>, focus formation<sup>25</sup>, recombination between sister chromatids<sup>27</sup> and NHEJ assays<sup>30</sup> were as described previously. CtIP downregulation was as previously reported<sup>17</sup>. Western blotting was by standard methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 21 February; accepted 26 June 2008.

Published online 20 August 2008.

- Shrivastav, M., De Haro, L. P. & Nickoloff, J. A. Regulation of DNA double-strand break repair pathway choice. *Cell Res.* **18**, 134–147 (2008).
- Aylon, Y., Liefshitz, B. & Kupiec, M. The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. *EMBO J.* **23**, 4868–4875 (2004).
- Caspari, T., Murray, J. M. & Carr, A. M. Cdc2-cyclin B kinase activity links Crb2 and Rqh1-topoisomerase III. *Genes Dev.* **16**, 1195–1208 (2002).
- Hinz, J. M., Yamada, N. A., Salazar, E. P., Tebbs, R. S. & Thompson, L. H. Influence of double-strand-break repair pathways on radiosensitivity throughout the cell cycle in CHO cells. *DNA Repair (Amst.)* **4**, 782–792 (2005).
- Ira, G. *et al.* DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature* **431**, 1011–1017 (2004).
- Karathanasis, E. & Wilson, T. E. Enhancement of *Saccharomyces cerevisiae* end-joining efficiency by cell growth stage but not by impairment of recombination. *Genetics* **161**, 1015–1027 (2002).
- Esashi, F. *et al.* CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature* **434**, 598–604 (2005).
- Aylon, Y. & Kupiec, M. DSB repair: the yeast paradigm. *DNA Repair (Amst.)* **3**, 797–815 (2004).

- Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. The *Saccharomyces cerevisiae* Sae2 protein promotes resection and bridging of double strand break ends. *J. Biol. Chem.* **280**, 38631–38638 (2005).
- McKee, A. H. & Kleckner, N. A general method for identifying recessive diploid-specific mutations in *Saccharomyces cerevisiae*, its application to the isolation of mutants blocked at intermediate stages of meiotic prophase and characterization of a new gene *SAE2*. *Genetics* **146**, 797–816 (1997).
- Neale, M. J., Pan, J. & Keeney, S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053–1057 (2005).
- Prinz, S., Amon, A. & Klein, F. Isolation of *COM1*, a new gene required to complete meiotic double-strand break-induced recombination in *Saccharomyces cerevisiae*. *Genetics* **146**, 781–795 (1997).
- Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Mol. Cell* **28**, 638–651 (2007).
- Baroni, E., Viscardi, V., Cartagena-Lirola, H., Lucchini, G. & Longhese, M. P. The functions of budding yeast Sae2 in the DNA damage response require Mec1- and Tel1-dependent phosphorylation. *Mol. Cell Biol.* **24**, 4151–4165 (2004).
- Mendenhall, M. D. & Hodge, A. E. Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **62**, 1191–1243 (1998).
- Penkner, A. *et al.* A conserved function for a *Caenorhabditis elegans* Com1/Sae2/CtIP protein homolog in meiotic recombination. *EMBO J.* **26**, 5071–5082 (2007).
- Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450**, 509–514 (2007).
- Uanschou, C. *et al.* A novel plant gene essential for meiosis is related to the human CtIP and the yeast *COM1/SAE2* gene. *EMBO J.* **26**, 5061–5070 (2007).
- Pommier, Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nature Rev. Cancer* **6**, 789–802 (2006).
- Chen, J., Saha, P., Kornbluth, S., Dynlacht, B. D. & Dutta, A. Cyclin-binding motifs are essential for the function of p21<sup>CIP1</sup>. *Mol. Cell Biol.* **16**, 4673–4682 (1996).
- Lobachev, K. S., Gordenin, D. A. & Resnick, M. A. The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **108**, 183–193 (2002).
- Sugawara, N. & Haber, J. E. Repair of DNA double strand breaks: *in vivo* biochemistry. *Methods Enzymol.* **408**, 416–429 (2006).
- Bishop, A. C. *et al.* A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407**, 395–401 (2000).
- Lazzaro, F. *et al.* Histone methyltransferase Dot1 and Rad9 inhibit single-stranded DNA accumulation at DSBs and uncapped telomeres. *EMBO J.* **27**, 1502–1512 (2008).
- Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699–713 (2004).
- Rattray, A. J., McGill, C. B., Shafer, B. K. & Strathern, J. N. Fidelity of mitotic double-strand-break repair in *Saccharomyces cerevisiae*: a role for *SAE2/COM1*. *Genetics* **158**, 109–122 (2001).
- Cortes-Ledesma, F. & Aguilera, A. Double-strand breaks arising by replication through a nick are repaired by cohesin-dependent sister-chromatid exchange. *EMBO Rep.* **7**, 919–926 (2006).
- Boulton, S. J. & Jackson, S. P. *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.* **15**, 5093–5103 (1996).
- Lee, K. & Lee, S. E. *Saccharomyces cerevisiae* Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes microhomology-mediated end joining. *Genetics* **176**, 2003–2014 (2007).
- Limbo, O. *et al.* Ctp1 is a cell-cycle-regulated protein that functions with Mre11 complex to control double-strand break repair by homologous recombination. *Mol. Cell* **28**, 134–146 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. P. Longhese, R. Rothstein, K. Lobachev, M. Lichten and M. Foiani for providing strains, and R. Driscoll, S. Gravel, K. Dry and K. Miller for helpful discussions and comments on the manuscript. P.H. is the recipient of a Long-Term EMBO Fellowship. A.A.S. is supported by a Swiss National Foundation Grant. The S.P.J. laboratory is supported by grants from Cancer Research UK and the European Community (Integrated Project DNA repair, grant LSHG-CT-2005-512113). The A.A. laboratory is supported by grants from the Spanish Ministry of Science and Education (BFU2006-05260 and CDS2007-0015) and Junta de Andalucía (CVI624).

**Author Contributions** A.A.S. identified the homology between Sae2 and CtIP, cloned *SAE2* into pGEX-4T1 and made the original *sae2-S267A* and *sae2-S267E* mutations. All the experiments shown were performed by P.H. and were conceived by P.H. and S.P.J., except those on SCR analyses that were performed by F.C.-L. and conceived by F.C.-L. and A.A. P.H. and S.P.J. wrote the paper. All authors discussed and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to S.P.J. ([s.jackson@gurdon.cam.ac.uk](mailto:s.jackson@gurdon.cam.ac.uk)).

## METHODS

**TAP-tagged Sae2 immunoprecipitation.** TAP complexes were purified by a variation of previously described methods<sup>31</sup>. Cultures (250 ml) of TAP-tagged Sae2 variants were collected by centrifugation at 4 °C and resuspended in 1 volume of 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.15% Nonidet P40 in the presence of protease inhibitor (Roche) and phosphatase inhibitors. Extracts were prepared with a One-Shot cell disruptor (Constant Systems) and centrifuged for 1 h at 3,000 r.p.m. (1,400g) and 4 °C. Next, samples were incubated for 2 h at 4 °C with IgG-Sepharose (Amersham) pre-equilibrated in the same buffer. The matrix was then packed in a column and washed with 50 ml of the same buffer at 4 °C. Next, the resin was resuspended in 100 µl of 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% Nonidet P40, 0.5 mM EDTA, 1 mM dithiothreitol, transferred to a microcentrifuge tube, incubated for 2 h at 16 °C and then overnight at 4 °C with 10 U of TEV protease (Qiagen) to release Sae2 complexes from the beads. Through this procedure, Sae2 retained half of the TAP tag that could then be detected with anti-TAP antibody (Open Biosystems). Samples were centrifuged for 1 min at maximum speed (2,900g) at 4 °C; the supernatant was transferred to a new tube and 100 µl of sample loading buffer was added followed by immunoblot analysis by SDS-PAGE with the following antibodies: anti-TAP, anti-γS267 (custom made; Eurogentec), anti-PGK1 (Molecular Probes), Cdc28, Clb2 and Clb3 (Santa Cruz).

**Human cell survival assays.** Human U2OS cells expressing siRNA-resistant wild-type or mutant GFP-CtIP fusions were downregulated for endogenous CtIP with a previously published siRNA<sup>17</sup>, and 72 h afterwards were exposed to doses of camptothecin for 1 h. Survivals represent the number of colonies formed after 12 days normalized with an unirradiated control.

**Sporulation efficiency.** Homozygous diploids were grown overnight in YPAD medium, washed twice with warm sporulation medium, left in sporulation medium for 24 h at 30 °C, then fixed with 50% ethanol and stained with 4,6-diamidino-2-phenylindole (DAPI). The percentage of sporulated cells was determined by microscopy<sup>10</sup>.

**DNA-end resection assay.** Cultures of *sae2Δ cdc28-as1 GAL1::HO* strain transformed with wild-type *SAE2*, *sae2-S267A*, *sae2-S267E* or empty vector were grown to mid-exponential phase in raffinose. Samples were taken at indicated times after inducing HO by the addition of galactose. DNA was isolated, of which 1 mg was blotted in neutral and denaturing conditions with a dot-blot manifold as described previously<sup>23</sup>, then hybridized with radioactively labelled probes against the *MAT* locus, 5 kb downstream of the *MAT* locus or *LEU2* locus. Signals were quantified with a FLA-5000 instrument (Fuji) and values obtained in neutral conditions were normalized to those obtained under denaturing conditions (see Supplementary Fig. 3a). When indicated, 2 h before the addition of galactose, the culture was split in two and dimethylsulphoxide or Cdc28-as1 inhibitor 1NM-PP1 (5 µM final concentration; Calbiochem) was added.

**Rad52 and Mre11 foci analyses.** Mre11-YFP Rad52-RFP *sae2Δ* strain transformed with *SAE2*, *sae2-S267A*, *sae2-S267E* or empty vector was irradiated (30 Gy) with a Faxitron (Faxitron X-ray Corporation). Samples were taken, fixed by the addition of 0.1 volume of formaldehyde, washed three times with PBS, sonicated for 10 s and mixed 1:1 with DAPI-containing mounting medium (Vector Laboratories Inc.). Microscopy was with a DeltaVision microscope (Applied Precision). A minimum of 50 G1 (unbudded) and 50 S/G2 (budded) cells were counted at each time point and for each sample.

**Survival of irradiation.** *sae2Δ* mutants transformed with wild-type *SAE2*, *sae2-S267A*, *sae2-S267E* or an empty vector were grown asynchronously or were arrested with  $\alpha$ -factor (G1) or nocodazole (G2); they were then irradiated with 300 Gy (Faxitron), kept for 6 h in the same cell-cycle stage and then plated. Colonies arising were normalized with respect to non-irradiated samples and plotted.

**NHEJ assays.** A pRS416 vector restricted with *Xho*I, *Sac*I or *Sma*I was transformed into cells harbouring various *sae2* mutations. The number of colonies formed after 3 days was normalized with the number of colonies obtained in a parallel transformation with a circular pRS416 plasmid. Plasmids from 50 independent clones of each strain transformed with a *Xho*I-restricted plasmid as described previously were isolated and sequenced.

31. Puig, O. *et al.* The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**, 218–229 (2001).



# Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease

Chun Tang<sup>1</sup>†, John M. Louis<sup>1</sup>, Annie Aniana<sup>1</sup>, Jeong-Yong Suh<sup>1</sup> & G. Marius Clore<sup>1</sup>

HIV-1 protease processes the Gag and Gag-Pol polyproteins into mature structural and functional proteins, including itself, and is therefore indispensable for viral maturation<sup>1,2</sup>. The mature protease is active only as a dimer<sup>3–5</sup> with each subunit contributing catalytic residues<sup>6</sup>. The full-length transframe region protease precursor appears to be monomeric yet undergoes maturation via intramolecular cleavage of a putative precursor dimer<sup>5,7–11</sup>, concomitant with the appearance of mature-like catalytic activity<sup>7,9</sup>. How such intramolecular cleavage can occur when the amino and carboxy termini of the mature protease are part of an intersubunit  $\beta$ -sheet located distal from the active site is unclear. Here we visualize the early events in N-terminal autoprocessing using an inactive mini-precursor with a four-residue N-terminal extension that mimics the transframe region protease precursor<sup>5,12</sup>. Using paramagnetic relaxation enhancement, a technique that is exquisitely sensitive to the presence of minor species<sup>13–16</sup>, we show that the mini-precursor forms highly transient, lowly populated (3–5%) dimeric encounter complexes that involve the mature dimer interface but occupy a wide range of subunit orientations relative to the mature dimer. Furthermore, the occupancy of the mature dimer configuration constitutes a very small fraction of the self-associated species (accounting for the very low enzymatic activity of the protease precursor), and the N-terminal extension makes transient intra- and intersubunit contacts with the substrate binding site and is therefore available for autocleavage when the correct dimer orientation is sampled within the encounter complex ensemble.

The regulation of HIV-1 protease autoprocessing is modulated by the N-terminal flanking transframe region (TFR) sequence (Fig. 1a)<sup>2</sup>. The catalytic activity of the monomeric protease precursor is approximately three orders of magnitude less than that of the mature protease dimer (which has a monomer–dimer equilibrium dissociation constant  $K_d < 10$  nM)<sup>2,5</sup>. The appearance of mature-like catalytic activity and stable dimer formation is directly correlated with a single rate-limiting step comprising intramolecular (first order) cleavage of a putative transient dimeric precursor species at the p6<sup>pol</sup>–protease (PR) junction<sup>7,9,10</sup>. Mutations within the latter that prevent cleavage lead to the production of an N-terminally extended 17-kDa protease precursor species, and cause a severe defect in Gag polyprotein processing and the complete loss of viral infectivity *in vivo*<sup>17,18</sup>. Subsequent cleavage at the C terminus of protease at the PR–reverse transcriptase (RT) junction (Fig. 1a) occurs via an intermolecular (second order) reaction catalysed by a fully active protease dimer<sup>19</sup>. Mutations within the PR–RT junction that block C-terminal cleavage do not significantly affect either enzymatic activity and dimerization of the protease *in vitro*<sup>19,20</sup> or processing of HIV-1 precursor proteins, virus maturation, viability and morphology *in vivo*<sup>20</sup>, indicating that the presence of the C-terminal reverse transcriptase sequence has negligible influence on the protease precursor<sup>19,20</sup>. Thus, only autoprocessing at the N

terminus of protease at the p6<sup>pol</sup>–PR junction is an absolute prerequisite for stable protease dimer formation, the appearance of mature catalytic activity and complete processing of viral precursors. Before cleavage at the p6<sup>pol</sup>–PR junction, intermediate precursor forms may be liberated by intramolecular cleavage at competing sites (for example, p2–NC and TFP–p6<sup>pol</sup>; see Fig. 1a) that become available for productive binding and hydrolysis<sup>11</sup>, but these precursors will show the same low catalytic activity as that of the p6<sup>pol</sup>–PR precursor<sup>9,10</sup>.

As little as a four-residue extension at the N terminus of protease, corresponding to the C-terminal residues of p6<sup>pol</sup>, in conjunction with a D25N mutation result in an effectively monomeric species<sup>5,12</sup>. Disruption of the native protease dimer by N-terminal extension is due to removal of the protons on the secondary amine of the N-terminal proline residue, disrupting the interstrand hydrogen bond between the amine of the N-terminal proline of one subunit and the C-terminal carbonyl oxygen of the second subunit<sup>6</sup>. C-terminal extension, however, does not have an impact on this interstrand hydrogen bond because the secondary amine of Pro 1 is preserved. Therefore, we made use of the mini-precursor, bearing only the N-terminal cleavage site, to visualize the early transient events involved in autoprocessing of the protease at the p6<sup>pol</sup>–PR junction that is required for the formation of a fully active, stable protease dimer.

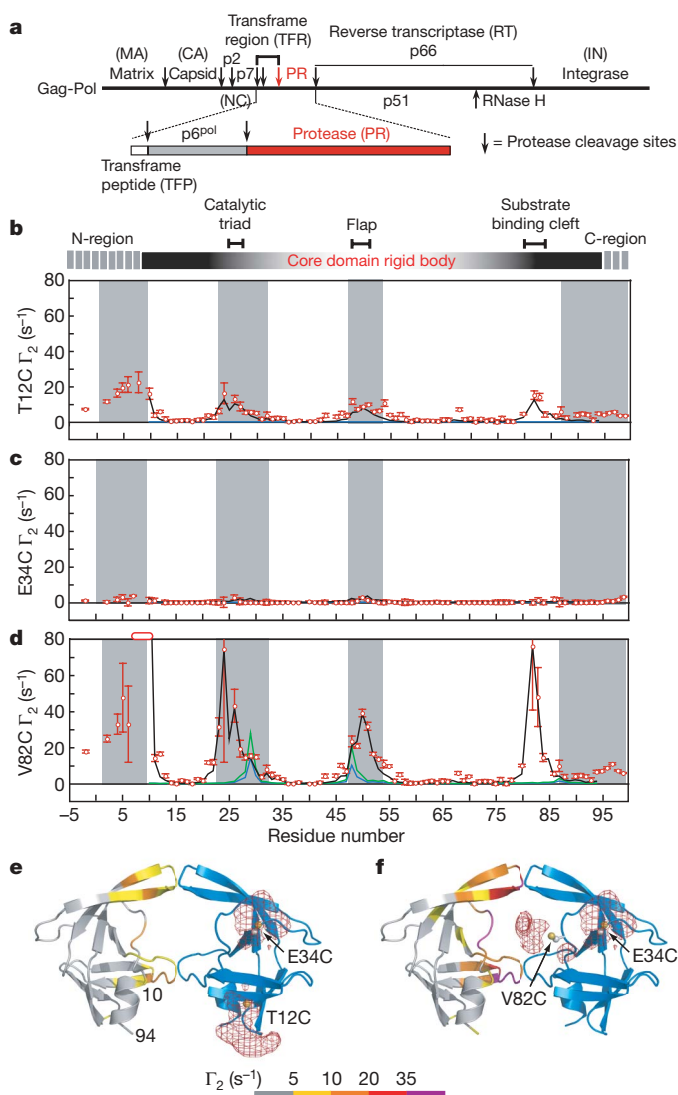
The optimized mini-precursor protease construct <sup>SFNF</sup>PR(D25N) comprises a four-residue N-terminal extension (Ser-Phe-Asn-Phe) derived from the TFR (Fig. 1a), a D25N mutation to abolish all residual catalytic activity, and C67A and C95A mutations to remove surface cysteines (Supplementary Fig. 1a)<sup>9,10,12</sup>. The corresponding active <sup>SFNF</sup>PR(D25) mini-precursor construct undergoes autoprocessing during expression to release the mature protease (see Methods). NMR analysis of <sup>SFNF</sup>PR(D25N) shows that it is monomeric (with an upper limit of  $\sim 10\%$  dimer from translational diffusion measurements); the secondary and tertiary structures of the mature protease are preserved with the exception of the N- and C-terminal strands which form an intersubunit four-stranded anti-parallel  $\beta$ -sheet in the mature dimer; and residues –4 to 9 and 95–99 are disordered and highly mobile (see Methods and Supplementary Fig. 1b–e).

Because enzymatically active protease is dimeric, and the rate-limiting step in autoprocessing is unimolecular<sup>7,9</sup>, transient self-association of the precursor must occur to initiate autoprocessing. To visualize this phenomenon we measured intermolecular paramagnetic relaxation enhancements (PREs) by introducing a spin label via conjugation to three engineered surface-exposed cysteine residues: T12C, E34C and V82C (one at a time). These sites are frequently mutated in viable HIV-1 variants<sup>2</sup>. T12C and V82C are located at the periphery of the substrate-binding cleft in the mature dimer, whereas E34C is relatively far removed from the dimer interface (Fig. 1e, f). In a rapidly exchanging system, the PRE  $^1\text{H}_N\text{-}\Gamma_2$  rates<sup>21</sup> are population-weighted averages of the PRE rates of the species present<sup>13,14</sup>. Because the PRE rate

<sup>1</sup>Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA. †Present address: Department of Biochemistry, University of Missouri, Columbia, Missouri 65211, USA.

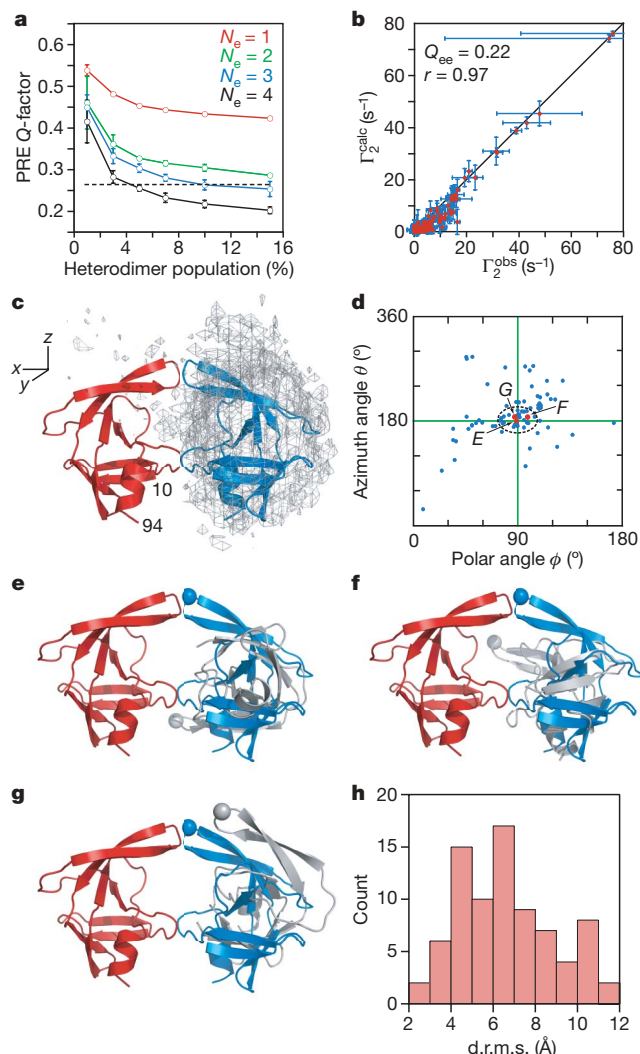
for a paramagnetic centre-proton pair is proportional to the  $\langle r^{-6} \rangle$  average of the distance between them, and the PRE effect is large owing to the high magnetic moment of an unpaired electron, the PRE in the fast exchange regime is very sensitive to the presence of lowly populated (<5%), highly transient species in solution providing there are paramagnetic centre-proton distances in the minor species that are shorter than in the predominant species<sup>13–15</sup>.

PREs were measured on a 1:1 mixture of 0.2 mM U- $[^2\text{H}/^{13}\text{C}/^{15}\text{N}]$ -labelled  $^{\text{SFNF}}\text{PR}(\text{D25N})$  and spin-labelled  $^{\text{SFNF}}\text{PR}(\text{D25N})$  at natural isotopic abundance. Because  $^1\text{H}_\text{N}$ - $\Gamma_2$  rates are measured using  $^1\text{H}$ - $^{15}\text{N}$  correlation-based experiments<sup>21</sup>, the observed  $^1\text{H}_\text{N}$ - $\Gamma_2$  rates arise solely from intermolecular interactions between the spin-labelled



**Figure 1 | Intermolecular PRE profiles.** **a**, Organization of the Gag-Pol polyprotein<sup>1,2</sup>. **b–d**, Intermolecular PREs observed on U- $[^2\text{H}/^{13}\text{C}/^{15}\text{N}]$ -labelled  $^{\text{SFNF}}\text{PR}(\text{D25N})$  originating from a spin label conjugated to T12C (**b**), E34C (**c**) and V82C (**d**) of  $^{\text{SFNF}}\text{PR}(\text{D25N})$  at natural isotopic abundance. Residues broadened beyond detection are denoted by open bars. Error bars represent 1 s.d.  $\Gamma_2$  rates back-calculated from the structure of the mature dimer (for the core residues 10–94) at populations of 1% and 2% are shown as blue and green lines, respectively. Average  $\Gamma_2$  rates derived from the top 20 structures of the  $N_e = 4$  simulated annealing calculations at a population of 5% heterodimer are shown as black lines. Grey shaded areas delineate residues that are buried at the dimer interface in the mature protease. **e**, **f**, Observed intermolecular PREs originating from the spin label attached to T12C (**e**) and V82C (**f**) colour-coded on a ribbon diagram of the mature dimer<sup>24</sup> (spin label attached to the blue subunit). Atomic probability density maps<sup>25</sup> (plotted at a threshold of 10% of maximum) showing the distribution of the spin-label oxygen radicals are shown as red meshes.

protein and the isotopically labelled protein (Fig. 1b–d). For the E34C spin label, no  $^1\text{H}_\text{N}$ - $\Gamma_2$  rates greater than  $5 \text{ s}^{-1}$  are observed (Fig. 1c); this sample therefore provides a negative control, excluding the existence of solvent PRE effects arising from diffusion and random elastic collisions, or from direct intermolecular interactions between the spin



**Figure 2 | Ensemble simulated annealing and the protease mini-precursor encounter complex ensemble.** **a**, PRE Q-factor as a function of ensemble size and population of heterodimer. Dashed line denotes the expected Q-factor when agreement between observed and calculated  $\Gamma_2$  rates is comparable to the experimental error in the measurements. **b**, Correlation between observed and calculated  $\Gamma_2$  rates for  $N_e = 4$  and a heterodimer population of 5%.  $Q_{\text{ee}}$  is the ensemble of ensemble average PRE Q-factor for the 20 calculated  $N_e = 4$  ensembles and  $r$  the correlation coefficient. Error bars in **a** and **b** represent 1 s.d. **c**, Atomic probability density map<sup>25</sup> (grey mesh, plotted at a threshold of 20% of maximum) showing the distribution of the spin-labelled subunit relative to the isotopically labelled subunit (red ribbon) in the  $^{\text{SFNF}}\text{PR}(\text{D25N})$  encounter complexes. The location of the second subunit in the mature dimer is shown as a blue ribbon. **d**, Orientations in spherical coordinates of the vector joining the centre of masses of the two interacting molecules in the encounter complexes relative to the coordinate system shown in **c** with the z axis corresponding to the  $C_2$  symmetry axis of the mature dimer. The  $\phi, \theta$  angles for the mature dimer are located at the crosshair. **e–g**, Representative encounter complexes (labelled and denoted by red dots in **d**) corresponding to the structures with the closest spherical angles (**e**), the smallest d.r.m.s. (**f**) and the smallest atomic r.m.s. displacement (**g**) relative to the mature dimer. The C $\alpha$  atom of Gly 51 at the tip of the flap is shown as a sphere to guide the eye. The isotopically labelled and spin-labelled subunits are shown in red and grey, respectively; the blue subunit corresponds to the orientation relative to the red subunit seen in the mature dimer. **h**, Histogram of the d.r.m.s. metric for the  $N_e = 4$  structures (total of  $20 \times 4 = 80$  conformers) at a population of 5% heterodimer.

label and the U- $^{2}\text{H}/^{13}\text{C}/^{15}\text{N}$ -labelled protein. The PRE profiles for the T12C (Fig. 1b) and V82C (Fig. 1d) spin labels are similar but the magnitude for the latter is 4- to 8- fold greater than for the former. Within the ordered core of the precursor (residues 10–94), large intermolecular PREs are observed for residues 21–30, 46–55 and 80–85 located at or close to the dimer interface. Residues 21–30 encompass the catalytic triad, residues 46–66 correspond to the flap region which gates the active site, and residues 80–81 and 83–84 are located in the substrate binding cleft (Fig. 1e, f). In addition, the N-terminal region experiences sizeable PREs from the T12C (Fig. 1b) and V82C (Fig. 1d) spin labels. These data demonstrate that transient self-association of the precursor involves residues located at the dimer interface in the mature dimer. A similar intermolecular PRE profile is observed from V82C spin-labelled, full-length TFR-PR(D25N) precursor to U- $^{2}\text{H}/^{13}\text{C}/^{15}\text{N}$ - $^{\text{SFNF}}$ PR(D25N), indicating that the transient dimerization interface is preserved on further N-terminal extension of the protease precursor (Supplementary Fig. 2a).

Back-calculation of the PREs from the structure of the mature dimer shows that almost zero PRE values are expected for the T12C and E34C spin labels at a population of 1–2% mature heterodimer (Fig. 1b, c). For the V82C label, small PRE values at a population of 1–2% mature heterodimer are predicted for residues 27–30 and 48–50 (Fig. 1d, blue line). The mature dimer does not predict the large observed PRE values observed for residues 20–26, 30–35 and 80–83. Furthermore, in the mature dimer residues 80–83 of one subunit are located on the opposite side of the dimer interface from residues 80–83 of the other subunit, and thus the large intermolecular PREs observed from the V82C spin label to residues 80–83 would require a  $\sim 180^\circ$  rotation of one subunit relative to its position in the mature dimer. Thus, the upper limit of the total population of mature dimer (heterodimer and homodimer) cannot exceed 2–4%.

Transient interactions between  $^{\text{SFNF}}$ PR(D25N) precursor monomers were visualized semi-quantitatively using rigid-body simulated annealing calculations<sup>14,16,22</sup> to optimize the agreement between observed and calculated  $\Gamma_2$  rates arising from the T12C, E34C and V82C spin labels simultaneously (see Methods). The flexible N- and C-terminal regions (residues –4 to 9 and 95–99, respectively) were excluded from the calculations. A single conformer representation ( $N_e = 1$ ) for the transient dimer does not account for the PRE data and even at a heterodimer population of 15% the PRE Q-factor<sup>23</sup> (see Methods for definition) has a value of greater than 0.4 (Fig. 2a). Thus, the dimeric  $^{\text{SFNF}}$ PR(D25N) precursor is an ensemble of multiple encounter complexes. For  $N_e \geq 2$ , the average PRE Q-factor decreases rapidly as the heterodimer population is increased above 1%, levelling off at a population of  $\sim 5\%$  (Fig. 2a). The best results are obtained with  $N_e = 4$ , and larger ensemble sizes are unjustified and would result in over-fitting the data. For  $N_e = 4$ , the PRE Q-factors at a heterodimer population of 3–5% are close to the expected PRE Q-factor based on experimental error (Fig. 2a), consistent with translational diffusion data (Methods and Supplementary Fig. 1d). Given a total protein concentration of 0.4 mM, the apparent  $K_d$  for self-association is therefore 3–6 mM. A comparison of the calculated and observed PRE profiles and a correlation plot of observed versus calculated  $\Gamma_2$  rates for  $N_e = 4$  at a heterodimer population of 5% are shown in Fig. 1b–d and Fig. 2b, respectively.

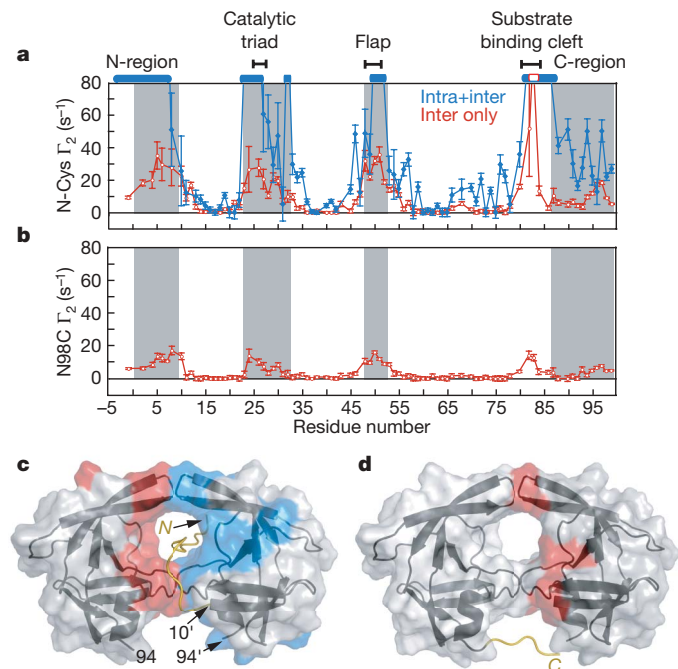
The distribution of the spin-labelled monomer relative to the isotopically labelled monomer in the computed ensemble of  $^{\text{SFNF}}$ PR(D25N) encounter complexes is shown in Fig. 2c. The predominant interactions between the two monomers involve the same residues that comprise the dimer interface in the mature dimer, and one subunit of the mature dimer is embedded within the ensemble distribution of the spin-labelled subunit. The orientation of the subunits in the encounter complex ensembles can be described by spherical angles describing the orientation of the vector joining the centre of masses of the two subunits to the coordinate axis frame. Many members within the calculated ensemble are clustered around the values corresponding to the mature dimer (Fig. 2d). This is reflected in the distribution of the distance root mean square (d.r.m.s.; see Methods) deviation metric where over one-half of the

ensemble members have d.r.m.s. values less than 6 Å (Fig. 2h). However, the structures with spherical angles close to the mature dimer (indicated by arrows in Fig. 2d) and low d.r.m.s. values have a widespread range of relative self-rotations, as illustrated by three examples comprising the ensemble members with the closest spherical angles to the mature dimer (Fig. 2e), the smallest d.r.m.s. (Fig. 2f) and the smallest C $\alpha$  atomic r.m.s. displacement (Fig. 2g). The difference from the mature dimer in rotation angle about the axis joining the centre of masses of the two subunits ranges from  $13^\circ$  (Fig. 2g) to  $135^\circ$  (Fig. 2e), with an intermediate rotation angle of  $70^\circ$  for the structure in Fig. 2f (see Supplementary Fig. 3 for definitions). One can therefore conclude that the actual occupancy of a structure within the encounter complex ensemble corresponding to the mature dimer is very small.

To probe the conformational space sample by the disordered N-terminal flanking sequence of the  $^{\text{SFNF}}$ PR(D25N) precursor we introduced a spin label on a Cys residue inserted immediately after the N-terminal serine ( $^{\text{S(C)FNF}}$ PR(D25N)). PRE measurements were carried out on a 1:1 mixture of 0.2 mM U- $^{2}\text{H}/^{13}\text{C}/^{15}\text{N}$ - $^{\text{SFNF}}$ PR(D25N) precursor and 0.2 mM spin-labelled, natural isotopic abundance  $^{\text{S(C)FNF}}$ PR(D25N) to detect intermolecular PREs, and on a sample of 0.2 mM spin-labelled, U- $^{2}\text{H}/^{13}\text{C}/^{15}\text{N}$ -labelled  $^{\text{S(C)FNF}}$ PR(D25N) to observe both inter- and intramolecular PRE effects. Although the overall PRE profiles for the two samples are similar (although differences in detail are apparent), the magnitude of the PREs for the second sample is much larger than for the first, reflecting the contribution from intramolecular PREs (Fig. 3a). The N-terminal residues –4 to 9, and residues comprising the active site, flap and substrate-binding cleft, display large inter- and intramolecular PREs (Fig. 3a, c). The intermolecular PREs involving residues 82–84 are fully consistent with the large intermolecular PREs observed on the N-terminal residues from spin-labelled V82C (Fig. 1d). These data indicate that the N-terminal tail can insert itself into the active site and make transient contact with both subunits in the encounter complex ensemble. The spin label is located four residues proximal to the scissile peptide bond, and the observation that large PREs are observed for both sides of the active site (see Fig. 3c) suggests that the tail shuttles back and forth within the substrate binding cleft formed by the two subunits in the context of a dimer. Such translational movement is a functional requirement, as the protease precursor cuts the N-terminal transframe region in two major locations before cleaving its C terminus (Fig. 1a)<sup>9,10</sup>. This is confirmed by the observation of a very similar intermolecular PRE profile from full-length TFR-PR(D25N) spin-labelled at position –44, four residues downstream from the TFP-p6<sup>pol</sup> cleavage site at residues –48/–49, to U- $^{2}\text{H}/^{13}\text{C}/^{15}\text{N}$ - $^{\text{SFNF}}$ PR(D25N) (Supplementary Fig. 2b).

The C-terminal region of the  $^{\text{SFNF}}$ PR(D25N) precursor was spin-labelled at N98C. The resulting intermolecular PREs are much smaller than those with the spin label at the N terminus, but the PRE profiles are similar (Fig. 3b, d). Thus, the C-terminal flexible region can also make intermolecular contacts with the active site and substrate-binding cleft in the context of the precursor encounter complex ensemble. Because the N- and C termini are highly mobile, intermolecular PREs between the N- and C termini will be significantly attenuated. Nevertheless, intermolecular PREs are observed on residues 95–97 from the spin label at the N terminus (Fig. 3a), and on residues 5–8 (Fig. 3b) from the spin label at the C terminus (Fig. 3b). Small intermolecular PREs are also observed from the N98C spin label to the C-terminal region (residues 95–99). These observations might suggest the existence of transient, loose interactions between the N- and C termini that may partially approximate a portion of the intersubunit  $\beta$ -sheet in the mature dimer.

The PRE data presented here demonstrate that although the HIV-1 protease precursor is predominantly monomeric, transient encounter complex dimers are formed using the same interface as that of the mature dimer but with a wide range of relative subunit orientations. Only a very small fraction of the encounter complexes adopt the same subunit orientation as in the mature protease, accounting for the very low enzymatic activity of the precursor. This small subset, which may be partially stabilized by transient, loose interactions involving the N- and



**Figure 3 | PRE profiles with spin labels attached at the N- and C termini of the  $^{SFNF}PR(D25N)$  mini-precursor.** **a**, Intermolecular PREs (red) observed for a 1:1 mixture (0.2 mM each) of N-terminal spin-labelled  $^{S(C)FNF}PR(D25N)$  at natural isotopic abundance and  $U\text{-}[^2H/^{13}C/^{15}N]\text{-}^{SFNF}PR(D25N)$ , and the sum of the inter- and intramolecular PREs (blue) observed for 0.2 mM N-terminal spin-labelled  $U\text{-}[^2H/^{13}C/^{15}N]\text{-}^{S(C)FNF}PR(D25N)$ . Residues broadened beyond detection are denoted by open bars. **b**, Intermolecular PREs observed for a 1:1 mixture (0.2 mM each) of  $U\text{-}[^2H/^{13}C/^{15}N]\text{-}^{SFNF}PR(D25N)$  and C-terminal spin-labelled (at N98C)  $^{SFNF}PR(D25N)$  at natural isotopic abundance. Grey shaded areas in **a** and **b** delineate residues that are buried at the dimer interface in the mature protease. Error bars in **a** and **b** represent 1 s.d. **c**, **d**, Inter- and intramolecular PREs with  $\Gamma_2$  rates  $>10\text{ s}^{-1}$  colour-coded in red and blue, respectively, onto the molecular surface of the mature protease dimer originating from the N-terminal (**c**) and the C-terminal (**d**) spin labels. The intramolecular PRE rates are given by the difference in PRE rates between the blue and red profiles in **a**. Cartoons of modelled N-terminal (residues -4 to 9) and C-terminal (residues 95–99) regions bearing the spin labels are included in **c** and **d**, respectively.

C-terminal regions, can accommodate transient insertion of the N-terminal region including the N-terminal cleavage site in the substrate binding cleft, thereby providing a structural model for autoprocessing at the N terminus of the protease leading to the formation of a stable dimer with mature catalytic activity.

## METHODS SUMMARY

**Sample preparation and NMR spectroscopy.** Protein expression, mutagenesis, purification and conjugation of engineered surface cysteine residues to 3-iodo-methyl-(1-oxy-2,2,5,5-tetramethylpyrrolidine) are described in the Methods. Samples for NMR were in 20 mM sodium phosphate buffer, pH 5.8. NMR experiments were collected at 20 °C at a  $^1H$  spectrometer frequency of 600 MHz.  $^1H_N$  PRE data were acquired using a two-dimensional  $^1H\text{-}^{15}N$  correlation-based pulse scheme with an interleaved two time-point measurement<sup>21</sup>. **Simulated annealing calculations.** Conjoined rigid-body/torsion angle dynamics simulated annealing calculations on the basis of the PRE data were carried out using Xplor-NIH<sup>22</sup> as described<sup>14</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 7 July; accepted 12 August 2008.

- Louis, J. M., Weber, I. T., Tozser, J., Clore, G. M. & Gronenborn, A. M. HIV-1 protease: maturation, enzyme specificity, and drug resistance. *Adv. Pharmacol.* **49**, 111–146 (2000).

- Louis, J. M., Ishima, R., Torchia, D. A. & Weber, I. T. HIV-1 protease: structure, dynamics and inhibition. *Adv. Pharmacol.* **55**, 261–298 (2007).
- Wlodawer, A. & Erikson, J. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* **62**, 543–585 (1993).
- Wlodawer, A. & Vondrasek, J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 249–284 (1998).
- Ishima, R., Torchia, D. A., Lynch, S. M., Gronenborn, A. M. & Louis, J. M. Solution structure of the mature HIV-1 protease monomer: insight into the tertiary fold and stability of a precursor. *J. Biol. Chem.* **278**, 43311–43319 (2003).
- Miller, M. *et al.* Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **246**, 1149–1152 (1989).
- Louis, J. M., Nashed, N. T., Parris, K. D., Kimmel, A. R. & Jerina, D. M. Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. *Proc. Natl Acad. Sci. USA* **91**, 7970–7974 (1994).
- Co, E. *et al.* Proteolytic processing mechanisms of a miniprecursor of the aspartic protease of human immunodeficiency virus type 1. *Biochemistry* **33**, 1248–1254 (1994).
- Louis, J. M., Wondrak, E. M., Kimmel, A. R., Wingfield, P. T. & Nashed, N. T. Proteolytic processing of HIV-1 protease precursor, kinetics and mechanism. *J. Biol. Chem.* **274**, 23437–23442 (1999).
- Louis, J. M., Clore, G. M. & Gronenborn, A. M. Autoprocessing of HIV-1 protease is tightly coupled to protein folding. *Nature Struct. Biol.* **6**, 868–874 (1999).
- Pettit, S. C., Everitt, L. E., Choudhury, S., Dunn, B. M. & Kaplan, A. H. Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. *J. Virol.* **78**, 8477–8485 (2004).
- Ishima, R., Torchia, D. A. & Louis, J. M. Mutational and structural studies aimed at characterizing the monomer of HIV-1 protease and its precursor. *J. Biol. Chem.* **282**, 17190–17199 (2007).
- Iwahara, J. & Clore, G. M. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227–1230 (2006).
- Tang, C., Iwahara, J. & Clore, G. M. Visualization of transient encounter complexes in protein-protein association. *Nature* **444**, 383–386 (2006).
- Volkov, A. N., Worall, J. A., Holtzmann, E. & Ubbink, M. Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proc. Natl Acad. Sci. USA* **103**, 18945–18950 (2006).
- Tang, C., Schwieters, C. D. & Clore, G. M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **449**, 1078–1082 (2007).
- Tessmer, U. & Kräusslich, H.-G. Cleavage of human immunodeficiency virus type 1 proteinase from the N-terminally adjacent p6\* protein is essential for efficient Gag polyprotein processing and viral infectivity. *J. Virol.* **72**, 3459–3463 (1998).
- Ludwig, C., Leiberer, A. & Wagner, G. Importance of protease cleavage sites within and flanking human immunodeficiency virus type 1 transframe protein p6\* for spatiotemporal regulation of protease activation. *J. Virol.* **82**, 4573–4584 (2008).
- Wondrak, E. M., Nashed, N. T., Haber, M. T., Jerina, D. M. & Louis, J. M. A transient precursor of the HIV-1 protease: isolation, characterization and kinetics of maturation. *J. Biol. Chem.* **271**, 4477–4481 (1996).
- Cherry, E. *et al.* Characterization of human immunodeficiency virus type-1 (HIV-1) particles that express protease-reverse transcriptase fusion proteins. *J. Mol. Biol.* **284**, 43–56 (1998).
- Iwahara, J., Tang, C. & Clore, G. M. Practical aspects of  $^1H$  transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn. Reson.* **184**, 185–195 (2007).
- Schwieters, C. D., Kuszewski, J. & Clore, G. M. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47–62 (2006).
- Iwahara, J., Schwieters, C. D. & Clore, G. M. Ensemble approach for NMR structure refinement against  $^1H$  paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J. Am. Chem. Soc.* **126**, 5879–5896 (2004).
- Spinelli, S., Liu, Q. Z., Alzari, P. M., Hirel, P. H. & Poljak, R. J. The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. *Biochimie* **73**, 1391–1396 (1991).
- Schwieters, C. D. & Clore, G. M. Reweighted atomic densities to represent ensembles of NMR structures. *J. Biomol. NMR* **23**, 221–225 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank R. Ishima for providing initial backbone assignments for the  $^{SFNF}PR(D25N)$  protease construct; C. Schwieters for many discussions; Y. Sheng for help with the CS-Rosetta calculations; Y. Kim for providing the code for structure clustering and d.r.m.s. calculations; and J. Sayer for MALDI measurements. This work was supported by funds from the Intramural Program of the NIH, NIDDK and the AIDS Targeted Antiviral program of the Office of the Director of the NIH (to G.M.C.).

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to G.M.C. ([mariusc@mail.nih.gov](mailto:mariusc@mail.nih.gov)).

## METHODS

**Vector construction and protein sample preparation for NMR studies.** Mutations T12C, E34C, V82C and N98C within the 99-amino-acid-long HIV-1 protease sequence and the mutation to insert a Cys in the flanking SFNF sequence (C-terminal residues of p6<sup>pol</sup> within the transframe region; see Fig. 1a and Supplementary Fig. 1) to generate <sup>S(C)FNF</sup>PR(D25N) were introduced in the <sup>SFNF</sup>PR(D25N) template<sup>5</sup> using the appropriate forward and reverse primers and the QuikChange kit and protocol (Stratagene). The <sup>S(C)FNF</sup>PR(D25N) construct was used because we were unable to obtain efficient spin-labelling of a precursor protein bearing an N-terminal cysteine. Mutations A(-44)C (fifth residue of p6<sup>pol</sup>) and V82C (in the protease sequence) were also introduced in the full-length TFR-PR(D25N) construct (that is, TFP-p6<sup>pol</sup>-PR(D25N); see Fig. 1a and Supplementary Fig. 1a) using the same protocol. (The TFR is 56 residues in length and adopts a random coil conformation.) The newly introduced mutations were verified both by DNA sequencing and mass spectrometry. (Note that the <sup>SFNF</sup>PR(D25N) template, in addition to the D25N mutation which eliminates all traces of catalytic activity, and the C67A and C95A mutations which remove all additional surface cysteine residues other than that to which the spin label is going to be attached, also contains three other mutations, Q7K, L33I and L63I; the latter three mutations restrict autoproteolysis of the mature protease dimer, and have been shown to have indiscernible effects on structure, stability and catalytic activity of the mature dimer<sup>10</sup>.)

*Escherichia coli* BL21 (DE3) host cells bearing the appropriate vector were grown in Luria-Bertani medium or in D<sub>2</sub>O-based minimal medium containing <sup>15</sup>N-NH<sub>4</sub>Cl and <sup>13</sup>C<sub>6</sub>,<sup>2</sup>H<sub>7</sub>-glucose as the sole nitrogen and carbon sources, respectively, at 37 °C, and induced for expression. Proteins were purified from inclusion bodies using an established protocol as described previously involving size-exclusion chromatography under denaturing conditions followed by reverse-phase HPLC<sup>9,26</sup>. Peak fractions (~0.5 mg ml<sup>-1</sup>) were stored in aliquots at -70 °C. Alternatively, two aliquots (2.5 mg) of the proteins were lyophilized and stored at -20 °C.

A total of 2.5 mg of the lyophilized protein was dissolved in 1.2 ml of 4 M guanidinium-HCl, 1.7 mM HCl, pH 1.6. Spin-label conjugation was carried out by dissolving 0.5 mg of 3-iodomethyl-(1-oxy-2,2,5,5-tetramethylpyrrolidine) (catalogue number I709500; Toronto Research Chemicals) in 10 μl of ethanol, followed by the addition of 140 μl of 1 M Tris-HCl, pH 8, and adding the resulting mixture to the protein solution. After incubation for 1 h at room temperature, 30 μl of 1 M dithiothreitol was added and the incubation continued for another 1.5 h. The sample was loaded onto a Superdex-75 column (1.6 × 60 cm, GE HealthCare) equilibrated in 4 M guanidinium-HCl, 20 mM sodium formate, pH 2.6, at a flow rate of 1.4 ml min<sup>-1</sup> at room temperature. Peak fractions were pooled and the concentration was estimated by measuring absorbance at 280 nm. The extent of labelling was 100% as determined by MALDI-TOF analysis on a Voyager-DE instrument (Perceptive Biosystems). Spin-labelling does not perturb the structure of the <sup>SFNF</sup>PR(D25N) mini-precursor as judged by NMR spectroscopy. The three mutations within the protein core, T12C, E34C and V82C, are frequently mutated in viable HIV-1 variants and are therefore not expected to alter significantly the catalytic properties of the protease<sup>2</sup>. It should be noted that V82C is located close to the substrate binding cleft comprising residues 80–81 and 83–85, but its side chain points outwards towards solvent. In the one instance where kinetic data are available for a mutation at position 82 (V82A), only a modest 10–15% decrease in *k*<sub>cat</sub>/*K*<sub>m</sub> relative to wild type is observed, and structural differences between wild-type protease and the V82A mutant are insignificant, with an r.m.s. deviation between the two crystal structures of only 0.12 Å for all main chain atoms<sup>27</sup>. Thus, the presence of a bulky spin label at position 82 would not be expected to result in a major perturbation in catalytic activity.

After extensive dialysis against 7 mM HCl, 1.4 mg each of the conjugated protein and the U-[<sup>2</sup>H/<sup>13</sup>C/<sup>15</sup>N]-labelled <sup>SFNF</sup>PR(D25N) protein were mixed and adjusted to a final concentration of 0.25 mg ml<sup>-1</sup> protein, 35% acetonitrile and 0.05% trifluoroacetic acid. The solution was dialysed against 2 l of 7 mM HCl and 4 l of 20 mM sodium phosphate, pH 5.8, each for a period of 1.5–2 h and concentrated to ~400 μM using Amicon Ultra-4 (10,000 MWCO) devices. Protein concentration (mg ml<sup>-1</sup>) was determined spectrophotometrically using  $\epsilon$  (0.1%) = 1.097 at 280 nm.

**Control active <sup>SFNF</sup>PR(D25) mini-precursor protease construct.** The <sup>SFNF</sup>PR(D25N) precursor construct does not undergo autoprocessing owing to the substitution of the active site Asp 25 by Asn. To verify that <sup>SFNF</sup>PR(D25N) represents a suitable model system we examined the autoprocessing activity of the corresponding <sup>SFNF</sup>PR(D25) precursor; that is, the precursor without the active site mutation. Most of the expressed protein undergoes maturation at the N terminus (between Phe-Pro) of the protease in the control <sup>SFNF</sup>PR(D25) precursor to produce the mature protease as expected. This was confirmed by

subjecting an aliquot of the purified (dissolved) inclusion bodies to electrospray-mass spectrometry. The measured mass of 10,728 Da clearly corresponds to the PR(D25) mature protease (calculated mass of 10,728.3 Da). Under identical conditions of analysis for <sup>SFNF</sup>PR(D25N), which is devoid of catalytic activity, only the full-length protein corresponding to a mass of 11,222 Da (calculated mass of 11,222.8 Da) is observed consistent with previous observations from studies using the inactive full-length TFR-PR(D25N) precursor, which does not undergo maturation<sup>5</sup>, as compared to the active TFR-PR(D25) precursor, which exhibits time-dependent processing at the p6<sup>pol</sup>-protease junction to release the mature protease<sup>9,10</sup>.

**NMR experiments.** All NMR data were acquired at 20 °C on a Bruker DRX600 spectrometer equipped with a z-gradient triple resonance cryoprobe.

Measurement of translational diffusion coefficients (*D*<sub>s</sub>) by pulse field gradient NMR<sup>28</sup> was carried out using the Watergate BPP-LED pulse scheme described previously<sup>29</sup>. The translational diffusion coefficient *D*<sub>s</sub> is derived from a linear least-squares fit to a plot of ln(*I*(*f*)/*I*(*f*<sub>0</sub>)) versus (*f*<sup>2</sup> - *f*<sub>0</sub><sup>2</sup>):

$$\ln[(I(f)/I(f_0))] = -(\gamma\delta G_{\max})^2(f^2 - f_0^2)(\Delta - \delta/3 - \tau/2)D_s$$

where *I*(*f*) and *I*(*f*<sub>0</sub>) are the intensities of the NMR signal at fractional gradient strengths of *f* and *f*<sub>0</sub>; *f*<sub>0</sub> is the fractional gradient strength of the reference spectrum (0.1); *f* is the fractional gradient strength with values of 0.2, 0.3, 0.4, 0.5 and 0.6 times *G*<sub>max</sub>, the maximum gradient strength (70 × 10<sup>-4</sup> T cm<sup>-1</sup>);  $\gamma$  is the gyromagnetic ratio of <sup>1</sup>H (2.6752 × 10<sup>8</sup> s<sup>-1</sup> T<sup>-1</sup>);  $\Delta$  = 15.4 ms;  $\delta$  = 5 ms (gradient duration); and  $\tau$  = 0.2 ms. The overall diffusion delay is 10 ms. The value of the scaling factor ( $\gamma\delta G_{\max})^2(\Delta - \delta/3 - \tau/2)$  is 1.19 × 10<sup>10</sup> s m<sup>-2</sup>. The values of *D*<sub>s</sub> were 9.3(±0.4) × 10<sup>-11</sup> and 12.9(±0.5) × 10<sup>-11</sup> m<sup>2</sup> s<sup>-1</sup> for the <sup>SFNF</sup>PR(D25N) precursor and the mature PR(D25N) dimer, respectively, at the same (0.4 mM) subunit concentration (Supplementary Fig. 1d). The ratio of the two *D*<sub>s</sub> values (0.72 ± 0.04) is fully consistent with the expected value of 0.75 for a *D*<sub>s</sub><sup>monomer</sup>/*D*<sub>s</sub><sup>dimer</sup> ratio<sup>28</sup>, placing an upper limit of about 10% for the population of dimeric species.

<sup>15</sup>N-<sup>1</sup>H heteronuclear NOE measurements were carried out using a flip-back scheme as described<sup>30</sup>. Residues -4 to 9 and 95–99 of <sup>SFNF</sup>PR(D25N) have heteronuclear <sup>15</sup>N-<sup>1</sup>H NOE values ranging from -1 to 0.5 indicating that they are disordered and highly mobile. Backbone assignments were derived using the following three-dimensional triple resonance experiments: HNCO, HN(CO)CA and CBCA(CO)NH<sup>31,32</sup>. The weighted mean backbone chemical shift difference between different constructs is given by  $[\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}25}^2 + \Delta\delta_{\text{C}\alpha}^2/4]^{1/2}$  as described previously<sup>33</sup>. A comparison of <sup>1</sup>H/<sup>15</sup>N/<sup>13</sup>C $\alpha$  chemical shifts reveals significant perturbations relative to the corresponding mature dimeric PR(D25N) for residues located at the dimer interface (Supplementary Fig. 1b), but only minor perturbations relative to the equivalent monomeric PR(1–95) construct obtained by deletion of the C-terminal four residues (Supplementary Fig. 1c)<sup>5</sup>. Analysis of the chemical shift index (based on <sup>13</sup>C $\alpha$ , <sup>13</sup>C $\beta$  and <sup>13</sup>C' shifts)<sup>34</sup> for <sup>SFNF</sup>PR(D25N) and PR(D25N) indicates that the secondary structure elements are preserved in the precursor with the exception of the N- and C-terminal strands which form an intersubunit four-stranded antiparallel  $\beta$ -sheet in the mature dimer (Supplementary Fig. 1e).

PRE <sup>1</sup>H<sub>N</sub>- $\Gamma_2$  rates are given by the difference in *R*<sub>2</sub> relaxation rates between the paramagnetic (spin-labelled) and diamagnetic states of the protein. *R*<sub>2</sub> rates were determined from a two-time-point interleaved two-dimensional <sup>1</sup>H-<sup>15</sup>N correlation-based experiment, as described previously<sup>21</sup>. The time interval between the two time points was 32 ms for the intermolecular PRE measurements and 4 ms for the intramolecular PRE measurements. The short time interval for the latter is used to minimize any errors in  $\Gamma_2$  rates introduced by any potential diamagnetic contamination (that is, spin-labelling less than 100%)<sup>21</sup>.

**Tertiary structure of <sup>SFNF</sup>PR(D25N).** To verify that the tertiary structure of the ordered region of <sup>SFNF</sup>PR(D25N) (that is, residues 10–94) is the same as that of an individual subunit of the mature protease, we made use of the CS-Rosetta chemical shift structure determination algorithm which uses a hybrid approach of chemical-shift-based fragment selection and ROSETTA Monte Carlo driven fragment assembly<sup>35</sup>. The resulting ten lowest energy models are essentially identical to the corresponding region of the mature dimer with a backbone r.m.s. deviation of only 1.3 ± 0.2 Å (Supplementary Fig. 1e).

**PRE calculations and ensemble refinement.** Because the electron relaxation rate  $\tau_e$  of the free radical is much longer than that of the protein rotational correlation time  $\tau_r$ <sup>21</sup>, the PRE correlation time  $\tau_c$  [ $= (\tau_r^{-1} + \tau_e^{-1})^{-1}$ ] for the calculation of intermolecular PRE rates was assumed to be the same as  $\tau_r$  (12 ns) for the mature protease dimer<sup>36</sup>. To account for the flexibility of the linker between the spin label and the protein backbone, a ten-conformer randomized ensemble was used to represent the conformational space sampled by the spin label. The randomized ensemble was generated by high-temperature simulated annealing and slow cooling in Xplor-NIH<sup>22</sup> subject to a target function compris-

ing stereochemical terms, a quartic van der Waals repulsion term to prevent atomic overlap between the spin label and the protein, and a multidimensional conformational database potential of mean force<sup>37</sup> describing the  $\phi/\psi/\chi_1$  conformational space available to the surface cysteine residue to which the spin label was conjugated. Note that overlap between the members of the Cys spin-label ensemble is permitted as the ten-member ensemble represents a distribution of states. To ensure full sampling of the conformational space available to the spin label a different ten-conformer randomized ensemble was used for each structure calculation. Agreement between observed and calculated  $\Gamma_2$  rates is given by the PRE  $Q$ -factor,  $Q_{\text{PRE}}$ :<sup>23</sup>

$$Q_{\text{PRE}} = \left[ \frac{\sum_i \left\{ \Gamma_{2,i}^{\text{obs}} - p \langle \Gamma_{2,i}^{\text{calc}} \rangle \right\}^2}{\sum_i \left( \Gamma_{2,i}^{\text{obs}} \right)^2} \right]^{1/2}$$

where  $\Gamma_{2,i}^{\text{obs}}$  and  $\langle \Gamma_{2,i}^{\text{calc}} \rangle$  are the observed and ensemble average calculated transverse  $\Gamma_2$  rates for residue  $i$ , respectively, and  $p$  is the overall population of the encounter complex species. All members of an ensemble of size  $N_e$  are weighted equally. For the average  $Q$ -factor  $\langle Q \rangle$  for all calculated  $n$  ensembles,  $\langle \Gamma_{2,i}^{\text{calc}} \rangle$  is averaged over the members of each  $N_e$  ensemble. For the ensemble of ensembles average PRE  $Q$ -factor,  $Q_{\text{ee}}$ ,  $\langle \Gamma_{2,i}^{\text{calc}} \rangle$  is averaged over all ensemble members and all ensembles<sup>14</sup>.

The coordinates used in the Xplor-NIH<sup>22</sup> calculations were taken from the X-ray structure of the unliganded mature HIV-1 protease dimer (Protein Data Bank accession code 1HHP)<sup>24</sup>. Residues 10–94 were treated as a rigid body, and the flexible N- and C-terminal residues were not included in the calculations. The coordinates of the isotopically labelled subunit were held fixed, the initial positions of the spin-labelled subunit (at natural isotopic abundance) were randomized, and rigid-body simulated annealing was carried out against the PRE data sets for the spin label conjugated to the T12C, E34C and V82C sites simultaneously. The target function comprises a PRE restraint term<sup>23</sup>, a quartic van der Waals repulsion term to prevent atomic overlap between the spin-labelled and isotopically labelled subunits, and a very weak radius of gyration term<sup>38</sup> to ensure that each member of the ensemble makes at least some intermolecular contacts<sup>14,39</sup>. Note that atomic overlap between ensemble members of spin-labelled subunits is permitted as these represent separate but rapidly interconverting configurations of the encounter complex species<sup>14,39</sup>. A grid search was performed varying the population of heterodimer and the ensemble size  $N_e$  used to represent the self-associated species<sup>14</sup>. For each ensemble size and population of encounter complex species, 100 calculations were carried out. Ensembles were ranked by PRE  $Q$ -factor and van der Waals repulsion energies, and the top 20 ensembles with the smallest PRE  $Q$ -factors were used for subsequent analysis<sup>39</sup>. Structures were rendered using PyMol (<http://www.pymol.org>) and re-weighted atomic probability density maps were generated using Xplor-NIH<sup>22</sup> as described<sup>25</sup>.

**d.r.m.s. metric.** One metric we used to compare the precursor encounter complexes with the mature dimer was the distance root mean square (d.r.m.s.) metric defined by<sup>40</sup>:

$$\text{d.r.m.s.} = \frac{1}{N} \sum_{i,j} \left| d_{i,j}^{\text{precursor}} - d_{i,j}^{\text{mature}} \right|$$

where  $N$  is the number of distinct residue pairs ( $i, j$ ), and  $d_{i,j}^{\text{precursor}}$  and  $d_{i,j}^{\text{mature}}$  are

the distance matrices in a calculated precursor encounter complex structure and the mature HIV-1 protease dimer structure, respectively.

**Spherical coordinate systems used to describe relative subunit orientation in the encounter complexes.** Two spherical coordinate systems are used to describe the relative orientation of the subunits in the precursor encounter complexes<sup>40</sup>. The first (polar angle  $\phi$  and azimuth angle  $\theta$ ) describes the orientation of the vector joining the centre of masses of the two subunits (shown as grey spheres in Supplementary Fig. 3a) to an external axis system with the  $z$  axis corresponding to the  $C_2$  symmetry axis of the mature dimer. The second (polar angle  $\alpha$  and azimuth angle  $\beta$ ) describes the orientation of a vector joining the centre of mass of the second subunit to an arbitrarily chosen atom of the same subunit (C $\alpha$  atom of Gly 51) relative to an axis system with the  $z'$  axis given by the vector joining the centre of masses of the two subunits (with the red subunit in Fig. 2 corresponding to the fixed reference subunit) (Supplementary Fig. 3a).

26. Wondrak, E. M. & Louis, J. M. Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease. *Biochemistry* **35**, 12957–12962 (1996).
27. Mahalingam, B. *et al.* Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site. *Eur. J. Biochem.* **271**, 1516–1524 (2004).
28. Altieri, A. S., Hinton, D. P. & Byrd, R. A. Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. *J. Am. Chem. Soc.* **117**, 7566–7567 (1995).
29. Chou, J. J., Baber, J. L. & Bax, A. Characterization of phospholipid mixed micelles by translational diffusion. *J. Biomol. NMR* **29**, 299–308 (2004).
30. Grzesiek, S. & Bax, A. The importance of not saturating H<sub>2</sub>O in protein NMR: application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594 (1993).
31. Clore, G. M. & Gronenborn, A. M. Two-, three- and four-dimensional NMR methods for obtaining larger and more precise three-dimensional structures of proteins in solution. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 29–63 (1991).
32. Clore, G. M. & Gronenborn, A. M. Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol.* **239**, 349–363 (1994).
33. Grzesiek, S., Stahl, S. J., Wingfield, P. T. & Bax, A. The CD4 determinant of downregulation by HIV-1 Nef directly binds to Nef: mapping of the Nef binding surface by NMR. *Biochemistry* **35**, 10256–10261 (1996).
34. Wishart, D. S. & Sykes, B. D. The <sup>13</sup>C chemical-shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J. Biomol. NMR* **4**, 171–180 (1994).
35. Shen, Y. *et al.* Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl Acad. Sci. USA* **105**, 4685–4690 (2008).
36. Katoh, E. *et al.* A solution NMR study of the binding kinetics and internal dynamics of an HIV-1 protease-substrate complex. *Protein Sci.* **12**, 1376–1385 (2003).
37. Clore, G. M. & Kuszewski, J.  $\chi_1$  rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force. *J. Am. Chem. Soc.* **124**, 2866–2867 (2002).
38. Kuszewski, J., Gronenborn, A. M. & Clore, G. M. Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J. Am. Chem. Soc.* **121**, 2337–2338 (1999).
39. Tang, C., Ghirlando, R. & Clore, G. M. Visualization of transient ultra-weak protein self-association in solution using paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **130**, 4048–4056 (2008).
40. Kim, Y. C., Tang, C., Clore, G. M. & Hummer, G. Replica exchange simulations of transient encounter complexes in protein-protein association. *Proc. Natl Acad. Sci. USA* **105**, 12855–12860 (2008).

# Biochemistry's new look

**Until now, metabolomics researchers have had to adapt technology developed mainly for proteomics. But there are now solutions designed with them in mind. Nathan Blow reports.**

Metabolomics — the comprehensive study of metabolic reactions — is gaining ground alongside its older siblings genomics and proteomics. “Unlike some of the other ‘omics’ that we have seen, metabolomics is going to produce a lot of useful information right from the start,” says Gary Siuzdak, professor of molecular biology at the Scripps Research Institute in La Jolla, California. He is one of a growing number of biologists using advanced technology to explore biochemical questions on a scale that would have seemed impossible a decade ago.

“The metabolome is the best indicator of an organism’s phenotype,” says David Wishart at the University of Alberta in Edmonton, Canada. Wishart was one of the instigators of the Human Metabolome Project, a US\$7.5-million effort funded by Genome Canada to systematically characterize the metabolites of the human body. He gives the example of a person holding their breath for five minutes. Although genomic or proteomic analysis would not provide any evidence of stress during this short period — even as the person turns blue — metabolite profiles would show dramatic changes within the body.

Unlike a genome or even a proteome, however, a metabolome is tricky to pin down. Wishart notes that although researchers know there are 3 billion base pairs in the human genome, if you ask biochemists how many small-molecule metabolites there are in the human body, they come back with numbers ranging from 3,000 to 100,000. And this poses a real challenge for metabolomics research, as both ends of the scale could be correct.

The Human Metabolome Project has pegged the number of endogenous metabolites in the human body at around 3,000 — which most researchers agree on. But humans also take in small molecules from the environment — preservatives in food, chemicals in the air, metabolites produced through the breakdown of drugs and toxins — making an exact figure hard to determine.

## Separation anxiety

With metabolites in such a state of flux, researchers do not have an easy task. Nevertheless, advances in chromatography, mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) are allowing them to make headway in defining different



The Pegasus 4D GCxGC MS TOF system enables multidimensional approaches to GC separation.

metabolomes and understanding how changes in the concentrations of metabolites relate to human health and disease.

One of the problems is that metabolites come in a variety of chemical forms. “I would say one of the real challenges of metabolomics is that each metabolite is its own unique puzzle,” says Trent Northen, a scientist at Lawrence Berkeley National Laboratory in California. And in most cases the first step in solving the puzzle is isolating the metabolite for analysis.

No one separation method works for all metabolites, so researchers rely on combinations of gas chromatography (GC), liquid chromatography (LC) and emerging capillary electrophoresis (CE).

Historically, GC separation has had the edge. “GC–MS technology may not be sexy,

but huge databases are available,” says Northen. These GC–MS databases, compiled over more than four decades, enable researchers to compare a wide range of spectra to arrive at a chemical identification.

Multidimensional GC, often called ‘GCxGC’ or two-dimensional GC, offers even better separation. “When people are doing GCxGC, they are trying to get more separation chroma-

tographically of very complex samples,” says Steven Fischer, a senior applications chemist at Agilent Technologies in Santa Clara, California. To achieve this, GCxGC uses two separation phases, such as a non-polar and a polar phase, in two capillary columns in series in the instrument.

Agilent recently introduced the 7890 GC system, which can perform multidimensional GC, and Thermo Fisher Scientific in Waltham Massachusetts, has developed the Trace GCxGC system. The Pegasus 4D GCxGC MS time-of-flight (TOF) system from LECO, based in St Joseph, Michigan, uses a thermal modulator placed between the two GC columns to collect effluent from the first column before going into the second phase of separation. The power of the multidimensional approach is starting to be reported. In May this year a group reported the use of GCxGC with LC–MS to generate a draft metabolic network for the single-celled alga *Chlamydomonas reinhardtii*<sup>1</sup>.

## Class action

GC is particularly useful for mixtures of volatiles, such as steroids, saccharides and sugar alcohols, which can be sent directly into the gas phase for separation. Metabolites in human biofluids and tissues therefore present a technical challenge for GC, as most are not volatile. Non-volatile metabolites either need complicated chemical transformation before GC or



Gary Siuzdak: developing new approaches to metabolite identification.

separation by other types of chromatography.

One of these is high-performance liquid chromatography (HPLC), a well-established lab workhorse. It uses a combination of solvents, pressure and matrix particle sizes to separate molecules on the basis of their retention times in a column packed with matrix. HPLC can separate a broad range of metabolites, including non-volatiles, and remains a favourite among metabolomics researchers.

Most advances in HPLC involve increases in the pressure applied and changes in matrix particle size. Ultra-performance liquid chromatography (UPLC), commercialized by Waters Corporation in Milford, Massachusetts, is becoming more widely used in the metabolomics community. It takes advantage of higher pressure (83 megapascals compared with 21 megapascals for HPLC) and smaller particles (less than 2 micrometres diameter compared with 3 micrometres for HPLC) to obtain faster separation times.

But like GC, HPLC has technical stumbling blocks. Reversed-phase HPLC (in which the stationary phase is non-polar) is often used for metabolomics analysis, but reversed-phase separation often fails with hydrophilic metabolites. These tend to be so water soluble that they interact poorly with the non-polar bonding phase and are

rapidly eluted, according to Phil Koerner, a senior technical manager from chromatography specialists Phenomenex in Torrance, California.

So in 2007, Phenomenex introduced the Luna HILIC column. "I like to refer to it as reverse reverse-phase chromatography," says Koerner. In the HILIC approach, the weak solvent, which is applied first, is a polar organic solvent (not water as in reversed-phase HPLC), and the strong solvent, applied second, is water. This causes the order of elution to be completely reversed, with the most hydrophilic compounds being eluted last. Although Koerner acknowledges that the HILIC approach is not new, it was the need to separate hydrophilic metabolites on the large scale required by metabolomics that led Phenomenex and other companies, such as Waters and Tosoh Bioscience of Stuttgart, Ger-

many, to start supplying a greater number and range of HILIC columns.

Capillary electrophoresis followed by MS (CE-MS) is not yet so popular with the metabolomics community as either GC or HPLC, but several developers are hoping to change this. "It can be very difficult to use this approach," acknowledges Ryuji Kanno, president of Human Metabolome Technologies based in Tokyo, Japan. This approach uses electrophoretic mobility to separate low-molecular-weight ionic compounds that are difficult to separate by GC or HPLC. The company has been working closely with Agilent to develop optimized reagents and capillary columns, and is providing training with Agilent's CE-qTOF MS system to make the CE approach more accessible to metabolomics researchers, says Kanno.

Mass spectrometry is not the only method

that can be used to detect metabolites once separated. Wishart and his colleagues recently compared MS and NMR to look at metabolites in cerebral spinal fluid<sup>2</sup>. They found little overlap in the metabolites detected by the two methods, and the conclusion was clear: "We do not have a single perfect metabolite detector," says Wishart.

MS and NMR each have their supporters. "One of the main strengths of NMR is that



HILIC columns are making the hunt for hydrophilic metabolites easier.

PHENOMENEX

## DARK MATTER

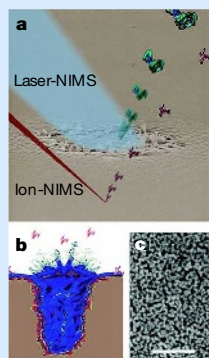
"There are a lot of small molecules that we do not even know about yet," says Arthur Castle, programme director for the Roadmap Metabolomics Technology development programme at the US National Institutes of Health in Bethesda, Maryland. Metabolomics has a good handle on analysing human primary metabolites, but when it comes to lipids, secondary metabolites, xenobiotics and the products of gut microflora, we are just scratching the surface, says Castle.

The problem is part technological, part informatics. Steven Fischer at Agilent Technologies in Santa Clara, California, points out that some compounds are not stable and undergo chemical transformation during separation. Researchers will probably still see these transformed molecules by mass spectrometry (MS), but they may

be misidentified, highlighting the need for follow-up experiments. And this is where metabolomics has an advantage over the other 'omics'. "There is so much knowledge of biochemistry that when we find a potential biomarker or a new drug mechanism we already know a lot about it," says Michael Milburn, chief scientific officer at Metabolon in Durham, North Carolina.

Trent Northen, now at the Lawrence Berkeley Laboratory in Berkeley, California, and Oscar Yanes, working in Gary Siuzdak's lab at the Scripps Research Institute in La Jolla, California, may have developed a new way to get at some of these 'unknown' molecules with

an ionization technique called nanostructure-initiator mass spectrometry (NIMS)<sup>4</sup>. The idea is to transfer a biomolecule into



NIMS: nanostructured surface releases material.

the gas phase from a nanostructured surface simply by making that surface disappear. "We came up with the idea of putting a wax underneath, so when the nanostructured surface was irradiated it would melt and vaporize, allowing the molecule to go into the gas phase," says Northen. They finally came up with a perfluorinate surface for the trapped initiator phase.

Siuzdak's group found that NIMS worked not only for proteins, but also for small molecules such as metabolites. "Perfluorinates do not ionize well, so it allows us to

see things in the lower mass region where metabolites like to hang out," says Siuzdak.

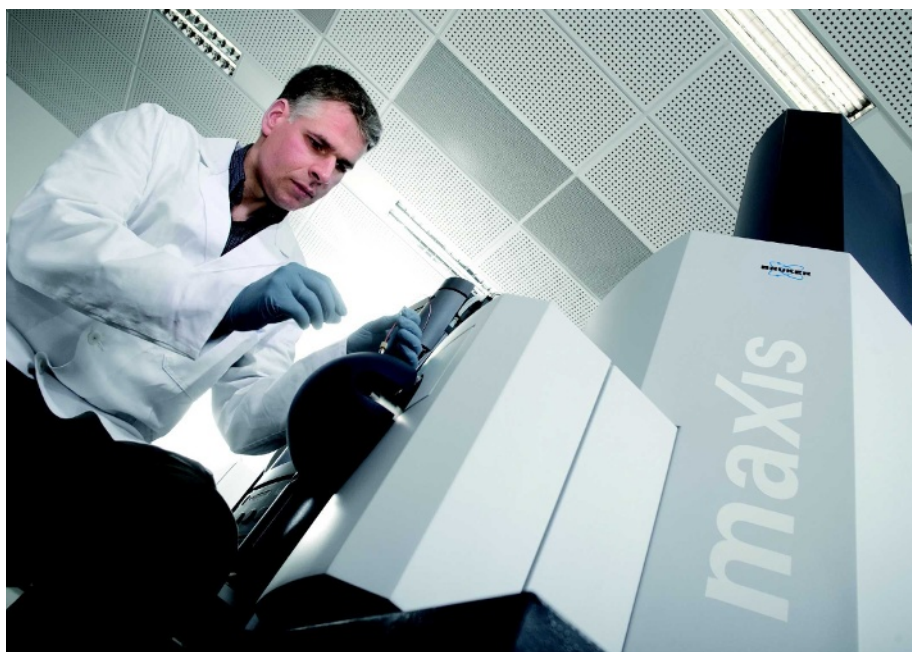
Using perfluorinated materials may have another advantage for metabolomics. "As these are highly hydrophobic surfaces, we can apply very dirty complex samples, such as blood, for direct analysis," says Yanes. If you put a drop of urine or blood on the NIMS chip the metabolites will attach, but all the salts and other chemicals that normally interfere with MS stay in solution. Yanes is now exploiting this property to follow drug metabolism by looking at uptake in blood, clearance in urine and tissue localization.

For Siuzdak, exploring the 'unknown' metabolite world is an exciting prospect. "We are getting involved in an area where we don't know what the molecules' structures are or what they do, so it is really just a fantastic area for discovery."

REF. 4

N.B.





The maXis system from Bruker Daltonics can use both UPLC and CE separation approaches.

Newton. He says researchers at Chenomx have performed many studies in which biologically meaningful differences between samples were easily captured with NMR, even though some compounds in the samples probably fell below the sensitivity limits of the instrument (see 'Dark matter').

MS, on the other hand, is a very sensitive method for metabolite identification and, unlike NMR, is easily coupled to upstream separation techniques. Siuzdak says his group can see thousands of molecules in an MS analysis — and that number can be doubled by changing from positive- to negative-ion mode. And by using both reversed-phase chromatography and HILIC columns, they are seeing more hydrophilic compounds in their analyses than before. "I would venture that we are now seeing over an order of magnitude more than what you would see with NMR," he says.

### Detector development

As researchers in the MS camp turn towards TOF and ion-trap MS instruments for metabolite analysis, developers are responding to their complex needs. Bruker Daltonics in Billerica, Massachusetts, has introduced the maXis ultra-high resolution (UHR)-TOF MS system, which can accommodate both UPLC and CE separation. Applied Biosystems in Foster City, California, in collaboration with MDS SCIEX in Toronto, Ontario, have the ion-trap system 4000 QTrap LC/MS/MS that can interface with Applied Biosystem's LightSight software for small-molecule analysis and identification.

it is an unbiased, universal detector," says Jack Newton, a product manager at Chenomx in Edmonton, Canada, which was co-founded by Wishart in 2000. This attribute, along with NMR's ability to determine structure and perform quantitative analysis is particularly attractive to metabolomics researchers who need a way to compare and exchange results between labs. "The move is afoot — people want to get

to that common language of compound names and concentrations," says Wishart, as this will make integrating data sets and obtaining systems-level views of cell physiology possible.

The challenge with NMR is instrument sensitivity — NMR is less sensitive than MS, often identifying far fewer metabolites in the same sample. "For us, the relevant question is how sensitive do you need to be," says

## WINE-OMICS

For Kirsten Skogerson at the University of California, Davis, wondering about how chemical composition affects the flavour and body of a wine took her from a degree in viticulture and enology into metabolomics research. When Skogerson arrived in Oliver Fiehn's lab as a postgrad she looked for a project that would marry Fiehn's expertise in metabolomics and her interest in wine.

"There are so many questions in wine science that you could start to answer by doing a global analysis," she says. A deeper understanding of the biochemistry of grape-juice fermentation could help the winemaking industry by complementing the arts of the traditional wine taster. So Skogerson and Fiehn set out to survey wine 'metabolomes', in search of key chemical components contributing to body.

Using proton nuclear magnetic resonance (NMR)

and gas chromatography-mass spectrometry (GC-MS), they looked at 17 different white wines with a wide range of body. For GC-MS analysis, they first removed the alcohol under reduced pressure and then ran samples on a LECO Pegasus IV GC TOF MS system and analysed the spectra using the BinBase program developed in Fiehn's lab. Each wine was also directly analysed on a Bruker Daltonics 600 MHz NMR instrument with the resulting peaks being compared to the commercially available Chenomx NMR database for metabolite identification. "When you think about it, you have the grape metabolome being acted on by the yeast, plus the added complexity from the yeast metabolome, so the metabolite profile of a wine is very complex," says Skogerson.

They found a total of 413 metabolites among the wines — probably only a small fraction of



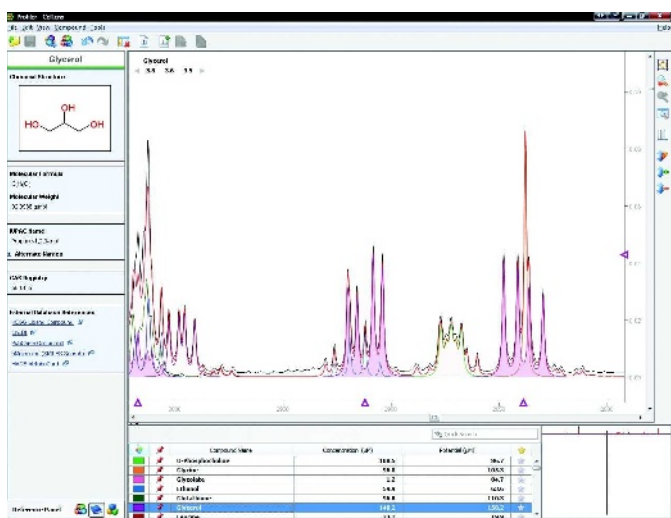
the wine metabolome — of which 108 could be positively identified. And in both data sets, the amino acid proline showed a positive correlation with body as assessed by trained wine tasters. How proline relates to body is not yet clear, however. "That is the hard part of

being in metabolomics — you get clues, but the follow-up is the real challenge," says Skogerson. Still, she thinks proline could be used as marker for a wine's viscosity.

Red-wine drinkers have not been forgotten. Bruker Daltonics in Billerica, Maryland, has profiled red wines for important polyphenolic secondary metabolites such as tannins, flavonoids and anthocyanins. This demonstration used the Acquity ultra-performance liquid chromatography system from Waters to separate red wine metabolites for analysis by Bruker's LC-ESI QTOF MS instrument as well as analysis by NMR coupled with Bruker's BioSpin Spectral Base analysis package.

Does knowing the chemistry behind that wonderful bottle of wine take away from the pleasure? Not according to Skogerson. "Science has the potential to bring the art of winemaking to a higher level."

N.B.



**Chenomx has developed a searchable NMR database for metabolomics.**

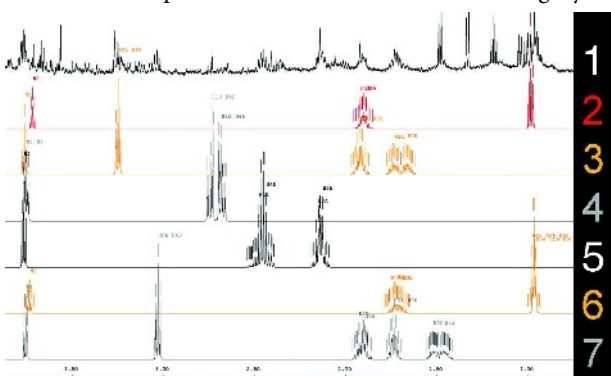
Both Agilent and Thermo Fisher Scientific also offer MS systems and software packages designed for metabolite analysis.

Some researchers and developers are designing platforms to bring the two camps closer together — incorporating NMR and MS instruments in a single system. Bruker BioSpin in Billerica, Massachusetts, has developed the Metabolic Profiler, a system that combines a liquid handler, the Avance III NMR spectrometer and an LC-electrospray ionization (ESI)-microTOF MS, all under the control of a single data-management and analysis system.

But what researchers dream of is a single detection ‘chip’ for all metabolites. “In my lab we have four platforms, and each platform looks at a certain part of the metabolome,” says Oliver Fiehn, a metabolomics researcher at the University of California, Davis. But he doubts that a single chip could ever become reality.

“The lack of such a technology is the Achilles heel of metabolomics,” says Wishart, noting that the most that researchers can analyse at any one time with current technologies is 10–15% of the entire metabolome — and even that’s stretching it.

“The big bottleneck is really compound identification,” says Fiehn. Unblocking it will need the addition of many more well-annotated reference spectra in the databases.



**Comparisons of samples to reference spectra databases help reveal the identity of metabolites.**

And that will take time. Chenomx was founded with the aim of developing a database for NMR analysis, and that has taken several years of intensive effort, says Newton. Different chemical environments can influence a compound’s NMR spectra, so researchers at Chenomx had to acquire spectra at ten pHs, ranging from 4 to 9, for each of the more than 300 reference compounds now in their proprietary database.

Metabolite data-

bases for MS have also been springing up as more researchers move into the field. One of the first was METLIN (<http://metlin.scripps.edu>), a publicly accessible database that was started in Siuzdak’s lab.

“We currently have 23,000 metabolites in there,” says Siuzdak, of which around 2,500 are identified endogenous metabolites. METLIN also contains a set of about 8,000 theoretical di- and tripeptides along with theoretical lipids, drugs and metabolites.

To expand the scope of METLIN, Agilent has collaborated with the Scripps Center for Mass Spectrometry to analyse chromatographic standards and add information about mass and retention time, with the intent of using these properties in addition to isotope pattern matching for identification. “Our goal is to get to the point where the most common metabolites encountered by researchers are easily identifiable,” says Agilent’s Fischer.

Gregory Stephanopoulos, a chemical engineer at the Massachusetts Institute of Technology in Cambridge, is taking a different approach to metabolite identification. Several years ago a student approached him with an interesting metabolomics project, but the catch was that the lab would first have to increase the number of reference spectra in its library to enable metabolite identification. Although Stephanopoulos liked the project, he did not like the idea of simply collecting spectra to fill a data-

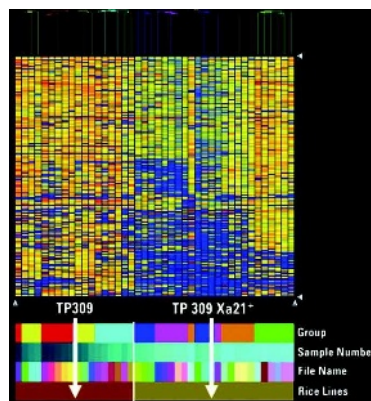
base. “I thought that there had to be a better way to deal with the issue,” he recalls.

The result is a web-based program called SpectConnect, which was launched in 2007 to help researchers identify important metabolites that might serve as biomarkers<sup>3</sup>. The SpectConnect algorithm tracks and catalogues GC–MS spectra that are conserved in multiple samples — an indication that these represent real compounds instead of noise in the sample, helping to guide researchers with their follow-up efforts at full metabolite identification.

**Numbers game**

The good news for metabolomics researchers is that NMR and MS metabolite databases are increasing in both number and size as new metabolomes are analysed. “One of the things that changed for us over the past 18 months is the places we are applying the technology,” says Michael Milburn, chief scientific officer at Metabolon in Durham, North Carolina. Metabolomic approaches are now addressing biological questions in areas ranging from drug discovery and cosmetics development to plant

science and winemaking (see ‘Wine-omics’). Publicly accessible databases include MassBank for high-resolution ESI mass spectra of metabolites ([www.massbank.jp](http://www.massbank.jp)), BinBase for processing and analysing of dissimilar MS spectra (<http://sourceforge.net/projects/binbase>), and MetWare (<http://msbi.ipb-halle.de/msbi/metware>) for the storage and analysis of metabolomic experiments. Commercial data-



**Samples can be grouped by the similarity of mass abundance profiles.**

bases include Metabolon’s, containing spectra of more than 6,000 reference metabolites, and Bio-Rad’s KnowITAll spectral database of more than 1.3 million entries, including MS and NMR references.

But there’s still a way to go before metabolite identification is as simple as ‘query and get a chemical name.’ “The database changes have been encouraging,” says Stephanopoulos. But not enough to change his mind about the need for tools such as SpectConnect.

Arthur Castle, programme director for the Roadmap Metabolomics Technology development programme at the US National Institutes of Health, has seen the pieces falling into place over the past couple of years. “The technology is very close to being there — it is just a question of putting it all together now,” he says.

**Nathan Blow is technology editor for Nature and Nature Methods.**

1. May, P. et al. *Genetics* **179**, 157–166 (2008).
2. Wishart, D. S. et al. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **15**, 164–173 (2008).
3. Styczynski, M. P. et al. *Anal. Chem.* **79**, 966–973 (2007).
4. Northen, T. R. et al. *Nature* **449**, 1033–1036 (2007).

separation by other types of chromatography.

One of these is high-performance liquid chromatography (HPLC), a well-established lab workhorse. It uses a combination of solvents, pressure and matrix particle sizes to separate molecules on the basis of their retention times in a column packed with matrix. HPLC can separate a broad range of metabolites, including non-volatiles, and remains a favourite among metabolomics researchers.

Most advances in HPLC involve increases in the pressure applied and changes in matrix particle size. Ultra-performance liquid chromatography (UPLC), commercialized by Waters Corporation in Milford, Massachusetts, is becoming more widely used in the metabolomics community. It takes advantage of higher pressure (83 megapascals compared with 21 megapascals for HPLC) and smaller particles (less than 2 micrometres diameter compared with 3 micrometres for HPLC) to obtain faster separation times.

But like GC, HPLC has technical stumbling blocks. Reversed-phase HPLC (in which the stationary phase is non-polar) is often used for metabolomics analysis, but reversed-phase separation often fails with hydrophilic metabolites. These tend to be so water soluble that they interact poorly with the non-polar bonding phase and are

rapidly eluted, according to Phil Koerner, a senior technical manager from chromatography specialists Phenomenex in Torrance, California.

So in 2007, Phenomenex introduced the Luna HILIC column. "I like to refer to it as reverse reverse-phase chromatography," says Koerner. In the HILIC approach, the weak solvent, which is applied first, is a polar organic solvent (not water as in reversed-phase HPLC), and the strong solvent, applied second, is water. This causes the order of elution to be completely reversed, with the most hydrophilic compounds being eluted last. Although Koerner acknowledges that the HILIC approach is not new, it was the need to separate hydrophilic metabolites on the large scale required by metabolomics that led Phenomenex and other companies, such as Waters and Tosoh Bioscience of Stuttgart, Ger-

many, to start supplying a greater number and range of HILIC columns.

Capillary electrophoresis followed by MS (CE-MS) is not yet so popular with the metabolomics community as either GC or HPLC, but several developers are hoping to change this. "It can be very difficult to use this approach," acknowledges Ryuji Kanno, president of Human Metabolome Technologies based in Tokyo, Japan. This approach uses electrophoretic mobility to separate low-molecular-weight ionic compounds that are difficult to separate by GC or HPLC. The company has been working closely with Agilent to develop optimized reagents and capillary columns, and is providing training with Agilent's CE-qTOF MS system to make the CE approach more accessible to metabolomics researchers, says Kanno.

Mass spectrometry is not the only method

that can be used to detect metabolites once separated. Wishart and his colleagues recently compared MS and NMR to look at metabolites in cerebral spinal fluid<sup>2</sup>. They found little overlap in the metabolites detected by the two methods, and the conclusion was clear: "We do not have a single perfect metabolite detector," says Wishart.

MS and NMR each have their supporters. "One of the main strengths of NMR is that



HILIC columns are making the hunt for hydrophilic metabolites easier.

PHENOMENEX

## DARK MATTER

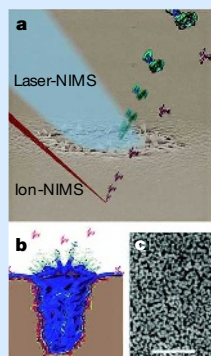
"There are a lot of small molecules that we do not even know about yet," says Arthur Castle, programme director for the Roadmap Metabolomics Technology development programme at the US National Institutes of Health in Bethesda, Maryland. Metabolomics has a good handle on analysing human primary metabolites, but when it comes to lipids, secondary metabolites, xenobiotics and the products of gut microflora, we are just scratching the surface, says Castle.

The problem is part technological, part informatics. Steven Fischer at Agilent Technologies in Santa Clara, California, points out that some compounds are not stable and undergo chemical transformation during separation. Researchers will probably still see these transformed molecules by mass spectrometry (MS), but they may

be misidentified, highlighting the need for follow-up experiments. And this is where metabolomics has an advantage over the other 'omics'. "There is so much knowledge of biochemistry that when we find a potential biomarker or a new drug mechanism we already know a lot about it," says Michael Milburn, chief scientific officer at Metabolon in Durham, North Carolina.

Trent Northen, now at the Lawrence Berkeley Laboratory in Berkeley, California, and Oscar Yanes, working in Gary Siuzdak's lab at the Scripps Research Institute in La Jolla, California, may have developed a new way to get at some of these 'unknown' molecules with

an ionization technique called nanostructure-initiator mass spectrometry (NIMS)<sup>4</sup>. The idea is to transfer a biomolecule into



NIMS: nanostructured surface releases material.

the gas phase from a nanostructured surface simply by making that surface disappear. "We came up with the idea of putting a wax underneath, so when the nanostructured surface was irradiated it would melt and vaporize, allowing the molecule to go into the gas phase," says Northen. They finally came up with a perfluorinate surface for the trapped initiator phase.

Siuzdak's group found that NIMS worked not only for proteins, but also for small molecules such as metabolites. "Perfluorinates do not ionize well, so it allows us to

see things in the lower mass region where metabolites like to hang out," says Siuzdak.

Using perfluorinated materials may have another advantage for metabolomics. "As these are highly hydrophobic surfaces, we can apply very dirty complex samples, such as blood, for direct analysis," says Yanes. If you put a drop of urine or blood on the NIMS chip the metabolites will attach, but all the salts and other chemicals that normally interfere with MS stay in solution. Yanes is now exploiting this property to follow drug metabolism by looking at uptake in blood, clearance in urine and tissue localization.

For Siuzdak, exploring the 'unknown' metabolite world is an exciting prospect. "We are getting involved in an area where we don't know what the molecules' structures are or what they do, so it is really just a fantastic area for discovery."

REF. 4

N.B.



The maXis system from Bruker Daltonics can use both UPLC and CE separation approaches.

Newton. He says researchers at Chenomx have performed many studies in which biologically meaningful differences between samples were easily captured with NMR, even though some compounds in the samples probably fell below the sensitivity limits of the instrument (see 'Dark matter').

MS, on the other hand, is a very sensitive method for metabolite identification and, unlike NMR, is easily coupled to upstream separation techniques. Siuzdak says his group can see thousands of molecules in an MS analysis — and that number can be doubled by changing from positive- to negative-ion mode. And by using both reversed-phase chromatography and HILIC columns, they are seeing more hydrophilic compounds in their analyses than before. "I would venture that we are now seeing over an order of magnitude more than what you would see with NMR," he says.

### Detector development

As researchers in the MS camp turn towards TOF and ion-trap MS instruments for metabolite analysis, developers are responding to their complex needs. Bruker Daltonics in Billerica, Massachusetts, has introduced the maXis ultra-high resolution (UHR)-TOF MS system, which can accommodate both UPLC and CE separation. Applied Biosystems in Foster City, California, in collaboration with MDS SCIEX in Toronto, Ontario, have the ion-trap system 4000 QTrap LC/MS/MS that can interface with Applied Biosystem's LightSight software for small-molecule analysis and identification.

it is an unbiased, universal detector," says Jack Newton, a product manager at Chenomx in Edmonton, Canada, which was co-founded by Wishart in 2000. This attribute, along with NMR's ability to determine structure and perform quantitative analysis is particularly attractive to metabolomics researchers who need a way to compare and exchange results between labs. "The move is afoot — people want to get

to that common language of compound names and concentrations," says Wishart, as this will make integrating data sets and obtaining systems-level views of cell physiology possible.

The challenge with NMR is instrument sensitivity — NMR is less sensitive than MS, often identifying far fewer metabolites in the same sample. "For us, the relevant question is how sensitive do you need to be," says

## WINE-OMICS

For Kirsten Skogerson at the University of California, Davis, wondering about how chemical composition affects the flavour and body of a wine took her from a degree in viticulture and enology into metabolomics research. When Skogerson arrived in Oliver Fiehn's lab as a postgrad she looked for a project that would marry Fiehn's expertise in metabolomics and her interest in wine.

"There are so many questions in wine science that you could start to answer by doing a global analysis," she says. A deeper understanding of the biochemistry of grape-juice fermentation could help the winemaking industry by complementing the arts of the traditional wine taster. So Skogerson and Fiehn set out to survey wine 'metabolomes', in search of key chemical components contributing to body.

Using proton nuclear magnetic resonance (NMR)

and gas chromatography-mass spectrometry (GC-MS), they looked at 17 different white wines with a wide range of body. For GC-MS analysis, they first removed the alcohol under reduced pressure and then ran samples on a LECO Pegasus IV GC TOF MS system and analysed the spectra using the BinBase program developed in Fiehn's lab. Each wine was also directly analysed on a Bruker Daltonics 600 MHz NMR instrument with the resulting peaks being compared to the commercially available Chenomx NMR database for metabolite identification. "When you think about it, you have the grape metabolome being acted on by the yeast, plus the added complexity from the yeast metabolome, so the metabolite profile of a wine is very complex," says Skogerson.

They found a total of 413 metabolites among the wines — probably only a small fraction of



the wine metabolome — of which 108 could be positively identified. And in both data sets, the amino acid proline showed a positive correlation with body as assessed by trained wine tasters. How proline relates to body is not yet clear, however. "That is the hard part of

being in metabolomics — you get clues, but the follow-up is the real challenge," says Skogerson. Still, she thinks proline could be used as marker for a wine's viscosity.

Red-wine drinkers have not been forgotten. Bruker Daltonics in Billerica, Maryland, has profiled red wines for important polyphenolic secondary metabolites such as tannins, flavonoids and anthocyanins. This demonstration used the Acquity ultra-performance liquid chromatography system from Waters to separate red wine metabolites for analysis by Bruker's LC-ESI QTOF MS instrument as well as analysis by NMR coupled with Bruker's BioSpin Spectral Base analysis package.

Does knowing the chemistry behind that wonderful bottle of wine take away from the pleasure? Not according to Skogerson. "Science has the potential to bring the art of winemaking to a higher level."

N.B.

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL
<b>Companies specializing in metabolomics applications</b>			
<a href="#">BIOCRAATES Life Sciences</a>	Services for the identification and quantification of metabolomic biomarkers	Innsbruck, Austria	<a href="http://www.biocrates.com">www.biocrates.com</a>
<a href="#">Chenomx</a>	NMR database development for metabolite analysis	Alberta, Canada	<a href="http://www.chenomx.com">www.chenomx.com</a>
<a href="#">Human Metabolome Technologies</a>	Developing CE separation approaches for metabolite analysis	Tokyo, Japan	<a href="http://www.humanmetabolome.com">www.humanmetabolome.com</a>
<a href="#">Ingenuity Systems</a>	Software for metabolomics and pathway analysis; Ingenuity Pathway Analysis; analysis services	Redwood City, California	<a href="http://www.ingenuity.com">www.ingenuity.com</a>
<a href="#">Metanomics Health</a>	Services for metabolite profiling	Berlin, Germany	<a href="http://www.metanomics-health.de">www.metanomics-health.de</a>
<a href="#">Metabolon</a>	Contract services provider in metabolomics	Durham, North Carolina	<a href="http://www.metabolon.com">www.metabolon.com</a>
<a href="#">Molecular Biometrics</a>	Using metabolomics approaches to develop accurate diagnostic tools	Chester, New Jersey	<a href="http://www.molecularbiometrics.com">www.molecularbiometrics.com</a>
<a href="#">Umetrics</a>	SIMCA-P+ software for the analysis of metabolomic data and biomarker discovery	Kinnelon, New Jersey	<a href="http://www.umetrics.com">www.umetrics.com</a>
<a href="#">Unilever</a>	Metabolomics research in food and flavour preservation	Vlaardingen, Netherlands	<a href="http://www.unilever.com">www.unilever.com</a>
<b>NMR instrumentation</b>			
<a href="#">Cryomagnetics</a>	NMR magnets and cryostat accessories	Oak Ridge, Tennessee	<a href="http://www.cryomagnetics.com">www.cryomagnetics.com</a>
<a href="#">JEOL</a>	NMR instrumentation, mass spectrometers and accessories	Tokyo, Japan	<a href="http://www.jeol.com">www.jeol.com</a>
<a href="#">LipoScience</a>	Using NMR to explore metabolism; NMR LipoProfile test	Raleigh, North Carolina	<a href="http://www.liposcience.com">www.liposcience.com</a>
<a href="#">Oxford Instruments</a>	Dilution refrigerators, superconducting magnets, optical and spectroscopy cryostats	Oxfordshire, UK	<a href="http://www.oxinst.com">www.oxinst.com</a>
<a href="#">Process Control Technology</a>	NMR instrumentation and accessories	Fort Collins, Colorado	<a href="http://www.pctnmr.com">www.pctnmr.com</a>
<a href="#">Process NMR Associates</a>	NMR services provider; laboratory method development for NMR	Danbury, Connecticut	<a href="http://www.process-nmr.com">www.process-nmr.com</a>
<a href="#">Wilmad-LabGlass</a>	NMR sample tubes and accessories; lab glassware	Buena, New Jersey	<a href="http://www.wilmad-labglass.com">www.wilmad-labglass.com</a>
<b>Chromatography</b>			
<a href="#">Advanced Separation Technologies</a>	Analytical and preparative chromatography products; HPLC and chiral chromatography	Bellefonte, Pennsylvania	<a href="http://www.astecusa.com">www.astecusa.com</a>
<a href="#">Alcott Chromatography</a>	HPLC equipment and software, ion chromatography, hydrochromatic chromatographs	Norcross, Georgia	<a href="http://www.alcottchromatography.com">www.alcottchromatography.com</a>
<a href="#">Biotage</a>	Chromatography columns and accessories; FLASH chromatography systems	Uppsala, Sweden	<a href="http://www.biotage.com">www.biotage.com</a>
<a href="#">Capital HPLC</a>	Suppliers of HPLC and CE columns and accessories	West Lothian, UK	<a href="http://www.capital-hplc.co.uk">www.capital-hplc.co.uk</a>
<a href="#">Cecil Instruments</a>	Modular HPLC instruments, ion chromatography, UV/visible spectrophotometers	Cambridge, UK	<a href="http://www.cecilinstruments.com">www.cecilinstruments.com</a>
<a href="#">DataApex</a>	Chromatography software	Prague, Czech Republic	<a href="http://www.dataapex.com">www.dataapex.com</a>
<a href="#">GE Healthcare</a>	AKTExpress automated liquid chromatography platform	Little Chalfont, UK	<a href="http://www.gehealthcare.com">www.gehealthcare.com</a>
<a href="#">Gilson</a>	Solid-phase extraction systems, liquid chromatography	Middleton, Wisconsin	<a href="http://www.gilson.com">www.gilson.com</a>
<a href="#">Hamilton</a>	Automated solid-phase extraction systems	Reno, Nevada	<a href="http://www.hamiltoncompany.com">www.hamiltoncompany.com</a>
<a href="#">Phenomenex</a>	Supplier of chromatography columns and accessories	Torrence, California	<a href="http://www.phenomenex.com">www.phenomenex.com</a>
<a href="#">Tosoh Bioscience</a>	Liquid chromatography columns, products and accessories	Stuttgart, Germany	<a href="http://www.tosohbioscience.com">www.tosohbioscience.com</a>
<b>Mass spectrometers</b>			
<a href="#">Agilent Technologies</a>	Mass spectrometry instruments, sample preparation and software	Santa Clara, California	<a href="http://www.agilent.com">www.agilent.com</a>
<a href="#">Applied Biosystems</a>	Mass spectrometry instruments, reagents and software; ion sources	Foster City, California	<a href="http://www.appliedbiosystems.com">www.appliedbiosystems.com</a>
<a href="#">Bruker</a>	Mass spectrometers	Billerica, Massachusetts and Bremen, Germany	<a href="http://www.bdal.de">www.bdal.de</a>
<a href="#">Hitachi High-Technologies</a>	LC systems; LC/MS systems	Tokyo, Japan	<a href="http://www.hitachi-hitec.com/global">www.hitachi-hitec.com/global</a>
<a href="#">LECO</a>	Mass spectrometry instrumentation and analysis tools	St Joseph, Michigan	<a href="http://www.leco.com">www.leco.com</a>
<a href="#">MDS SCIEX</a>	Mass spectrometry systems and technology; HPLC systems	Sunnyvale, California	<a href="http://www.mdsciex.com">www.mdsciex.com</a>
<a href="#">Shimadzu</a>	Laboratory instruments including mass spectrometers and data-management systems	Kyoto, Japan	<a href="http://www.shimadzu.com">www.shimadzu.com</a>
<a href="#">Thermo Fisher Scientific</a>	Analytical instruments including mass spectrometers, lab equipment, software, services, consumables and reagents	Waltham, Massachusetts	<a href="http://www.thermo.com">www.thermo.com</a>
<a href="#">Varian</a>	Analytical instruments including nuclear magnetic resonance and magnetic resonance imaging systems, mass spectrometers, Fourier transform-infrared and X-ray crystallography	Palo Alto, California	<a href="http://www.varianinc.com">www.varianinc.com</a>
<a href="#">Waters</a>	Liquid chromatography systems; chromatography columns and chemistry products; mass spectrometry systems; laboratory informatics solutions	Millford, Massachusetts	<a href="http://www.waters.com">www.waters.com</a>

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL
<b>General</b>			
<a href="#">Alexis Biochemicals</a>	Reagents for molecular- and cell-biology research	Lausanne, Switzerland	<a href="http://www.alexis-corp.com">www.alexis-corp.com</a>
<a href="#">Attagene</a>	Transcription-factor profiling system; software	Research Triangle Park, North Carolina	<a href="http://www.attagene.com">www.attagene.com</a>
<a href="#">BD Biosciences</a>	Research reagents, bioimaging systems, instrumentations	San Diego, California	<a href="http://www.bd.com">www.bd.com</a>
<a href="#">Biomol</a>	Services for chemical synthesis, cell culture and antibody production	Hamburg, Germany	<a href="http://www.biomol.de">www.biomol.de</a>
<a href="#">Bio-Rad</a>	Products, instruments and software for life-sciences research	Hercules, California	<a href="http://www.bio-rad.com">www.bio-rad.com</a> ●
<a href="#">BMG LABTECH</a>	Microplate and array readers and handling systems	Offenburg, Germany	<a href="http://www.bmglabtech.com">www.bmglabtech.com</a>
<a href="#">Brinkmann Instruments</a>	Laboratory instruments; consumables	Westbury, New York	<a href="http://www.brinkmann.com">www.brinkmann.com</a>
<a href="#">Cambrex</a>	Products for molecular- and cell-biology research	East Rutherford, New Jersey	<a href="http://www.cambrex.com">www.cambrex.com</a>
<a href="#">Cole-Parmer</a>	Instruments and reagents	Vernon Hills, Illinois	<a href="http://www.coleparmer.com">www.coleparmer.com</a>
<a href="#">EMD</a>	Calbiochem, Novabiochem and Novagen product lines	San Diego, California	<a href="http://www.emdbiosciences.com">www.emdbiosciences.com</a>
<a href="#">Enzo Life Sciences</a>	Consumables and assays for molecular biology, gene expression and genomic analysis	New York, New York	<a href="http://www.enzo.com">www.enzo.com</a>
<a href="#">Geneservice</a>	Genomic and proteomic resources; contract services for DNA sequencing, microarray analysis and SNP genotyping	Cambridge, UK	<a href="http://www.geneservice.co.uk">www.geneservice.co.uk</a>
<a href="#">Harvard Apparatus</a>	Instruments and equipment for electrophysiology and cell biology	Holliston, Massachusetts	<a href="http://www.harvardapparatus.com">www.harvardapparatus.com</a>
<a href="#">Horiba Jobin Yvon</a>	Spectroscopy systems and accessories including Raman, atomic emission, and UV spectroscopy	Edison, New Jersey	<a href="http://www.jobinyvon.com">www.jobinyvon.com</a>
<a href="#">INTEGRA Biosciences</a>	Equipment for sterilization, liquid handling, cell culture and sample storage	Baar, Switzerland	<a href="http://www.integra-biosciences.com">www.integra-biosciences.com</a>
<a href="#">Irvine Scientific</a>	Defined media for cell-culture applications; custom media services	Santa Ana, California	<a href="http://www.irvinesci.com">www.irvinesci.com</a>
<a href="#">Lonza</a>	Molecular biology reagents and systems; advanced chemical synthesis	Basel, Switzerland	<a href="http://www.lonza.com">www.lonza.com</a> ●
<a href="#">Merck</a>	Chemicals, kits and reagents for chemistry, molecular and cell biology-related research	Darmstadt, Germany	<a href="http://www.merck.de">www.merck.de</a>
<a href="#">Molecular Devices</a>	Liquid-handling and microplate processing equipment; imaging instruments	Sunnyvale, California	<a href="http://www.moleculardevices.com">www.moleculardevices.com</a> ●
<a href="#">MP Biomedicals</a>	Reagents and chemicals for research	Aurora, Ohio	<a href="http://www.mpbio.com">www.mpbio.com</a>
<a href="#">New England Biolabs</a>	Molecular-biology-related reagents, kits and enzymes	Ipswich, Massachusetts	<a href="http://www.neb.com">www.neb.com</a>
<a href="#">Nalge Nunc International</a>	Labware	Rochester, New York	<a href="http://www.nalgenunc.com">www.nalgenunc.com</a>
<a href="#">PerkinElmer</a>	Instruments, reagents and kits for life sciences	Waltham, Massachusetts	<a href="http://las.perkinelmer.com">las.perkinelmer.com</a>
<a href="#">Promega</a>	Chemicals for mass spectrometry analysis	Madison, Wisconsin	<a href="http://www.promega.com">www.promega.com</a> ●
<a href="#">PREMIER Biosoft International</a>	Software for life-sciences research	Palo Alto, California	<a href="http://www.premierbiosoft.com">www.premierbiosoft.com</a>
<a href="#">Princeton Separations</a>	DNA purification columns and reagents, fluorescent protein labelling kits,	Adelphia, New Jersey	<a href="http://www.prinsep.com">www.prinsep.com</a>
<a href="#">Stratagene</a>	Tools and reagents for molecular biology, genomics and proteomics	La Jolla, California	<a href="http://www.stratagene.com">www.stratagene.com</a>
<a href="#">Takara Bio</a>	Reagents, kits and consumables for molecular biology	Shiga, Japan	<a href="http://www.takara-bio.com">www.takara-bio.com</a>
<a href="#">Tocris Bioscience</a>	Chemicals for life-science research; contract research services	Avonmouth, UK	<a href="http://www.tocris.com">www.tocris.com</a>
<a href="#">Wako Chemicals USA</a>	Speciality chemicals supplier; clinical diagnostic reagents	Richmond, Virginia	<a href="http://www.wakousa.com">www.wakousa.com</a> ●
<a href="#">USB</a>	Chemicals and reagents for molecular biology	Cleveland, Ohio	<a href="http://www.usbweb.com">www.usbweb.com</a>

● see advertisement

# naturejobs

**THE CAREERS  
MAGAZINE FOR  
SCIENTISTS**

**T**he decline in the number of young junior faculty members at US universities could make it difficult to fill leadership positions in the future, according to a study by the American Council on Education. Just 3% of tenure-track faculty members at institutions that run full undergraduate courses are aged 34 or younger, so finding experienced people to be chairs, deans and university presidents will be difficult in the next generation, according to the council's report, *Too many Rungs on the Ladder? Faculty Demographics and the Future Leadership of Higher Education*.

Higher education has been affected by three main issues. First, faculty members are now generally older than they were when mandatory retirement was abolished in 1994, and are reluctant to retire. Second, universities rely more on part-time and non-tenure-track positions. And third, students are completing their doctorates later in life. The trend towards more and longer postdoc courses means that people have less time to gain the experience needed to advance into leadership positions.

The situation has both positive and negative aspects. On one hand, young faculty will have less competition and more positions available when they finally manage to move through the ranks. On the other hand, they might not have the chance to do so until they reach retirement age. So, a professor who is currently 45 years old could well still have the chance to become a dean — or even a university president — but perhaps not until he or she reaches the age of 70.

There are no easy solutions. Universities are unlikely to force ageing faculty off the top rung of the ladder or to create new positions for young professors. However, they may be forced to remove some rungs from the ladder — or else place some closer together, so that today's junior faculty members can progress from chair to dean more quickly. Perhaps today's administration can help mentor tomorrow's university leaders — or, at least, convince them that such positions are worth the wait.

**Paul Smaglik moderates the Naturejobs career forum on Nature Networks.**

## CONTACTS

**Editor:** Gene Russo

**European Head Office, London**  
The Macmillan Building,  
4 Crinan Street, London N1 9XW, UK  
Tel: +44 (0) 20 7843 4961  
Fax: +44 (0) 20 7843 4996  
e-mail: [naturejobs@nature.com](mailto:naturejobs@nature.com)

**European Sales Manager:**  
Andy Douglas (4975)  
e-mail: [a.douglas@nature.com](mailto:a.douglas@nature.com)

**Natureevents:**  
Ghizlaine Ababou (+44 (0) 20 7014 4015)  
e-mail: [g.ababou@nature.com](mailto:g.ababou@nature.com)

**UK Corporate:**  
Nils Moeller (4953)

**Southwest UK/RoW:**  
Alexander Ranken (4944)

**Northeast UK/Ireland:**  
Matthew Ward (+44 (0) 20 7014 4059)

**France/Switzerland/Belgium:**  
Muriel Lestringuez (4994)

**Scandinavia/Spain/Portugal/Italy:**  
Evelina Rubio-Hakansson (4973)

**North Germany/The Netherlands/Eastern Europe:** Reya Silao (4970)

**South Germany/Austria:**  
Hildi Rowland (+44 (0) 20 7014 4084)

**Advertising Production Manager:**

Stephen Russell  
To send materials use London address above.  
Tel: +44 (0) 20 7843 4816  
Fax: +44 (0) 20 7843 4996  
e-mail: [naturejobs@nature.com](mailto:naturejobs@nature.com)

**Naturejobs web development:** Tom Hancock  
**Naturejobs online production:** Dennis Chu

**US Head Office, New York**  
75 Varick Street, 9th Floor,  
New York, NY 10013-1917  
Tel: +1 800 989 7718

Fax: +1 800 989 7103  
e-mail: [naturejobs@natureny.com](mailto:naturejobs@natureny.com)

**US Sales Manager:** Peter Bless

**India**  
Vikas Chawla (+91 1242881057)  
e-mail: [v.chawla@nature.com](mailto:v.chawla@nature.com)

**Japan Head Office, Tokyo**  
Chiyoda Building, 2-37 Ichigayatamachi,  
Shinjuku-ku, Tokyo 162-0843  
Tel: +81 3 3267 8751  
Fax: +81 3 3267 8746

**Asia-Pacific Sales Manager:**  
Ayako Watanabe (+81 3 3267 8765)  
e-mail: [a.watanabe@natureasia.com](mailto:a.watanabe@natureasia.com)  
**Business Development Manager, Greater China/Singapore:**  
Gloria To (+852 2811 7191)  
e-mail: [g.to@natureasia.com](mailto:g.to@natureasia.com)

# MOVERS

**Stephen Brandt, Director, Oregon Sea Grant, Corvallis, Oregon**



**1997-2008** Director, NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan

**1994-97** Director, Great Lakes Center for Environmental Research and Education and Professor of Biology, State University of New York College at Buffalo, New York

As the new director of Oregon Sea Grant, Stephen Brandt has eagerly accepted a daunting task: helping the US Pacific coastal regions address fisheries declines and prepare for climate change. It's his latest interaction with Sea Grant, the coastal science programme of the US National Oceanic and Atmospheric Organization (NOAA), which has been a staple of his career since he was conducting graduate research.

California Sea Grant director Russ Moll says that Brandt's background will boost ecosystem-based management efforts. "Stephen is one of those rare folks with the skills to look at the big picture in oceans — which we need as we struggle with ecosystem-wide concerns such as ocean acidification," says Moll.

Brandt started his science career with a mathematics degree at the University of Wisconsin in Madison. But the outdoorsman decided to get a second degree in zoology, and spent weekends conducting field work on Wisconsin's freshwater lakes. That led to a graduate project applying sonar to study fish dynamics, then a PhD using underwater acoustics to see how temperature affects habitat preference in Great Lakes fish. But instead of accepting a tenure-track position there, he joined the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Marine Laboratories in Australia. "I sought adventure when, at 28, I took the Australia position — and I got a full-blooded marine experience," he says.

At the CSIRO, he refined acoustic approaches to investigate how Australia's vast, warm eddies might serve as nursery grounds for fish in the open sea. After four years, Brandt returned to the United States to study the Great Lakes' evolving salmon fishery with Sea Grant's programme at the State University of New York in Syracuse. Later, he studied the largest US estuary at the University of Maryland's Chesapeake Biological Laboratory.

When Sea Grant's Great Lakes Center for Environmental Research and Education was created in 1994, Brandt jumped at the chance to direct it. Four years later, he was overseeing the NOAA's Great Lakes Environmental Research Laboratory in Ann Arbor, Michigan, where he created a single 'science' branch to strengthen the cross-disciplinary work that bolsters their now-leading role in ecosystem forecasting. Moll says that Brandt's past success with region-wide projects will help the west coast to tackle the effects of climate change, including organism range shifts and increased storminess. ■

**Virginia Gewin**

## NETWORKS & SUPPORT

### Masters of professional science

Some people have expressed concern about the legitimacy, usefulness and costs of professional science master's (PSM) degrees, a relatively new US entity that could promise additional career avenues for fledgling scientists (see *Nature* **454**, 547; 2008). As advocates and purveyors of the PSM, we would like to address those concerns and endorse it as an option for scientists seeking management and science training in just a few years.

First, it is not intended for students considering a PhD, although a few PSM graduates do continue on to the PhD after becoming excited by the research to which they have been exposed. Rather, it is designed for students and science professionals who want to work in non-academic sectors, in interdisciplinary fields and in emerging areas. Science professionals looking to gain a competitive edge, re-enter the workforce or refine their skills may also find it worthwhile.

Second, the PSM is a relatively new degree; there are approximately 2,100 graduates nationwide. Not enough data exist yet to declare it a clear success. But the data we have are promising. A recent survey showed that in two years alone, the number of programmes grew by at least 20% (2006-08), and enrolment increased

54% (2004-06). Placement data are also encouraging: almost 70% of 2006 graduates who were not already working full-time found employment in business, government and non-profit sectors, either before they graduated or immediately after.

Cost is an issue. In contrast to PhDs, the master's degree is usually funded by the student. However, the highly competitive salaries PSM graduates can expect make it a worthwhile investment. A recent report from the National Research Council (NRC) notes a strong and growing current demand for master's-level science professionals and healthy growth in the salaries of master's degree-holders in science and engineering — salaries that have grown faster during the past ten years than those of PhD holders. The NRC advocated financial aid for PSM students.

There is good evidence to date that the PSM is a worthwhile investment, and that it benefits the institution and the employer. We are confident that forthcoming data will support the PSM even more strongly. ■

**Eleanor Babco is co-project director, Professional Master's Initiatives. Carol Lynch is senior scholar in residence and director of Professional Master's Programs at the Council for Graduate Schools.**

#### POSTDOC JOURNAL

### The coming challenge

In 2009, I will start yet another project: a baby. It is a terrifying prospect. But many seem surprised at my financial anxiety, given the Singapore government's policies aimed at boosting a birthrate in decline. Incentives range from financial bonuses to the creation of a fund to encourage family-friendly work practices. Some even say that childbirth is a woman's 'national service' or duty.

The tax breaks and additional days of childcare leave provide welcome relief, but as a researcher, many family-friendly options are not feasible, such as extended maternity leave or working part-time or from home. Given the high expectations of employers, a career break might mean career suicide. We are evaluated according to productivity, which is inevitably affected by parenthood. One non-scientist relative of mine was told that her maternity leave cost her a promotion. The competitive environment may be exacerbated by single people and childless couples who are upset by policies they perceive as discriminatory.

A newspaper article here recently profiled two successful female senior researchers, citing them as role models. One is single, the other divorced. Being successful may come at the cost of one's marriage. As I prepare to start a family, I must re-evaluate my priorities. I will soon discover for myself how Singapore's biomedical research community defines 'work-life balance'. ■

**Amanda Goh is a postdoctoral fellow in cell biology under the Agency of Science, Technology and Research in Singapore.**



naturejobs

More than  
just jobs!

Movers

Careers

Focuses

Regions

Podcasts

Prospects

Spotlights

Career fairs

Job channels

Announcements

Special reports

Visit

naturejobs.com

nature publishing group National University of Ireland, Galway  
*Ollscoil na hÉireann, Gaillimh*

## ESTABLISHED PROFESSOR OF PLANT SCIENCE

The National University of Ireland Galway is a student-centred, research intensive University of international standing, with a staff committed to excellence. The University is committed to delivering a stimulating learning environment and providing leadership in key scientific, economic, social, and cultural areas.

To further enhance the University's mission, applications are now invited for the post of Established Professor of Plant Science within the School of Natural Sciences at NUI Galway. The holder of the Chair will provide academic leadership to the Discipline of Botany/Plant Science within the School.

The Discipline of Botany/Plant Science is a key component of a newly established School, within the College of Science, that also includes Biochemistry, Earth & Ocean Science, Microbiology and Zoology. The various Disciplines within the School have a long-established international reputation for both teaching and research, led by more than 50 permanent members of academic staff. The School offers four-year BSc (Honours) undenominated degree programmes in each of its five cognate disciplines, as well as in several interdisciplinary, denominated areas. Its members are also active in taught and research Graduate programmes that lead to MSc and PhD Degrees, as well as Higher Diplomas.

All current academic members of staff in the Discipline of Botany/Plant Science are actively engaged in scholarship and research. Within the School, research in biological and environmental sciences span all levels of organisation, from molecular and cell biology, through research at the organismal level, to the study of populations, communities and ecosystems. These activities are aligned with the University's strategic, cross-cutting research areas in biomedical research, energy, environment and the marine. Members of the School and Discipline have strong interactions with the University's Research Institutes, including the Environmental Change Institute, the National Centre for Biomedical Engineering Science and the Martin Ryan Marine Science Institute and also with other Schools across the University. Further details can be obtained in the supporting documentation and web sites listed therein.

Applicants for the Chair should have relevant expertise in Molecular/Cell Science or in the area of the Environment (Terrestrial, Freshwater or Marine). It will also be an advantage if a candidate's interests are a good fit to the strategic goals of the School and University and further strengthen interactions with some of the University's cognate Research Institutes.

*The successful candidate shall possess:*

- The capacity to provide dynamic academic leadership in the development of the discipline and generally in the promotion of teaching and research in the plant sciences
- An internationally-recognised track record for high-quality research in a relevant area as evidenced by an exceptional publication record in top-ranking, peer-reviewed international journals, a proven ability to develop innovative research together with demonstrated success in competing for funding at national and international levels
- A strong commitment to excellence in teaching and learning at all levels within the University as evidenced by significant experience of course delivery and design
- The capacity to represent effectively the Discipline and the School, inside and outside the University
- The ability to potentially act as Head of School and, in particular, demonstrated excellent interpersonal, communication and other relevant skills appropriate to that role
- A willingness to participate in the overall life of the University.

**Additional information** is available at: <http://www.nuigalway.ie/botany>

**For informal discussion contact:** Dr. Gerry Morgan, Dean of the College of Science, 091-493615 or email [dean.science@nuigalway.ie](mailto:dean.science@nuigalway.ie)

W171017R

[www.nuigalway.ie](http://www.nuigalway.ie)



## Science Foundation Ireland

# Ireland

funds great research...  
...maybe it's your turn!

### Science Foundation Ireland, (SFI)

the national foundation for excellence in scientific research is investing in academic researchers and research teams who are most likely to generate new knowledge, leading edge technologies, and competitive enterprises.

SFI has a flexible grants and awards portfolio and several times a year issues calls for proposals from scientists and engineers. SFI's award programmes include:

#### Principal Investigator Programme

for outstanding researchers, normally ranging between €50,000 - €1 million per year and may be up to five years in duration.

#### Research Professor Recruitment Awards

for outstanding researchers, with particularly distinguished international reputations, awards normally ranging up to €500,000 per annum for up to two years.

#### E.T.S. Walton Visitor Awards

supporting leading international scientists who visit Ireland to undertake research for up to one year, normally ranging up to €200,000. **Call opens 6 October 2008**

#### President of Ireland Young Researcher Awards (PIYRA)

attracting to Ireland and supporting Irish researchers within five years of completing their PhD, normally up to €1 million over five years.

### Career Opportunities at SFI

SFI's culture is innovative, highly motivated, and dynamic. We employ people from science, business, technology, and academic backgrounds who can help strengthen an outstanding Irish research environment.

For details on current employment opportunities at SFI please visit [www.sfi.ie](http://www.sfi.ie) or contact [hr@sfi.ie](mailto:hr@sfi.ie)

### Current employment opportunities on SFI funded research projects

#### University College Dublin

Further Details on UCD Posts:

[www.ucd.ie/jobopportunities](http://www.ucd.ie/jobopportunities)

#### UCD School of Electrical, Electronic and Mechanical Engineering UCD Engineering and Materials Science Centre

Charles Parsons Award Researcher - Fixed Term 6 Year Post

#### UCD School of Agriculture, Food Science & Veterinary Medicine Reproductive Biology Research Cluster (RBRC)

Postdoctoral Fellow – Fixed Term Post

#### UCD School of Biomolecular and Biomedical Science

Postdoctoral Fellow - Cell and Molecular Biology - Fixed Term Post

Postdoctoral Fellow - Molecular Biology & Bioinformatics - Fixed Term Post

Further Details on UCD Studentships:

[www.ucd.ie/graduatestudies/opportunities](http://www.ucd.ie/graduatestudies/opportunities)

#### UCD School of Medicine and Medical Sciences Conway Institute of Biomolecular and Biomedical Research

Phd Studentships to Study the Contribution of the IL-13 Receptor Complex to the Development of Pulmonary Fibrosis

#### UCD School of Agriculture, Food Science & Veterinary Medicine Reproductive Biology Research Cluster (RBRC)

Phd Projects available include the Regulation of MHC-1 Expression during Preimplantation Embryo Development and Molecular Analysis of Mammalian Oocyte Maturation

**Tyndall National Institute**Further Details: [www.tyndall.ie/careers](http://www.tyndall.ie/careers)

Phd position - Theory of Next Generation Epitaxial Quantum Dot Materials and Devices

Phd position - Design of Molecular and Nanoscale Electronic Devices

Phd position - In-situ Spectroscopic Reaction Monitoring

Phd position - Design and Synthesis of Novel Precursors

Phd position - Kinetic Monte Carlo Simulations

Phd position - Electrical Characterisation of Dielectric/Gate Structures

Phd position - Inelastic Electron Tunneling Spectroscopy

Phd position - First Principles Modelling

Phd position - Nano-casting Routes

Phd position - Interconnect Materials

Phd position - Advanced Thin Film Growth by CVD and ALD

Phd position - Development of Large-Scale Colloidal Crystallisation Methods for the Production of Photonic Crystals

Post-doc position - Experimental Physics of Photonic Materials and Devices

Post-doc position - Theory of Extreme Semiconductor Alloys

Phd position - Avalanche Photodetectors

Phd position - Miniaturised Systems for the Built Environment

Phd position - Develop energy harvesting mechanisms for the next generation, wearable wireless sensor network systems

Phd position - Investigate &amp; Develop Photonics Systems Research

Phd position - Long Wavelength, Site Controlled Pyramidal Quantum Wires

Phd position - Biology of novel vascular progenitor cells and the implications for this field by developing platform diagnostic devices for cardiovascular disease

Phd position - Solid State Electronic Structure Theory

Phd position - Development of Electrical Contacting Strategies for Template-Based Nanostructures and Nanostructured Arrays

Phd position - Synthesis and Characterisation of Optically Active Nanostructured Oxides

Phd position - Harmonic Mode Locking in MultiWavelength Fabry-Perot Lasers

Phd position - Nonlinear Optics of Refractive Index Profiles with Tailored Group Velocity

Post-doc position - Optical Signal Processing for High Capacity Networks

Post-doc position - All-Optical Switching Techniques for Optical Networks

Post-doc position - Wearable Wireless Sensor Network Systems for Health

Post-doc position - Energy Aware Hardware for Wireless Sensor Network Systems

Post-doc position - Deployment Engineer for Wireless Systems

Post-doc position - Advanced Semiconductor Laser, Simulation, Design, Characterisation and Test

Post-doc position - Packaging Engineer

Post-doc position - Electron Microscopy Technician

Post-doc position - Generation and Characterisation of Ferroelectric and High-K Metal Oxide Dielectric Nanostructures

Senior Post-doc position - Photonic Communication System Research

Post-doc position - Energy Harvesting for Wireless Sensors in the Built Environment

**National University of Ireland, Maynooth**Further Details: [www.cs.nuim.ie/~tนาughton/](http://www.cs.nuim.ie/~tนาughton/)

Post-doc position (12 months) Department of Computer Science - Image processing and analysis of digital holograms of real-world three-dimensional objects

**Trinity College Dublin**Further Details: [www.tcd.ie/vacancies](http://www.tcd.ie/vacancies)**School of Genetics and Microbiology  
Smurfit Institute of Genetics**

2 Postdoctoral Fellowships in Bioinformatics and Genome Evolution

Postdoctoral Fellowship

PhD Studentship

**School of Medicine  
Department of Clinical Medicine**

Phd Studentship (3-year contract)

**School of Engineering  
Trinity Centre for Bioengineering**

Phd Studentship in Computation Studies in Mechanobiology

**CRANN (Centre for Research on  
Adaptive  
Nanostructures and  
Nanodevices)**

Postdoctoral Researcher (1-year contract)

**National University of Ireland,  
Galway**

Further Details:

[www.nuigalway.ie/vacancies](http://www.nuigalway.ie/vacancies)

Postdoctoral Researcher in Nuclear Organisation and Gene Expression

**Dublin City University**

Further Details:

[www.dcu.ie/vacancies/current.shtml](http://www.dcu.ie/vacancies/current.shtml)

Postdoctoral Research in Optical Communications and High Speed Opto-Electronics - 3 year contract RINCE

**Biomedical Diagnostics Institute  
(BDI)**

Further Details:

[www.bdi.ie/about\\_bdi/careers.html](http://www.bdi.ie/about_bdi/careers.html)

Postdoctoral Researcher - Microfluidics &amp; Polymer Microfabrication - 12 month contract

Postdoctoral Researcher - Surface Science of Bioassay Devices - 12 month contract

Research Assistant (position based at Royal College of Surgeons in Ireland) - Immunoassay/Protein Biochemist - 12 month contract

**Science Foundation Ireland**Wilton Park House, Wilton Place, Dublin 2, Ireland. tel +353 1 607 3200 fax +353 1 607 3201 email [info@sfi.ie](mailto:info@sfi.ie)Apply for an SFI award or learn more about our programmes at [www.sfi.ie](http://www.sfi.ie)

National Development Plan 2007 - 2013

## Professorships and Lectureships in Science, Engineering and Computing

Dublin City University ([www.dcu.ie](http://www.dcu.ie)) is Ireland's youngest university and our rapid growth and leadership has helped to transform many of the traditional practices and assumptions of higher education. Through our current strategic plan we aim to distinguish ourselves through our engagement with industry and business in building internationally competitive research teams, which will succeed in converting the results of high quality basic research into significant social and economic impacts.

DCU wishes to invite **Expressions of Interest** from high quality candidates at both **Lecturer and Professor** levels, in context of the preparation of the DCU submission to the Science Foundation Ireland 'Stokes Professorship and Lecturer Programme' 2009 ([www.sfi.ie](http://www.sfi.ie)), with a specialty in one or more of the following fields:

### Biological/Health Sciences

- Biomedical Diagnostics
- Drug Discovery and Development
- Cancer Research
- Therapeutics

### Engineering

- Image processing & Analysis Network Innovations
- High Speed Devices and Systems and
- Content Analysis in Context
- Innovative Manufacturing
- Biomedical Engineering Sustainable Technology

### Computing

- Software Engineering, Language Technologies, Scientific
- Computing Information Security

### Physical/Chemical Sciences

- Photonic Sciences
- Plasma Science and Technology
- Environmental Monitoring

Further details on the application process are available at: [www.dcu.ie/vacancies/current.shtml](http://www.dcu.ie/vacancies/current.shtml)


**Expressions of Interest should be emailed to: [hr.applications@dcu.ie](mailto:hr.applications@dcu.ie) by 5pm on Friday October 24th.**

W171220R

**Is your work-life  
out of balance?**



**Keep up-to-date with all  
the latest issues affecting  
scientists and their  
working environment on  
[naturejobs.com](http://naturejobs.com)**

nature publishing group 

# Mars is the wrong colour

No longer seeing red.

**Ian Randal Strock**

There's a difference between a conspiracy and what Margaret Mead called "a small group of thoughtful, committed people". The former will always fail, because someone will give it away. The latter, as Mead said, "can change the world. Indeed, it is the only thing that ever has."

Many ventures fail not for a bad idea or poor planning, but because the wrong group of people is involved. I'd learned that time and again in my organizing efforts: fan groups, idealistic businesses and even social organizations that sputtered along but eventually collapsed because the people involved didn't mesh well, or didn't have the right skill sets.

This time, I vowed, it would be different. This time, my goal dwarfed all the others, and required far more commitment and cohesion among the people I'd be gathering to carry out the plan. So I went slowly. I didn't start with public pronouncements or marketing campaigns or mass appeals to everyone I knew.

Instead, I found people of like mind through quiet, one-on-one conversations. I found my potential collaborators at scientific conferences, through friends and business acquaintances and, in one instance, a call out of the blue to an author who wrote something in a story that resonated with me.

It was a large group for a conspiracy, but a small number of people to carry out such a grandiose plan. And this one time, finally, I knew before I asked that they would all say yes. I convinced my double handful of potential co-conspirators to get together at a quiet resort during the off-season. They all said yes to the first date I proposed.

After the introductions were out of the way, I think they all realized I had something larger in mind than a social gathering. I stood up, and everyone else fell silent, looking at me with anticipation.

"Mars is the wrong colour," I said. "After so many missions looking for life, I think we can be fairly certain that there is none on Mars. There may once have been, but there's nothing now, and the environment won't permit anything we'd recognize as life to grow."

There were nods of agreement around the room, and an encouraging, "Yes, go on."

"We can keep going as we have been, sending probes to Mars every two years to investigate smaller and smaller possibilities, or we can gather a small group of people," I looked around the room with purpose, and knew they were the right crowd, because every one of them met my eyes and smiled, "and admit there's no life, but that it's time to start seeding it."

"NASA is launching the Firebird in a year. It will be a successor to the Phoenix, with a soil laboratory, a digging arm and mobility."



It was definitely the right crowd. I didn't have to finish the concept before someone else said: "The Flora experiment. We'll need to co-opt it, make it smaller and lighter than the programme integrators expect."

"Then we can take the extra space and put what in it?"

"Extremophiles. Some sort of microbes that will be happy in the current Martian environment, and which will excrete —"

"The oxygen and ozone we need to transform the environment."

"Someone will notice."

"By then, it'll be too late to stop the process. If word doesn't get out, the rest of the world may assume it's a natural process."

"Or a miracle."

"But we'll know better."

"We won't have a living Mars to visit, nor will our grandchildren."

"No, but *their* grandchildren will. Governments move too slowly, and don't think far enough ahead."

"Better a second habitable planet in a bunch of generations than none at all."

"If we're caught ..."

"What? No country I know has a law against terraforming Mars."

"But they'll come up with something. Theft of government services? Deceit? Lying to a federal agency?"

"Then we'll pay a fine or go to jail."

"But Mars will live."

"What do we call the project?"

"If there's a name, there will be something to give us away. Something to let slip."

"Don't use your government e-mail accounts to talk about this with anyone."

We broke up into working groups. One to decide which microbes to send. Another to build their flight compartment. Yet another to redesign Flora to make room for our unannounced package. And those who would be inspecting the craft before launch, who were necessary to overlook our modifications. By the end of the weekend, we came back together, and we had a workable plan. The scheduling was going to be tight, but that only meant less time for our plot to be discovered and stopped.

We all left that weekend vowing to carry out our plan without telling a soul: not our spouses, not our children, and leave not a word in our wills. What we were doing, we were doing for all humanity.

I organized the group: that was my contribution. I'm not a scientist, not a biologist, I don't work for NASA, and I had nothing more to do with any of the project members. But ego is a funny thing. As the organizer, I feel a certain responsibility to them, to let the world know that there is hope; the hope of another planet for our grandchildren's grandchildren, if only we can keep ourselves alive long enough for Mars to be ready. Firebird will land on Mars very soon; there's nothing anyone can do now to stop the process, short of nuking Mars, but that's not going to happen. So I've written this story.

Keep looking up, and tell your children. There will come a time when Mars is no longer the red planet. Watch for the white clouds, the blue water and the green life. It's coming. ■

**Ian Randal Strock** (<http://ianrandalstrock.livejournal.com>) is the editor of *SFScope.com*, the online trade journal of the science-fiction fields. He is also the author of *The Presidential Book of Lists* (Random House/Villard, October 2008).

JACEY