

Impermanence of dendritic spines in live adult CA1 hippocampus

Alessio Attardo^{1,2†*}, James E. Fitzgerald^{1†*} & Mark J. Schnitzer^{1,2,3}

The mammalian hippocampus is crucial for episodic memory formation¹ and transiently retains information for about 3–4 weeks in adult mice and longer in humans². Although neuroscientists widely believe that neural synapses are elemental sites of information storage³, there has been no direct evidence that hippocampal synapses persist for time intervals commensurate with the duration of hippocampal-dependent memory. Here we tested the prediction that the lifetimes of hippocampal synapses match the longevity of hippocampal memory. By using time-lapse two-photon microendoscopy⁴ in the CA1 hippocampal area of live mice, we monitored the turnover dynamics of the pyramidal neurons' basal dendritic spines, postsynaptic structures whose turnover dynamics are thought to reflect those of excitatory synaptic connections^{5,6}. Strikingly, CA1 spine turnover dynamics differed sharply from those seen previously in the neocortex^{7–9}. Mathematical modelling revealed that the data best matched kinetic models with a single population of spines with a mean lifetime of approximately 1–2 weeks. This implies ~100% turnover in ~2–3 times this interval, a near full erasure of the synaptic connectivity pattern. Although *N*-methyl-D-aspartate (NMDA) receptor blockade stabilizes spines in the neocortex^{10,11}, in CA1 it transiently increased the rate of spine loss and thus lowered spine density. These results reveal that adult neocortical and hippocampal pyramidal neurons have divergent patterns of spine regulation and quantitatively support the idea that the transience of hippocampal-dependent memory directly reflects the turnover dynamics of hippocampal synapses.

The hypothesis that synaptic connectivity patterns encode information has profoundly shaped research on long-term memory. In the hippocampus, synapses in basal CA1 mainly receive inputs from hippocampal area CA3, and the CA3 → CA1 projection has been widely studied regarding its plasticity and key role in memory. As in the neocortex, dendritic spines in the hippocampus are good proxies for excitatory synapses¹², motivating time-lapse imaging of spines as a means of monitoring synaptic turnover^{7–10}.

Previous work has illustrated *in vivo* imaging of CA1 spines in acute and recently also in chronic preparations^{13–15}. We tracked spines for up to ~14 weeks by combining microendoscopes of diffraction-limited resolution¹⁴ (0.85 NA), a chronic mouse preparation for time-lapse imaging in deep brain areas⁴, and *Thy1-GFP* mice that express green fluorescent protein (GFP) in a sparse subset of CA1 pyramidal neurons (Fig. 1 and Extended Data Fig. 1). Histological analyses confirmed that this approach induced minimal activation of glia (Extended Data Fig. 2), as shown previously^{4,16}.

A major concern was that it is not possible to distinguish two or more spines spaced within the resolution limit of two-photon microscopy. This issue is critical for studies of hippocampal spines, which are more densely packed than neocortical spines¹⁷. To gauge how commonly the appearances of adjacent spines merged together in optical images, we examined tissue slices from *Thy1-GFP* mice, using both

two-photon microendoscopy and stimulated-emission depletion (STED) microscopy. STED microscopy offered super-resolution (~70 nm full width at half maximum (FWHM) lateral resolution), an optical resolution nearly nine times finer than that of two-photon microendoscopy¹⁴ (~610 nm), permitting tests comparing pairs of images of the same CA1 dendrites (Fig. 2a and Extended Data Fig. 3).

As expected, we saw nearby spines in STED images that appeared merged in the two-photon images (Fig. 2a). Of 151 spines that appeared unitary in the two-photon images, $23 \pm 3.6\%$ (standard error of the mean (s.e.m.)) were actually two spines and $6.0 \pm 1.6\%$ were actually three spines ($n = 12$ dendrites) (Fig. 2b). Distances between merged spines in the two-photon images ($0.51 \pm 0.14 \mu\text{m}$; mean \pm standard deviation (s.d.)) were below the ($0.61 \mu\text{m}$) resolution limit¹⁴ (Fig. 2c). Clearly, merging can induce illusory spine stability, since two or more real spines must vanish for a merged spine to disappear.

To treat merging effects quantitatively, we developed a mathematical framework that permits systematic examination of turnover dynamics across different kinetic models and investigation of how the merging of spines on imaging alters the manifestations of these dynamics in two-photon imaging data (Supplementary Information and Extended Data Figs 4–7). We used computer simulations to study how the density and apparent kinetics of merged spines vary with geometric variations of individual spines, spine density, resolution and spine kinetics (Supplementary Information and Extended Data Fig. 8). We also checked experimentally whether fluctuations in spine angle and length, and the radius of the dendrite near the spine, might impact measures of spine turnover (Extended Data Figs 6 and 9). By simulating time-lapse image series, we scored and analysed synthetic data across a broad range of optical conditions, spine densities, geometries and turnover kinetics (Fig. 2d and Extended Data Figs 4 and 5).

The simulations and mathematical modelling confirmed that naive analyses of two-photon data are inappropriate at the spine densities in CA1, owing to merging and the resulting illusion of increased stability (Supplementary Information and Extended Data Figs 4 and 7). For stability analyses, we followed previous studies^{7–9} in our use of the survival fraction curve, $S(t)$, the fraction of spines appearing in the initial image acquired at $t = 0$ that also appeared in the image acquired at time t . The shape and asymptotic value of $S(t)$ provide powerful constraints on kinetic models of turnover and the fraction of spines that are permanent (that is, a rate constant of zero for spine loss)^{7–9} (Supplementary Information). Strikingly, visual scoring of simulated images (Fig. 2d and Extended Data Fig. 5) yielded underestimates of spine density (Fig. 2e) and patent overestimates of the lifetimes of spines (Fig. 2f). But crucially, our treatment accurately predicted the relationship between the actual density and the visually determined underestimate (Fig. 2e), and properly explained the apparent turnover dynamics, $S(t)$, in terms of the actual kinetics (Fig. 2f and Extended Data Figs 5c, 7). Overall, the simulations showed that face-value interpretations of two-photon images from CA1 are untrustworthy, but

¹James H. Clark Center for Biomedical Engineering & Sciences, Stanford University, Stanford, California 94305, USA. ²Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA. ³CNC Program, Stanford University, Stanford, California 94305, USA. [†]Present addresses: Max Planck Institute of Psychiatry, 80804 Munich, Germany (A.A.); Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA (J.E.F.).

*These authors contributed equally to this work.

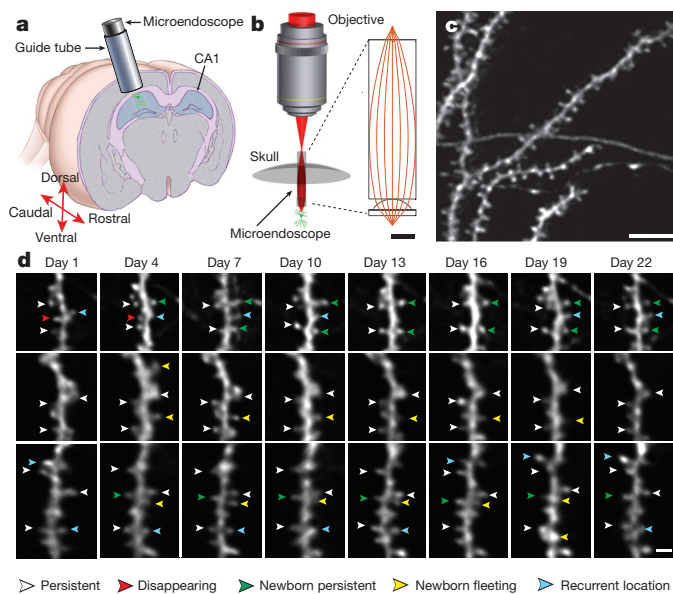


Figure 1 | Dendritic spines are dynamic in CA1 hippocampus of the adult mouse. **a**, A sealed, glass guide tube implanted dorsal to CA1 allows time-lapse *in vivo* imaging of dendritic spines. **b**, A doublet microendoscope projects the laser scanning pattern onto the specimen plane in the tissue. Inset: red lines indicate optical ray trajectories. **c**, CA1 dendritic spines in a live *Thy1-GFP* mouse. **d**, Time-lapse image sequences. Arrowheads indicate spines that either persist across the sequence (white arrowheads), disappear midway (red), arise midway and then persist (green), arise midway and then later disappear (yellow), or disappear and then later appear at an indistinguishable location (cyan). Scale bars: 500 μm (**b**, inset); 10 μm (**c**); 2 μm (**d**).

that it is possible to make quantitatively correct inferences about spine kinetics provided that one properly accounts for the optical resolution. Using the same framework, we next analysed real data.

Initial analyses focused on four mice in which we acquired image stacks of CA1 pyramidal cells and tracked spines every 3 days for 21 days (60 dendrites total; 50 ± 7 (mean \pm s.d.) per session) (Fig. 3a). Whenever individual spines appeared at indistinguishable locations on two or more successive sessions, we identified these as observations of the same spine. Overall, we made 4,903 spine observations (613 ± 71 (mean \pm s.d.) per day). Spine densities were invariant over time (Fig. 3b) (Wilcoxon signed-rank test; $n = 16$ –50 dendrites per comparison of a pair of days; significance threshold = 0.0018 after Dunn–Šidák correction for 28 comparisons; $P > 0.047$ for all comparisons), as were spine volumes ($P = 0.87$; $n = 43$ spines, Kruskal–Wallis analysis of variance (ANOVA)) (Extended Data Fig. 10) and the turnover ratio, the fraction of spines arising or vanishing since the last session⁷ (Fig. 3b) (Wilcoxon signed-rank test; 14–40 dendrites; significance threshold = 0.0025 after correction for 20 comparisons; $P > 0.31$ for all comparisons). Fewer than 50% of spines ($46 \pm 2\%$; mean \pm s.e.m.; $n = 4$ mice) were seen throughout the experiment (Extended Data Fig. 1d, e), although our simulations had shown that this naive observation overestimated spine stability.

The time invariance of spine densities and turnover ratios implied that through our mathematical framework we could determine the underlying kinetic parameters governing turnover. $S(t)$ curves for the total and newborn spine populations ostensibly resembled those reported for the neocortex^{7–9} (Fig. 3c). However, unlike in the neocortex, the 46% of spines seen in all sessions differed notably from the odds that a spine was observed twice in the same location across two distant time points, as quantified by the asymptotic value of $S(t)$ ($73 \pm 3\%$). This discrepancy suggested that, as our modelling had indicated, many CA1 spines might vanish and reappear in an ongoing way at indistinguishable locations.

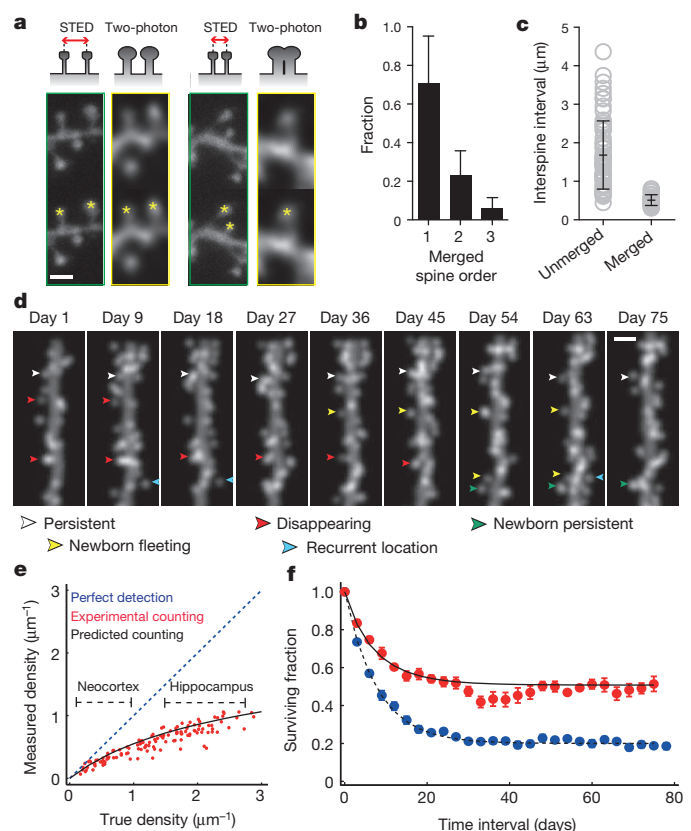


Figure 2 | A simple kinetic model is sufficient to describe CA1 pyramidal cell spine dynamics. **a**, Two-photon microendoscopy and STED imaging of the same dendrites *in vitro*. Top, two-photon images depict spines closer than the resolution limit as merged entities. Bottom, asterisks mark example visually scored spines, showing cases in which nearby spines do (right) or do not (left) merge. **b**, Fraction of spines ($n = 151$ total) seen by two-photon imaging that were one, two or three spines as determined by STED imaging. Black bars show mean \pm s.d. for 12 dendrites. **c**, Separations between adjacent unmerged spines and pairs of spines that appeared merged by two-photon imaging. Open grey circles mark individual results from each of $n = 150$ spines. Black bars show mean \pm s.d. **d**, Example computer-simulated, time-lapse image sequence used to quantify how resolution limits impact measured spine densities and dynamics. **e**, Computational modelling predicts the underestimation of spine density due to the finite optical resolution. Blue diagonal line: perfect detection of all spines. Black horizontal dashed lines: typical ranges of spine densities on pyramidal cells in neocortex and hippocampus. Red data: results from visually scoring simulated images of dendrites of varying spine densities. Black curve: prediction from the scoring model using 600 nm as the minimum separation between two spines correctly distinguished. **f**, Modelling predicts the overestimation of spine stability due to merging of adjacent spines in resolution-limited images. Blue data: survival fraction values (mean \pm s.e.m.) for actual spine turnover in computer simulations (spine density: $2.56 \mu\text{m}^{-1}$). Red data: apparent turnover for these same simulated dendrites, as scored from simulated two-photon images. Black curves: theoretical predictions for spine survival based on the scoring model. Scale bars: 1 μm (**a**); and 2 μm (**d**).

We next tested whether a prolonged environmental enrichment would alter spine turnover. Previous data from rats have indicated that basal CA1 spine density can rise $\sim 10\%$ after environmental enrichment¹⁸. Data from mice are limited to CA1 apical dendrites and have yielded contradictory results^{19,20} (Supplementary Discussion). In three mice we imaged a total of 55 basal dendrites (39 ± 14 (s.d.) per day) across 16 sessions, 3 days apart (Fig. 3d). After session 8 we moved the mice to an enriched environment, where they stayed throughout sessions 9–16. We made 8,727 spine observations in total (545 ± 216 (s.d.) per day).

Comparisons of baseline and enriched conditions revealed no differences in spine density or turnover (Fig. 3e) ($n = 10$ –53 dendrites;

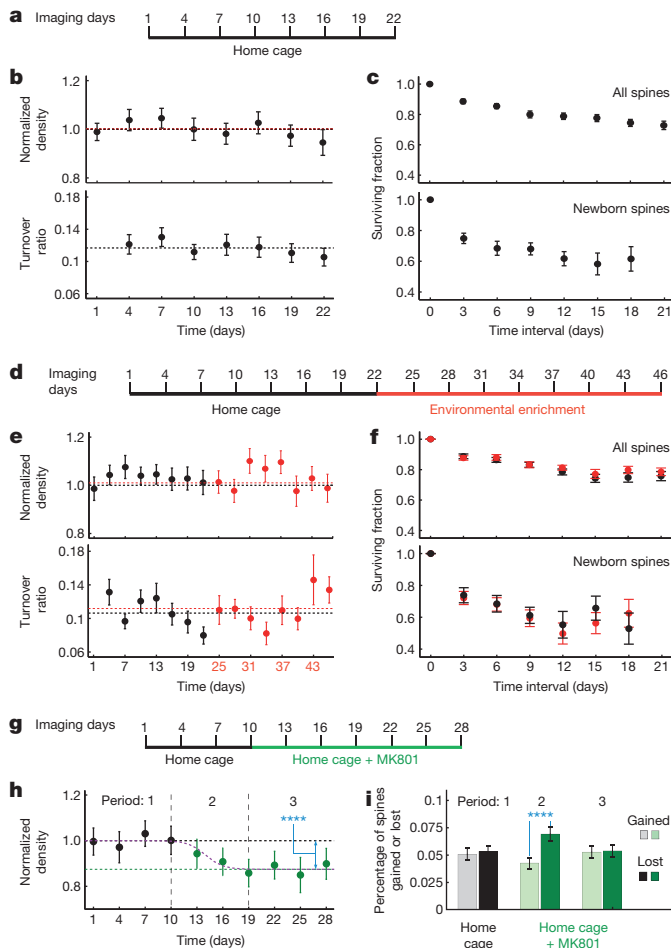


Figure 3 | NMDA receptor blockade, but not environmental enrichment, altered spine turnover dynamics. **a**, Schedule of baseline imaging sessions. **b**, Neither measured spine densities (top) nor turnover ratios (bottom) varied for mice in their home cages. Horizontal lines: mean spine density and turnover ratio. **c**, Spine survival (top) and newborn spine survival (bottom). **d**, Schedule for study on environmental enrichment. **e**, **f**, No significant differences existed between baseline (black points) and enriched conditions (red) regarding spine density (**e**, top), turnover (**e**, bottom), survival (**f**, top), or newborn spine survival (**f**, bottom). **g**, Schedule for the study on NMDA receptor blockade. **h**, MK801 caused a significant decline in spine density. Data are from mice imaged four times before (black) and six times during (green) MK801 administration. Black and green horizontal lines respectively indicate mean densities during baseline and on the last 4 days of MK801 treatment. The density decrease was highly significant (Wilcoxon signed-rank test; $n = 29$ dendrites; $P = 0.0007$). **i**, The decline in spine density early in MK801 treatment arose from a transient, highly significant difference between the rates at which spines were lost (darker bars) and gained (lighter bars) (Wilcoxon signed-rank test; $n = 29$ dendrites; $P = 0.0008$). Greyscale and green-shaded bars represent percentages of spines gained and lost for sessions before and during MK801 dosage. **** $P < 0.001$. We normalized spine densities to their mean values in baseline conditions, which were $1.03 \mu\text{m}^{-1}$ (**b**, top), $0.90 \mu\text{m}^{-1}$ (**e**, top) and $0.78 \mu\text{m}^{-1}$ (**h**). All error bars are s.e.m. for dendrites.

$P > 0.10$, 8 paired comparisons of density; $P > 0.039$, 7 paired comparisons of turnover ratio; Wilcoxon signed-rank tests with significance thresholds of 0.006 and 0.007, respectively, after corrections for multiple comparisons). Neither were there differences in spine survival ($P > 0.057$; 7 time points; $n = 10$ –49 dendrites; Wilcoxon signed-rank test; significance threshold of 0.007 after Dunn–Šidák correction), nor in newborn spine survival ($P > 0.29$; 6 time points; $n = 5$ –35 dendrites; significance threshold of 0.008; Fig. 3f). Thus, in mice, continuous enrichment does not substantially alter spine dynamics on CA1 basal dendrites. Nevertheless, mean volumes of stable spines underwent a

slight ($7 \pm 3\%$ (s.e.m.)) but significant decline upon enrichment (Wilcoxon signed-rank test; $P = 0.007$; 60 spines tracked for 16 sessions) (Extended Data Fig. 10d). Data on the structural effects of long-term potentiation (LTP) suggest an explanation of these findings; CA1 spine densities rise transiently after LTP induction but return to baseline values 2 h later²¹, implying that continual enrichment would cause no net change in spine densities (Supplementary Discussion).

We next examined whether blockade of NMDA glutamate receptors impacts spine turnover. These receptors are involved in multiple forms of neural plasticity, including in the CA1 area²². In the neocortex, NMDA receptor blockade stabilizes spines by slowing their elimination while keeping their formation rate unchanged¹⁰. We tracked CA1 spines across 10 sessions at 3-day intervals in mice receiving the NMDA receptor blocker MK801 beginning after session 4 and onward (Fig. 3g). We examined 26 dendrites (25 ± 1.4 (s.d.) per session), made 5,020 spine observations (502 ± 32 (s.d.) per day), and found that MK801 induced a significant decline ($12 \pm 3\%$ (s.e.m.)) in spine density (Wilcoxon signed-rank test; 25 dendrites; $P = 0.0007$) (Fig. 3h). This stemmed from a transient disparity in the rates of spine loss versus gain (loss rate was $215 \pm 42\%$ (s.e.m.) of the rate of gain; Wilcoxon signed-rank test; 25 dendrites; $P = 0.0008$) (Fig. 3i). These results indicate that the survival odds of CA1 spines depend on NMDA receptor function, illustrate our ability to detect changes in spine dynamics, and show that CA1 and neocortical spines have divergent responses to NMDA receptor blockade.

To ascertain the underlying time constants governing spine turnover, we compared the $S(t)$ curves of visually scored spines to predictions from a wide range of candidate kinetic models (Fig. 4a). In each model there was a subset (0–100%) of permanent spines; the remaining spines were impermanent, with a characteristic lifetime, τ . Since environmental enrichment left the observed spine dynamics unchanged, we pooled the baseline and enriched data sets to extend the analyses to longer time-scales (Fig. 4b). By varying the actual spine density, fraction of stable spines, and characteristic lifetime for the unstable fraction, we identified the model that best fit the $S(t)$ curves, using a maximum likelihood criterion (Supplementary Information). This model had 100% impermanent spines, with an actual density of $2.6 \mu\text{m}^{-1}$ and τ of ~ 10 days (Fig. 4a, b). This is twice the ~ 5 -day lifetime reported for the transient subset of neocortical spines^{7–9}. There were also models with both permanent and impermanent spines that gave reasonable, albeit poorer fits to the CA1 data (Fig. 4a, b). Crucially, our analysis identified all models whose fits were significantly worse than the best model (white regions in Fig. 4a; $P < 0.05$, likelihood-ratio test); we regarded these as unsatisfactory in accounting for the CA1 data (Supplementary Information).

Our results pointing to a single population of unstable spines in CA1 contrast markedly with findings in adult neocortex, where $>50\%$ of spines seem permanent^{7–9}. To make even-handed comparisons, we used our framework to re-analyse published data acquired in the mouse somatosensory neocortex that had supported this conclusion⁷. Owing to the lower density of neocortical spines, merging is far less of a concern (Fig. 4c), and our modelling confirmed that $\sim 60\%$ of neocortical spines are stable over very long timescales, supporting past conclusions⁷.

Nevertheless, we found very significant differences between CA1 and neocortical spine turnover dynamics (Fig. 4a; $P = 0.01$, likelihood ratio test). Models with only impermanent spines, which well described CA1, were insufficient ($P < 10^{-62}$, likelihood ratio test) for neocortex (Fig. 4a, d). The discrepant lifetimes of impermanent spines in the two areas (~ 10 days (CA1) versus ~ 5 days (neocortex)) posed further incompatibility (Fig. 4a, e). Conversely, models that explained neocortical spine turnover were incompatible with the CA1 data ($P = 0.01$, likelihood ratio test). Modelling alone cannot eliminate the possibility that CA1 basal dendrites have some permanent spines, but if any such spines exist they compose a far smaller fraction than in the neocortex. Hence, CA1 and the neocortex have distinct spine dynamics, percentages of impermanent spines, and turnover time constants.

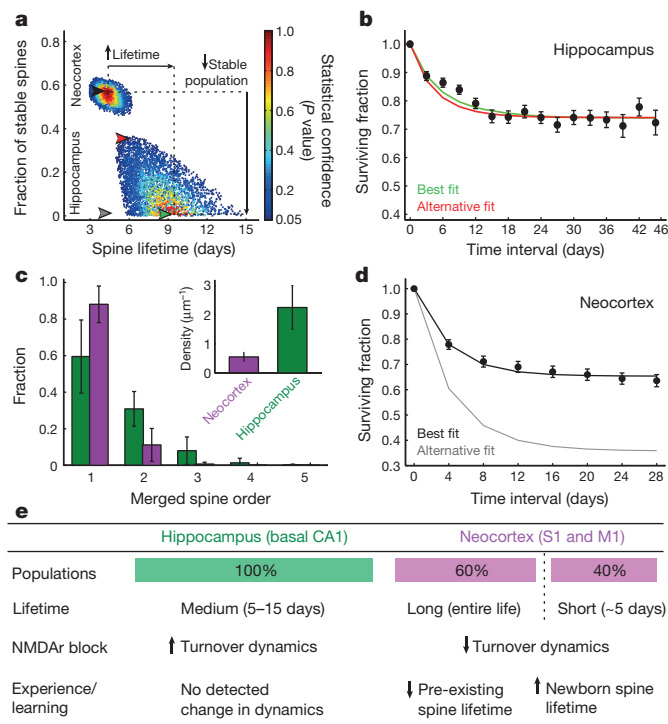


Figure 4 | CA1 and neocortical spines exhibit distinct turnover kinetics.

a, Multiple kinetic models are consistent with data on spine survival. Each model considered had two subpopulations: permanent and impermanent spines. Abscissa: mean lifetime for impermanent spines. Ordinate: fraction of spines that are permanent. Each datum is for an individual model; colour denotes the level of statistical significance at which the model could be rejected. Red points denote models that best fit the data. No results are shown for models incompatible with the data ($P < 0.05$). Models that best fit data from CA1 have ~100% impermanent spines, with a ~10 day lifetime (green arrowhead). There are also models with permanent subpopulations that cannot be statistically rejected (for example, red arrowhead). Models that best fit patterns of neocortical spine turnover (black arrowhead), from mice age- and gender-matched⁷ to those used here, have ~50–60% permanent spines and a shorter lifetime (~5 days) for impermanent spines than in CA1. Models lacking permanent spines poorly fit the neocortical data; grey arrowhead marks the model for the grey curve in **d**. The four arrowheads indicate the models that generated the colour-corresponding curve fits in **b**, **d**. **b**, Empirically determined survival curve for CA1 spines (black data: mean \pm s.e.m.; data set of Fig. 3e, f) over 46 days, compared to predictions (solid curves) from two of the models in **a** (green and red arrowheads in **a**). Green curve: best-fitting model, which has no permanent spines. Red curve: an example model with both stable and unstable spines. **c**, Owing to the higher density of spines in CA1 than in the neocortex (inset), optical merging is far more common in CA1. Given what appears to be one spine, vertical bars represent the probability as determined from the computational model that the observation is actually of 1–5 spines. Probabilities were calculated using spine density values of the inset, $0.75 \mu\text{m}^{-1}$ as the minimum separation, L , needed to distinguish adjacent spines. Error bars: range of results for L within 0.5 – $1.0 \mu\text{m}$. **d**, Empirically determined survival curves for neocortical spines (data set from ref. 7) over 28 days, compared with predictions (solid curves) from two different models for spine turnover in **a** (grey and black arrowheads in **a**). Black curve: best fit attained with a stable subpopulation of spines. Grey curve: a model lacking permanent spines, poorly fitting the data. **e**, Spines in CA1 and neocortex differ substantially in proportions of permanent versus impermanent spines, spine lifetimes, effects of NMDA receptor blockade, and learning or novel experience.

Further distinguishing CA1 and the neocortex are the contrary roles of NMDA receptor blockade (Figs 3h, i and 4e). In the neocortex, MK801 promotes stability by decreasing spine loss¹⁰ and blocking spine addition¹¹; conversely, NMDA receptor activation may speed turnover via addition of new spines and removal of pre-existing neocortical spines supporting older memories. In CA1, MK801 speeds turnover and promotes instability (Fig. 3h, i), suggesting that NMDA

receptor activation may transiently slow turnover and stabilize spines. Indeed, LTP induction in CA1 is associated with stabilization of existing spines and growth of new spines^{5,23,24}.

A natural interpretation is that spine dynamics may be specialized by brain area to suit the duration of information retention. The neocortex, a more permanent repository, might need long-lasting spines for permanent information storage and shorter-lasting ones ready to be stabilized if needed^{8,9}. The hippocampus, an apparently transient repository of information, might only require transient spines. The ~1–2-week mean lifetime for the ~100% impermanent CA1 spines implies a near full erasure of synaptic connectivity patterns in ~3–6 weeks, matching the durations that spatial and episodic memories are hippocampal-dependent in rodents¹² (but see also ref. 25). The ensemble place codes of CA1 neurons also refresh in ~1 month¹⁶, which could arise from turnover of the cells' synaptic inputs. Since 75–80% of CA3 \rightarrow CA1 inputs are monosynaptic²⁶, CA1 spine impermanence probably implies a continuous re-patterning of CA3 \rightarrow CA1 connectivity throughout adulthood, which, owing to the sheer number of synapses and sparse connections is unlikely to assume the same configuration twice.

Supporting these interpretations, artificial neural networks often show a correspondence between synapse lifetime and memory longevity²⁷, although in some models spine turnover and memory erasure can be dissociated^{28,29}. Computational studies also show that elimination of old synapses can enhance memory capacity^{29,30}. More broadly, networks that can alter synaptic lifetimes, not just connection strengths, can more stably store long-term memories while rapidly encoding new ones²⁷.

The data described here are consistent with a single class of CA1 spines, but future studies should examine both the finer kinetic features and cellular or network mechanisms of turnover²³. By using fluorescence tags to mark spines undergoing plastic changes, *in vivo* imaging might help relate connection strengths, spine lifetimes and memory performance. Finally, researchers should investigate spine turnover in animals as they learn to perform a hippocampal-dependent behaviour, to build on the results here by looking for direct relationships between CA1 spine stability and learning.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 August 2014; accepted 10 April 2015.

Published online 22 June 2015.

- Squire, L. R. & Zola-Morgan, S. The medial temporal lobe memory system. *Science* **253**, 1380–1386 (1991).
- Frankland, P. W. & Bontempi, B. The organization of recent and remote memories. *Nature Rev. Neurosci.* **6**, 119–130 (2005).
- Frey, U. & Morris, R. G. Synaptic tagging and long-term potentiation. *Nature* **385**, 533–536 (1997).
- Barretto, R. P. *et al.* Time-lapse imaging of disease progression in deep brain areas using fluorescence microendoscopy. *Nature Med.* **17**, 223–228 (2011).
- Engert, F. & Bonhoeffer, T. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature* **399**, 66–70 (1999).
- Maletic-Savatic, M., Malinow, R. & Svoboda, K. Rapid dendritic morphogenesis in CA1 hippocampal dendrites induced by synaptic activity. *Science* **283**, 1923–1927 (1999).
- Holtmaat, A. J. *et al.* Transient and persistent dendritic spines in the neocortex *in vivo*. *Neuron* **45**, 279–291 (2005).
- Xu, T. *et al.* Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* **462**, 915–919 (2009).
- Yang, G., Pan, F. & Gan, W. B. Stably maintained dendritic spines are associated with lifelong memories. *Nature* **462**, 920–924 (2009).
- Zuo, Y., Yang, G., Kwon, E. & Gan, W. B. Long-term sensory deprivation prevents dendritic spine loss in primary somatosensory cortex. *Nature* **436**, 261–265 (2005).
- Yang, G. *et al.* Sleep promotes branch-specific formation of dendritic spines after learning. *Science* **344**, 1173–1178 (2014).
- Harris, K. M. Structure, development, and plasticity of dendritic spines. *Curr. Opin. Neurobiol.* **9**, 343–348 (1999).
- Mizrahi, A., Crowley, J. C., Shtoyerman, E. & Katz, L. C. High-resolution *in vivo* imaging of hippocampal dendrites and spines. *J. Neurosci.* **24**, 3147–3151 (2004).

14. Barretto, R. P., Messerschmidt, B. & Schnitzer, M. J. *In vivo* fluorescence imaging with high-resolution microlenses. *Nature Methods* **6**, 511–512 (2009).
15. Gu, L. *et al.* Long-term *in vivo* imaging of dendritic spines in the hippocampus reveals structural plasticity. *J. Neurosci.* **34**, 13948–13953 (2014).
16. Ziv, Y. *et al.* Long-term dynamics of CA1 hippocampal place codes. *Nature Neurosci.* **16**, 264–266 (2013).
17. Harris, K. M. & Stevens, J. K. Dendritic spines of CA 1 pyramidal cells in the rat hippocampus: serial electron microscopy with reference to their biophysical characteristics. *J. Neurosci.* **9**, 2982–2997 (1989).
18. Moser, M. B., Trommald, M. & Andersen, P. An increase in dendritic spine density on hippocampal CA1 pyramidal cells following spatial learning in adult rats suggests the formation of new synapses. *Proc. Natl Acad. Sci. USA* **91**, 12673–12675 (1994).
19. Rampon, C. *et al.* Enrichment induces structural changes and recovery from nonspatial memory deficits in CA1 NMDAR1-knockout mice. *Nature Neurosci.* **3**, 238–244 (2000).
20. Sanders, J., Cowansage, K., Baumgartel, K. & Mayford, M. Elimination of dendritic spines with long-term memory is specific to active circuits. *J. Neurosci.* **32**, 12570–12578 (2012).
21. Bourne, J. N. & Harris, K. M. Coordination of size and number of excitatory and inhibitory synapses results in a balanced structural plasticity along mature hippocampal CA1 dendrites during LTP. *Hippocampus* **21**, 354–373 (2011).
22. Huerta, P. T., Sun, L. D., Wilson, M. A. & Tonegawa, S. Formation of temporal memory requires NMDA receptors within CA1 pyramidal neurons. *Neuron* **25**, 473–480 (2000).
23. Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J. & Kasai, H. Principles of long-term dynamics of dendritic spines. *J. Neurosci.* **28**, 13592–13608 (2008).
24. Toni, N., Buchs, P. A., Nikonenko, I., Bron, C. R. & Müller, D. LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. *Nature* **402**, 421–425 (1999).
25. Goshen, I. *et al.* Dynamics of retrieval strategies for remote memories. *Cell* **147**, 678–689 (2011).
26. Sorra, K. E. & Harris, K. M. Stability in synapse number and size at 2 hr after long-term potentiation in hippocampal area CA1. *J. Neurosci.* **18**, 658–671 (1998).
27. Fusi, S., Drew, P. J. & Abbott, L. F. Cascade models of synaptically stored memories. *Neuron* **45**, 599–611 (2005).
28. Abraham, W. C. & Robins, A. Memory retention—the synaptic stability versus plasticity dilemma. *Trends Neurosci.* **28**, 73–78 (2005).
29. Wu, X. E. & Mel, B. W. Capacity-enhancing synaptic learning rules in a medial temporal lobe online learning model. *Neuron* **62**, 31–41 (2009).
30. Poirazi, P. & Mel, B. W. Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* **29**, 779–796 (2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Li, J. Lecoq and E. T. W. Ho for technical assistance, R. P. J. Barretto and B. Messerschmidt for advice, and A. Holtmaat for providing published data sets. The National Institute of Mental Health, The National Institute on Aging, and the Ellison Foundation provided grants (M.J.S.). Fellowships from the National Science Foundation and The Stanford Center for Mind, Brain, and Computation provided graduate research support (J.E.F.).

Author Contributions A.A. and M.J.S. designed experiments. A.A. performed experiments. A.A. and J.E.F. analysed data. J.E.F. designed and performed the modelling. A.A., J.E.F. and M.J.S. wrote the paper. M.J.S. supervised the research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.J.S. (mschnitz@stanford.edu).

METHODS

Animals and surgical preparation. Stanford University's Administrative Panel on Laboratory Animal Care approved all procedures. We imaged neurons in mice expressing GFP driven by the *Thy1* promoter³¹ (heterozygous males 10–12 weeks old, GFP-M lines on a C57BL/6 × F1 background). We did not perform any formal randomization in the assignment of mice to specific groups, but we informally selected mice in a random manner without use of any exclusion criteria. We performed surgeries as previously published⁴ but with a few modifications. We anaesthetized mice using isoflurane (1.5–3% in O₂) and implanted a stainless steel screw into the cranium above the brain's right hemisphere. We performed a craniotomy in the left hemisphere (2.0 ± 0.3 mm posterior to bregma, 2.0 ± 0.3 mm lateral to midline) using a 1.8-mm-diameter trephine and implanted the optical guide tube with its window just dorsal to, but not within, area CA1, preserving the alveus.

Guide tubes and microlenses. Guide tubes were glass capillaries (1.5 mm (ID), 1.8 mm (OD) or 2 mm (ID), 2.4 mm (OD); 2–3 mm in length). We attached a circular coverslip, matched in diameter to that of the capillary's outer edge, to one end of the guide tube by using optical epoxy (Norland Optical Adhesive 81). We used 1.0-mm-diameter micro-optical probes of diffraction-limited resolution (0.8 NA, 250 μ m working distance in water) that were encased in a 1.4-mm-diameter sheath¹⁴.

In vivo two-photon imaging. We used a modified commercial two-photon microscope (Prairie Technologies) equipped with a tuneable Ti:Sapphire laser (Chameleon, Coherent). We tuned the laser emission to 920 nm and adjusted the average illumination power at the sample (~ 5 –25 mW) for consistency in signal strength across imaging sessions in each mouse. For microendoscopy we used a 20×0.8 NA objective (Zeiss, Plan-Apochromat) to deliver illumination into the microlenses. In some cases we imaged directly through the glass cannula using a Olympus LUMPlan F/IR 0.8 NA $\times 40$ water immersion objective lens and confirmed that the optical resolution in all three spatial axes was identical between the two approaches. Beginning at 15–18 days after surgery, we imaged mice every 3 days under isoflurane anaesthesia (1.5% in O₂) for a total of 8–16 sessions each lasting 60–90 min. We imaged some mice at irregular intervals up to 80 days.

MK801 treatment. In mice subject to the protocol of Fig. 2d, after the fourth imaging session we administered MK801 (Tocris Bioscience; 0.25 mg g⁻¹ body weight; dissolved in saline) in two intraperitoneal injections each day (8–10 h apart) as described previously¹⁰.

Enriched environment. Animals given an enriched environment had a larger cage (42 (length) \times 21.5 (width) \times 21.5 (height) cm³) that contained a running wheel, objects of various colours, textures and shapes, plastic tunnels, and food with different flavours. We changed the objects, as well as their placements within the cage, every 3–4 days to encourage exploration and maintain novelty. We provided food and water *ad libitum*.

Histology. At the end of *in vivo* experimentation, we deeply anaesthetized mice with ketamine (100 mg kg⁻¹) and xylazine (20 mg kg⁻¹). We then perfused PBS (pH 7.4) into the heart, followed by 4% paraformaldehyde in PBS. We fixed brains overnight at 4 °C and prepared floating sections (50 μ m) on a vibrating microtome (VT1000S, Leica). Before *in vitro* imaging of GFP fluorescence, we washed the fluorescent sections with PBS buffer several times and quenched them by incubation in 150 mM glycine in PBS for 15 min. After three washes in PBS, we mounted sections with Fluoromount-G (Southern Biotech). We inspected the sections using either two-photon fluorescence imaging or a STED microscope (Leica TCS STED CW, equipped with a Leica HCX PL APO 100 \times 1.40 NA oil-immersion objective.).

For immunostaining sections were washed with PBS buffer several times before quenching and permeabilization (15 min incubation in 0.1% Triton-X in PBS). Sections were incubated in blocking solution (1% Triton X-100, 2% BSA, 2% goat serum in PBS) for 4 h. Primary antibodies (rat anti-CD68, FA11-ab5344, Abcam, 1:100 dilution; mouse anti-GFAP, MAB3402, Millipore, 1:500 dilution) were diluted in blocking solution and sections were incubated overnight in this solution. The following day sections were washed with PBS and incubated in diluted secondary antibody: Cy5-conjugated goat anti-mouse IgG, A10524 and Cy5-conjugated goat anti-rat IgG, A10525; both Molecular Probes; both 1:1,000 dilution, in blocking solution for 3 h. All staining procedures were done at room temperature. After three washes in PBS, sections were mounted with Fluoromount-G (Southern Biotech). Histological specimens were inspected on a confocal fluorescence microscope (Leica SP2 AOBs).

Imaging sessions. We mounted the isoflurane-anaesthetized mice on a stereotaxic frame. To perform microendoscopy we fully inserted the microendoscope probe into the guide tube such that the probe rested on the guide tube's glass window. To attain precise and reliable three-dimensional alignments across all imaging sessions of the brain tissue undergoing imaging, we used a laser-based alignment method. We positioned a laser beam such that when the mouse's

head was properly aligned, the beam reflected off the back surface of the microendoscope and hit a designated target. This ensured that at each imaging session the long axis of the microendoscope was perpendicular to the optical table to within ~ 1 angular degree. Otherwise we inserted a drop of water in the cannula and imaged using the water immersion objective. As previously described¹⁴, we experimentally confirmed that the resolution limits of the two approaches were essentially identical.

Image acquisition. During the first imaging session, we selected several regions of brain tissue for longitudinal monitoring across the duration of the time-lapse experiment. Each of these regions contained between 1 and 7 dendritic segments visibly expressing GFP. In each imaging session, we acquired 6–8 image stacks of each selected regions using a voxel size of $0.0725 \times 0.0725 \times 0.628 \mu\text{m}^3$.

Image pre-processing. To improve the visual saliency of fine details within each image stack, initial pre-processing of the image stacks involved a blind deconvolution based on an expectation-maximization routine (Autodeblur from Autoquant). We then aligned all individual images acquired at the same depth in tissue using the TurboReg plug-in routine for ImageJ. Finally, we averaged pixel intensities across the aligned stacks, yielding a single stack that we used in subsequent analyses.

Scoring of dendritic spines. We scored spines using a custom MATLAB interface that supported manual labelling of spines using the computer mouse, measurements of dendrite length and spine position, and alignments of time-lapse sets of image stacks. For each region of tissue monitored, we loaded all the image stacks acquired across time, such that the temporal sequence of the stacks was preserved but the experimenter was blind to their dates of image acquisition during spine scoring. We excluded images whose quality was insufficient to score spines.

We scored spines similarly to as described previously³² but with a few modifications. We labelled protrusions as dendritic spines only if they extended laterally from the dendritic shaft by $>0.4 \mu\text{m}$ (Extended Data Fig. 6a, b). We did not include protrusions of $<0.4 \mu\text{m}$ in the analysis (Extended Data Fig. 6c, d). When a spine first appeared in the time-lapse image data we assigned it a unique identity. We preserved the spine's identity across consecutive time points if the distances between the spine in question and two or three of its neighbouring spines were stable. In ambiguous cases, which were hardly the norm, we required stability to $<2 \mu\text{m}$.

The surviving fraction, $S(t)$, at time t was defined as the fraction of spines present on the first imaging day that were also present a time t later. For the deliberate purpose of attaining conservative estimates of (for example, lower bounds on) the proportion of impermanent spines, in the calculation of $S(t)$ we handled the 18% of spines in the recurrent-location category (Fig. 1d and Extended Data Fig. 1d, e) in the following way. When checking pairs of images acquired an interval t apart, we deliberately did not distinguish between whether the second image contained the original spine or its replacement spine at the same location. This approach thereby underestimated spine turnover as inferred from analyses of $S(t)$, implying that our conclusion of CA1 spine impermanence is not only mathematically conservative but also robust to any scoring errors in which we might have erroneously missed a spine that had in fact persisted to subsequent imaging sessions.

Turnover ratio was defined as the sum of spines gained and lost between two consecutive time points normalized by the total number of spines present at these time points. Spines lost or gained were defined as the number of spines lost or gained between two consecutive time points, respectively, normalized by the total number of spines present at these time points.

To make coarse estimates of spine volumes, for stable spines we determined each spine's fluorescence within a manually drawn region of interest (ROI) in the axial section in which the spine head appeared at its biggest diameter. We normalized this value by the fluorescence value attained by moving the ROI to within the nearby dendritic shaft (as in ref. 7).

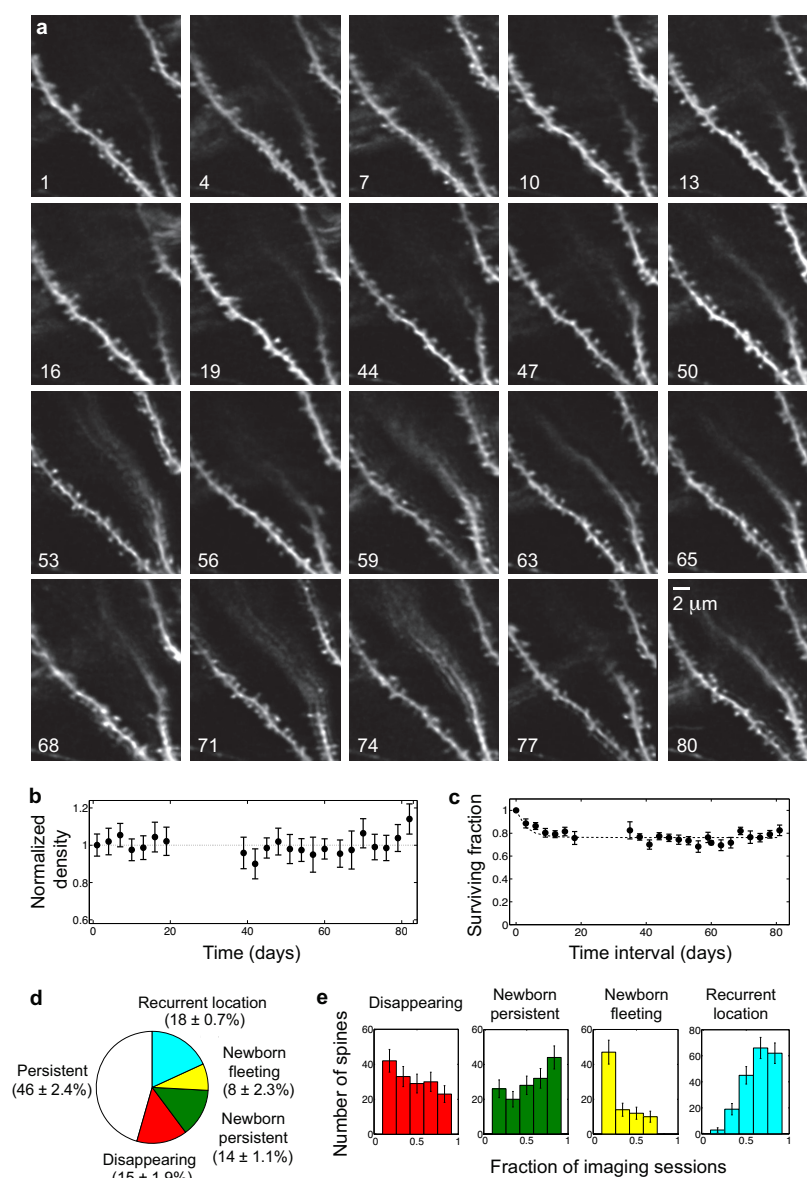
Statistical analysis. To test for differences in spine densities, turnover ratios and surviving fractions either over time or between different groups, we used non-parametric two-sided statistical testing (Wilcoxon signed-rank, Mann-Whitney U and Kruskal-Wallis ANOVA tests) to avoid assumptions of normality and Dunn-Sidak correction for multiple comparisons. Sample size was chosen to match published work^{7–9}.

To compare experimentally measured spine survival to the theoretical predictions from kinetic modelling, we assessed the goodness-of-fit for each model by using both the reduced chi-squared statistic and the log-likelihood function (Supplementary Methods). Both the mean and covariance of the surviving fraction depended on the model parameters and influenced the goodness-of-fit (Supplementary Methods). We also required that the parameter describing the minimal separation needed to resolve two spines was 0.5–1 μm (other values for this minimal separation are implausible) (Supplementary Methods).

Simulated data sets. We modelled the microscope's optics on the basis of prior measurements¹⁴ and tuned the kinetics of spine turnover, spine geometries and

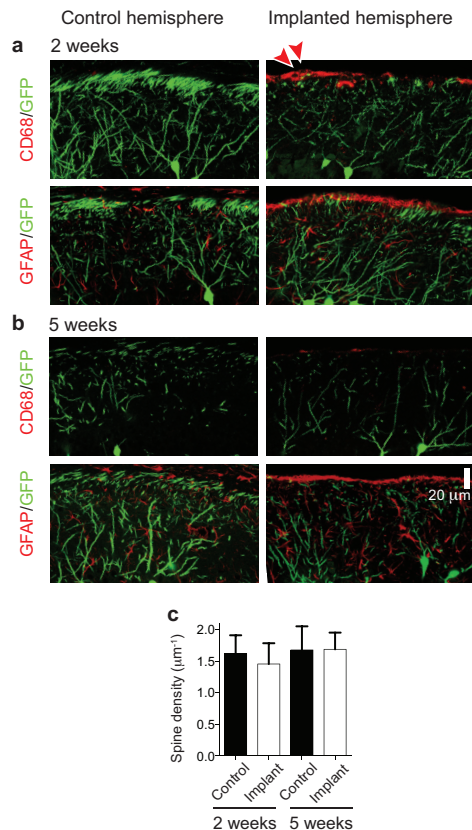
dendrite geometries to produce simulated image sequences that the data analyst judged to be similar to the actual data (Supplementary Information and Extended Data Fig. 5). In some data sets, we matched the simulated spine kinetics to those inferred from our *in vivo* measurements.

31. Feng, G. *et al.* Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000).
32. Holtmaat, A. *et al.* Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. *Nature Protocols* **4**, 1128–1144 (2009).
33. Dombeck, D. A., Harvey, C. D., Tian, L., Looger, L. L. & Tank, D. W. Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nature Neurosci.* **13**, 1433–1440 (2010).
34. Mishchenko, Y. *et al.* Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron* **67**, 1009–1020 (2010).

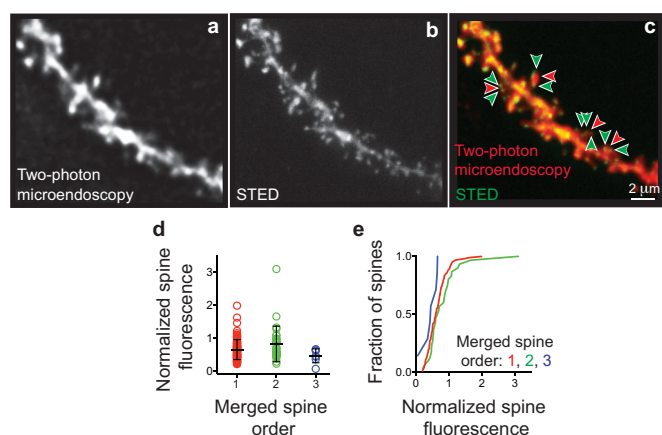


Extended Data Figure 1 | *In vivo* imaging of CA1 spine dynamics over extended time intervals. **a**, *In vivo*, 80-day-long time-lapse image data set sampled at variable intervals. Each image shown is the maximum projection of 4–8 images acquired at adjacent *z*-planes. Scale bar, 2 μm . **b**, **c**, Direct empirical determinations of spine density (**b**) and spine survival (**c**) across the 80 days. A normalized spine density of one corresponds to a measured spine density of $1.16 \mu\text{m}^{-1}$. Data points are mean \pm s.e.m. for 16 dendrites. **d**, *In vivo*, 22-day-long time-lapse image data set sampled every 3 days. Over 50% of the

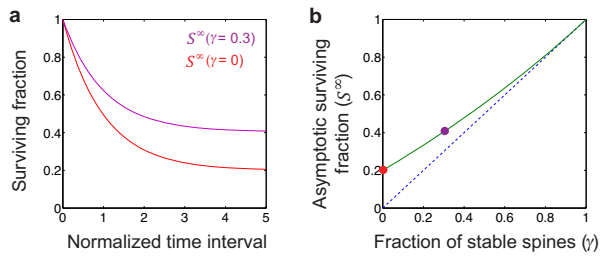
spines underwent visually noticeable dynamic changes. Pie chart shows the proportions of spines that were persistent or exhibited different patterns of turnover ($n = 1,075$ total spines from 4 mice). Colour coding is the same as in Fig. 1d. Error bars are s.e.m. for four mice. **e**, Histograms show the distributions, for each class of spines, of the fraction of imaging sessions in which each spine was observed within the same 22-day data set of **d**. Colour coding is the same as in Fig. 1d. Error bars represent s.d. estimated as counting errors.



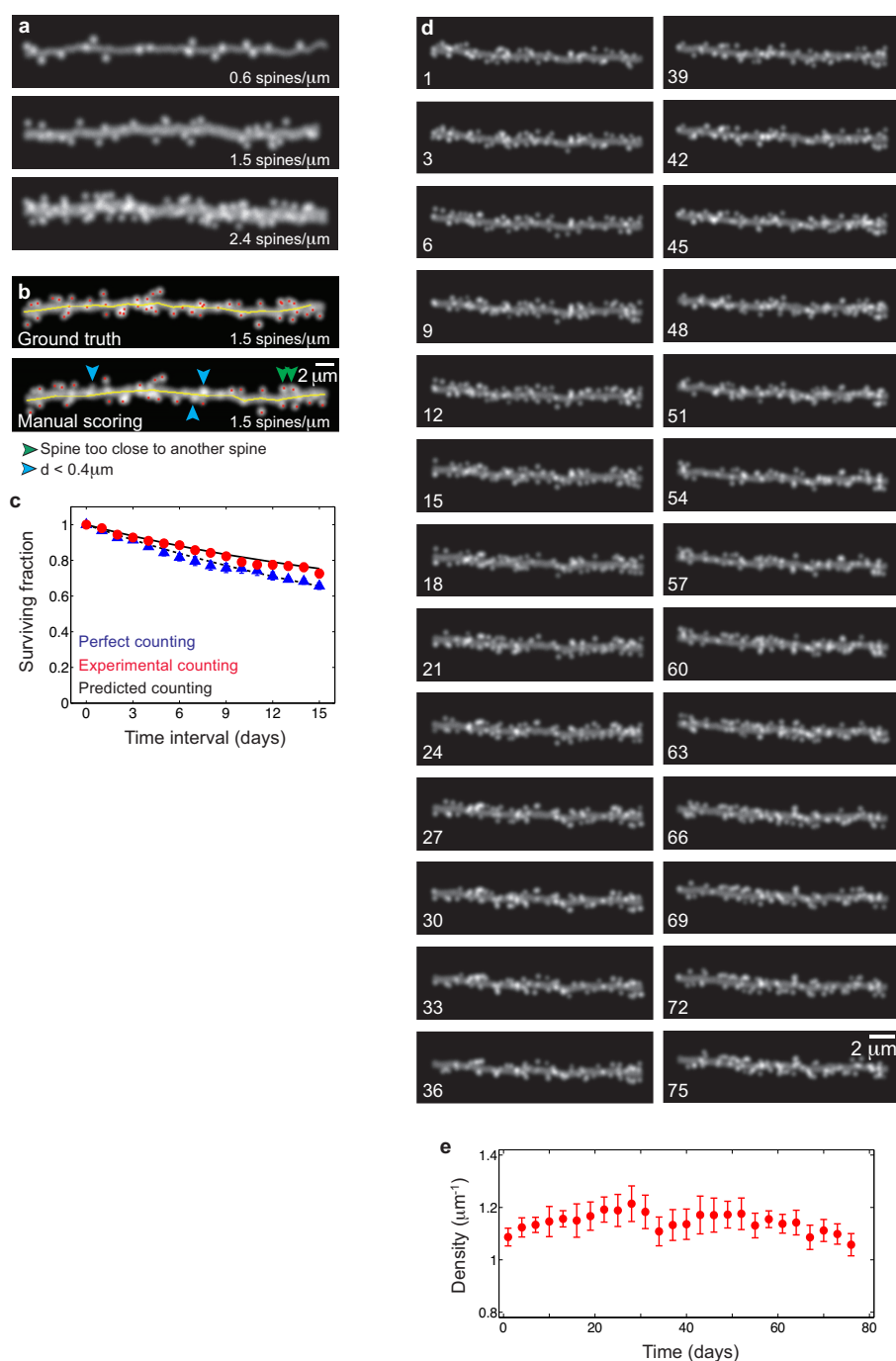
Extended Data Figure 2 | The chronic CA1 preparation induces a minimal, 5- μm -thick layer of glial activation and does not affect spine density. **a–c,** In two groups of mice, each comprising two *Thy1-GFP* transgenic animals (8–10 weeks old), we implanted the imaging guide tube just dorsal to hippocampal area CA1 following established procedures^{4,16,33}. **a, b,** We euthanized, sliced and stained the first group after two further weeks (**a**), and the second group after five further weeks (**b**). Confocal fluorescence images of the stained tissue slices revealed activated microglia (CD68 staining, top, red), astrocytes (GFAP staining, bottom, red), GFP-expression pyramidal neurons (green), and permitted quantifications of spine density in CA1 regions both ipsilateral and contralateral to the implant. **a,** Two weeks after implantation, confocal photomicrographs (maximum intensity projection of four separate z-planes, axially spaced 0.2 μm apart) revealed a limited presence of activated microglia (red arrowheads indicate single cells) on the implanted hemisphere (top right) but were virtually non-existent in the contralateral control CA1 from the same mice (top left). Staining for astrocytes on the implanted hemisphere (bottom right) was almost indistinguishable from the control hemisphere (bottom left), except for a 5–10- μm -thick layer of astrocyte label abutting the optical surface of the imaging guide tube. **b,** Five weeks after implantation, confocal photomicrographs (maximum intensity projection of four separate z-planes, axially spaced 0.2 μm apart) revealed an almost undetectable presence of activated microglia on the implanted hemisphere (top right), comparable to the contralateral control hemisphere (top left). Staining for astrocytes in the implanted hemisphere (bottom right) was almost indistinguishable from the control hemisphere (bottom left), except for a 5–10- μm -thick layer of astrocyte label abutting the optical surface of the imaging guide tube. **c,** Mean density of spines on pyramidal cell basal dendrites in the implanted CA1 (white columns) was statistically indistinguishable from the contralateral control CA1 (black columns), at 2 weeks ($P = 0.21$; Mann–Whitney *U*-test; $n = 11$ dendrites) and at five weeks ($P = 0.98$; Mann–Whitney *U*-test; $n = 11$ dendrites) after implantation. Error bars are s.d.



Extended Data Figure 3 | Two-photon and STED imaging of the same CA1 spines in fixed tissue reveals that nearby spines can merge in two-photon images. **a–c,** Example two-photon microendoscopy (**a**), STED (**b**), and overlay (**c**) images of the same CA1 basal dendrite, acquired in a fixed brain slice from a *Thy1-GFP* mouse. Nearby spines that are clearly distinguishable in the STED image (green arrowheads) but within the diffraction-resolution limit of two-photon microendoscopy (0.85 NA) appear as single, merged entities within the two-photon image (red arrowheads). The two-photon image shown is the maximum intensity projection of three optical sections axially spaced 0.6 μm apart. The STED image is the maximum intensity projection of six optical sections spaced 0.3 μm apart. Scale, 2 μm. **d, e,** To attain an approximate measure of spine volume, we quantified each spine's fluorescence in manually drawn regions of interest (ROIs) and normalized it by the fluorescence value in the nearby dendritic shaft, within an ROI of identical shape and size within a single axial section of the two-photon image stack. To ascertain whether each of the spines scored in the two-photon images was actually a merged spine or not, we consulted the STED images of the same dendrite. Plotted are the normalized fluorescence values (**d**) for unitary spines as well as doublet and triplet merged spines (black lines: mean values \pm s.d.; coloured points: data from individual spines), and the cumulative distributions of these measurements (**e**). The distributions of normalized fluorescence were statistically indistinguishable ($P > 0.06$; Kruskal–Wallis ANOVA; $N = 100, 30$ and 7, respectively, for unitary, doublet and triplet spines), probably reflecting the substantial range of CA1 spine geometries (Extended Data Fig. 9).

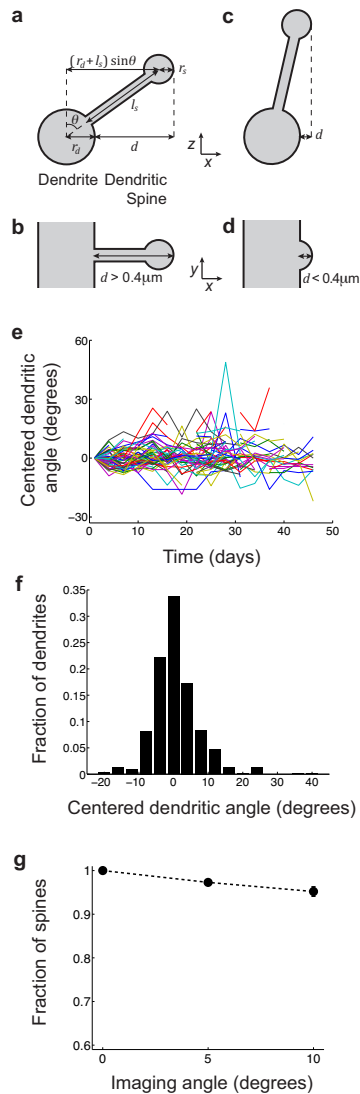


Extended Data Figure 4 | The asymptotic value of the surviving fraction of spines exceeds the fraction of permanent spines. **a**, Surviving fraction curves for models in which the fraction of permanently stable spines is $\gamma = 0$ (red curve) and $\gamma = 0.3$ (purple curve) (Supplementary Methods). The timescale of spine survival was $\tau = 1$, and the filling fraction value was $f = 0.2$. The surviving fraction asymptotes to a value, S^∞ , that encodes the fraction of stable spines. (Supplementary Information has a list of all mathematical variables used in this work, and their definitions). **b**, The time asymptotic value of the surviving fraction (green curve) exceeds the fraction of stable spines (dashed blue line). Coloured circles correspond to the surviving fraction curves plotted in **a**.

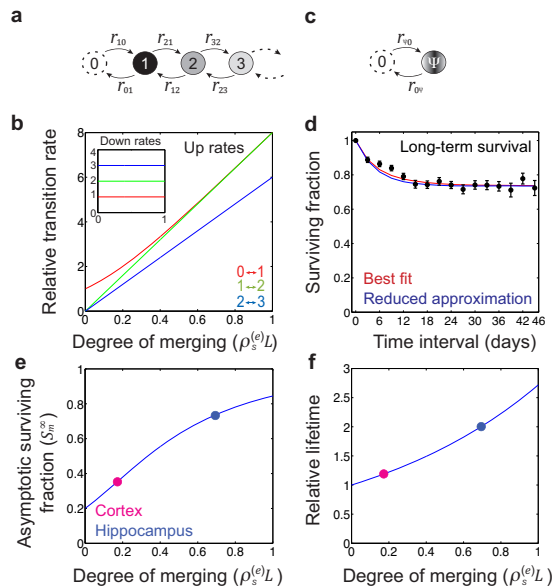


Extended Data Figure 5 | Examples of simulated imaging data sets and their scoring. **a**, Example simulated images (Supplementary Methods) of dendrites for which the spine density was $0.6 \mu\text{m}^{-1}$ (top), $1.5 \mu\text{m}^{-1}$ (middle) and $2.4 \mu\text{m}^{-1}$ (bottom). **b**, Ground truth and manual scoring of a simulated dendrite for which the spine density was $1.5 \mu\text{m}^{-1}$. Green arrowheads indicate counting errors originating from the optical merging of spines (Supplementary Methods). Blue arrowheads indicate counting errors that occur when the spine's projection into the optical plane is too short (Supplementary Methods). Scale bar, $2 \mu\text{m}$. **c**, The visually scored surviving fraction (red circles) differed from the true spine surviving fraction (blue triangles), but the departures were

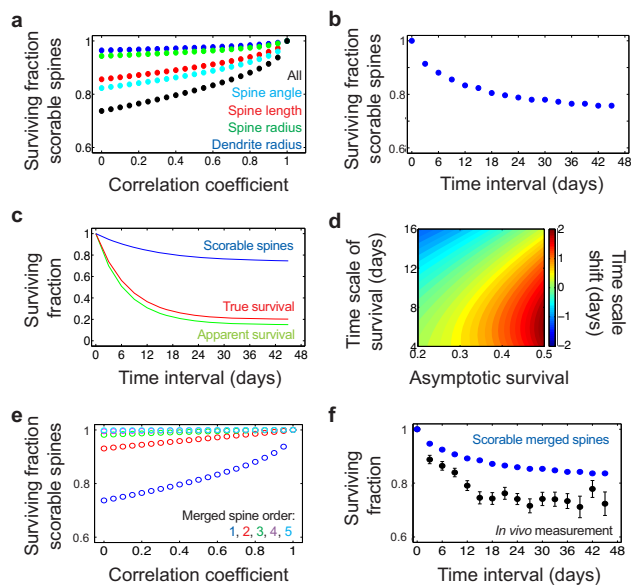
well predicted by the kinetic model (solid black curve) (Supplementary Methods). **d**, We simulated and scored a long-term lapse imaging data set with kinetic parameters that matched the best-fit model of Fig. 4b (Supplementary Methods). Even though the data set lacked stable spines, many simulated spines appeared to persist for long time intervals. Scale bar, $2 \mu\text{m}$. **e**, Although the spine density in the simulated data was $2.56 \mu\text{m}^{-1}$, visual scoring yielded a lower spine density. We used the measured and true spine densities to estimate the extent of merging and the counting resolution (Supplementary Methods). Data points are mean \pm s.e.m for 20 simulated dendrites (**c**) or 10 simulated dendrites (**e**).



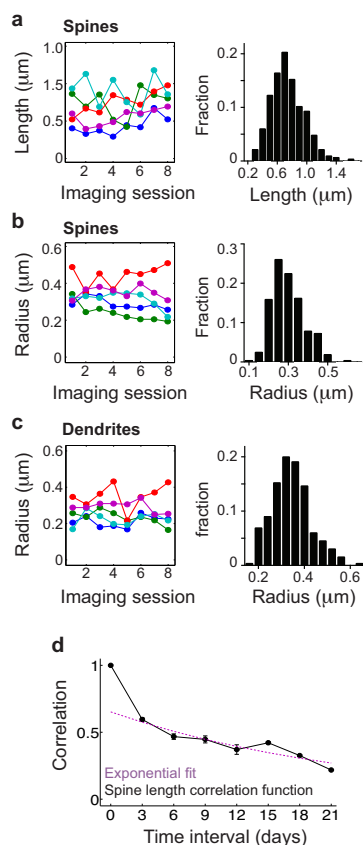
Extended Data Figure 6 | Variability of imaging angles has virtually no impact on determinations of spine turnover. **a–g**, We examined empirically whether variations in dendritic angle across different imaging sessions might impact determinations of spine turnover. However, the variations in dendritic angle that were actually present in our data sets were insufficient to cause illusory turnover. **a, b**, Dendritic spines can be detected when the angle (θ) between a spine and the normal vector is large (Supplementary Methods). View of a dendrite and spine in the (x, z) (**a**) and optical (x, y) (**b**) planes. **c, d**, Dendritic spines cannot be detected when the angle between a spine and the normal vector (θ) is small (Supplementary Methods). View of a dendrite and spine in the (x, z) (**c**) and optical (x, y) (**d**), planes. **e**, For every dendrite and time point, we estimated the dendrite's angle with respect to the optical plane using the three-dimensional coordinates of two manually labelled points on the dendrite chosen to flank the region of dendrite containing the scored spines. Over time, individual dendrites varied about their initial angle ($n = 55$ dendrites tracked over 16 sessions; data set of Fig. 3d). **f**, Distribution of the fluctuations in angle, pooled across the 55 dendrites, relative to the initial angle as seen in the first imaging session. The average magnitude of an angular fluctuation was 4.5° , and 90% of angular fluctuations were $<10^\circ$ in magnitude. Thus, a 5° fluctuation was typical in our data set, whereas a 10° fluctuation was atypically large. **g**, To determine if variability in the imaging angle might impact determinations of spine turnover, we imaged 18 dendrites in fixed slices while deliberately tilting the imaging plane by 0° , 5° and 10° . We made a total of 989 spine observations. Over 95% of spines scored in the 0° condition were also correctly scored when the specimen was tilted by 5° or 10° . Overall, the level of angular fluctuations in the *in vivo* imaging data has virtually no impact on turnover scores.



Extended Data Figure 7 | Kinetic modelling well describes how optical merging affects spine turnover dynamics as monitored with finite optical resolution. **a**, Diagram of the kinetic scheme used to describe merged spine dynamics (Supplementary Methods). Each state is labelled by the number of actual spines that have merged in appearance to a single spine (a quantity that we call the merged spine order; Supplementary Methods). State '0' indicates the absence of a spine, and state '1' indicates a spine that is truly unitary. Transitions occur between adjacent states in the kinetic ladder diagram with rate constants r_{mn} . **b**, The rate constants governing increases in merged spine order depend on two parameters (Supplementary Methods): (1) the initial state or merged spine order; and (2) the overall degree of merging in the spine image data set, which is proportional to the product of the spine density and the shortest resolvable interspine interval (denoted L) (Supplementary Methods). By contrast, the rate constants governing decreases in merged spine order (inset) depend only on the initial merged spine order (Supplementary Methods). **c**, In the case when all spines are labile, a collapsed kinetic scheme in which a single state (Ψ) combines all merged spine orders above zero approximates the complete model (**d**) and can be solved mathematically (Supplementary Methods). **d**, The surviving fraction curve generated from the collapsed kinetic scheme (blue curve) fits the empirically observed surviving fraction (black data points) as well as the best-fit model (red). **e**, The asymptotic value of the surviving fraction is a function of the degree of merging (Supplementary Methods). A large degree of merging (as in CA1, blue circle) produces a larger asymptotic value of the merged spine surviving fraction than a small degree of merging (as in the neocortex, purple circle). **f**, The estimated lifetime of merged spines is a function of the degree of merging (Supplementary Methods). A large degree of merging (as in the hippocampal CA1, blue circle) produces a longer relative lifetime of merged spines than a small degree of merging (as in the neocortex, purple circle).

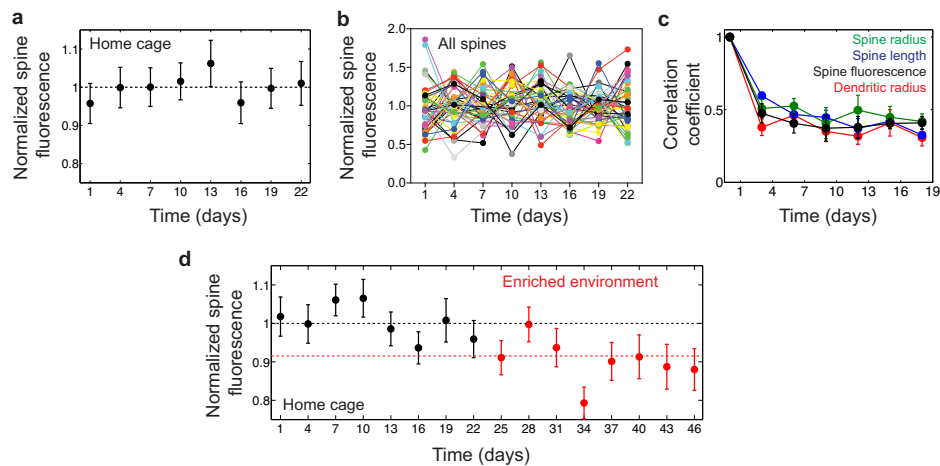


Extended Data Figure 8 | Dynamic spine geometries induce modest levels of apparent spine turnover that cannot explain the turnover measured *in vivo*. **a–f.** To study potential effects of fluctuations in spine geometry, we used values for the means and variances of dendrite radius, spine length and spine radius that were determined by electron microscopy^{17,34}. We then computationally examined how time-dependent fluctuations in these parameters would affect determinations of spine surviving fraction (Supplementary Methods). **a.** We examined how fluctuations in dendrite radius, spine radius, spine length and spine angle—individually (coloured data points) and all together (black data)—affect the spine surviving fraction when the fluctuating geometric parameters are chosen stochastically in each of two imaging sessions, as a function of the parameter's time correlation between the two sessions (Supplementary Methods). As expected, when the two sessions involved image pairs that were perfectly correlated, the surviving fraction reached 100%. Fluctuations in all four parameters had greater effects than fluctuations in individual geometric parameters. **b.** To estimate the time dependence of the surviving fraction of scorable spines from **a**, we assumed all geometric parameters evolved according to the time-correlation function that we empirically determined from *in vivo* imaging data (Extended Data Fig. 9d). **c.** The apparent surviving fraction is the product of the true surviving fraction and the surviving fraction of scorable spines (Supplementary Methods). For the best-fit kinetic model, the apparent surviving fraction is very close to the true surviving fraction. **d.** The difference between the fitted timescale of the apparent surviving fraction and the true survival timescale is small across the range of model parameters consistent with the *in vivo* data (Supplementary Methods). **e.** The graph plots the lower bound of the surviving fraction of scorable merged spines as a function of the time-correlation function shared by all four geometric parameters, for different merged spine orders. As this lower bound increases rapidly with the merged spine order, artefactual turnover due to unscorable spines is unlikely when spine merging is common (Supplementary Methods). **f.** We combined Fig. 4c and Extended Data Fig. 8e to bound the turnover that could result from unscorable spines (Supplementary Methods). As the empirically measured surviving fraction falls below the lower bound obtained for the surviving fraction of scorable merged spines, ongoing changes in the geometric parameters of spines cannot account for the observed spine turnover.



Extended Data Figure 9 | Dynamics of spine geometries measured *in vivo*.

a, Time courses of the spine length, measured from the border of the dendritic shaft to the centre of the spine, for five example spines tracked over eight imaging sessions (left). Distribution of spine lengths (right; $n = 344$ spine observations). **b**, Time courses of the spine radius, measured from the border to the centre of the spine, for five example spines tracked over eight imaging sessions (left). Distribution of spine radii (right; $n = 344$ spine observations). **c**, Time courses of the dendritic radius, measured from the border to the centre of the dendrite, at the location of five example spines tracked over eight imaging sessions (left). Distribution of all dendritic radii (right; $n = 344$ dendrite observations). **d**, Experimental spine length time-correlation function and its exponential fit (Supplementary Methods).



Extended Data Figure 10 | Volumes of stable spines fluctuate minimally over time. **a–d**, To attain an approximate measure of spine volume for stable spines, we quantified each spine's fluorescence in manually drawn regions of interest (ROIs) and normalized it by the fluorescence value in the nearby dendritic shaft, as determined within a ROI of identical shape and size in a single z-section image acquired by two-photon microendoscopy. In addition, each spine's fluorescence value at each time point is normalized to its own mean over the entire experiment. **a**, **b**, Mean (\pm s.e.m.) fluorescence intensities of all spines (**a**), and individual spines (**b**), from a set of 43 stable spines, across a

21-day span during which mice ($n = 4$) were in their home cages (same data as for Fig. 3a). Dashed black line in **a** indicates the mean over all imaging sessions. **c**, The correlation functions of spine radius (green), length (blue), fluorescence (black) and dendritic radius (red) are indistinguishable from each other. **d**, Mean (\pm s.e.m.) fluorescence intensities of 61 stable spines across a 46-day span during which mice ($n = 3$) initially resided in their home cages (black data points) but later moved to an enriched environment (red points) (same data set as for Fig. 3d). Dashed black and red lines respectively denote the mean values over the baseline and enriched periods.