Supplementary Materials for

Sustained rescue of prefrontal circuit dysfunction by antidepressant-induced spine formation

R. N. Moda-Sava*, M. H. Murdock*, P. K. Parekh*, R. N. Fetcho, B. S. Huang, T. N. Huynh, J. Witztum, D. C. Shaver, D. L. Rosenthal, E. J. Alway, K. Lopez, Y. Meng, L. Nellissen, L. Grosenick, T. A. Milner, K. Deisseroth, H. Bito, H. Kasai, C. Liston†

*These authors contributed equally to this work.
†Corresponding author. Email: col2004@med.cornell.edu

Published 12 April 2019, *Science* **364**, eaat8078 (2019)
DOI: 10.1126/science.aat8078

This PDF file includes:

Materials and Methods

Animals. All experiments used young adult male mice (9-20 weeks old): Thy1/YFP-H transgenic mice (Jackson Labs) for spine imaging experiments and C57BL/6J mice (Jackson Labs) for all other experiments. Mice had *ad libitum* access to food and water (except during restraint stress; see below), and were group housed (2-5 mice per cage) under a 12 h light/dark cycle (lights on 6:00). All procedures were overseen by and adherent to the rules set forth by the Weill Cornell Medical College Institutional Animal Care and Use Committee. Sample sizes for each experiment were determined using G*power analysis software, based on anticipated effect sizes that were estimated from previously published reports whenever they were available, and were powered to detect moderate, biologically meaningful effect sizes.

Chronic CORT, chronic stress, and ketamine treatment protocols. For chronic CORT exposure experiments, mice were exposed to CORT in the drinking water for either 10 or 21 days as described in the text. CORT (Sigma) was dissolved in limiting amounts of 100% ethanol (Sigma) and mixed with animal facility-provided drinking water to a final concentration of 0.1 mg/ml CORT and 1% ethanol. A control group received regular drinking water without CORT. For repeated restraint stress experiments, mice were restrained in wire mesh tubes for ~4 hours per day for 21 days. Except for these 4 hours, they had *ad libitum* access to food and water for the remainder of the day. A control group received gentle handling for ~15 minutes per day for 21 days. At the end of each of these stress paradigms (chronic CORT x 10 days, chronic CORT x 21 days, repeated restraint x 21 days), subjects were injected with either ketamine or a saline vehicle control. Ketamine was acquired from Weill Cornell Veterinary Services, and a single dose of 10 mg/kg dissolved in saline was administered via intraperitoneal injection.

Viral vectors. AAV1-Syn-GCaMP6s was obtained from the University of Pennsylvania Vector Core. The viral titer was typically 6 x 10^{13} particles/mL. AAV1-SARE-Arc-PSDdelta1.2-mVenus-PARac1.SV40 was custom-prepared by the University of Pennsylvania Vector Core. The viral titer was typically 3.54 x 10^{12} particles/mL. All viruses were subdivided into 5 microliter aliquots, and stored at –80º C until they were used.

Surgery: Stereotactic injections. Anesthesia was induced using isoflurane (induction, 5%; maintenance, 1-2%). We administered dexamethasone (1 mg/kg, i.p.) immediately prior to surgery to reduce brain swelling, and metacam (1 mg/kg, i.p.) as a prophylactic analgesic. After animals were placed in a stereotactic frame (Kopf Instruments), sterile eye lubricant (Puralube, FischerSci) was administered to prevent corneal drying, and a microwavable heating pad (Snugglesafe) was used to maintain body temperature.

Scalp fur was trimmed, and the skull surface was exposed with a midline scalp incision. Bupivicaine (0.05 mL, 5 mg/ml) was administered topically as a second prophylactic analgesic. The target site was identified using stereotactic coordinates (see below). A high-speed dental drill (Model EXL-M40, Osada Inc) and 0.5-mm burr were used to open a small (~0.5 mm) craniotomy. The virus was injected at a rate of 100-150 nL/min using a Nanofil syringe (World Precision Instruments) with a 33 G beveled needle (World Precision Instruments) and pump (World Precision Instruments) mounted onto the stereotactic frame. 500 nL of AAV1-Syn-GCaMP6s was injected for two-photon calcium imaging experiments, 300 nL of AAV1-Syn-GCaMP6s was injected for single fiber photometry experiments, and 1500 nL of a 1:2 ratio of AAV1-Syn-GCaMP6s and AAV1-SARE-Arc-PSDdelta1.2-mVenus-PARac1.SV40 was injected bilaterally for the dual fiber photometry experiments. After each injection, the needle was kept in place for two minutes to allow time for diffusion of the virus prior to removing the needle from the brain slowly over a two-minute period. Prefrontal infusion coordinates for two-photon imaging through a microprism were (from bregma): anterior-posterior (AP) +1.7 mm; mediolateral (ML) 0.35; and dorsal-ventral (DV) –1.2. Ventromedial prefrontal infusion coordinates for fiber photometry experiments in vmPFC were (from bregma): AP +1.7 mm; ML 0.35; and DV -2.5 or –3.0.

Surgery: Microprism implantations. Microprism implantations were performed as previously described (*13*). Briefly, the animal was prepared for craniotomy as above, and a small circular section of skin (~1 cm in diameter) was excised using surgical scissors (Fine Science Tools). The periosteum was bluntly dissected away and bupivacaine (0.05 mL, 5 mg/kg) was topically applied. The region to be imaged (medial prefrontal cortex) was identified using stereotaxic

coordinates (AP +1.7 mm, ML 0.35 mm). A custom-made circular titanium headplate was attached to the skull using dental cement (C&B Metabond, Parkell Inc). The titanium head plate was then screwed into a custom-built fork fixed to a solid metal base.

A ~4-mm circular craniotomy centered over the midline (AP +1.5 mm) was opened using a 0.5-mm burr (Fine Science Tools) and a high-speed hand dental drill (Osada), taking great care not to compress brain tissue or damage the sagittal venous sinus. Gentle irrigation with phosphate-buffered saline (137 mM NaCl, 27 mM KCl, 10 mM phosphate buffer, VWR) was used to clear debris at regular intervals. The dura beneath the craniotomy was delicately removed using fine forceps (Fine Science Tools). Sugi swabs (John Weiss & Son, Ltd) were used to absorb trace bleeding.

The microprism to be implanted was a square (1.5 mm), right-angle prism made of borosilicate glass with a reflective aluminum coating on the hypotenuse with a silicon dioxide protective coating (Optosigma). We secured the microprism to a 3-mm glass coverslip (Warner Instruments) using ultraviolet light-activated optical adhesive (Norland, Thorlabs), positioning the prism face at the center of the coverslip and taking care to leave the prism face free of adhesive material. The prism was slowly lowered into the brain using a digital micromanipulator over the course of ~10 minutes, until the prism face was flush with the midline and the prism was submerged in the brain. To aid in this process, and to ensure micrometer-precision of the position of the prism, we used a stereotaxic micromanipulator (Kopf) with the prism attached to a central vacuum line via an 18G needle. Veterinary adhesive (Vetbond, Fisher Scientific) was used to form a seal between the coverslip and the skull. A layer of Metabond was then applied for added durability. Metacam (1 mg/kg, i.p.) was administered as an analgesic 24 hours after surgery, and as needed thereafter. We (*13*) and others (*16*) have shown that this preparation is well tolerated. We found there is minimal reactive gliosis beyond 50 micrometers from the prism face; therefore, we limited our analyses to images acquired from a position > 50 μm from the prism. After allowing ~2 weeks for recovery, this preparation allowed us to image dendritic spines in the anterior cingulate (ACg) and prelimbic (PL) regions of the intact, contralateral mPFC, through a light path reflected off the hypotenuse of the prism and across the midline (as described below). We found that this protocol yielded high-quality images that could be reliably relocated at all three time points studied here and were unobscured by bleeding or excessive gliosis in >80% of surgeries.

Surgery: Optical fiber implantations for fiber photometry. Optical fiber implantations were performed following the stereotaxic viral injection described above. After opening a craniotomy slightly larger than the optical fiber, a 400 μm diameter multimodal optical fiber (Thorlabs) was lowered using a stereotaxic micromanipulator (from Bregma: AP +1.7, ML 0.35, DV -2.5 or -2.8, just dorsal to the viral injection coordinates). For experiments involving dual fibers, we used custom dual core implantable fiber optic cannulae (fiber core diameter: 400 μm; fiber pitch, center to center: 800 μm; fiber length: 3 mm) (Thorlabs). Vetbond was used to create a seal between the exposed craniotomy and the optical fiber, and Metabond was used to stabilize the position of the optical fiber. Mice were randomly assigned to experimental groups prior to the start of the experiment.

Histology, Immunohistochemistry, and Related Analyses. To confirm the spatial extent of GCaMP6s expression in medial PFC, mice were anesthetized with Euthasol (200-300ul per 25g mouse, 3.9 mg/ml, i.p.) at the end of each experiment and transcardially perfused with 1x PBS followed by 4% paraformaldehyde (PFA) diluted in PBS. The brains were removed and fixed overnight in 4% PFA, then subsequently washed and stored in 30% sucrose (diluted in PBS) for 2-3 days. The brains were then frozen in embedding medium (Tissue-Tek). 30 μm coronal sections were cut on a Leica CM3050 S Vibratome and placed on coverslips. After drying, the slides were washed three times with 1x PBS, and mounted with coverslips using Vectashield mounting media with DAPI (Vector Laboratories). Images were acquired using either an Olympus confocal scanning laser microscope or Leica DFC360 FX epifluorescent microscope.

To quantify chronic CORT and ketamine effects on synapse density and spine density in various mPFC subregions (fig. S6), Thy1-YFP transgenic mice were randomized to three groups receiving either 1) chronic CORT in the drinking water x 10 days + vehicle injection on Day 10; 2) chronic CORT in the drinking water x 10 days + ketamine injection on Day 10; or 3) a control group receiving regular drinking water (no CORT) and a vehicle injection on Day 10. On Day 11, they were deeply anesthetized with 150 mg/kg i.p. sodium pentobarbital and transcardially perfused at a rate of 9mL/min with 5mL saline/heparin solution (Henry Schein, Melville, NY) followed by 30mL of 3.75% acrolein (Polysciences, Warrington, PA) in 2% paraformaldehyde

(PFA; Electron Microscopy Sciences, Fort Washington, PA). PFA was prepared with 0.2M phosphate buffer (PB; VWR, Radnor, PA). Brains were harvested into a 1.87% acrolein/2% PFA solution, placed on a shaker for 30 minutes, and then rinsed and transferred to 0.1M PB and stored at 4°C until sectioning. 40μm-thick brain sections were obtained on a vibrating microtome (Leica Microsystems, Deerfield, IL) and stored in labeled 24-well plates containing cold 0.1M PB. Sections were transferred to a tissue storage solution (30% sucrose and 30% ethylene glycol in 0.1M PB) and kept at –20°C until punch coding and immunohistochemical processing.

Tissue was processed for GFP, PSD-95 and DAPI labeling. Floating sections were rinsed 3 times (10min each) in 1x Tris-buffered saline (TBS, Sigma Aldrich, Burlington, MA) on a rotary shaker and then blocked for 1 hour in TBS containing 4% Horse whole serum, 0.2% BSA, and 0.3% Triton-X 100. Because a mouse monoclonal antibody was chosen against PSD-95, tissue was treated with Mouse on Mouse (M.O.M.) reagent to block endogenous mouse IgG staining (Vector Biolabs, Malvern, PA) prior to incubation with primary antibodies. A mixture of primary antibodies containing goat polyclonal to GFP 1:1000 (VWR) and mouse monoclonal to PSD-95 1:200 (Sigma Aldrich) were prepared in TBS with 0.3% Triton-X and incubated on a nutator for 48 hours at 4°C. Following incubation, sections were rinsed 3 times (10 minutes each) in TBS and a secondary antibody mixture was prepared containing donkey-anti-mouse Alexa Fluor 546 1:400 and donkey-anti-goat Alexa Fluor 488 1:400 (Invitrogen, Carlsbad, CA) in 0.3% Triton-TBS. Sections were kept in the dark and incubated at room temperature for 2.5 hours, then rinsed once more and mounted in serial order on super-plus glass slides (VWR), air-dried in the dark and cover-slipped with Hard-set mounting medium containing DAPI (Vector Labs, Burlingame, CA).

Images for PSD95-stained sections were acquired on an Olympus FV1000 confocal microscope. YFP-Alexa 488 fluorescence was imaged using 488 nm laser excitation and a 500–545nm bandpass filter. PSD95-Alexa 546 was imaged using 559 nm laser excitation and a 575-675nm bandpass filter. These excitation/emission settings, in addition to acquiring channels sequentially, were chosen to minimize fluorescence bleeding from the green channel to the red channel. All images were acquired at 512 x 512 resolution using 2x Kalman line averaging. Acquisition settings remained constant between all images acquired at the same magnification. Specific ROIs were chosen for to sample the mPFC at approximately the infralimbic, prelimbic, and anterior cingulate cortex. High-resolution z-stacks of each ROI were acquired using a 60x

oil-immersion lens (Plan Apo, Olympus, NA = 1.4) with 2x optical zoom, 100μm pinhole, and at a step size of 0.5μm over 10μm total depth.

Analysis. Spine density analyses were performed by manually counting dendritic spines on individual dendritic branch segments of approximately 20–30 μm in length, involving 50+ dendritic branch segments per mPFC subregion per subject. Likewise, PSD-95 immunofluorescence density was quantified by counting immunofluorescent puncta in confocal image Z stacks. The data acquisition and quantifications were performed by investigators blind to the experimental status of the subject. The PSD-95 and spine density data in fig. S6 were expressed in terms of Z scores normalized to the mean densities observed in the control condition.

AS-PaRac1 Experiments. For the validation experiments reported in fig. S13, we injected a viral vector driving AS-PaRac1 expression (AAV1-SARE-Arc-PSDdelta1.2-mVenus-PARac1) into the medial PFC and implanted a microprism for visualizing and quantifying AS-PaRac1 expression patterns. After two weeks to allow time for surgical recovery and viral vector expression, we used 2P imaging to visualize AS-PaRac1 expression patterns before and one day after ketamine treatment. (For imaging details, see the following section: Data Acquisition: Two-Photon Imaging). As in our analyses of PSD-95 expression, AS-PaRac1 expression patterns were quantified in fig. S13D-E by manually counting AS-PaRac1+ fluorescent puncta in 2P image Z stacks. For the analyses in fig. S13F-J, in which we tested whether AS-PaRac1+ fluorescent puncta were spatially clustered, we used the 3D object counter function in the Image J software package to automatically detect AS-PaRac1+ fluorescent puncta and record the location of their centroids in X, Y, and Z coordinates. For fig. S13H, we measured the mean distance between each AS-PaRac1+ punctum and its nearest neighbor. Next, we tested whether the mean distance between pairs of AS-PaRac1+ puncta was significantly different from what would be expected by chance in simulated data. Here, the observed number of AS-PaRac1+ puncta in each 3D image Z stack was randomly distributed in a simulated space of equal volume, and the distance between nearest neighbor pairs was calculated for each simulated AS-PaRac1+ punctum. This process was repeated 1,000 times, enabling us to calculate a 99.9% confidence interval for the mean distance between AS-PaRac1+ puncta that would be expected by chance if they were

randomly distributed throughout the volume. Similarly, for fig. S13I-J, we classified each AS-PaRac1+ punctum as clustered if the distance to its nearest neighbor was < 5.8 μm, a threshold determined by the distribution of AS-PaRac1+ puncta in the simulation described above, in which the probability of observing a pair of puncta separated by < 5.8 μm was $P < 0.05$. Triplets and quadruplets of clustered AS-PaRac1+ puncta occurred when the two or three nearest neighbors all occurred within < 5.8 μm of a given punctum, respectively. In fig. S13J, we calculated 99.9% confidence intervals for the occurrence of 2 clustered puncta and 3+ clustered puncta (triplets and quadruplets) by testing how often they occurred across the 1,000 iterations of randomly distributed puncta described above.

For the experiments reported in Fig. 6, in which we tested the effects of activating AS-PaRac1 on behavior and calcium transient frequency in the tail suspension test, experimental and control animals were stimulated with a 470 nm LED (Thorlas, M470F3) with an intensity of 20 mW at the tip of the fiber. The light was controlled by an arduino and pulsed for 150 ms at 1 Hz for a period of 1 hour. For fiber photometry experiments, the light was delivered through a 0.48 NA, 400 μm core branching fiberoptic patchcord (Doric). For spine-imaging experiments, a 0.48 NA, 400 μm core optic fiber patch cord (Doric patchcord) was interfaced with the cranial glass coverslip using Kwik-Sil adhesive (WPI) while the mouse was head fixed.

Data Acquisition: Two-Photon Imaging. All images were acquired using a commercial two-photon laser-scanning microscope (Olympus RS) equipped with a scanning galvanometer and a Spectra-Physics Mai Tai DeepSee laser tuned to 920 nm. To allow for high-resolution imaging through a 1.5-mm microprism, we used a 25X, 1.00 numerical aperture water immersion objective with a 4-mm working distance (Olympus). Fluorescence was recorded through gallium arsenide phosphide (GaAsP) detectors using the Fluoview acquisition software (Olympus) using a green light emission bandpass filter (Semrock). All imaging experiments began by obtaining a low-magnification *z* stack (no digital zoom) to aid in relocating the same sites repeatedly over time, in conjunction with vascular landmarks and the contours of the prism. For calcium imaging experiments, we acquired time-lapse images (512 x 512 pixels, 350 frames, ~1 frame per second) spanning an area of mPFC measuring approximately 508 μm by 508 μm. All calcium imaging experiments occurred in awake mice. For spine imaging and AS-PaRac1 imaging experiments,

we acquired *z* stacks (512 x 512 pixels, 2-μs pixel dwell time, 0.75–1 μm step size) with 3X digital zoom through up to 250 μm of tissue in *z*. Spine imaging experiments occurred under KX anesthesia (ketamine 100 mg/mL and xylazine 10 mg/mL, at dosages of 0.1 mL/10 g body weight), except during the ketamine treatment experiments. Although previous studies suggest that the antidepressant-like effects of ketamine on behavior and spine density occur only after subanesthetic doses of ketamine (and not after anesthetic doses) (*4*), we sought to avoid this confound entirely by imaging awake mice that were previously habituated to the imaging procedure for all imaging experiments occurring before and after antidepressant-dose ketamine treatment. If we observed substantial movement of the mouse during image acquisition, we obtained a second image stack at the same position. Images contaminated by movement (easily detectable upon visual inspection) were excluded from our spine imaging analyses.

Data Acquisition: Fiber Photometry. Fiber photometry was performed to measure calcium dependent activity dynamics during tail suspension behavior (*18*, *88*). To excite GCaMP6s, light from a 470 nm LED (Thorlabs, M470F3) modulated at a frequency of 521 Hz was passed through a filter (Semrock, FF02-472/30), reflected by a dichroic (Semrock, FF495-Di03), and coupled to a 0.48 NA, 400 μm core optic fiber patch cord (Doric) for single fiber experiments or a 0.48 NA, 400 μm core dual fiberoptic patchcord with guiding socket (Doric) for dual fiber experiments. Emitted fluorescence traveled back through the patch cord, passed through the dichroic, a filter (Semrock, FF01-535/50), and was focused onto a photodetector (Newport, Model 2151). The modulated signal passed from the photodetector to a RP2.1 real-time processor (Tucker Davis Technologies) where it was demodulated and low-pass filtered using a corner frequency of 15 Hz. TTL pulses denoting the start of tail suspension trials were passed to this processor in real time for alignment of calcium signals to behavioral measures. Note that the photometry experiment in Fig. 6D involved mice that were expressing GCaMP6s plus mVenus-tagged AS-PaRac1. Although the emission spectra of GCaMP6s and mVenus do overlap, the co-expression of mVenus had no significant effect on our ability to record the fluctuating GCaMP6s signal, presumably because photometry signal is dominated by GCaMP6s fluorescence, because the mVenus signal is much weaker due to the fact that AS-PaRac1 is expressed so sparsely. We confirmed this by comparing their dynamic ranges: there was no significant difference in

dynamic range in mice co-expressing GCaMP6s and AS-PaRac1 (mean=9.9% +/– 2.0%) compared to mice expressing GCaMP6s alone (mean=7.7% +/– 2.3%; t=1.59, P=0.13).

Data Analysis: Dendritic Spine Remodeling. For all analyses of dendritic spine remodeling in Thy1/YFP transgenic mice, individual dendritic branch segments were randomly selected from a given imaging volume by an investigator blind to the experimental condition, provided that dendritic spines could be clearly resolved on that segment, as per previously established methods (Yang et al., 2009). We used ImageJ to quantify spine remodeling by comparing the same dendritic branch segments in pairs of images obtained over a given imaging interval. Formed spines (absent in image 1, present in image 2) and eliminated spines (present in image 1, absent in image 2) were quantified as a percent of the total number of spines identified in the initial image, typically 150–200 spines per subject. Spine images were scored by two investigators blinded to the experimental condition.

Statistics: Wilcoxon rank-sum tests (suitable for smaller sample sizes that may not be normally distributed) were used to compare spine elimination and formation rates across experimental conditions. We also examined the distribution of spine elimination and spine formation across individual dendritic branch segments, and used linear mixed effects modeling (fixed factor: experimental group, random factor: subject) to test for effects of experimental condition on spine elimination by dendritic branch segment (Fig. 1D: chronic CORT vs. control) or spine formation by dendritic branch segment (Fig. 1F chronic CORT + ketamine vs. chronic CORT + vehicle).

To test for spatially clustered spine elimination (Fig. 1E), we calculated the distance between each eliminated spine and its nearest eliminated neighbor. (Spines that did not have an eliminated neighbor on a contiguous dendritic branch segment were excluded from this analysis.) Next, we tested whether the mean distance between pairs of eliminated spines was significantly different from what would be expected by chance in simulated data. Here, the observed number of eliminated spines was randomly distributed along a single "synthetic dendrite" as in Ref. (*89*) with a length equal to the sum of the lengths of all analyzed dendritic branches, and the distance between nearest neighbor pairs was calculated for each eliminated spine. This process was repeated 1,000 times, enabling us to calculate a 99.9% confidence interval for the mean distance

between eliminated spines. An analogous analysis was carried out to test for spatially clustered spine formation in Fig. 1G.

To test whether ketamine selectively restores spines that were lost during chronic CORT exposure (Fig. 2), we classified each spine that formed in the 24 hour interval after ketamine (or vehicle) treatment as 1) a "restored formed spine" if it was located < 2 μm from a previously eliminated spine and extending from the same side of the dendrite; or 2) a "de novo formed spine" if it was located > 2 μm from a previously eliminated spine or extended from the opposite side of the dendrite. Wilcoxon rank sum tests were used to compare the proportion of formed spines that were restored spines (Fig. 2B: restored formed spines / all formed spines) and the fraction of eliminated spines that were rescued (Fig. 2E: restored formed spines / all eliminated spines) in the ketamine- vs. vehicle-treated groups. To test whether the restored spine rate was significantly different than would be expected by chance given the observed rate of ketamine-induced spine formation and the observed rate and spatial distribution of CORT-induced spine elimination (Fig. 2B), we calculated a 99.9% confidence interval for this value using the method described above for Figs. 1E and 1G

Data Analysis: Two-Photon Calcium Imaging. *Preprocessing.* For *x-y* motion correction, we used the ImageJ plugin "Image Stabilizer" created by Kang Li and Steven Kang from Carnegie Mellon University, available for download online. Image time series were segmented into individual cells using custom MATLAB scripts based on an established sorting algorithm combining independent components analysis and image segmentation based on threshold intensity, variance, and skewness in the *x-y* motion corrected data set, and upper and lower limits for cell size (*90–92*). Image segmentation results were manually inspected for quality control purposes. A small proportion of segmented ROIs that did not resemble neuronal soma in their shape or size (< ~5% and typically representing neuropil or apical dendrites in cross section) were excluded from further analyses. Fluorescence signal time series (ΔF/F: change in fluorescence divided by baseline fluorescence) were calculated for each individual neuronal segment: a 40-second sliding window was used to calculate the baseline fluorescence for each cell, accounting for both differences in GCaMP expression and de-trending for slow time-scale changes in fluorescence as described elsewhere (*91*). The onset of a calcium transient was

identified as occurring when ΔF/F exceeded two standard deviations above the baseline fluorescence, and the termination of a calcium transient was identified as occurring when ΔF/F fell below 0.5 standard deviation above the baseline fluorescence. For the purposes of calculating the correlation matrices depicted in Fig. 3D-E, all other values in the time series were set to zero. Cells that did not exhibit any statistically significant calcium transients during a given recording session were excluded from all analyses.

Statistics. To test for changes in functional connectivity between cells after chronic CORT and ketamine treatment, we calculated correlation matrices (Spearman's rho) between each cell and every other cell in a given field of view. Linear mixed effects modeling (fixed factor: experimental time point [baseline, chronic CORT, chronic CORT + ketamine], random factor: subject) was used to test for changes in the distribution of correlation coefficients after chronic CORT exposure and after ketamine treatment (Figs. 3D-E), accounting for shared variance within subjects. Statistically significant multicellular ensemble events were identified as occurring when the number of simultaneously active cells exceeded the number expected by chance given the observed level of activity across cells in the field of view. This threshold was calculated separately for each calcium imaging time series (i.e. for each subject on each session) by 1) shuffling the offset of each cell's time series with respect to every other cell's time series (i.e. adding a random lag); 2) calculating the number of simultaneously active cells at each time point; and 3) repeating this process 1,000 times to calculate a 99.9% confidence interval for this value. A statistically significant multicellular ensemble event was identified as occurring when the number of simultaneously active cells exceeded this value (typically 25-35% of all cells). In panel 3F, the frequency of multicellular ensemble events was estimated for each subject at each experimental time point (baseline, post-chronic CORT, post-chronic CORT+ketamine) in two three-minute windows, and Friedman's test (a non-parametric test suitable for repeated measures in small samples) was used to test for changes in ensemble event frequency over time. In panel 3G, the maximum proportion of participating cells (expressed as a percent of all cells in the field of view) was calculated for each ensemble event, and a linear mixed effects model (factor: experimental time point [baseline, chronic CORT, chronic CORT + ketamine], random factor: subject) was used to test for changes in the proportion of participating cells over time.

Data Analysis: Fiber Photometry. Data were analyzed using custom MATLAB (Mathworks) scripts. Fiber photometry data were recorded continuously during tail suspension. Fluorescence signals were first normalized (transformed to ΔF/F) by calculating the median signal across the recording period, subtracting this from each data point, and then dividing by the median signal. Significant calcium transient events were identified time points in which ΔF/F surpassed 0.5 median absolute deviations (MADs) as established elsewhere (*18*). The onset time of each transient event was defined as the time at which the derivative of the signal first rose above one MAD of the derivative and then remained above one MAD of the derivative until the signal reached 0.5 MADs (i.e., the onset of an event was defined as the first time point in a rising-slope that leads to a ΔF/F of at least 0.5 MADs, and the termination of the event was defined as the time point when ΔF/F fell below 0.5 MADs). The amplitude of a calcium transient event was defined as the maximum value of ΔF/F reached during the duration of the event (from event onset until the time at which signal fell below 0.5 MADs).

Statistics. The timing of struggling-related vmPFC activity was assessed by calculating the mean change in signal (ΔF/F) time-locked to the onset of struggling, averaged across all bouts of struggling across all subjects (Fig. 4B). The latency to onset of struggling (Fig. 4C) for each calcium transient event that occurred during a bout of immobility was defined as the time (in s) from the onset of the calcium transient to the onset of the next bout of struggling (defined below). To test whether vmPFC activity during bouts of immobility was predictive of shifts to struggling behavior (Fig. 4D), Monte Carlo simulation was used to estimate a 99.99% confidence interval for the median latency that would be expected by chance given the observed frequency of calcium transients and shifts to immobility by shuffling the offset of the fiber photometry recording time series with respect to the behavioral time series (i.e. adding a random lag) for each subject and repeating this procedure 10,000 times. To test whether the latency from the onset of a calcium transient to the onset of struggling varied with the amplitude of the calcium transient (Fig. 4D, all transients were sorted into quartiles, and a linear mixed effects model (fixed factor: amplitude quartile; random factor: subject) was used to test for a main effect of transient amplitude on latency to onset of struggling. Kruskal-Wallis ANOVA was used to test for an effect of experimental condition (control, chronic CORT, chronic CORT+ket) on total duration of immobility during tail suspension (Fig. 4E) and the frequency of vmPFC calcium transients (events per s; Fig. 4F). To test whether vmPFC calcium transients accumulated more

slowly over time during bouts of immobility in the chronic CORT-exposed group, the latency from the onset of immobility to the onset of each calcium transient was calculated, and a linear mixed effects model (fixed factor: experimental condition, random factor: subject) was used to test for between-group differences (Fig. 4G).

Behavioral Data. *Tail Suspension.* Subjects were video recorded for six minutes while suspended 15 cm from the top of a compartment measuring 55 cm in height using lab tape (VWR). A small plastic tube was placed around the base of the tail to prevent tail climbing. The behavioral apparatus was thoroughly cleaned with 70% ethanol between animals. Time spent completely immobile, except for respiratory movement, was determined by an observer blind to the experimental condition of the animals. Inter-rater reliability was assessed by comparing onset times for 127 bouts of immobility in a subset of mice scored by three independent raters blind to experimental condition. Estimates of the error in identifying the onset of immobility were assessed by repeatedly scoring the same bouts of immobility (fig. S8).

Elevated Plus Maze (EPM). The EPM task was conducted as described previously (*93*). Briefly, the maze was constructed of gray Plexiglas raised 70 cm above the floor and consisted of two closed arms with 14 cm high opaque walls and two opposite open arms of the same size (29 × 6 cm). All animals were habituated to the behavior room before their test sessions for one hour. The task was conducted under dim light (~6 Lux in center of maze), and the movements of the animals were captured by a digital camera from above and live-tracked using Ethovision XT 11 software (Noldus). To begin a trial, the animal was placed in the center of the maze facing an open arm, and allowed to freely navigate the maze for 300 s. Open-arm exploration time was assessed by computing the percentage of time spent in the open arms out of 300 s total. Between sessions, the maze was thoroughly cleaned with 70% ethanol to remove any residual odors. All behavioral sessions were conducted during the active (dark) phase of the animal's circadian cycle.

Sucrose Preference. All subjects were habituated to two leak-resistant water bottles in their home cage for three days prior to testing (group-housed as above). On test day, they were water restricted for eight hours prior to testing. At the onset of the sucrose preference test, each subject was temporarily single-housed in a cage with two leak-resistant water bottles, one

containing standard animal facility water and the other containing a 3% sucrose-in-water solution. Each subject had libitum access to the two water bottles for four hours. At the end of four hours, they were returned to their home cages. Water consumption and sucrose water consumption were calculated by weighing the water bottles before and after the four-hour testing session, and sucrose preference was calculated by calculating the fraction (%) of sucrose water consumption divided by total water consumption.

Statistics. All investigators were blind to experimental status during the acquisition and analysis of all three behavioral experiments. Kruskal-Wallis ANOVA and post-hoc linear tests were used to test for differences in immobility, open arm exploration, and sucrose preference across experimental conditions. All P values for all experiments (behavior and otherwise) are two-sided.

Supplementary Figure Legends

Fig. S1. Chronic CORT induces depression-related behavioral effects. Chronic CORT **(A)** reduced open-arm exploration in the elevated plus maze ($P=1.1e^{-4}$, Wilcoxon W=326.0, n=15 chronic CORT and n=16 controls), **(B)** decreased sucrose preference ($P=3.2e^{-4}$, Wilcoxon W=304.0, n=16 chronic CORT and n=14 controls), and **(C)** increased immobility in the tail suspension test ($P=7.2e^{-5}$, Wilcoxon W=121.0, n=16 chronic CORT and n=14 controls). Throughout all boxplot figures (unless otherwise noted), boxes denote the median and interquartile range (IQR), and whiskers denote the full range of the data excluding outliers (1.5 X IQR). In all figures, asterisks denote significantly different groups at $P<0.05$.

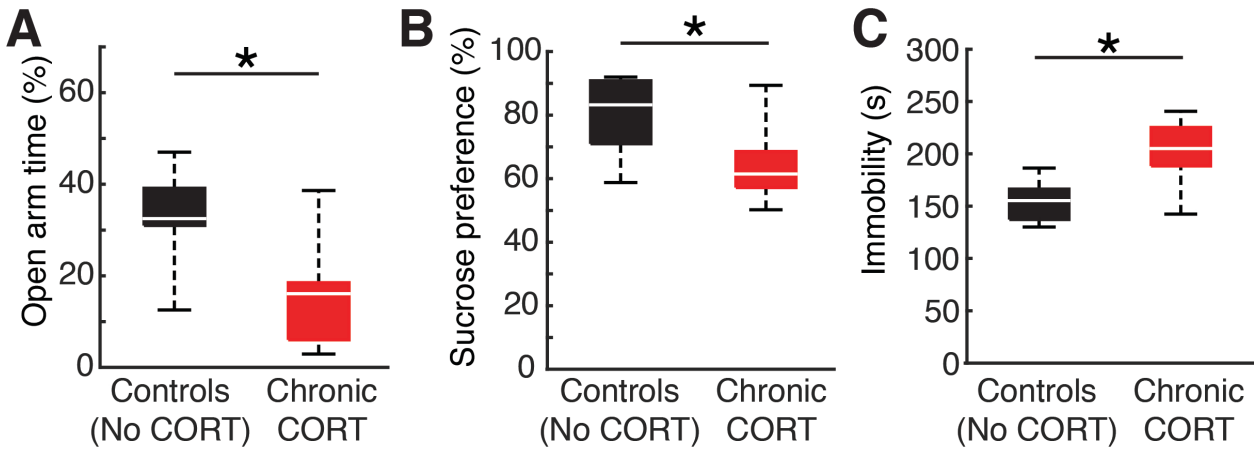

Fig. S2. Time-course and stability of CORT-induced spine remodeling over 10 and 21 days. **(A)** Schematic of experimental timeline. Two-photon imaging of YFP-expressing projection neurons before CORT exposure and after 10 and 21 days of exposure to CORT in the drinking water. **(B)** There was a modest decrease in the formation of new spines after 10 days of CORT exposure (N=6), relative to untreated controls (N=6), but this difference was not statistically significant (Wilcoxon W=49.5, P=0.10). **(C)** There was a significant increase in spine elimination after 10 days of CORT exposure (Wilcoxon W=23.5, P=0.011). **(D)** To test the longitudinal stability of CORT effects on spine elimination, we classified spines that were eliminated at the 10-day time point as "stably eliminated" if they were absent at both the 10- and 21-day time points or as "spontaneously restored" if they were absent on Day 10 but present on Day 21. The vast majority (mean = 88.8% in n=6 mice) of spines that were eliminated on Day 10 were also absent on Day 21.

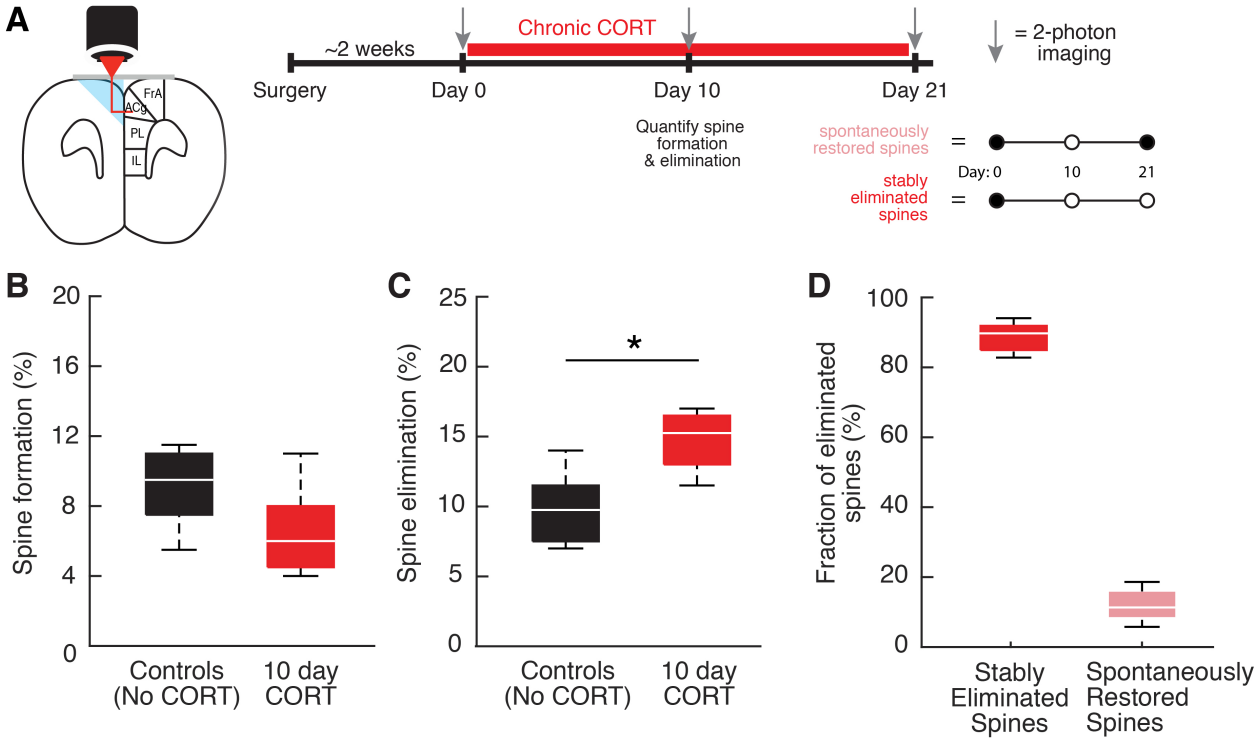

Fig. S3: Effects of repeated restraint stress and ketamine treatment on spine remodeling and behavior. (A) Experimental timeline and schematic. Mice were randomized to one of four groups receiving 1) repeated restraint stress x 21 days; 2) a gentle handling control group x 21 days; 3) restraint stress + ketamine; or 4) restraint stress + vehicle. To avoid habituation in the behavioral experiments in panels B-D, each mouse was tested only once on each behavior. **(B)** Repeated restraint stress decreased open arm exploration time in the elevated plus maze (N=16 mice per group, Z=2.79, P=0.0053), and ketamine reversed this effect (N=16 mice per group; Z=3.69, P=3.18e-4). **(C)** Repeated restraint stress decreased sucrose preference, and ketamine reversed this effect (N=16 mice per group; Kruskal Wallis ANOVA, $\chi^2 = 9.96$, P = 0.019). **(D)** Repeated restraint stress increased immobility in the tail suspension test, and ketamine reversed this effect (N=16 mice per group; Kruskal Wallis ANOVA, $\chi^2 = 28.0$, P = $3.63e^{-6}$). **(E-F)** Repeated restraint stress decreased spine formation (N=6 unstressed controls, N=10 stressed mice, Z=77.0, P=0.003, Wilcoxon) and increased spine elimination (Z=22.5, P=$7.5e^{-4}$, Wilcoxon). **(G)** Spatially clustered spine elimination after repeated restraint stress. The mean distance (red circle = 2.97 μm) between an eliminated spine and its nearest eliminated neighbor was significantly reduced ($P<0.001$) relative to what would be expected by chance given the observed rate of spine elimination, in simulations denoted by the boxplot (see Supplementary Methods). (**H-I)** Ketamine increased spine formation (N=6 mice with restraint stress + ketamine, N=4 mice with restraint stress + vehicle, Z = 10.0, p = 0.0095, Wilcoxon) and had no effect on spine elimination (Z=18.5, P=0.524, Wilcoxon). **(J)** Spatially clustered spine formation after ketamine treatment: analogous to panel G, the mean distance (blue circle, 3.05 μm) between each formed spine and its nearest formed spine neighbor was significantly reduced ($P<0.001$) relative to what would be expected by chance in simulations denoted by the boxplot. **(K)** Restored spines (blue) were defined as spines that formed in the 24-hour period after ketamine treatment in a position < 2 μm from a spine eliminated during restraint stress. **(L)** 44.5% of newly formed spines were located < 2 μm from a spine lost during repeated restraint stress (N=6 mice), significantly greater than the 15.5% observed in vehicle-treated controls (N=4 mice, Wilcoxon W=10.0, P=0.0095). **(M)** The observed fraction of restored spines (blue circle, 43.3%) was significantly (P<0.001) greater than expected by chance given the observed rate of spine formation in simulated data in which formed spines were randomly distributed (boxplot; see Supplementary Methods).

Fig. S3: Effects of repeated restraint stress and ketamine treatment on spine remodeling and behavior.

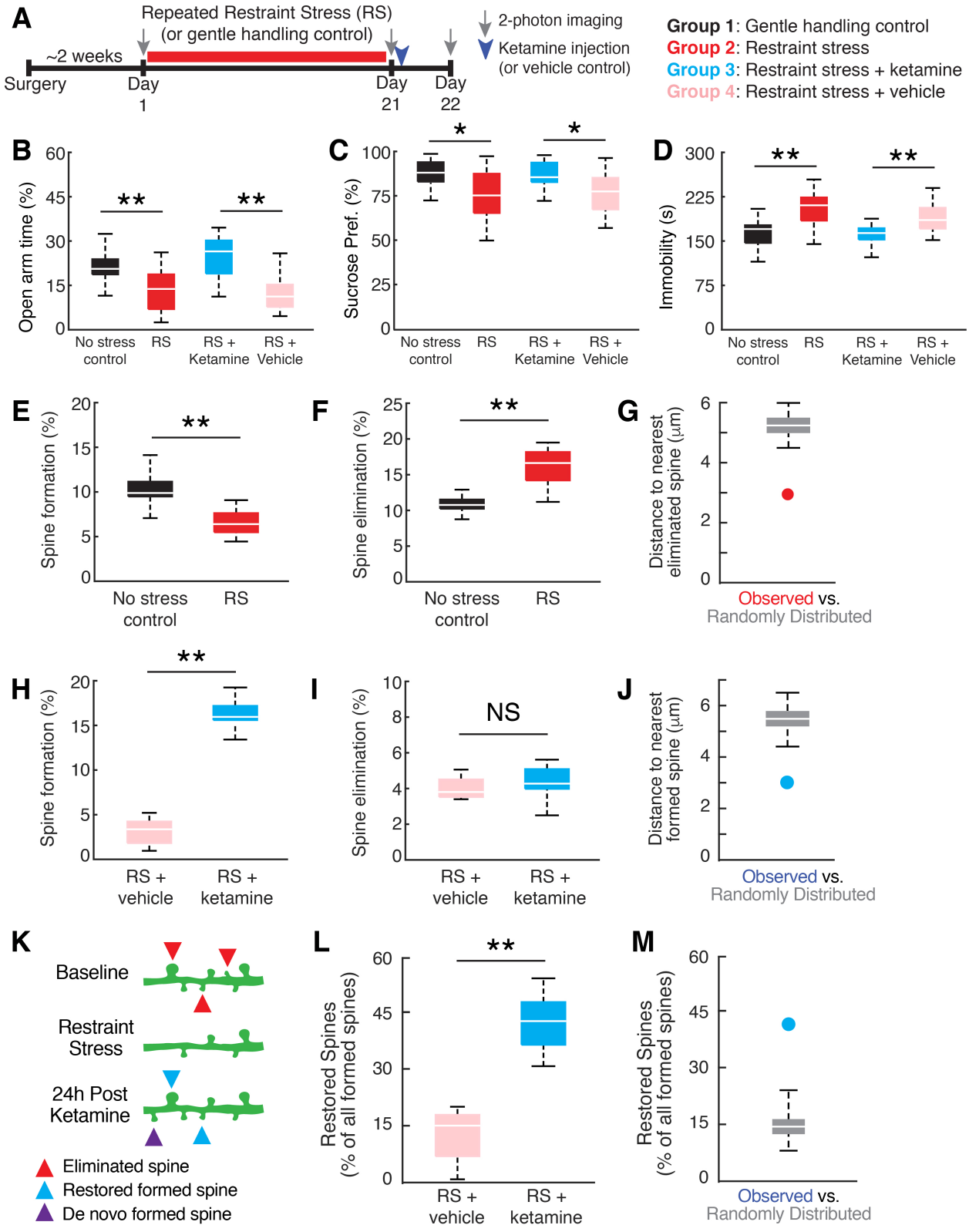

Fig. S4. Ketamine rescues depression-related behaviors. Ketamine rescued chronic CORT effects on **(A)** open-arm exploration in the elevated plus maze (P=9.7e-5, Wilcoxon W=138.0, n=15 chronic CORT and n=15 controls), **(B)** sucrose preference (P=0.007, Wilcoxon W=167.0, n=15 chronic CORT and n=15 controls), and **(C)** immobility in the tail suspension test (P=1.4e-4, Wilcoxon W=325.0, n=15 chronic CORT and n=15 controls).

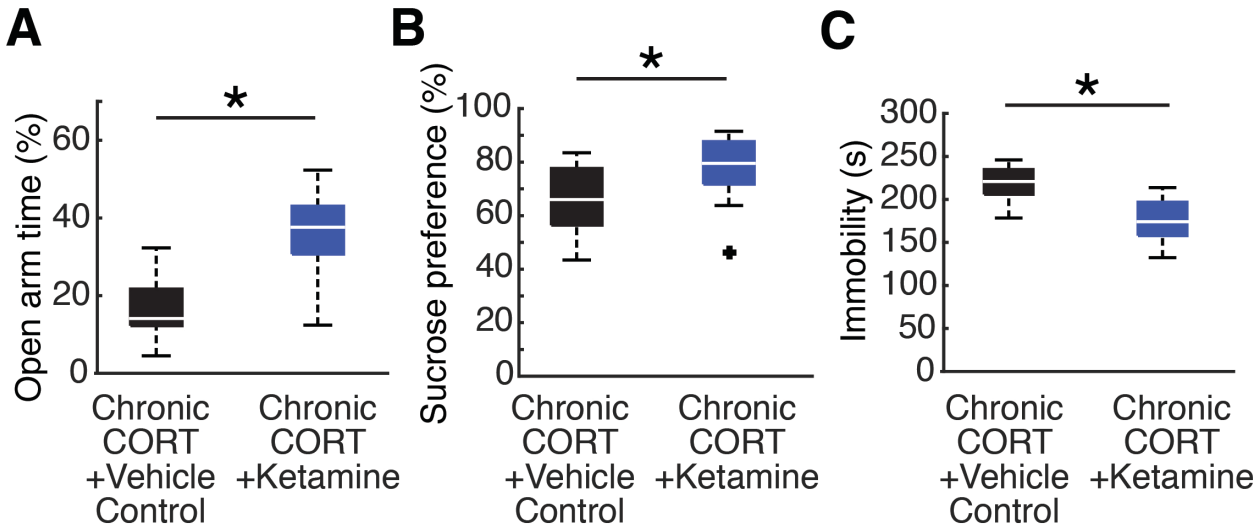

Fig. S5. Ketamine generates new synapses. (A) Schematic for quantifying the proportion of spines formed in the 24 hours after ketamine treatment and persisting for at least four days thereafter. Previous work has shown that essentially all of these persistent formed spines contain functional synapses (*41,42*). **(B)** Ketamine significantly increased the formation of spines persisting for at least four days (n=4 ketamine-treated mice, n=4 vehicle-treated controls, Wilcoxon W=10.0, P=0.029), which have been shown to contain functional synapses in previous studies (*41*, *42*). **(C)** Restored spines (n=93 spines located < 2 μm from a previously eliminated spine) were significantly more likely to survive for at least four days than de novo formed spines (n=106 spines, P=0.0018, X^2=9.75). **(D)** Ketamine (n=4) also increased the fraction of previously eliminated spines that were restored and persisted for at least four days compared to vehicle-treated controls (n=4, Wilcoxon W=10.0, P=0.029).

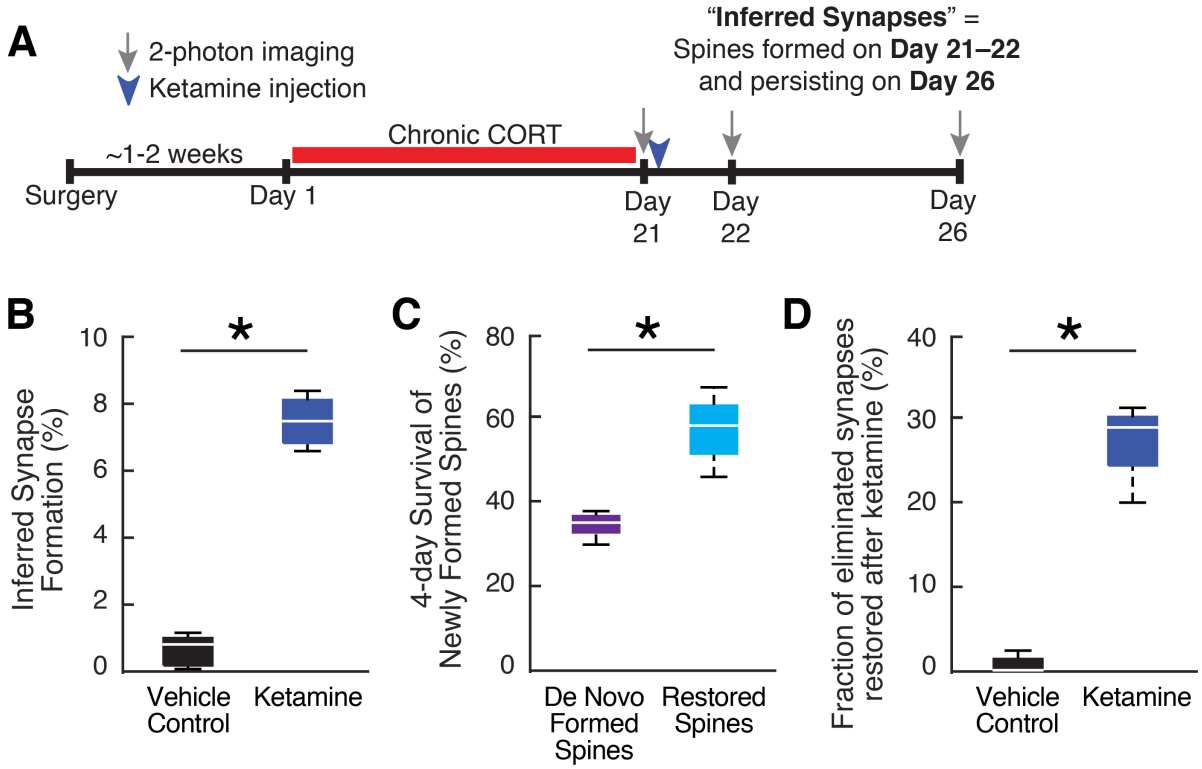

Fig. S6. Chronic CORT and ketamine effects on spine and synapse density in mPFC subregions. (A) Experimental timeline schematic. Thy1/YFP transgenic mice were randomized to three groups (N=6 per condition) receiving either 1) chronic CORT in the drinking water x 10 days + vehicle injection on Day 10; 2) chronic CORT in the drinking water x 10 days + ketamine injection on Day 10; or 3) a control group receiving regular drinking water (no CORT) and a vehicle injection on Day 10. **(B)** Representative image (max projection through a confocal Z-stack) of PSD-95 immunofluorescence in the anterior cingulate cortex. Confocal image stacks were also obtained from the prelimbic (PL) and infralimbic (IL) subregions of mPFC. Scale bar = 25 μm. **(C)** Two-factor, mixed effects ANOVA (subject = random effect) showed that chronic CORT decreased PSD-95 density and ketamine rescued this effect (main effect of experimental group: $F(2,45)=10.39$, $P=0.0019$). There was no effect of mPFC subregion ($F(2,45)=1.14$, $P=0.328$) and no group-by-subregion interaction ($F(4,45)=0.67$, $P=0.614$). PSD-95 density (see Supplementary Methods for details on quantification) is depicted in terms of Z scores normalized to the mean densities observed in the control condition. **(D-F)** PSD-95 density plotted separately for anterior cingulate, prelimbic, and infralimbic cortex, showing qualitatively similar effects across subregions. Post-hoc contrasts: ** = $P<0.01$, * = $P<0.05$, Holm-Bonferroni corrected for multiple comparisons. **(G)** At left, a representative image (max projection through a confocal Z-stack) of YFP-expressing projection neurons in ACg. The numbered, dashed boxes denote the dendrites depicted in the enlarged images at right. Scale bar = 25 μm. **(H)** Two-factor, mixed effects ANOVA (subject = random effect) showed that chronic CORT decreased spine density and ketamine rescued this effect (main effect of experimental group: $F(2,45)=10.26$, $P=0.00021$). There was also a main effect of mPFC subregion ($F(2,45) = 5.04$, $P=0.0.011$), but no group-by-subregion interaction ($F(4,45)=0.53$, $P=0.717$), indicating equivalent effects of chronic CORT and ketamine one all three mPFC regions. Spine density (see Supplementary Methods for details on quantification) is depicted in terms of Z scores normalized to the mean densities observed in the control condition. **(I-K)** Spine density plotted separately for anterior cingulate, prelimbic, and infralimbic cortex, showing overall lower spine densities in IL vs. ACg (which may be attributable in part to regional differences in Thy1-driven YFP expression) but qualitatively similar effects of chronic CORT and ketamine across subregions. Post-hoc contrasts: ** = $P<0.01$, * = $P<0.05$, after Holm-Bonferroni correction for multiple comparisons.

Fig. S6. Chronic CORT and ketamine effects on spine and synapse density in mPFC.

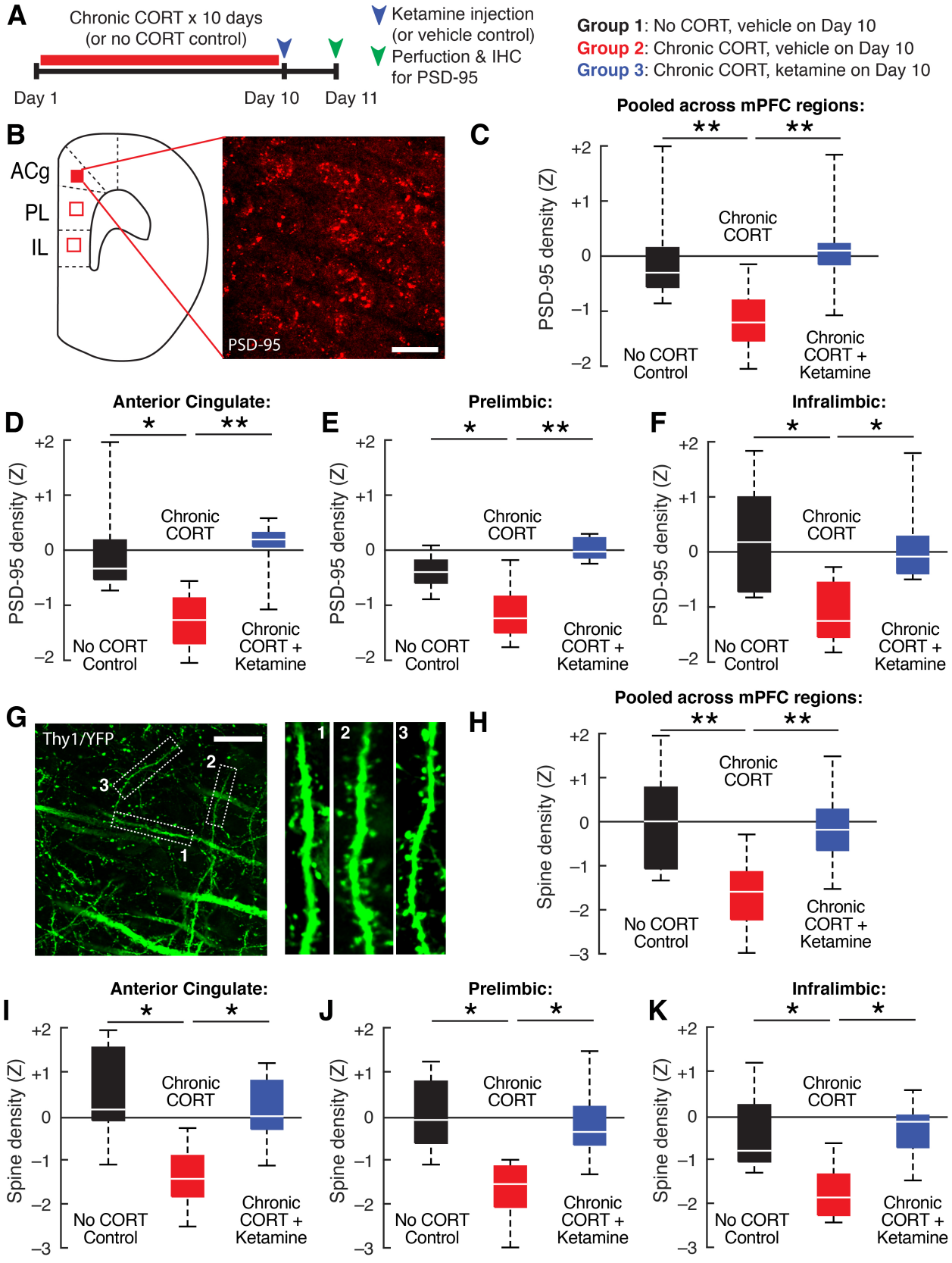

Fig. S7. PFC microcircuit function is stable over time in controls. (A) Cumulative distribution plot of cell:cell correlations for all cells across all animals in vehicle-treated controls (n=4 mice) imaged at the same time points as the CORT- and ketamine-treated mice depicted in Fig. 4. A linear mixed effects model (fixed factor: imaging time point [i.e. Day 0, Day 10, Day 11], random factor: subject) showed that there was no significant change in the mean correlation over time (t=1.68, P=0.095). **(B)** The frequency of multicellular ensemble events (expressed as the probability of an ensemble event occurring in any given time frame) was stable over time (X^2=0.64, P=0.725, Friedman's Test). **(C)** A linear mixed effects model showed that the proportion of cells participating in a multicellular ensemble event was also stable over time (t=0.496, P=0.621 for the main effect of time in n=131 multicellular ensemble events from n=4 mice).

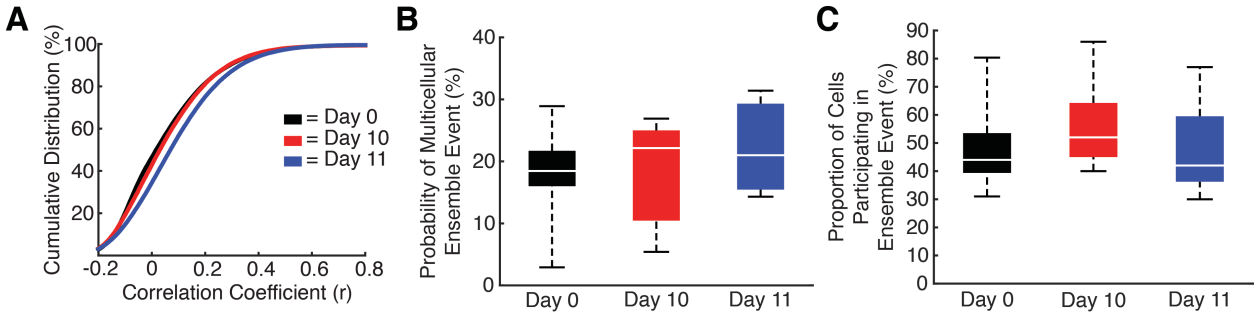

Fig. S8. Fiber photometry histology. (A) Confocal image showing GCaMP6s expression in ventromedial prefrontal cortex in a representative subject from the fiber photometry experiment depicted in Fig. 4. The dashed line in the inset depicts the location of the optical fiber. FrA = frontal association cortex; ACg = anterior cingulate cortex; PL = prelimbic cortex. IL = infralimbic cortex. **(B)** Across the photometry experiments in Figs. 4 and 6, the placement of the fiber tip along the dorsal-ventral axis varied from approximately –2.4 to –3.0 mm ventral to Bregma, an area denoted in green. (The experiments in Fig. 6 used dual-core fibers for bilateral light delivery, while the experiments in Fig. 4 used single-core fibers targeting just one hemisphere.) A representative 400-µm diameter optical fiber located in the same position as in the example in (A) is included for scale.

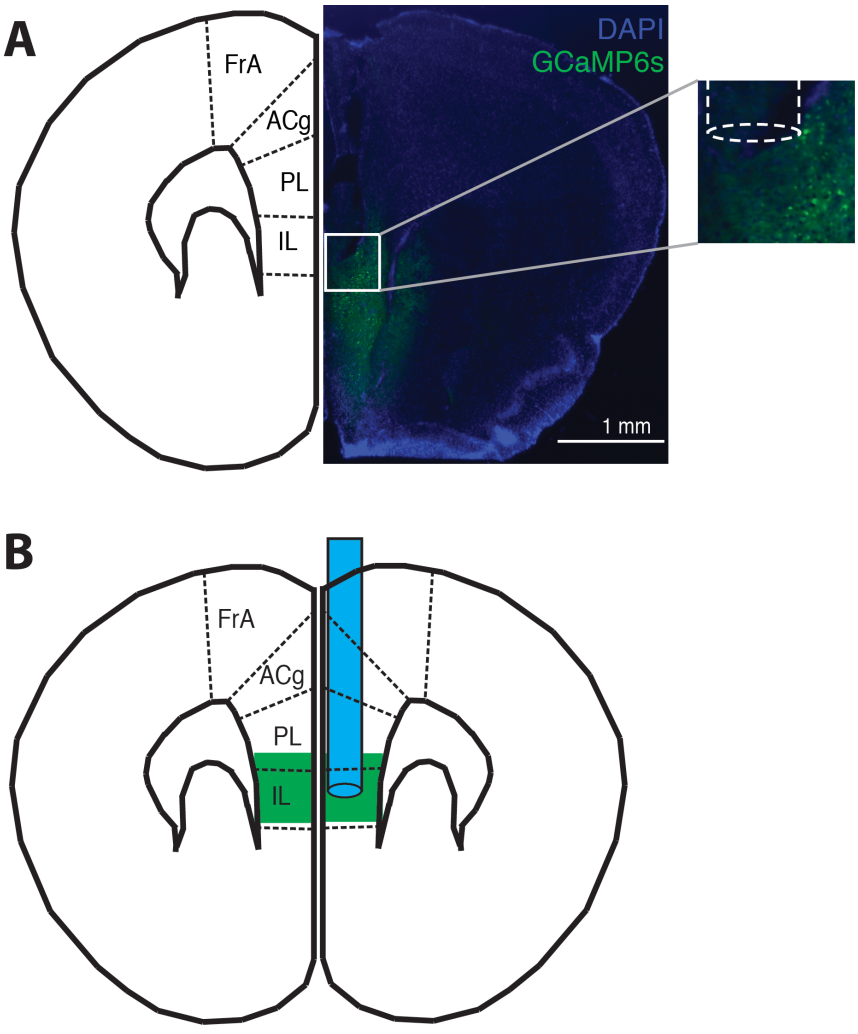

Fig. S9. Ventromedial PFC activity (mean ΔF/F over time) was significantly elevated during bouts of struggling relative to bouts of immobility (Wilcoxon rank sum test, Z=4.61, P=4.09e-6, N=29 mice).

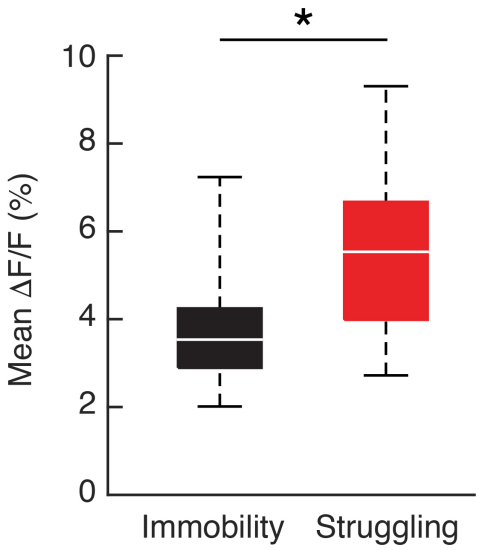

Fig. S10. No systematic delays in recording onset of struggling. To assess the accuracy of our scoring of the timing of the onset of struggling, three raters scored *N*=127 bouts of struggling. The boxplot depicts the median and interquartile range of the inter-rater difference, with the whiskers denoting the maximum and minimum values. The median inter-rater difference was 0.07 s (max 0.64 s), indicating that the observed increase in vmPFC activity prior to the onset of struggling (average latency 0.99 s denoted by the green data point) was not attributable to systematic delays in our manual scoring of the onset of struggling.

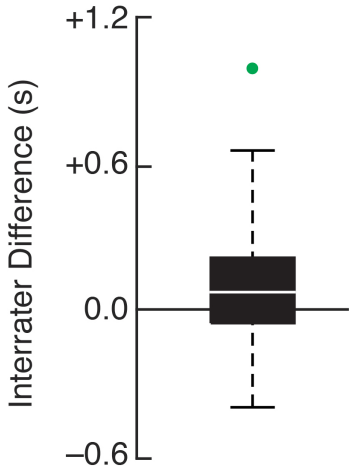

Fig. S11. No association between acute effects of ketamine on TST behavior and spine formation. (A) Schematic of experimental time course for panels B–C. As in Fig. 5A, to avoid behavioral habituation in the TST due to repeated testing on the same day, each subject was tested only once on this assay (i.e. at only one of the specific time points; between subjects design, N=12 subjects/time point/group). "Restored spines" were defined as spines that were present at baseline, eliminated after chronic CORT exposure, and restored at the specified time point after ketamine treatment in a position < 2 μm from their original position. **(B)** There was no correlation between immobility on the TST and the rate of lost spine restoration (Spearman's rho=0.105, P=0.744). **(C)** Because the rate of spine restoration was so low at the +3- and +6h-post-treatment time points, we also tested for an association between immobility and spine formation rate (i.e. the rate of newly formed spines after ketamine treatment, regardless of whether the new spine was located < 2 μm from a previously lost spine). There was no significant correlation (Spearman's rho = –0.191, P=0.552).

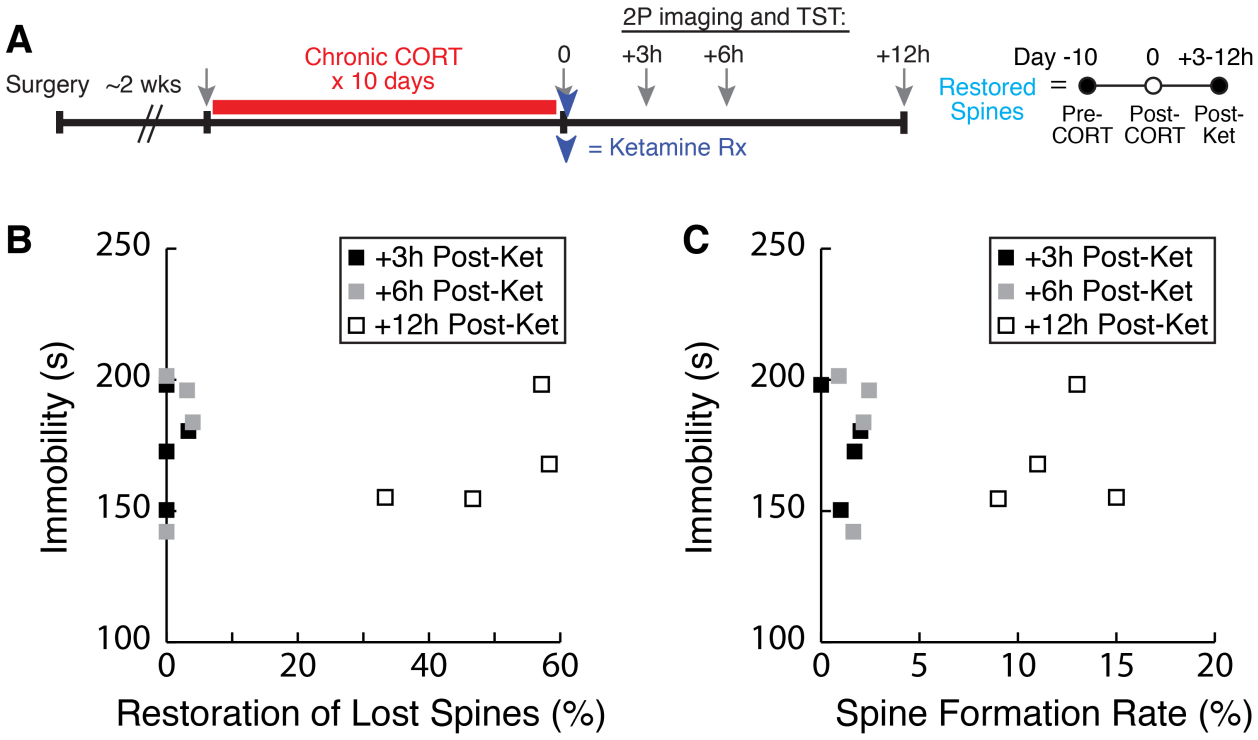

Fig. S12. Time course of ketamine effects on correlated activity in the PFC. Cumulative distribution plot of cell-cell correlations for all cells across all animals in each condition (*N*=95,418 correlations at baseline, *N*=112,020 after chronic CORT, and *N*=131,472, *N*=131,472, and *N*=131,472 correlations at +3h, +12h, and +24 after ketamine treatment in n=3 mice). There was a reduction in functional connectivity (mean correlation) after chronic CORT (mean r = 0.111 vs. r = 0.172 at baseline), and ketamine rescued this effect within 3 hours of treatment (r = 0.180, 0.226, and 0.221 at +3h, +12h, and +24h, respectively; N=5–8 samples per condition from n=3 subjects; mixed effects ANOVA, subject = random effect; main effect of time: F(4,32) = 3.05, p=0.032).

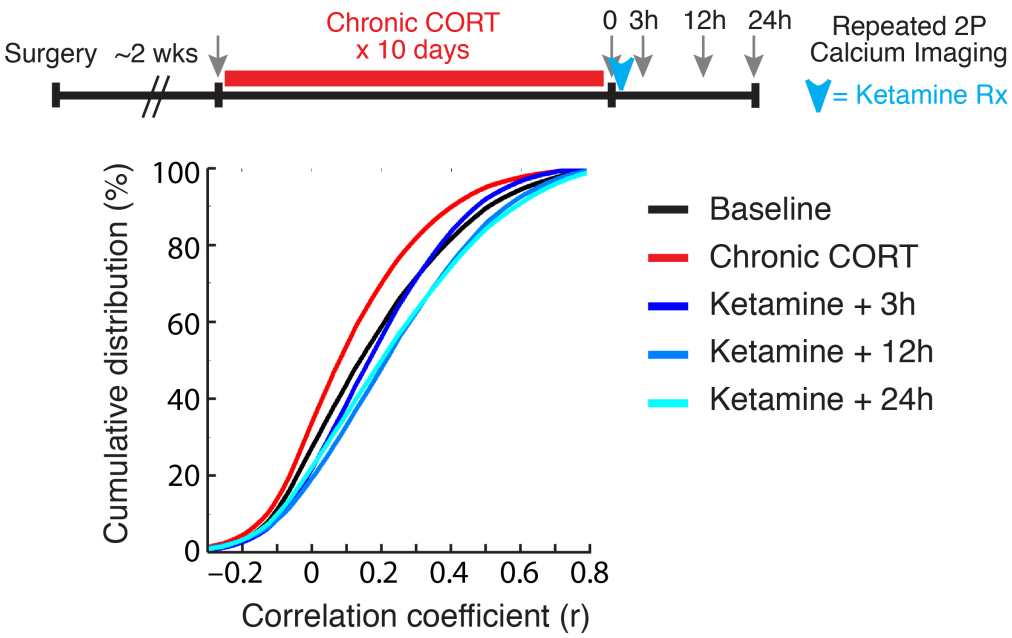

Fig. S13. Characterization of AS-PaRac1 expression. (A) Representative image of AS-PaRac1 expression in PFC (left) acquired in vivo via two-photon imaging through a chronically implanted microprism (right). Dashed lines denote blood vessels. The green line in the schematic at right denotes the approximate location of the imaged area with respect to the prism face. Scale bar = 100 μm. AS-PaRac1 expression was sparse and punctate, as previously observed in Ref. 20. **(B)** Experimental timeline. To characterize ketamine effects on AS-PaRac1 expression, mice were exposed to chronic CORT for 10 days and then imaged before and after ketamine treatment (N=34 two-photon image stack samples from n=4 mice). **(C)** Schematic representation of the location of AS-PaRac1+ puncta in a two-photon image Z stack (max projection through 75 μm) through the area depicted in panel A before (left) and one day after ketamine treatment (right). Scale bars = 100 μm. **(D)** Mixed effects ANOVA (subject = random effect) showed that the density of AS-PaRac1+ puncta increased after ketamine treatment (main effect of time: $F(1,66) = 65.4$, $P = 0.0009$. Blue data points denote the mean number of AS-PaRac1+ puncta per sample pre- and post-ketamine treatment. **(E)** Boxplot of the percent change in AS-PaRac1+ puncta pre- vs. one day post-ketamine treatment. **(F)** Representative high-resolution image (scale bar = 30 μm) of depicting examples of spatially clustered AS-PaRac1+ puncta one day after ketamine treatment. **(G)** Schematic representation of the location of AS-PaRac1+ puncta for the image in panel F. Blue dots denote spatially clustered pairs of puncta, and red dots denote spatially clustered triplets. Black dots denote puncta that were not spatially clustered. We classified each AS-PaRac1+ punctum as clustered if the distance to its nearest neighbor was < 5.8 μm. This threshold was determined based on simulations in which the observed number of puncta in each 3D image Z stack was randomly distributed in a simulated space of equal volume, and the distance between nearest neighbor pairs was calculated for each simulated AS-PaRac1+ punctum. This process was repeated 1,000 times, enabling us to calculate a 99.9% confidence interval for the mean distance between AS-PaRac1+ puncta that would be expected by chance if they were randomly distributed throughout the volume. The probability of observing a pair of puncta separated by < 5.8 μm was $P < 0.05$. **(H)** The mean distance between each AS-PaRac1 puncta and its nearest neighbor (blue dot) was significantly less than expected by chance in the simulations described in (G), indicating that the observed degree of spatial clustering was statistically significant. **(I)** Relative proportions unclustered puncta, clustered pairs, and clustered triplets and quadruplets across all samples, as defined by the classification system described in panel G. **(J)** The observed frequency of spatially clustered pairs (left) and triplets and quadruplets (right) was significantly greater than expected by chance in the simulations described in panel G. The grey boxplots denote the interquartile range and 99.9% confidence interval for these proportions across 1,000 simulations in which the observed number of AS-PaRac1 were randomly distributed throughout the volume. See Supplementary Methods for additional details.

Fig. S13. Characterization of AS-PaRac1 expression.