



Published in final edited form as:

*Mol Psychiatry*. 2022 March ; 27(3): 1502–1514. doi:10.1038/s41380-021-01364-y.

## Dopamine D2 receptors modulate the cholinergic pause and inhibitory learning

Eduardo F. Gallo, Ph.D.<sup>1</sup>, Julia Greenwald<sup>2,6</sup>, Jenna Yeisley<sup>1</sup>, Eric Teboul<sup>2,6</sup>, Kelly M. Martyniuk<sup>3,6</sup>, Joseph M. Villarín, M.D, Ph.D.<sup>2,6</sup>, Yulong Li, Ph.D.<sup>4</sup>, Jonathan A. Javitch, M.D, Ph.D.<sup>2,5,6</sup>, Peter D. Balsam, Ph.D.<sup>2,6,7,8</sup>, Christoph Kellendonk, Ph.D.<sup>2,5,6</sup>

<sup>1</sup>Department of Biological Sciences, Fordham University, Bronx, NY, USA

<sup>2</sup>Department of Psychiatry, Columbia University, New York, NY, USA

<sup>3</sup>Department of Neuroscience, Columbia University, New York, NY, USA

<sup>4</sup>State Key Laboratory of Membrane Biology, Peking University School of Life Sciences, Beijing, China

<sup>5</sup>Department of Molecular Pharmacology and Therapeutics, Columbia University, New York, NY, USA

<sup>6</sup>Division of Molecular Therapeutics, New York State Psychiatric Institute, New York, NY, USA

<sup>7</sup>Department of Psychology, Columbia University, New York, NY, USA

<sup>8</sup>Division of Developmental Neuroscience, New York State Psychiatric Institute, New York, NY, USA

### Abstract

Cholinergic interneurons (CINs) in the striatum respond to salient stimuli with a multiphasic response, including a pause, in neuronal activity. Slice physiology experiments have shown the importance of dopamine D2 receptors (D2Rs) in regulating CIN pausing yet the behavioral significance of the CIN pause and its regulation by dopamine *in vivo* is still unclear. Here, we show that D2R upregulation in CINs of the nucleus accumbens (NAc) lengthens the pause in CIN activity *ex vivo* and enlarges a stimulus-evoked decrease in acetylcholine (ACh) levels during behavior. This enhanced dip in ACh levels is associated with a selective deficit in the learning to inhibit responding in a Go/No-Go task. Our data demonstrate, therefore, the importance of CIN D2Rs in modulating the CIN response induced by salient stimuli and point to a role of this response in inhibitory learning. This work has important implications for brain disorders with altered striatal dopamine and ACh function, including schizophrenia and attention-deficit hyperactivity disorder (ADHD).

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:

**Corresponding authors:** Eduardo F. Gallo, Ph.D. [egallo4@fordham.edu](mailto:egallo4@fordham.edu); Christoph Kellendonk, Ph.D. [ck491@cumc.columbia.edu](mailto:ck491@cumc.columbia.edu).  
**Author contributions.** E.F.G., J.G., J.Y., E.T., K.M.M. performed the experiments. E.F.G., J.G., C.K., and P.D.B. analyzed the data. E.G. and C.K. designed the experiments and supervised the project, with input from P.D.B. and J.A.J.; E.F.G. and C.K. wrote the manuscript with input from J.G., J.A.J., Y.L., and P.D.B.

**Conflict of interest:** The authors declare that they have no conflicts of interest.

## Introduction

Cholinergic interneurons (CINs) account for less than 3% of the neuronal population of the nucleus accumbens (NAc), a region of the ventral striatum critically involved in motivated behavior<sup>1–5</sup>. While sparse, these neurons possess extensive axonal networks that allow them to exert widespread cholinergic influence over striatal neurons<sup>1</sup>. CINs regulate synaptic plasticity and excitability in the more abundant spiny projection neurons (SPNs), as well as in other interneurons<sup>6–10</sup>. In addition, CIN activity regulates local striatal dopamine (DA) release through complex signaling via nicotinic and muscarinic receptors<sup>11–14</sup>.

Therefore, NAc CINs are well positioned to be key regulators of reward-related behaviors, as supported by accumulating evidence from cell-targeted approaches in rodents. For example, immunotoxin-mediated ablation of rat NAc CINs increases sensitivity to the rewarding effects of cocaine<sup>15</sup>, while temporally discrete optogenetic silencing of NAc CINs blocks cocaine conditioned place preference<sup>7</sup> and reduces its extinction<sup>8</sup>. NAc CIN involvement in reward extends beyond addictive-like behaviors, influencing the hedonic impact of natural rewards<sup>16</sup> as well as the flexibility of reward-seeking strategies<sup>17</sup>. Pharmacogenetic inhibition of NAc CINs can also increase the motivational influence of appetitive cues on instrumental actions<sup>18</sup>. However, the dynamic cellular mechanisms modulating NAc CIN function, which may underlie the observed behavioral diversity, remain poorly understood.

Electrophysiological recordings in caudate and putamen of awake non-human primates (NHPs) revealed early on that tonically active neurons (TANs)—broadly believed to be CINs—respond to reward-associated stimuli and reward outcomes with multiphasic alterations in their firing patterns<sup>19, 20</sup>. This multiphasic change in activity includes a brief decrease or pause in CIN firing that can be flanked by pre-pause activation and rebound excitation<sup>19, 20</sup>. The pause response, which occurs following presentation of a brief reward-predictive cue, has been recorded primarily in NHPs<sup>19–23</sup>. The pause is not homogeneous across behavioral tasks and striatal subregions and can be triggered by both aversive and salient stimuli<sup>24–27</sup>. Currently, it remains unclear what information is conveyed by the pause in different behavioral contexts<sup>28, 29</sup>. Moreover, relatively little is known about its behavioral significance in rodents<sup>30, 31</sup>. Because ACh is generally thought to inhibit SPN activity<sup>7, 10</sup> the pause may be important for selecting, invigorating or inhibiting actions in response to cues<sup>17, 18, 32</sup>. However, to our knowledge, so far no study has attempted to directly manipulate the endogenous pause during behavior to determine how it affects behavior.

The pause has received considerable attention as a possible reinforcement learning signal. The pause develops over the course of training in individual CINs and across CIN populations<sup>19, 21</sup>. It is maintained following long intermissions in training, and it is sensitive to behavioral extinction<sup>19, 21</sup>. Supporting a role in reinforcement learning is the observation that, under certain behavioral conditions, CIN pauses are associated with changes in DA neuron activity<sup>21, 23</sup>. For example, salient stimuli evoke pauses whose latency coincides with increased phasic activity in midbrain DA neurons<sup>23, 27, 33</sup>. In addition, both 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced lesions of striatal dopaminergic innervation and local DA D2 receptor (D2R) blockade abolish the

generation of the CIN pause in a Pavlovian conditioning task in NHPs<sup>21</sup>. This classical study suggested that DA is necessary for native pauses in CIN activity triggered by rewarding stimuli.

Despite this finding, the mechanistic origin of the pause is still debated. Slice physiology data support a role for DA, as pharmacological blockade of D2Rs eliminates CIN pauses induced by DA neuron optogenetic stimulation and by DA uncaging<sup>34–37</sup>. Pauses evoked by DA or electrical striatal stimulation are similarly abolished in brain slices from mice with CIN-specific deletion of D2Rs<sup>37, 38</sup>. Complicating matters, CINs also receive inputs from motor cortical and centromedial and parafascicular thalamic areas (CM-Pf). *In vivo* cortical and thalamic electrical stimulation induces multiphasic responses in CINs, including pausing<sup>39, 40</sup>. Like for the DA lesion reported above, CM-Pf inhibition in behaving NHPs suppresses the TAN pause in response to reward-associated stimuli<sup>41</sup>, suggesting that thalamic input is centrally involved in this physiological response. Thalamic afferent stimulation in striatal slices evokes a similar burst-pause response in CINs in which the pause is abolished by D2R blockade<sup>42</sup>. This evidence supports coordinated involvement of DA with other neurotransmitter systems in pause generation. However, recent work proposes that receding excitatory input to CINs, via activation of Kv7.2/7.3 channels, is the main trigger of the pause<sup>40</sup>. In this same study, computational modeling suggests that D2R activation has a minor, if any, role in regulating the pause *in vivo*, and that its effects are limited to the late phase of the pause<sup>40</sup>.

In the ventral striatum, optogenetic stimulation studies in brain slices and *in vivo* point to yet another possible origin of the pause: GABAergic neurons of the ventral tegmental area (VTA). Light-evoked stimulation of VTA GABA projections to the NAc induced a pause-rebound response in CINs, and facilitated the learning of an aversive stimulus-outcome association<sup>43</sup>. In this context, the VTA GABA-evoked pause in CINs was DA-independent.

Thus, the neural origins of the pause remain an open question. Much of its mechanistic interrogation has come either from slice physiology or *in vivo* optogenetics, yet both approaches are inherently limited to artificial neuronal stimulation. Therefore, the cellular and circuit mechanisms that influence the induction and maintenance of behaviorally evoked pauses remain to be determined. This is especially true in the ventral striatum, where the duration and magnitude of the CIN pause in response to rewarding stimuli has been shown to be more prominent than in dorsal regions<sup>26, 30</sup>. Furthermore, to understand how DA regulates CIN activity during behavior, the DA-dependent component of the pause needs to be selectively manipulated *in vivo*.

Because D2Rs have been shown to be important for pause generation in rodent brain slices and *in vivo* in NHPs<sup>21, 38</sup>, we decided to directly target D2Rs in CINs using a cell-selective viral-based strategy. Specifically, we upregulated D2Rs in NAc CINs of adult mice, with the hypothesis that this should enhance the DA-induced pause in CIN activity. We further postulated that by enhancing the CIN pause, D2R upregulation should prolong behaviorally evoked phasic decreases in ACh activity. Indeed, using slice physiology we found D2R upregulation in CINs results in a significant prolongation of the pause in response to DA terminal optogenetic stimulation without affecting baseline firing. Furthermore, *in vivo* fiber

photometric analysis of signals emitted by a genetically encoded ACh sensor revealed a pause-like decrease in NAc ACh levels following lever presentation during a continuous reinforcement (CRF) task. This “pause” developed over the course of several training days. D2R upregulation in CINs led to an earlier appearance of this pause and was associated with an increase in pause amplitude and duration. We then determined whether the prolonged reduction in ACh activity induced by D2R upregulation following a cue might facilitate associative learning as proposed by studies using artificial stimulation<sup>18, 43</sup>. Surprisingly, D2R upregulation did not facilitate or impair performance on various associative learning tasks. To address whether enlarged pauses contribute to learning to suppress actions, we further analyzed the mice in a Go/No-Go task. We found that D2R upregulation delayed learning to withhold a learned response under No-Go conditions. Moreover, in control mice, Go and No-Go trial cues evoked distinct ACh signals, but such distinctions were largely absent following CIN D2R upregulation. These findings suggest that DA signaling via D2Rs expressed in NAc CINs regulates cue-evoked ACh levels and shapes inhibitory learning.

## Materials and Methods

### Mice

Adult male and female ChAT-Cre mice (GM60, GENSAT) were generated by backcrossing onto C57BL/6J background. Double-transgenic were generated by crossing ChAT-Cre (GM60, GENSAT) to DAT-IRES-Cre<sup>44</sup> (JAX stock #006660) mice. Mice were housed 3-5 per cage for most experiments on a 12-hr light/dark cycle, and all experiments were conducted in the light cycle. All experimental procedures were conducted following NIH guidelines and were approved by Institutional Animal Care and Use Committees of the New York State Psychiatric Institute and Fordham University.

### Surgical procedures

Under ketamine-induced anesthesia, mice (8 weeks old) received bilateral infusions (440 nL/side) of Cre-dependent double-inverted open reading frame (DIO) adenoassociated viruses (AAVs) into the nucleus accumbens (NAc) using stereotactic Bregma-based coordinates: AP, +1.70 mm; ML,  $\pm$ 1.20 mm; DV, -4.1 mm (from dura). For electrophysiology or behavior experiments, these include: AAV2/1-hSyn-DIO-D2R(L)-IRES-mVenus<sup>45, 46</sup>, AAV2/9-EF1a-DIO-D2R(S)-P2A-EGFP (constructed in-house; packaged by Virovek), or AAV2/5-hSyn-DIO-EGFP (UNC Vector Core, Chapel Hill, NC). We infused AAV2/5-FLEX-ChR2-mCherry (UNC Vector Core, Chapel Hill, NC) into the VTA (440 nL/side) using the following coordinates: AP, -3.5 mm; ML  $\pm$  0.5 mm, DV, -4.3 mm (from dura). For fiber photometry experiments, mice were anesthetized with isoflurane and received a 1:1 mixture (375 nL/side) of AAV2/9-hSyn-ACh3.0<sup>47</sup> and AAV2/1-hSyn-DIO-D2R(L)-IRES-mCherry (constructed in-house; packaged by Vector Biolabs) or AAV2/5-DIO-mCherry (UNC Vector Core, Chapel Hill, NC) at AP, +1.70 mm; ML,  $\pm$ 1.20 mm; and three DV sites, -4.2, -4.1, -4.0 mm (from dura). For dLight experiments, AAV2/5-hSyn-dLight1.2 (Addgene) was used. Following virus injection, a 400- $\mu$ m fiber optic cannula (Doric, Quebec, Canada) was carefully lowered to a depth of -4.1 mm cannula and fixed in place to the skull with dental cement anchored to machine

mini-screws. Groups of mice used for experiments were first assigned their AAV-genotype in a counterbalanced fashion that accounted for sex, age, home cage origin.

## Histology

Mice were transcardially perfused with ice-cold 4% paraformaldehyde (Sigma, St. Louis, MO) in PBS under deep anesthesia. Brains were harvested, post-fixed overnight and washed in PBS. Free-floating 30- $\mu$ m coronal sections were cut using a Leica VT2000 vibratome (Richmond, VA). After incubation in blocking solution (10% fetal bovine serum, 0.5% bovine serum albumin in 0.5% TBS-Triton X-100) for 1h at room temperature, sections were labeled overnight at 4°C with primary antibodies against GFP (chicken; 1:1000; AB13970 Abcam, Cambridge, MA), ChAT (goat; 1:100; AB144P Millipore, Burlington, MA), DsRed (rabbit; 1:500, 632496 Takara), TH (mouse, 1:750, 22941 Immunostar, Hudson, WI). Sections were incubated with corresponding fluorescent secondary antibodies for 2h at RT. Sections were then mounted on slides and coverslipped with Vectashield containing DAPI (Vector, Burlingame, CA). Digital images were acquired using a Nikon epifluorescence microscope or Leica TSP8 laser scanning confocal microscope and processed with NIH Image J software and Adobe Photoshop.

## Fluorescent in situ hybridization (FISH)

Labeling of Chat and Drd2 mRNAs was performed via single molecule fluorescent ISH (smFISH). Brains from six ChAT-Cre mice (three injected with AAV5-hSyn-DIO-GFP and three with AAV1-DIO-D2-IRES-mVenus, at coordinates targeting the NAc, as described above) were rapidly harvested and snap-frozen in OCT by immersion in isopentane chilled on dry ice and then stored at  $-80^{\circ}\text{C}$  until use. Coronal sections (20- $\mu$ m) were collected directly onto Superfrost Plus slides (Fisherbrand). RNAscope Fluorescent Multiplex labeling kit (ACDBio Cat No. 320850) was used to perform the smFISH assay according to the manufacturer's recommendations. The probes used for staining were mm-Chat-C2 (ACDBio Cat No. 408731-C2) and mm-Drd2-C3 (ACDBio Cat No. 406501-C3). After incubation with fluorescent-labeled probes, slides were mounted with VECTASHIELD HardSet Antifade Mounting Medium with DAPI (Vector Labs, H-1500-10). Fluorescent images were captured using sequential laser scanning confocal microscopy (Leica SP8). Mean fluorescence intensities resulting from hybridization with Chat and Drd2 probes were both quantified within masks taken of individual cell bodies of ChAT+ ( $n = 20$  per brain) and nearby ChAT- Drd2-expressing neurons ( $n = 20$  per brain) of the ventral striatum using ImageJ software. Multiple means were compared using Kruskal-Wallis tests with multiple comparisons testing. Investigators were blinded to the genotype of samples during experimentation and analysis.

## Slice preparation and patch clamp recording

Four weeks after surgery, brains were harvested into ice-cold, oxygenated ACSF containing (in mM): 1.25  $\text{NaH}_2\text{PO}_4$ , 2.5 KCl, 10 glucose, 26.2  $\text{NaHCO}_3$ , 126 NaCl, 2  $\text{CaCl}_2$  and 2  $\text{MgCl}_2$  (pH 7.4, 300–310 mOsm). Coronal striatal slices (200  $\mu$ m) were cut on a vibratome in ice-cold, oxygenated ACSF and immediately incubated at  $32^{\circ}\text{C}$  for 30 min followed by 1h at room temperature prior to recording. GFP-positive CINs within the NAc core were identified under IR-DIC optics and epifluorescence microscopy. Voltage- and current-

clamp whole-cell recordings were performed using standard techniques at 30–32°C, using an internal solution consisting of (in mM): 140 K<sup>+</sup>-gluconate, 10 HEPES, 0.1 CaCl<sub>2</sub>, 2 MgCl<sub>2</sub>, 1 EGTA, 2 Mg<sup>+</sup>-ATP, and 0.1 Na<sup>+</sup>-GTP (pH 7.3, 280 mOsm). Electrodes were pulled from 1.5 mm borosilicate-glass pipettes on a P-97 puller (Sutter Instruments). Electrode resistance was ~ 3–6 MΩ when filled with internal solution. Recordings were obtained with a Multiclamp 700B amplifier, digitized at 10 kHz using a Digidata 1440A acquisition system with Clampex 10, and analyzed with pClamp 10 (Molecular Devices). Only cells that maintained a stable access resistance (< 20MΩ) throughout the entire recording were analyzed. Membrane properties were extrapolated from current–voltage relationships obtained by injecting 500 ms currents ranging from –140 and +40 pA currents in 20 pA steps. Voltage clamp recordings for I<sub>h</sub> determination were performed by applying hyperpolarizing steps (–60 to –150 mV) from a holding potential of –50 mV. I<sub>h</sub> was calculated as the difference between the “late” or steady-state current and the “early” or instantaneous current, as done by others<sup>48</sup>. The early current was determined by fitting an exponential function to the current response and finding the value of this curve at the onset of the pulse, while the late current’s value was extrapolated from the final value of the current response at the offset of the pulse<sup>49</sup>. Cell-attached recordings were conducted at 30–32°C using ACSF as internal solution. Following a 3-min period of gap-free recording, optogenetic burst stimulation was applied to activate ChR2-mCherry-expressing DA terminals, as previously reported<sup>34</sup>. Briefly, ChR2 responses were evoked using field illumination (470 nm, 2.3 mW) through a 40x objective with a PE-100 CoolLED illumination system delivered in a 20-Hz train of five 5-ms pulses across 10 trials, each separated by 30 s. The interspike interval (ISI) before the stimulus was used to determine baseline spike frequency (Hz) and the pause was measured as the 10-trial average of the first ISI following the stimulus<sup>38, 42</sup>. Peristimulus histograms were made from ten consecutive traces (0.1 s bin).

### Operant apparatus

Sixteen operant chambers (model Env-307w; Med-Associates, St. Albans, VT) equipped with liquid dippers were used. Each chamber was inside a light- and sound-attenuating cabinet. The experimental chamber interior (22 × 18 × 13 cm) had flooring consisting of metal rods placed 0.87 cm apart. A feeder trough was centered on one wall of the chamber. Head entries into the trough were recorded with an infrared photocell detector. Raising of the dipper inside the trough delivered a drop of evaporated milk. Two retractable levers were mounted on either side of the feeder trough, with LED lights above them. A house light located on wall opposite to trough illuminated the chamber throughout all sessions.

### In vivo fiber photometry

Fiber photometry equipment was set up using a 4-channel LED Driver (DC4104, ThorLabs) connected to both a 405-nm LED and a 465-nm LED (Thorlabs, cLED\_405 and cLED\_465). The 405-nm LED was passed through a 410–10 nm bandpass filter (Thorlabs, FB405-10), while the 465-nm LED was passed through a GFP excitation filter (Thorlabs, MF469-35). Both LEDs were then coupled to a 425-nm long pass dichroic mirror (Thorlabs, DMLP 425) and subsequently a GFP dichroic filter (Thorlabs, MD498). A low-autofluorescence patch cord (400 μm/0.48NA, Doric) was attached to the cannula on

the mouse's head and used to collect fluorescence emissions. These signals were filtered through a 525-39 GFP emission filter (MF525,39, Thorlabs) coupled to a tube lens with a wavelength range of 425-675 nm (Edmund Optics, #62-561-INK) and subsequently a photoreceiver (Newport, model 2151; gain set to *DC Low*). Signals were sinusoidally modulated, using Synapse® software and RX8 and RZ5P Multi I/O Processors (Tucker-Davis Technologies), at 210 Hz and 330 Hz (405nm and 465nm, respectively) to allow for low-pass filtering at 3 Hz via a lock-in amplification detector.

Cannula-implanted mice began behavioral training 6-7 weeks after surgery. Behavior tasks were conducted under food restriction (85-90% of basal body weight) and began dipper training to retrieve a milk reward as previously described<sup>45</sup>. In this session, 20 dipper presentations were separated by a variable inter-trial interval (ITI) and ended after 20 rewards were earned, or after 30 min, whichever occurred first. Mice reached criterion when head entries were made during 20 dipper presentations in one session. In the second training session, mice were habituated to fiber optic patch cord tethering, and criterion was reached when mice made at least 40 rewarded head entries in 60 minutes. This was followed by training to lever press using a CRF schedule. Each CRF trial began with extension of the lever, which when first pressed would lead to a 5-s dipper presentation. At the end of the 5 s, the dipper was lowered (“dipper off”) and the lever was simultaneously retracted, marking the end of the trial. A variable ITI (mean 42 s; 5 s minimum) was used. The first two days of CRF training consisted of 30 trials, ending when mice earned 30 reinforcements. While mice were not imaged on Day 1 because tethering the mice to the photometry equipment impaired initial acquisition of the CRF task, they were imaged starting on Day 2 and on subsequent CRF sessions (Days 3-7), which consisted of 60 trials.

All photometry and CRF data utilized custom in-house Python analysis scripts, unless stated otherwise. Photometry signals were analyzed as time-locked events aligned to the lever extension of each trial. The 405-nm channel was used to control for potential noise/movement artifacts and the 465-nm channel was used to detect the conformational modulation of the ACh3.0 sensor by ACh. Both demodulated signals were extracted as a 20-s window surrounding the event, which was denoted as time = 0,  $t_0$ . Both signals were downsampled by a factor of 10 using a moving window mean. The change in fluorescence,  $F/F$  (%), was defined as  $(F-F_0)/F_0 \times 100$ , where  $F$  represents the fluorescent signal (465 nm) at each time point.  $F_0$  was calculated by applying a least-squares linear fit to the 405 nm signal to align with the 465 nm signal<sup>50, 51</sup>. To normalize signals across animals and sessions, we calculated a single baseline fluorescence value for each trial using the average of the 5-s period preceding the event ( $t_{-5}$  to  $t_0$ ) and subtracted that from the signal. The daily average ACh3.0 traces were calculated using session average traces from individual mice. Peak and dip amplitudes were calculated by taking the maximum value between 0 to 1 s, or minimum value 0 to 2 s of the session average traces, respectively. Dip duration was calculated using  $t_b - t_a$ , where  $t_b$  was the timestamp at the zero-crossing following the dip, and  $t_a$  was the timestamp at the zero-crossing preceding the dip. If there was no zero-crossing following the dip within 5 s of the lever extension,  $t_b$  was set to 5 s, so that  $t_b$  5 s. If there was no zero-crossing preceding the dip after  $t=0$ , then  $t_a$  was set to the 0 s. In cases where no zero-crossings were found, dip duration was set to 0. The A.U.C values were restricted to a 0 to 5 s window. Single-trial  $F/F$  (%) traces were used for correlation

analysis. Average peak onset was determined by identifying the maximum peak of the day average traces and calculating the latency to the preceding local minimum.

### **Pavlovian conditioning and Pavlovian-to-instrumental transfer (PIT)**

Mice began behavioral training at least four weeks after AAV surgery. Mice were weighed daily and food-restricted to 85-90% of baseline weight; water was available ad libitum. Prior to beginning Pavlovian conditioning, mice underwent one session of dipper training as described above. We then used an appetitive conditioning protocol<sup>52</sup> in which mice received 25 presentations of a feeder light conditioned stimulus (CS) that was followed by a milk dipper (unconditioned stimulus, US) in each of 16 daily sessions. CS duration was fixed at 8 s with variable ITI (mean of 80 s). Head entries during the CS and during the last 8 s of the ITI prior to the CS were recorded.

We also used a general PIT protocol, adapted from Collins et al<sup>53</sup>, where mice received 7 days of Pavlovian training in which an auditory CS<sup>+</sup> (either a tone or white noise) was paired with a 20% sucrose liquid reward. The CS<sup>+</sup>, which lasted 2 min, was presented 6 times with a variable ITI (mean 5 min). Sucrose dippers were given on a random-time 30-s schedule and were raised for 5 s. This was followed by training to lever press in a CRF schedule, as above, with the exception that levers remained out once extended. The reward consisted of raising the dipper for 5 s, and the session ended when the mouse earned 30 reinforcers, or 30 min elapsed, whichever occurred first. Sessions were repeated until mice obtained 30 reinforcers. Mice then received 2-3 days each of random ratio 5 (RR5), RR10 and RR20 schedules in the absence of the CS<sup>+</sup>. After a Pavlovian “reminder” session, mice were given a session where no rewards were given and in which they were exposed to the CS that was not initially chosen as CS<sup>+</sup> (“CS<sup>0</sup>”). Following a 30-min session of lever press extinction, in which no CSs were presented and lever pressing was not rewarded, the following day mice underwent a PIT test. The PIT test began with an 8-min extinction period, where lever pressing was not rewarded. The CS<sup>+</sup> and the CS<sup>0</sup> were then presented four times each in the following order: (noise = n, tone = t: n-t-t-n-t-n-t). Each stimulus lasted 2 min followed by a 3-min fixed ITI, and no rewards were given.

### **Go/No-Go**

We used a symmetrical Go/No-Go paradigm in which both Go and No-Go cues predict reward but signal different behavioral responses<sup>54</sup>. The first phase of training consisted of 60 Go trials. The 60 Go trials were signaled by the presence of a house light and lever extension. Mice received a reward if they pressed the lever within 5 s of its extension. Mice were trained on 5 s go-only trials for 8 days. In the second phase, 30 Go trials were intermixed with 30 No-Go trials, and presented pseudorandomly to have an equal number of both trial types in every block of 10 trials. In No-Go trials, mice learned to withhold presses of the same lever when the house light was turned off and an LED light turned on above the lever being extended. A reward was given in No-Go trials if mice did not press the lever for 5 s. All failures to correctly respond in either trial type, would initiate a new trial (average 40 s ITI). Mice were run for 30 days, and the hits (% correct Go trials/total number of Go trials) and false alarms (% incorrect No-Go trials/total number of No-Go trials) were calculated. Mice that did not reach criteria of >50% correct performance on No-Go trials



in at least 5 days were excluded from the analysis. For fiber photometry during Go/No-Go task, mice first received dipper training as well as lever press training on CRF schedule as above. Once mice achieved 60 reinforcements in 60 trials, they began the Go/No-Go task. Following completion of Phase 1, fiber photometry signals were recorded every third day of Phase 2, for a total of 11 recordings sessions (31 days of training). On the two intervening days, mice continued training and were attached to dummy cables but not recorded.

## Data analysis

Sample sizes were determined by performing statistical power analyses based on effect sizes observed in preliminary data or on similar work in the literature. Statistical analyses were performed using GraphPad Prism 5.01 or 9 (GraphPad), SPSS 25 software (IBM), MATLAB (MathWorks), or Python (SciPy.Stats). Data are generally expressed as mean  $\pm$  standard error of the mean (SEM). Paired and unpaired two-tailed Student's t-tests were used to compare 2-group data, as appropriate. Multiple comparisons were evaluated by one-, two-, or three-way ANOVA and Bonferroni's *post hoc* test, when appropriate. Kruskal-Wallis tests were used when multiple comparison samples did not meet the normality assumption. In rare cases of values missing randomly in repeated measures samples, the data were analyzed by fitting a mixed effects model, as implemented by Prism 9. Photometry correlation analyses were performed using Pearson's correlation coefficients. A p-value of  $< 0.05$  was considered statistically significant. Behavioral and electrophysiological findings were replicated with mice from different litters, ages, or sexes. Investigators were blinded to the genotype of mice during behavioral assays as well as throughout the data analysis. Computer code for data analysis is available upon request.

## Results

### D2R upregulation in NAc CINs does not interfere with intrinsic excitability or basal firing.

To test the role of D2Rs in CIN physiology, we selectively targeted NAc CINs by bilaterally injecting Cre-dependent adeno-associated viruses (AAVs) expressing either D2Rs or EGFP (control) into the NAc of choline acetyltransferase (ChAT)-Cre mice (Fig. 1A). Throughout the study we used either of two double-floxed inverse orientation (DIO) AAVs to overexpress the long or short variant of D2R, both of which are robustly expressed in CINs<sup>55</sup>. We used a D2R-IRES-mVenus AAV, which encodes the long isoform of the D2R gene and the YFP variant mVenus separated by an internal ribosome entry site (IRES) for bicistronic expression. We have previously shown this vector to lead to a three-fold increase in D2R binding in NAc membranes when targeting SPNs with D2-Cre mice<sup>45, 46, 56</sup>. Here, using single molecule fluorescent *in situ* hybridization (sm-FISH), we found a three-fold increase in *Drd2* mRNA in D2R-OE<sub>NAcChAT</sub> mice compared to ChAT+ neurons in EGFP<sub>NAcChAT</sub> mice that was selective to ChAT+ neurons and not observed in SPNs (Supplementary Fig. S1). We further generated an AAV encoding the short form of the D2R gene followed by a P2A linker sequence and EGFP (D2R-P2A-EGFP). Four weeks following AAV infusion into ChAT-Cre NAc core, either D2-P2A-EGFP or D2-IRES-mVenus were selectively expressed in large, spindle-shaped neurons with sparsely branched dendrites, typical of CIN morphology<sup>1</sup> (Fig. 1B–D). We confirmed the cholinergic identity of these neurons by co-immunolabeling with antibodies against ChAT (Fig. 1C–

D). Quantification of viral expression 5 months after AAV infusion showed that a high proportion of NAc ChAT-positive neurons expressed the D2-P2A-EGFP (80.96  $\pm$  2.545%, n = 5 mice) and the EGFP control vectors (81.94  $\pm$  2.887%, n = 7 mice). Moreover, D2R AAV expression was strongly enriched in the NAc core compared to NAc shell and anterior CPu (Supplementary Fig. S2).

We first sought to determine whether CIN-selective D2R upregulation altered intrinsic CIN membrane properties in adult brain slices. We performed whole-cell recordings from fluorescent CINs in the NAc core expressing either EGFP or D2-IRES-mVenus (Fig. 1E). Current clamp recordings showed typical CIN physiological responses to current injections<sup>57</sup>. As reported by others, depolarizing current injection led to regular, non-adaptive firing, whereas negative current injection produced an initial hyperpolarization followed by a depolarizing sag in membrane potential<sup>57</sup> (Fig. 1F). We found no significant alterations in resting membrane potential, input resistance or action potential threshold following CIN D2R upregulation (Fig. 1G–I).

The hyperpolarization-activated cation current  $I_h$ , which is prominent in CINs, contributes to this depolarizing sag<sup>49</sup> and has been shown to be sensitive to DA and D2R agonists<sup>48</sup>. Therefore, we measured  $I_h$  by holding the membrane potential at  $-50$  mV and using a series of hyperpolarizing commands to evoke this time- and voltage-dependent inward current (Fig. 1J). However, as shown by the current-voltage plots in Fig. 1K and Supplementary Fig. S3, D2R upregulation did not alter  $I_h$  amplitude at baseline or after quinpirole (10  $\mu$ M). In addition, cell-attached recordings revealed spontaneous firing activity in CINs expressing either EGFP or D2R (Fig. 1L). However, D2R upregulation did not affect firing rates (Fig. 1M).

### D2R upregulation in NAc CINs increases pause duration in slices.

Several studies in *ex vivo* slices have shown that bath application of the D2R antagonist sulpiride attenuates or eliminates the CIN pause in firing induced by DA in dorsal and ventral striatal regions<sup>34–36</sup>. Therefore, we sought to determine whether selective upregulation of D2Rs in CINs of the NAc core alters DA-evoked pausing. To this end, we first generated a double-transgenic mouse line (ChAT-Cre x DAT-IRES-Cre) that would enable expression of channelrhodopsin-2-mCherry (ChR2-mCherry) in midbrain DA neurons and overexpression of D2Rs in NAc CINs (Fig. 2A). Four weeks after viral infusions, we observed robust ChR2-mCherry expression in tyrosine hydroxylase (TH)-positive somas within the VTA and substantia nigra (SN) (Fig. 2B, i–ii). Importantly, we also observed widespread ChR2-mCherry expression in afferent fibers surrounding D2R or EGFP-expressing CINs in NAc (Fig. 2B, iii). No ChR2-positive terminals were observed in NAc of ChAT-Cre mice injected in VTA, confirming that ChR2+ terminals do not arise from possible midbrain cholinergic neurons in the double transgenic mouse (Supplementary Fig. S4). We stimulated these ChR2-positive terminals to elicit DA-evoked CIN pauses in NAc slices using an optogenetic strategy like that employed by Chuhma et al<sup>34</sup>. Specifically, we applied train photostimulation to DA afferents while recording from fluorescent CINs of the NAc core. This stimulation protocol (5 pulses at 20 Hz) has been used previously to simulate the DA neuron burst firing associated with reward-related stimuli<sup>34</sup>. Given

published work<sup>34</sup>, we expected train photostimulation to lead to a reduction in tonic firing in control NAc core CINs. In addition, because D2R activation in other neurons, such as DA neurons, leads to long-lasting hyperpolarization via  $G\alpha_i$ -mediated mechanisms<sup>58, 59</sup>, we hypothesized that D2R upregulation in NAc CINs would result in prolonged DA-elicited pauses. As expected, EGFP-expressing CINs showed a consistent pause or reduction in tonic firing, defined as the first ISI following photostimulation<sup>38, 42</sup> (Fig. 2C). Compared to EGFP, expression of both D2-IRES-mVenus and D2-P2A-EGFP resulted in a significantly increased average pause duration (Fig. 2D and Supplementary Fig. S5). This pause elongation was not associated with changes in the average ISI before stimulation, suggesting a specific role for CIN D2Rs in regulating the DA-evoked pause (Fig. 2E). These effects on the pause were also observed in peristimulus histograms showing average firing from all cells recorded (Fig. 2G). In addition, the pause elongation was reversed in CINs treated with sulpiride (10  $\mu$ M) prior to photostimulation (Fig. 2F, G). These results suggest that increased expression of D2Rs, either the short or long isoforms, in NAc core CINs results in a robust and consistent increase in DA-induced pause duration, without altering basal firing.

### D2R upregulation in NAc CINs alters ACh levels during reinforcement learning

Next, we sought to determine whether CIN D2R upregulation would lead to alterations in CIN function *in vivo*. To this end, we turned to fiber photometry and measured bulk NAc acetylcholine (ACh) levels. We used an optimized genetically-encoded GPCR-Activation Based ACh sensor (GRAB<sub>ACh3.0</sub> or ACh3.0)<sup>47</sup>. ACh3.0 generates a sensitive fluorescence signal when activated by physiological ACh levels in mouse brain<sup>47</sup>. To determine whether D2R upregulation altered ACh-related signals, we co-infused ACh3.0 with either AAV-DIO-mCherry or AAV-DIO-D2R-IRES-mCherry and implanted an optic fiber into the NAc core (Fig. 3A). Since D2R are selectively expressed in CINs but ACh3.0 expression is not Cre-dependent, we did not expect differences in sensor expression, which we verified with immunofluorescence (Supplementary Fig. S6). The mCherry-expressing constructs were generated to avoid potential interference between our GFP/YFP-based D2R constructs and the similar excitation/emission spectra of ACh3.0. ACh3.0 signals were obtained using 465-nm LED excitation through the implanted optic fiber. Signal traces obtained using 405-nm channel were subtracted from the 465-nm signal traces to minimize movement-related artifacts<sup>50, 51</sup> (Supplementary Fig. S7).

Mice were trained on a CRF schedule over 7 daily sessions. Mice were trained on Day 1 without tethering. ACh3.0 signals were recorded over the next 6 consecutive daily sessions (Days 2-7). Each of 30 or 60 CRF trials in a session began with extension of a lever, which would yield a reward when first pressed. With training, lever extension becomes a reward-predicting cue that leads to NAc DA release<sup>60</sup>. This was confirmed in this task using the genetically encoded DA sensor dLight1.2<sup>61</sup>, which showed a training-dependent increase in signals aligned to lever extension (Fig. 3B). The mean latency to press the lever upon its extension was not different between the two groups, suggesting that D2R upregulation does not alter responsiveness to the lever (Fig. 3C).

Previous electrophysiology findings have shown that CINs respond with pausing to reward predicting cues<sup>19–23</sup>. We therefore aligned the ACh3.0 signals to lever extension to determine whether it induced a reduction in ACh levels and whether this reduction is prolonged by D2R upregulation. Fig. 3D shows the average fluorescence signals for ACh3.0 when aligned to lever extension for training Days 2–7. F/F traces were baselined to the average of the 5 s preceding lever extension. Thus, the resulting fluorescent signal reflected task-evoked changes in baseline ACh levels, normalizing for variable baselines across animals and sessions.

As can be observed in Fig. 3D, both groups initially responded to lever extension with a brief increase in ACh levels [mean onset: mCherry, 0.15 s (0.076 – 0.24 s range); D2-ires-mCherry, 0.15 s (0.007 – 0.23 s range)]. With daily training, the amplitude of this ACh “peak” decreased in both groups. However, we found a significant reduction in peak amplitude in D2R-OE<sub>NacChAT</sub> mice (Fig. 3E). With daily training we also found that the peak was followed by a sustained “dip” below baseline ACh, reminiscent of the CIN pause. In both groups, the dip amplitude increased across days, yet D2R-OE<sub>NacChAT</sub> mice showed a significantly larger dip than controls that was already present by Day 2 (Fig. 3F).

To further examine the magnitude of the ACh3.0 signals evoked by lever extension, we measured dip duration, as well as the area under the curve (A.U.C.) above and below baseline, up to 5 seconds after lever extension (Fig. 3G–I). D2R upregulation was associated with a smaller positive A.U.C., a larger (more negative) negative A.U.C and a trend towards a longer dip duration ( $p = 0.053$ ). Together these results suggest that D2R upregulation biased the response to lever extension towards larger reductions in ACh levels.

To determine to what extent the alterations in ACh3.0 signals are due to D2R activation during the task, we treated the same mice with the D2R antagonist haloperidol (0.25 mg/kg i.p.). Following a break of 1–2 days, mice were imaged again for 3 consecutive days after receiving a vehicle injection (Veh 1 day), haloperidol (Hal day) and a second vehicle injection (Veh 2 day). As expected, haloperidol increased the press latency in both groups, but had a more pronounced effect on press latency after CIN D2R upregulation (Supplementary Fig. S8A). Figures Supplementary Fig. S8B, C show the lever extension-aligned ACh3.0 signals after the Veh 1, Hal and Veh 2 days. Peak amplitude, but not dip amplitude, was reduced by D2R upregulation and by haloperidol but there was no significant virus x treatment interaction (Supplementary Fig. S8D, E). Haloperidol significantly reduced dip duration and negative A.U.C in both groups (Supplementary Fig. S8F–H). Although haloperidol blocks D2Rs in all D2R-expressing cells, these findings suggest that ongoing D2R activation contributes to peak amplitude and to dip duration.

We also sought to determine whether ACh3.0 signals measured in response to lever extension correlated with various task-related events. Correlation plots for Days 2, 4, and 7 show the trial-by-trial relationship between ACh3.0 signals and task-related features such as press latency, head entries (while lever was available, during reward presentation or during ITIs), presses per trial, and the preceding ITI duration (Supplementary Fig. S9). We observed a weak, but significant negative correlation between press latency and either dip

amplitude or negative AUC, suggesting that a larger ACh dip is associated with an earlier lever press.

### **D2R upregulation in NAc CINs does not alter Pavlovian conditioning or the motivational influence of Pavlovian cues.**

The pause in CINs has been suggested to be important for learning of cue-reward associations<sup>19, 21, 28, 29, 43</sup>, yet whether the pause in NAc CINs plays a causal role in associative conditioning is unknown. The CIN pause has been hypothesized to reduce nicotinic receptor modulation of DA release and thereby to provide a permissive window for dopaminergic firing activity to shape learning<sup>12, 62</sup>. Therefore, given our findings that D2R upregulation lengthens the pause in NAc CIN firing *in vivo*, we hypothesized that additional D2Rs in NAc CINs would result in enhanced associative learning. To test this hypothesis, we trained mice expressing either EGFP or D2R-P2A-EGFP on a 16-session protocol of appetitive conditioning involving 25 presentations of an 8-s feeder light followed by a milk reward<sup>52</sup> (Supplementary Fig. S10A). We measured anticipatory head entry responses occurring during this CS. As both control mice and D2R-OE<sub>NAcChAT</sub> mice progressed through the sessions, the rate of responding during the CS increased and then became stable (Supplementary Fig. S10B). Responding during the preceding ITI, on the other hand, decreased over the sessions but did not differ between groups. Similar results were obtained when Pavlovian responding was expressed as a difference score by subtracting pre-CS ITI responding from the CS responding (Supplementary Fig. S10C), where responding significantly increased over sessions but was not different between groups. These results, therefore, suggest that D2R-OE<sub>NAcChAT</sub> mice learn this simple Pavlovian association, and that the level of overall responding to predictive cues is not changed.

In addition to predicting whether a reward will occur, a fixed duration CS enables animals to learn when a reward will occur<sup>52</sup>. Consistent with this, the latency to the first head entry increased with training but was not affected by D2R upregulation (Supplementary Fig. S10D). To gain a more accurate indication of the timing of conditioned responding during the 8-s CS, we analyzed the effect of D2R upregulation on head entry rates in each of the four quartiles of the CS across training session blocks (Supplementary Fig. S10E). We not only found a significant increase in responding across session blocks, but also increased responding throughout the CS, indicative of temporal control. Comparing response rates for both AAV groups over all CS quartiles did not yield a statistically significant AAV x quartile interaction ( $F_{(3,20)} = 0.412$ ,  $p = 0.20$ ).

Pavlovian cues can also invigorate instrumental responding for a reward, a process that was recently shown to be enhanced by inhibition of NAc core CIN activity<sup>18</sup>. Therefore, we tested whether D2R upregulation would lead to enhanced cue-motivated behavior as measured in a classical Pavlovian-to-instrumental transfer (PIT) task<sup>53</sup> (Supplementary Fig. S10F). In this task, mice expressing either EGFP or D2R AAVs first underwent a 7-day Pavlovian training phase involving presentation of a 2-min auditory stimulus (CS<sup>+</sup>) during which they were given a milk reward. The mice were also given one session in which a different, neutral CS was presented without reward delivery (CS<sup>0</sup>). Following Pavlovian training, the mice learned to press a lever to obtain the same milk reward

without CS presentations (instrumental phase). In the final transfer phase, lever press rates were measured following pseudorandom exposure to the CS<sup>+</sup> and CS<sup>0</sup> in the absence of reinforcement. Higher lever press rates during the CS<sup>+</sup> compared to CS<sup>0</sup> or the ITI reflect cue-induced invigoration of responding.

We found no impact of D2R upregulation on Pavlovian responding to CS<sup>+</sup> or ITI presentation (Supplementary Fig. S10G) or on instrumental responding on a random ratio schedule (Supplementary Fig. S10H). During the transfer phase, we found a significant increase in lever press rate during CS<sup>+</sup> compared to CS<sup>0</sup> and ITI, suggesting that PIT was successfully expressed (Supplementary Fig. S10I). However, D2R-OE<sub>NacChAT</sub> mice showed similar patterns of responding when compared to EGFP controls. Furthermore, PIT was not significantly altered in mice lacking CIN D2Rs (CIN-D2RKO) compared to controls (Supplementary Fig. S11). These results indicate that CIN D2R upregulation nor downregulation affect cue-induced invigoration of responding for a food reward.

### **D2R upregulation in NAc CINs impairs No-Go responding.**

Striatal ACh regulates the activity of SPNs, which are important for action selection and movement initiation<sup>63</sup>. The pause in CIN activity has been shown to enhance SPN activity<sup>7</sup> (but see Zucca et al<sup>9</sup>). We therefore hypothesized that pause enhancement, as seen in D2R-OE<sub>NacChAT</sub> mice, would impair the ability to suppress responding to obtain reward. To address this hypothesis, we used a Go/No-Go task which measures an animal's ability to withhold from responding and has been shown to elicit phasic DA release in the NAc core<sup>64</sup>. As shown in Fig. 4A, mice were first trained to press a lever within 5 s to obtain a reward, if lever extension occurred in the context of house light illumination. Training over 7 days (60 Go trials/day) improved the performance of mice in both groups to a similar degree (Fig. 4B). Following this Go-only phase, mice were then trained in sessions containing 30 Go-trials and 30 No-Go trials, randomly presented. While Go trials were the same as in the Go-only phase, No-Go trials were signaled by the simultaneous presentation of the lever with two cues (the house light turning off and LED lights above the lever turning on) (Fig. 4C). In No-Go trials, mice were required to withhold from pressing the lever for 5 s to obtain a reward. As seen in Fig. 4D, accuracy on Go trials continued to be unaltered by D2R upregulation during Go/No-Go training. We then analyzed the percent incorrect responses or false alarm rates during No-Go trials (Fig. 4E). Both groups exhibited similarly high false alarm rates early on in training and improved their performance over the 30 training days. However, we found that D2R upregulation significantly delayed the reduction in false alarm rates. Overall, these findings suggest that enhancing D2R levels and the CIN pause specifically impairs the learning to restrain actions, without affecting Go responding.

### **CIN D2R upregulation reduces the ACh signaling contrast between Go and No-Go cues.**

To determine whether the observed behavioral effects of CIN D2R upregulation were associated with alterations in NAc ACh dynamics, we used fiber photometric analysis during the Go/No-Go task. ACh3.0 signals were recorded in D2R-OE<sub>NacChAT</sub> mice and mCherry-expressing mice every third day over 31 daily sessions. Signals were aligned to the start of Go and No-Go trials, enabling measurements of phasic ACh activity in response to the different trial type cues. ACh activity following Go trial onset featured an initial peak,

followed by a dip and a post-dip rebound (Fig. 5). D2R upregulation resulted in increased dip amplitude and duration compared to control, particularly with further training, but did not alter peak amplitude or rebound (Supplementary Fig. S12). ACh responses after No-Go onset also featured a peak, dip and a rebound (Fig. 5), none of which were altered by D2R upregulation (Supplementary Fig. S12).

We then measured within-group effects of Go and No-Go onset on ACh responses. We hypothesized that if the ACh signaling contributes to differentiating between Go and No-Go trials, then the onset of these different trial types should generate contrasting ACh responses. If so, we further predicted that such contrasts in ACh signaling should be less pronounced following D2R upregulation. Indeed, control mice showed significantly greater peak and dip amplitudes in No-Go trials compared to Go trials (Fig. 5C, D). We found no trial type effect on either peak or dip amplitudes in D2R-OE<sub>NacChAT</sub> mice (Fig. 5F, G). This evidence suggests that D2R upregulation reduces the contrast in ACh peak and dip amplitude between trial types. Both control and D2R-OE<sub>NacChAT</sub> mice exhibited ACh rebounds that were significantly larger in No-Go trials compared to Go trials (Fig. 5H), suggesting that this measure is likely not related to the behavioral effects of D2R upregulation.

We then determined whether the different overall Go and No-Go ACh responses seen primarily in control mice in Fig. 5 were influenced by lever presses. We hypothesized that if lever pressing dominates the divergence between Go and No-Go trials, then ACh responses in Go-Incorrect and No-Go-Correct trials — where lever pressing is absent — should be similar. However, control mice still exhibited significantly larger peaks, dips, and rebounds in No-Go-Correct relative to Go-Incorrect (Supplementary Fig. S13), an effect that was not observed in D2R-OE<sub>NacChAT</sub> mice. These results are consistent with the notion that CIN D2R upregulation diminishes the contrast between ACh responses to Go and No-Go instructional cues, irrespective of lever press execution.

We next sought to determine whether ACh responses tracked the accuracy of Go and No-Go responding. In control mice, dip amplitude was significantly larger in Go-Correct relative to Go-Incorrect trials (Supplementary Fig. S14A–C). Similarly, in No-Go trials, control mice showed a larger ACh rebound A.U.C. in No-Go-Incorrect relative to No-Go-Correct trials (Supplementary Fig. S14D). Interestingly, these differences were not observed in D2R-OE<sub>NacChAT</sub> mice (Supplementary Fig. S14E–H), suggesting that D2R upregulation is associated with more equivalent ACh responses in correct and incorrect trials regardless of trial type. No other variables diverged significantly based on performance within each trial type in either group.

## Discussion

We have found that selective D2R upregulation in CINs lengthens the CIN pause evoked by DA terminal stimulation in NAc slices without altering basal CIN spiking or membrane properties. Moreover, we present *in vivo* evidence of multiphasic NAc ACh responses to a predictive cue during reinforcement learning, including a cue-evoked rise followed by a sustained decrease in ACh reminiscent of the CIN pause. D2R upregulation altered these responses, dampening the rise while enlarging the pause-like dip in ACh levels. This

manipulation did not alter the learning of Pavlovian cues or their motivational influence. Rather, we found that D2R upregulation in NAc CINs was associated with a delay in learning to inhibit responding in a Go/No-Go task. Analysis of ACh dynamics during this task revealed that Go and No-Go cues elicited distinct ACh responses, yet such distinctions were diminished by D2R upregulation. Our data suggest that D2Rs in NAc CINs regulate cue-evoked ACh levels and inhibitory learning.

The pause elongation observed in NAc CINs in our slice recordings following D2R upregulation is consistent with the reverse effect on CIN pausing previously reported with bath-applied D2R antagonists<sup>34, 35, 42</sup>. The effect is also in line with recent slice physiology studies showing that dorsal striatal CINs from mice lacking D2Rs in ChAT-expressing cells lack pausing but show no gross alterations in CIN firing<sup>37, 38</sup>. However, a lack of D2Rs since early in development could give rise to a wide range of unidentified adaptations. Therefore, our genetically targeted approach in adult NAc core provides evidence that increasing D2Rs in adult CINs is sufficient to enhance the pause in CIN firing evoked by phasic DA.

Reward-predicting stimuli are known to induce a pause in CIN and TAN activity in rodents and NHPs. The degree to which this response is regulated by DA has long been debated<sup>21, 23, 27</sup> (see Zhang and Cragg for review<sup>29</sup>). In order to measure the effect of D2R upregulation on behaviorally induced changes in ACh, we took advantage of the recent development of genetically-encoded neurotransmitter sensors with sub-second resolution<sup>47, 61</sup>.

In the CRF schedule, we detected dynamic multiphasic responses in ACh levels associated with lever presentation at trial onset (Fig. 3). The first signature was a brief peak in ACh above baseline that progressively decreased in size with daily training. The cholinergic peak most likely reflects cortical or thalamic excitation of CINs triggered by the trial onset cues<sup>39, 41</sup>. As discussed above, CM/Pf thalamic function is particularly critical for the cue-driven pause and the rebound, but perhaps less so for the initial excitation<sup>41</sup>. Furthermore, thalamic afferent stimulation in striatal slices evokes a burst-pause CIN response where only the pause is blocked by D2R antagonism<sup>42</sup>. Therefore, we were surprised to find that D2R upregulation altered peak amplitude. One possibility is that D2R upregulation in CINs elicits postsynaptic alterations that affect summation of excitatory inputs and/or ACh release. One candidate target is the N-type Ca<sup>2+</sup> current, which is a key contributor to both of these functions in CINs and is rapidly inactivated by D2R agonists<sup>55</sup>.

In addition to the initial ACh peak, we found that lever presentation evoked a subsequent dip in ACh below baseline levels, which lasted for up to several seconds. This is significantly longer than *in vivo* electrophysiological reports where the reward-related pause in tonic firing is typically in the hundreds of milliseconds range<sup>19–23, 27</sup>. Our results imply that cue-evoked reductions in ACh may persist beyond the resumption of CIN activity. Such an effect could be supported by the rapid and highly efficient clearing of ACh from the synaptic cleft by acetylcholinesterase<sup>9</sup>. In both groups, the dip amplitude increased over days, which could reflect increased synchronization of the NAc CIN population with training. A training-related increase in the number of neurons that pause has been previously recorded in NHPs<sup>19, 21</sup>. In D2R-OE<sub>NAcChAT</sub> mice, the dip was observed earlier in CRF



training and was also of larger amplitude and duration than in controls. The larger dip following D2R upregulation is possibly due to the enhanced inhibition of CINs by DA that we measured in the slice following optogenetic stimulation of DA terminals. Treatment with a D2R antagonist shortened pause duration, as expected due to blockade of CIN D2Rs, but surprisingly also reduced the initial ACh peak. This counterintuitive finding may be due to the action of systemic haloperidol treatment on D2Rs on other cell types, which somehow decrease excitatory inputs to the striatum.

Despite reports showing that reward-related or salient sensory stimuli can induce a pause in CINs, or that artificially induced pauses can alter behavior, there is still no causal evidence for a behavioral role of the native cue-evoked pause. Cholinergic pausing has been suggested to increase with repeated associative training,<sup>19, 21, 65</sup> and recent work has shown that silencing NAc CINs during the transfer phase of a PIT task enhances cue-driven invigoration of instrumental responding<sup>18</sup>. Because D2R-OE<sub>NAcChAT</sub> mice responded to cue presentation with a larger NAc ACh dip, we anticipated that these mice would exhibit enhanced performance in tasks involving Pavlovian cues. However, we did not observe changes in PIT following D2R upregulation, suggesting that enlargement of the native pause is not sufficient to alter cue-motivated behavior. While it is possible that the lack of a PIT enhancement following D2R upregulation is due to a behavioral ceiling effect, our data from CIN-D2RKO mice show no decrease in PIT. We also found that neither acquisition nor maintenance of the conditioned approach in a Pavlovian task was affected by D2R upregulation. Thus, while NAc core DA transmission has been implicated in cue-reward learning<sup>66, 67</sup>, our data indicate that CIN D2Rs in this region do not appear to be critical mediators of Pavlovian associations. This is consistent with a recent observation that ventral striatal CIN lesions do not alter initial learning of task contingencies but impair responding when novel contingencies are introduced (see below)<sup>17</sup>.

In the Go/No-Go task, D2R-OE<sub>NAcChAT</sub> mice exhibited accurate responding to the Go cue, like controls. In contrast, acquisition of the No-Go response was delayed. This result could be consistent with enhanced impulsive-like behavior. While little is known about the role of D2Rs in this specific task, reduced — but not enhanced — NAc D2R expression has been associated with higher trait impulsivity in rats<sup>68</sup>. Because of the widespread expression of D2Rs in NAc, however, it is unclear which D2Rs population(s) are involved. In contrast to our findings here, our recent work has shown that D2R upregulation in NAc D2R-expressing SPNs does not alter No-Go performance<sup>69</sup>, suggesting that No-Go learning is more sensitive to alterations in D2R levels in CINs.

D2R-OE<sub>NAcChAT</sub> mice eventually performed as well as controls in No-Go trials, arguing against a general increase in impulsive action. Alternatively, the effect of CIN D2R upregulation could be linked to deficits in behavioral flexibility. Manipulations of CINs or ACh in the striatum do not affect initial learning, but instead impact learning in conditions where animals must adapt their behavior to new task rules. In the dorsomedial striatum this has been shown for place and instrumental reversal learning<sup>32, 70–72</sup>. In the ventral striatum, a selective lesion of CINs increased perseverative errors when a visual stimulus was introduced as a new directional cue<sup>17</sup>. Our Go/No-Go task incorporates similar changes in contingencies in that a novel light above the lever indicates the new rule (not to press).

Therefore, the deficit in the Go/No-Go task may arise from a delay in acquiring the new task contingencies when the predicting cue is novel.

How can a larger pause lead to a specific deficit in adaptive learning? CINs are thought to inhibit SPNs via nicotinic activation of local interneurons or via muscarinic  $M_2/M_4$ -mediated inhibition of corticostriatal inputs<sup>7, 10, 73, 74</sup>. A larger pause may, therefore, lead to a more pronounced disinhibition of SPNs, which would favor activity-dependent plasticity of corticostriatal synapses supporting the currently prevailing action selection.

Such a model has been proposed by Franklin and Frank<sup>75</sup> and tested using a neuronal network model. Strikingly, when the authors varied the pause duration in their model, this affected reversal learning. Shorter pauses allowed for a faster reversal in a probabilistic reversal learning task than larger pauses. This finding is consistent with our data in D2R- $OE_{NacChAT}$  mice, where a longer pause is associated with a delay in switching strategies between the Go and No-Go trials. Note, however, that the model used a probabilistic reversal learning task and therefore it will need to be formally tested using the same task.

Our *in vivo* ACh activity findings during the Go/No-Go task suggest that besides increasing pause size, D2Rs may have a more nuanced contribution to the observed behavioral alterations. In control mice, we found distinctive patterns of cue-evoked phasic ACh signaling in Go versus No-Go trials, including larger peaks, dips, and post-dip rebounds in No-Go conditions. Intriguingly, D2R- $OE_{NacChAT}$  mice show comparable peak and dip amplitudes after onset of both trial types. Therefore, it is conceivable that a marked ACh signaling differential between Go and No-Go trials facilitates encoding of specific cue information. In addition, these data suggest that *reduced* contrast in ACh peaks and dips, as seen following D2R upregulation, is associated with deficits in learning to appropriately suppress responding.

How D2R upregulation leads to a reduction in ACh signaling differential (i.e. peak and dip size) in this task remains to be resolved. We speculate that increased CIN D2R function may limit new plasticity of excitatory and inhibitory inputs onto CINs when No-Go trials are introduced. This could delay the encoding and/or updating of novel No-Go cues as different from the Go cues. Similar deficits in new goal-directed learning have been seen following  $M_2/M_4R$  activation in dorsomedial striatum<sup>32</sup>. Alternatively, D2R upregulation could lead to comparable degrees of CIN synchronization, regardless of cue type, effectively equalizing and maintaining similar peak and dip responses to either Go or No-Go cues.

We found that ACh rebound levels were greater in No-Go vs. Go trials. We originally hypothesized that if the ACh rebound is related to action suppression in No-Go trials, then No-Go trials in which animals correctly withheld lever pressing would be associated with larger rebounds than No-Go trials with premature pressing. Surprisingly, we observed the opposite; the rebound was significantly greater in incorrect No-Go (press) relative to correct No-Go (withhold) trials in control mice, with a similar trend in D2R- $OE_{NacChAT}$  mice ( $p = 0.07$ ). This may suggest that larger rebounds are not related to action suppression but could potentially provide feedback about action errors in No-Go conditions.

In conclusion, we have shown that D2Rs in NAc CINs regulate the stimulus-evoked multiphasic ACh response during reinforced behaviors. Most notably, we have shown that enhancement of the native pause response, as well as reduced contrast in ACh responses to different predictive cues, are associated with a delay in learning to suppress a previously learned response to obtain the same reward. Abnormalities in striatal DA and ACh have been observed in Parkinson's disease and in neuropsychiatric disorders like schizophrenia and ADHD, where cognitive deficits and behavioral inflexibility are core symptoms. Thus, further dissection of the complex interactions between these neurotransmitter systems will not only provide a better mechanistic understanding of reward-related flexible learning in these disorders but will also shed light on improved treatment strategies.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements:

We would like to thank Dr. Lin Tian for providing dLight1.2, and Christine Lim, Julianna Cavallaro, Daphne Baker and Natalie Zarrelli for assistance with histology. Also, many thanks to Dr. Nao Chuhma and Dr. Steven Rayport for advice on slice physiology, and Joseph Floeder, Dr. Marie Labouesse and Dr. Mark Ansorge for assistance with fiber photometry.

## Funding:

C.K., J.G., K.M.M., J.A.J., and P.D.B. were supported by R01 MH093672 and R01 MH124858-01A1. E.F.G, J.Y. and E.T. were supported by K01 MH107648 and a Faculty Research Grant (Fordham University). J.M.V was supported by the Leon Levy Fellowship in Neuroscience.

## References:

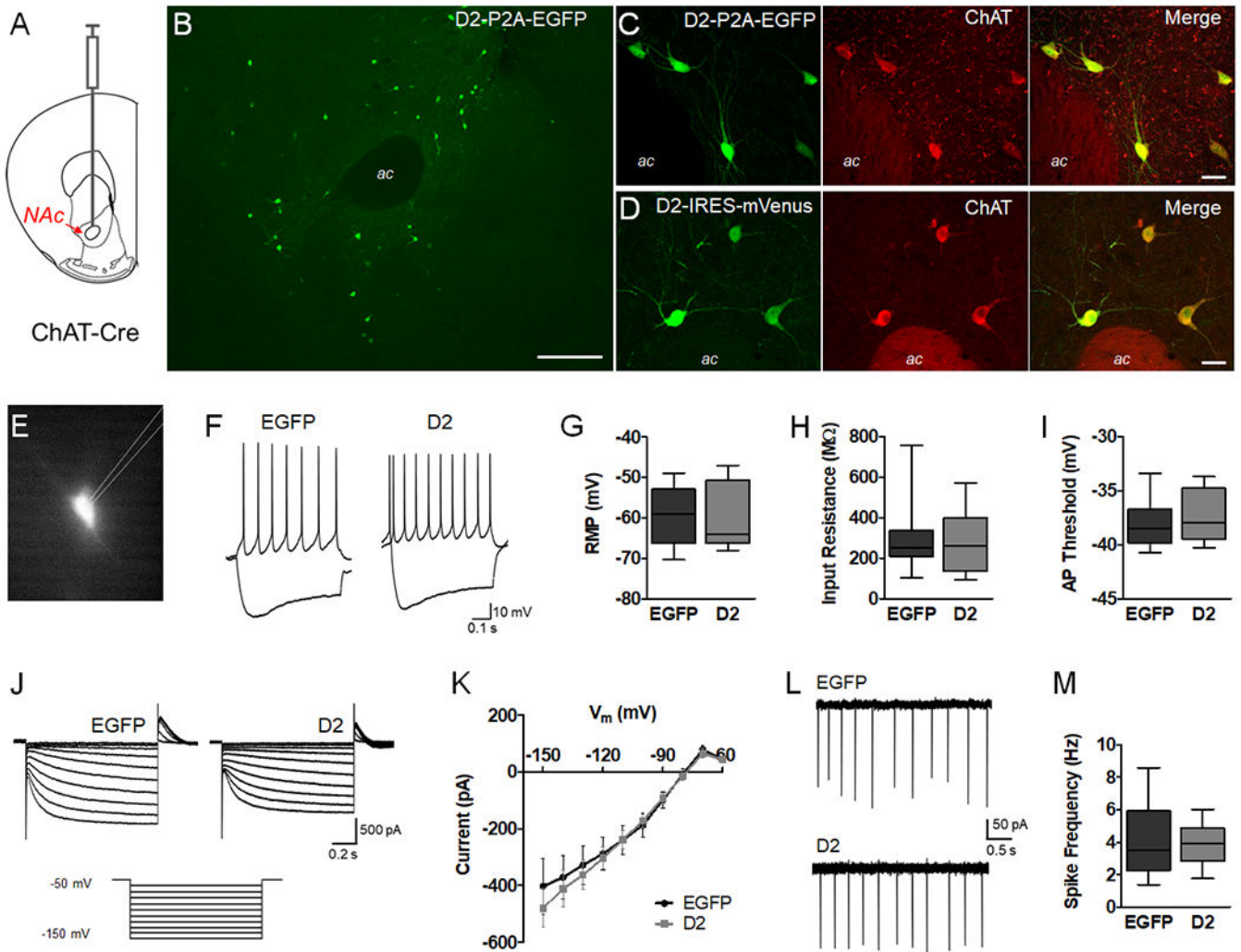
1. Bolam JP, Wainer BH, Smith AD. Characterization of cholinergic neurons in the rat neostriatum. A combination of choline acetyltransferase immunocytochemistry, Golgi-impregnation and electron microscopy. *Neuroscience* 1984; 12(3): 711–718. [PubMed: 6382048]
2. Matamales M, Gotz J, Bertran-Gonzalez J. Quantitative Imaging of Cholinergic Interneurons Reveals a Distinctive Spatial Organization and a Functional Gradient across the Mouse Striatum. *PLoS One* 2016; 11(6): e0157682. [PubMed: 27314496]
3. Mogenson GJ, Jones DL, Yim CY. From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* 1980; 14(2-3): 69–97. [PubMed: 6999537]
4. Kelley AE. Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. *Neurosci Biobehav Rev* 2004; 27(8): 765–776. [PubMed: 15019426]
5. Salamone JD, Correa M, Farrar A, Mingote SM. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* 2007; 191(3): 461–482. [PubMed: 17225164]
6. Wang Z, Kai L, Day M, Ronesi J, Yin HH, Ding J et al. Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* 2006; 50(3): 443–452. [PubMed: 16675398]
7. Witten IB, Lin SC, Brodsky M, Prakash R, Diester I, Anikeeva P et al. Cholinergic interneurons control local circuit activity and cocaine conditioning. *Science* 2010; 330(6011): 1677–1681. [PubMed: 21164015]
8. Lee J, Finkelstein J, Choi JY, Witten IB. Linking Cholinergic Interneurons, Synaptic Plasticity, and Behavior during the Extinction of a Cocaine-Context Association. *Neuron* 2016; 90(5): 1071–1085. [PubMed: 27210555]

9. Zucca S, Zucca A, Nakano T, Aoki S, Wickens J. Pauses in cholinergic interneuron firing exert an inhibitory control on striatal output in vivo. *eLife* 2018; 7.
10. English DF, Ibanez-Sandoval O, Stark E, Tecuapetla F, Buzsáki G, Deisseroth K et al. GABAergic circuits mediate the reinforcement-related signals of striatal cholinergic interneurons. *Nat Neurosci* 2011; 15(1): 123–130. [PubMed: 22158514]
11. Cachepe R, Mateo Y, Mathur BN, Irving J, Wang HL, Morales M et al. Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. *Cell reports* 2012; 2(1): 33–41. [PubMed: 22840394]
12. Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* 2012; 75(1): 58–64. [PubMed: 22794260]
13. Exley R, Clements MA, Hartung H, McIntosh JM, Cragg SJ.  $\alpha 6$ -Containing Nicotinic Acetylcholine Receptors Dominate the Nicotine Control of Dopamine Neurotransmission in Nucleus Accumbens. *Neuropsychopharmacology* 2007; 33: 2158. [PubMed: 18033235]
14. Rice ME, Cragg SJ. Nicotine amplifies reward-related dopamine signals in striatum. *Nat Neurosci* 2004; 7(6): 583–584. [PubMed: 15146188]
15. Hikida T, Kaneko S, Isobe T, Kitabatake Y, Watanabe D, Pastan I et al. Increased sensitivity to cocaine by cholinergic cell ablation in nucleus accumbens. *Proc Natl Acad Sci U S A* 2001; 98(23): 13351–13354. [PubMed: 11606786]
16. Warner-Schmidt JL, Schmidt EF, Marshall JJ, Rubin AJ, Arango-Lievano M, Kaplitt MG et al. Cholinergic interneurons in the nucleus accumbens regulate depression-like behavior. *Proc Natl Acad Sci U S A* 2012; 109(28): 11360–11365. [PubMed: 22733786]
17. Aoki S, Liu AW, Zucca A, Zucca S, Wickens JR. Role of Striatal Cholinergic Interneurons in Set-Shifting in the Rat. *J Neurosci* 2015; 35(25): 9424–9431. [PubMed: 26109665]
18. Collins AL, Aitken TJ, Huang IW, Shieh C, Greenfield VY, Monbouquette HG et al. Nucleus Accumbens Cholinergic Interneurons Oppose Cue-Motivated Behavior. *Biol Psychiatry* 2019.
19. Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M. Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J Neurosci* 1994; 14(6): 3969–3984. [PubMed: 8207500]
20. Kimura M, Rajkowski J, Evarts E. Tonically discharging putamen neurons exhibit set-dependent responses. *Proc Natl Acad Sci U S A* 1984; 81(15): 4998–5001. [PubMed: 6589643]
21. Aosaki T, Graybiel AM, Kimura M. Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. *Science* 1994; 265(5170): 412–415. [PubMed: 8023166]
22. Apicella P, Legallet E, Trouche E. Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. *Exp Brain Res* 1997; 116(3): 456–466. [PubMed: 9372294]
23. Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 2004; 43(1): 133–143. [PubMed: 15233923]
24. Ravel S, Legallet E, Apicella P. Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. *J Neurosci* 2003; 23(24): 8489–8497. [PubMed: 13679417]
25. Blazquez PM, Fujii N, Kojima J, Graybiel AM. A Network Representation of Response Probability in the Striatum. *Neuron* 2002; 33(6): 973–982. [PubMed: 11906702]
26. Marche K, Martel AC, Apicella P. Differences between Dorsal and Ventral Striatum in the Sensitivity of Tonically Active Neurons to Rewarding Events. *Front Syst Neurosci* 2017; 11: 52. [PubMed: 28790898]
27. Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H. Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J Neurosci* 2008; 28(45): 11673–11684. [PubMed: 18987203]
28. Apicella P The role of the intrinsic cholinergic system of the striatum: What have we learned from TAN recordings in behaving animals? *Neuroscience* 2017; 360: 81–94. [PubMed: 28768155]

29. Zhang YF, Cragg SJ. Pauses in Striatal Cholinergic Interneurons: What is Revealed by Their Common Themes and Variations? *Front Syst Neurosci* 2017; 11: 80. [PubMed: 29163075]
30. Benhamou L, Kehat O, Cohen D. Firing pattern characteristics of tonically active neurons in rat striatum: context dependent or species divergent? *J Neurosci* 2014; 34(6): 2299–2304. [PubMed: 24501368]
31. Atallah HE, McCool AD, Howe MW, Graybiel AM. Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. *Neuron* 2014; 82(5): 1145–1156. [PubMed: 24908491]
32. Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW. The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. *Neuron* 2013; 79(1): 153–166. [PubMed: 23770257]
33. Apicella P, Deffains M, Ravel S, Legallet E. Tonically active neurons in the striatum differentiate between delivery and omission of expected reward in a probabilistic task context. *Eur J Neurosci* 2009; 30(3): 515–526. [PubMed: 19656171]
34. Chuhma N, Mingote S, Moore H, Rayport S. Dopamine neurons control striatal cholinergic neurons via regionally heterogeneous dopamine and glutamate signaling. *Neuron* 2014; 81(4): 901–912. [PubMed: 24559678]
35. Straub C, Tritsch NX, Hagan NA, Gu C, Sabatini BL. Multiphasic modulation of cholinergic interneurons by nigrostriatal afferents. *J Neurosci* 2014; 34(25): 8557–8569. [PubMed: 24948810]
36. Wieland S, Du D, Oswald MJ, Parlato R, Köhr G, Kelsch W. Phasic dopaminergic activity exerts fast control of cholinergic interneuron firing via sequential NMDA, D2, and D1 receptor activation. *J Neurosci* 2014; 34(35): 11549–11559. [PubMed: 25164653]
37. Augustin SM, Chancey JH, Lovinger DM. Dual Dopaminergic Regulation of Corticostriatal Plasticity by Cholinergic Interneurons and Indirect Pathway Medium Spiny Neurons. *Cell reports* 2018; 24(11): 2883–2893. [PubMed: 30208314]
38. Kharkwal G, Brami-Cherrier K, Lizardi-Ortiz JE, Nelson AB, Ramos M, Del Barrio D et al. Parkinsonism Driven by Antipsychotics Originates from Dopaminergic Control of Striatal Cholinergic Interneurons. *Neuron* 2016; 91(1): 67–78. [PubMed: 27387649]
39. Doig NM, Magill PJ, Apicella P, Bolam JP, Sharott A. Cortical and thalamic excitation mediate the multiphasic responses of striatal cholinergic interneurons to motivationally salient stimuli. *J Neurosci* 2014; 34(8): 3101–3117. [PubMed: 24553950]
40. Zhang YF, Reynolds JNJ, Cragg SJ. Pauses in Cholinergic Interneuron Activity Are Driven by Excitatory Input and Delayed Rectification, with Dopamine Modulation. *Neuron* 2018; 98(5): 918–925.e913. [PubMed: 29754751]
41. Matsumoto N, Minamimoto T, Graybiel AM, Kimura M. Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J Neurophysiol* 2001; 85(2): 960–976. [PubMed: 11160526]
42. Ding JB, Guzman JN, Peterson JD, Goldberg JA, Surmeier DJ. Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron* 2010; 67(2): 294–307. [PubMed: 20670836]
43. Brown MT, Tan KR, O'Connor EC, Nikonenko I, Muller D, Luscher C. Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. *Nature* 2012; 492(7429): 452–456. [PubMed: 23178810]
44. Bäckman CM, Malik N, Zhang Y, Shan L, Grinberg A, Hoffer BJ et al. Characterization of a mouse strain expressing Cre recombinase from the 3' untranslated region of the dopamine transporter locus. *Genesis* 2006; 44(8): 383–390. [PubMed: 16865686]
45. Gallo EF, Meszaros J, Sherman JD, Chohan MO, Teboul E, Choi CS et al. Accumbens dopamine D2 receptors increase motivation by decreasing inhibitory transmission to the ventral pallidum. *Nat Commun* 2018; 9(1): 1086. [PubMed: 29540712]
46. Gallo EF, Salling MC, Feng B, Moron JA, Harrison NL, Javitch JA et al. Upregulation of Dopamine D2 Receptors in the Nucleus Accumbens Indirect Pathway Increases Locomotion but Does Not Reduce Alcohol Consumption. *Neuropsychopharmacology* 2015.
47. Jing M, Li Y, Zeng J, Huang P, Skirzewski M, Kljatic O et al. An optimized acetylcholine sensor for monitoring in vivo cholinergic activity. *Nat Methods* 2020; 17(11): 1139–1146. [PubMed: 32989318]

48. Deng P, Zhang Y, Xu ZC. Involvement of I(h) in dopamine modulation of tonic firing in striatal cholinergic interneurons. *J Neurosci* 2007; 27(12): 3148–3156. [PubMed: 17376976]
49. Wilson CJ. The Mechanism of Intrinsic Amplification of Hyperpolarizations and Spontaneous Bursting in Striatal Cholinergic Interneurons. *Neuron* 2005; 45(4): 575–585. [PubMed: 15721243]
50. Barker DJ, Miranda-Barrientos J, Zhang S, Root DH, Wang HL, Liu B et al. Lateral Preoptic Control of the Lateral Habenula through Convergent Glutamate and GABA Transmission. *Cell reports* 2017; 21(7): 1757–1769. [PubMed: 29141211]
51. Calipari ES, Bagot RC, Purushothaman I, Davidson TJ, Yorgason JT, Peña CJ et al. In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc Natl Acad Sci U S A* 2016; 113(10): 2726–2731. [PubMed: 26831103]
52. Ward RD, Gallistel CR, Jensen G, Richards VL, Fairhurst S, Balsam PD. Conditioned [corrected] stimulus informativeness governs conditioned stimulus-unconditioned stimulus associability. *J Exp Psychol Anim Behav Process* 2012; 38(3): 217–232. [PubMed: 22468633]
53. Collins AL, Aitken TJ, Greenfield VY, Ostlund SB, Wassum KM. Nucleus Accumbens Acetylcholine Receptors Modulate Dopamine and Motivation. *Neuropsychopharmacology* 2016; 41: 2830. [PubMed: 27240658]
54. Nautiyal KM, Wall MM, Wang S, Magalong VM, Ahmari SE, Balsam PD et al. Genetic and Modeling Approaches Reveal Distinct Components of Impulsive Behavior. *Neuropsychopharmacology* 2017; 42(6): 1182–1191. [PubMed: 27976680]
55. Yan Z, Song WJ, Surmeier J. D2 dopamine receptors reduce N-type Ca<sup>2+</sup> currents in rat neostriatal cholinergic interneurons through a membrane-delimited, protein-kinase-C-insensitive pathway. *J Neurophysiol* 1997; 77(2): 1003–1015. [PubMed: 9065864]
56. Donthamsetti P, Gallo EF, Buck DC, Stahl EL, Zhu Y, Lane JR et al. Arrestin recruitment to dopamine D2 receptor mediates locomotion but not incentive motivation. *Mol Psychiatry* 2020; 25(9): 2086–2100. [PubMed: 30120413]
57. Bennett BD, Wilson CJ. Spontaneous activity of neostriatal cholinergic interneurons in vitro. *J Neurosci* 1999; 19(13): 5586–5596. [PubMed: 10377365]
58. Lacey MG, Mercuri NB, North RA. Dopamine acts on D2 receptors to increase potassium conductance in neurones of the rat substantia nigra zona compacta. *J Physiol* 1987; 392: 397–416. [PubMed: 2451725]
59. Uchida S, Akaike N, Nabekura J. Dopamine activates inward rectifier K<sup>+</sup> channel in acutely dissociated rat substantia nigra neurones. *Neuropharmacology* 2000; 39(2): 191–201. [PubMed: 10670414]
60. Singer BF, Bryan MA, Popov P, Scarff R, Carter C, Wright E et al. The sensory features of a food cue influence its ability to act as an incentive stimulus and evoke dopamine release in the nucleus accumbens core. *Learn Mem* 2016; 23(11): 595–606. [PubMed: 27918279]
61. Patriarchi T, Cho JR, Merten K, Howe MW, Marley A, Xiong W-H et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science (New York, NY)* 2018; 360(6396): eaat4422.
62. Cragg SJ. Meaningful silences: how dopamine listens to the ACh pause. *Trends Neurosci* 2006; 29(3): 125–131. [PubMed: 16443285]
63. Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM et al. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* 2013; 494(7436): 238–242. [PubMed: 23354054]
64. Syed ECJ, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME. Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nat Neurosci* 2016; 19(1): 34–36. [PubMed: 26642087]
65. Bertran-Gonzalez J, Laurent V, Chieng BC, Christie MJ, Balleine BW. Learning-Related Translocation of  $\delta$ -Opioid Receptors on Ventral Striatal Cholinergic Interneurons Mediates Choice between Goal-Directed Actions. *The Journal of Neuroscience* 2013; 33(41): 16060–16071. [PubMed: 24107940]
66. Fligel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I et al. A selective role for dopamine in stimulus-reward learning. *Nature* 2011; 469(7328): 53–57. [PubMed: 21150898]

67. Heymann G, Jo YS, Reichard KL, McFarland N, Chavkin C, Palmiter RD et al. Synergy of distinct dopamine projection populations in behavioral reinforcement. *Neuron* 2020; 105(5): 909–920. e905. [PubMed: 31879163]
68. Dalley JW, Fryer TD, Brichard L, Robinson ESJ, Theobald DEH, Lääne K et al. Nucleus Accumbens D2/3 Receptors Predict Trait Impulsivity and Cocaine Reinforcement. *Science* 2007; 315(5816): 1267–1270. [PubMed: 17332411]
69. Martyniuk KM, Dandeneau M, Balsam PD, Kellendonk C. Dopamine D2R upregulation in ventral striatopallidal neurons does not affect Pavlovian or go/no-go learning. *Behav Neurosci* 2021; 135(3): 369–379. [PubMed: 34264690]
70. Ragozzino ME, Mohler EG, Prior M, Palencia CA, Rozman S. Acetylcholine activity in selective striatal regions supports behavioral flexibility. *Neurobiol Learn Mem* 2009; 91(1): 13–22. [PubMed: 18845266]
71. Matamales M, Skrbis Z, Hatch RJ, Balleine BW, Götz J, Bertran-Gonzalez J. Aging-Related Dysfunction of Striatal Cholinergic Interneurons Produces Conflict in Action Selection. *Neuron* 2016; 90(2): 362–373. [PubMed: 27100198]
72. Okada K, Nishizawa K, Fukabori R, Kai N, Shiota A, Ueda M et al. Enhanced flexibility of place discrimination learning by targeting striatal cholinergic interneurons. *Nature Communications* 2014; 5(1): 3778.
73. Faust TW, Assous M, Shah F, Tepper JM, Koós T. Novel fast adapting interneurons mediate cholinergic-induced fast GABAA inhibitory postsynaptic currents in striatal spiny neurons. *Eur J Neurosci* 2015; 42(2): 1764–1774. [PubMed: 25865337]
74. Pakhotin P, Bracci E. Cholinergic interneurons control the excitatory input to the striatum. *J Neurosci* 2007; 27(2): 391–400. [PubMed: 17215400]
75. Franklin NT, Frank MJ. A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife* 2015; 4: e12029. [PubMed: 26705698]



**Figure 1. D2R upregulation does not alter intrinsic properties or firing in NAc CINs.**

**A.** Schematic representation depicting injection of AAV into the NAc core of adult ChAT-Cre mice. **B.** Low magnification image of AAV-DIO-D2-P2A-EGFP expression in the NAc core 4 weeks after viral injection. Scale = 200  $\mu\text{m}$ . **C, D.** Double-immunolabeling of AAV-DIO-D2-P2A-EGFP or AAV-DIO-D2-IRES-mVenus expression and the cholinergic cell marker ChAT. Scale = 20  $\mu\text{m}$ . **E.** Representative epifluorescence image of *ex vivo* slice preparations from adult NAc, showing a visually identified EGFP-positive CIN. **F.** Current clamp recordings in whole-cell mode showing the voltage responses to  $-140$  and  $+40$  pA currents. **G-I.** Box plots (bars, min/max values; box, lower/upper quartile; line, median) showing resting membrane potential ( $t = 0.4814$ ,  $p = 0.6341$ ,  $n = 14-15$  cells/group), input resistance ( $t = 0.3712$ ,  $p = 0.7134$ ,  $n = 14-15$  cells/group) and action potential threshold ( $t = 0.7209$ ,  $p = 0.4774$ ,  $n = 14-15$  cells/group) were not altered by D2R upregulation. Data was analyzed using unpaired  $t$  tests (two-tailed). **J.** Representative voltage clamp recordings showing currents induced by hyperpolarizing voltage steps from a holding potential of  $-50$  mV ( $-60$  to  $-150$  mV). **K.**  $I_h$  was not altered by D2R upregulation ( $F_{(1,26)} = 0.117$ ,  $p = 0.7353$ ,  $n = 14-15$  cells/group). **L, M.** Cell-attached recordings to measure spontaneous



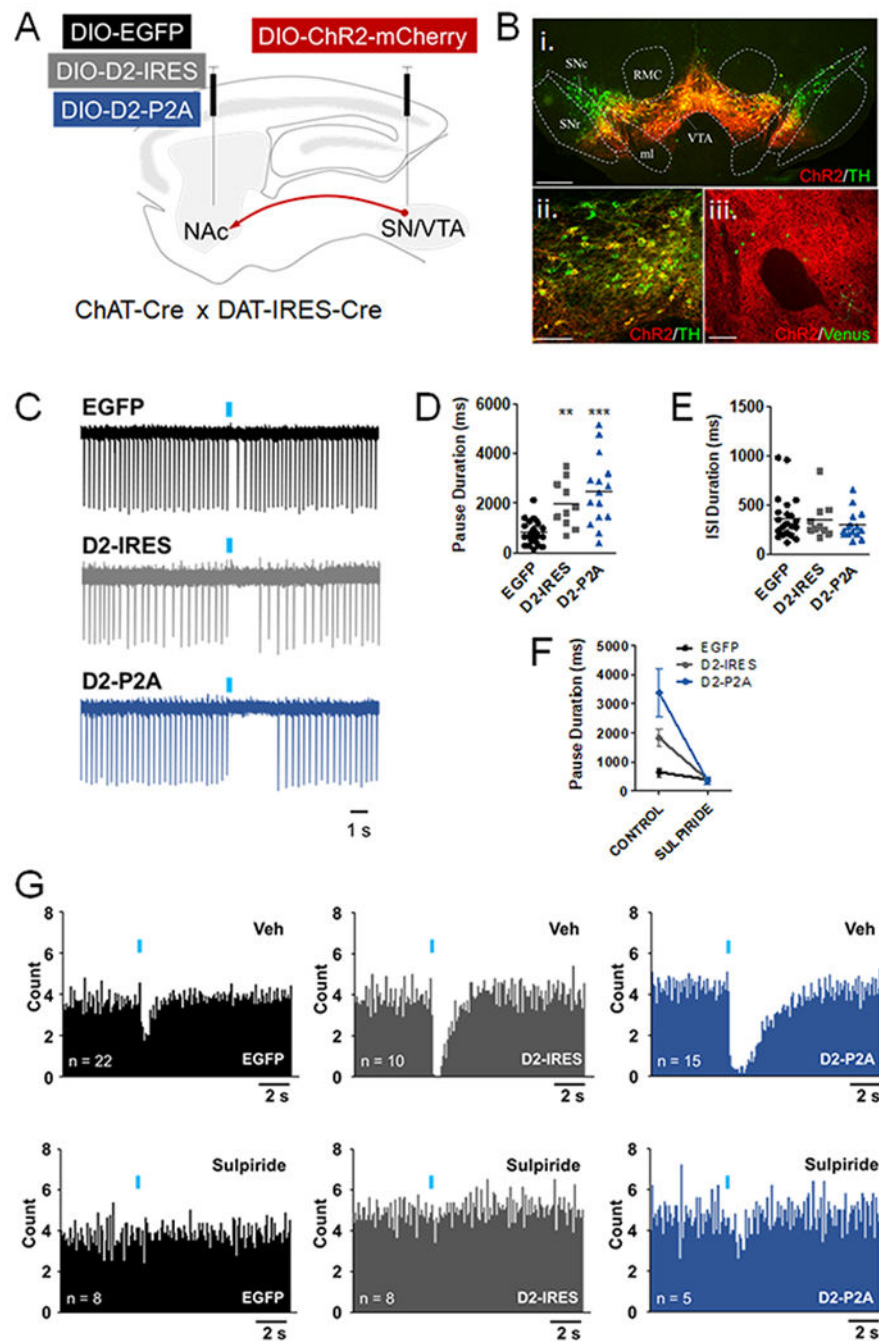
CIN activity revealed no difference in spike frequency ( $t = 0.1134$ ,  $p = 0.9108$ ,  $n = 10-13$  cells/group).

Author Manuscript

Author Manuscript

Author Manuscript

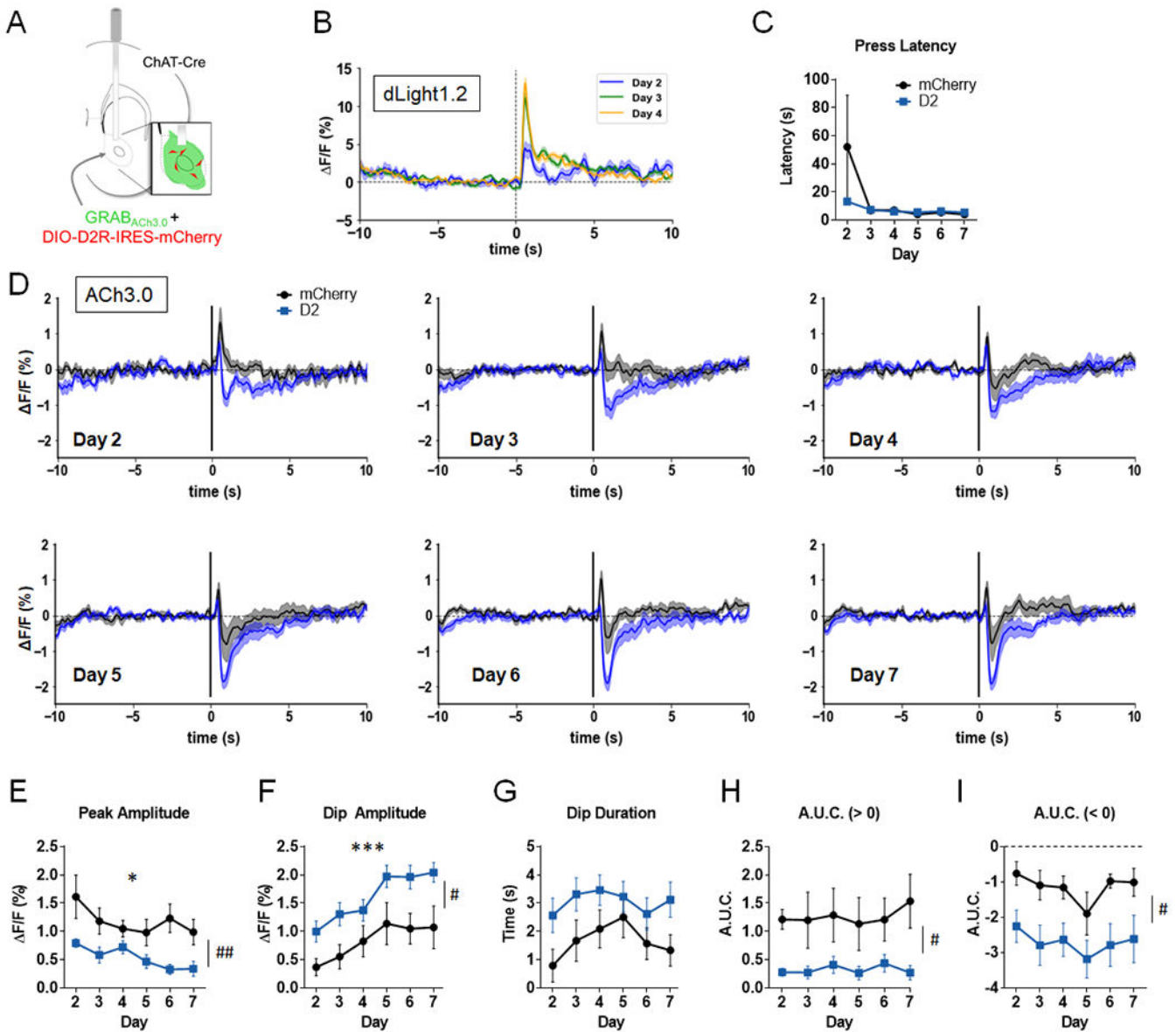
Author Manuscript



**Figure 2. D2R upregulation in NAc CINs increases pause duration.**

**A.** ChAT-Cre x DAT-IRES-Cre mice were injected into the VTA/SN with AAV-DIO-ChR2-mCherry and with either AAV-DIO-EGFP or AAV-DIO-D2-IRES-mVenus or AAV-DIO-D2-P2A-EGFP into the NAc. Red arrow represents the ChR2-positive afferents contacting the NAc. **B,i-ii.** Double immunolabeling showing co-localization between of Chr2-mCherry and TH in a midbrain section. Scale = 150 and 50  $\mu$ m. *SNc*, substantia nigra pars compacta; *SNr*, substantia nigra pars reticulata; *ml*, medial lemniscus; *VTA*, ventral tegmental area; *RMC*, red nucleus, magnocellular. **B,iii.** Sparse AAV-DIO-D2-IRES-mVenus-positive CINs

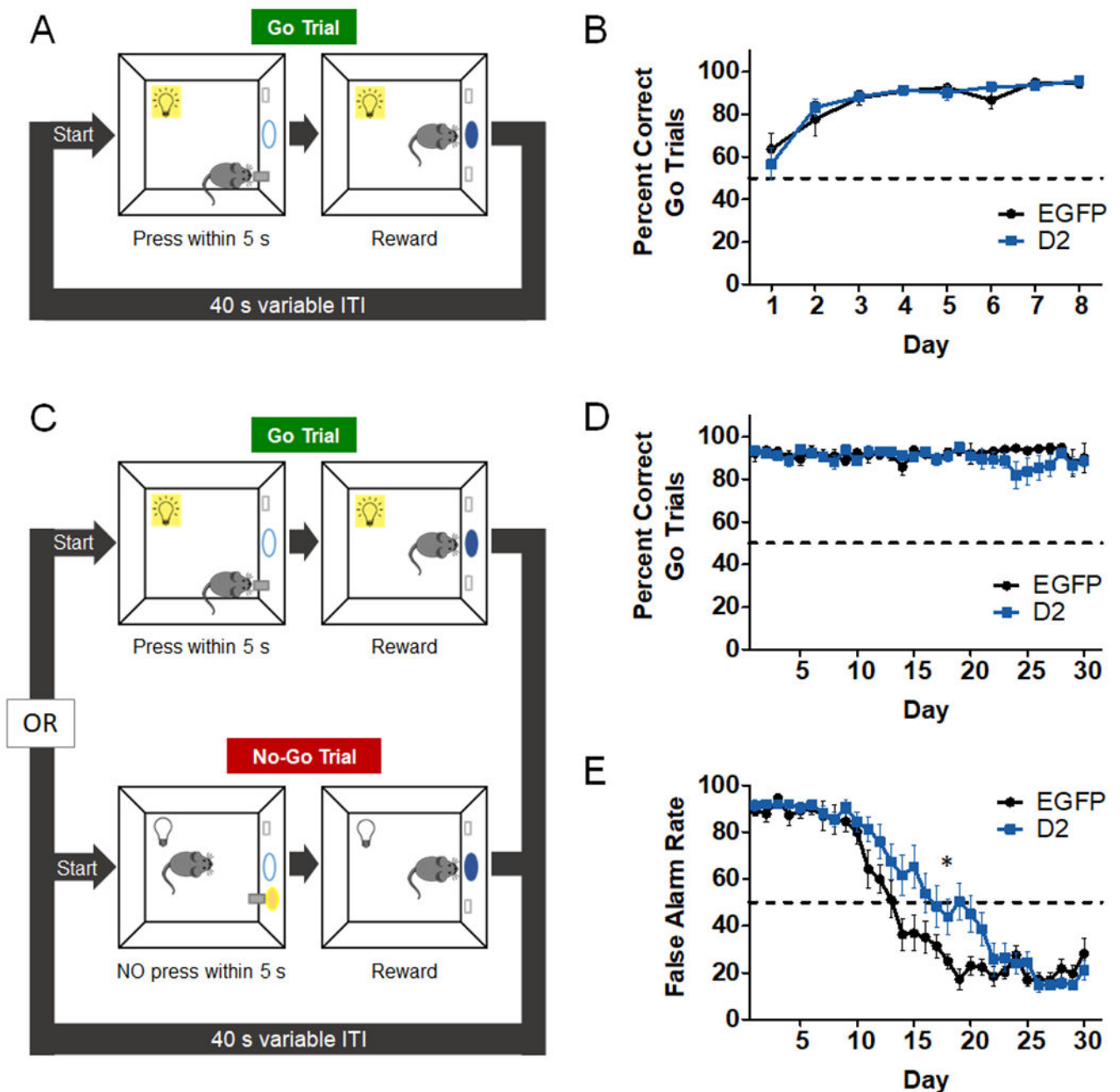
in the NAc core (green) surrounded by Chr2-positive afferents (red) from the midbrain. Scale = 50  $\mu$ m. **C.** Sample cell-attached recording traces following one trial of light-evoked burst stimulation (blue bars, 5 x 5 ms pulses, 20 Hz). **D.** Pause duration, measured as the average duration of the interspike interval (ISI) immediately following the stimulus across 10 trials, was significantly increased in cells expressing either of the D2R AAVs ( $F_{(2,46)} = 15.77$ ,  $***p < 0.0001$ . Bonferroni *post hoc* test: EGFP vs D2-IRES,  $**p < 0.001$ ; EGFP vs D2-P2A,  $p < 0.0001$ ; D2-IRES vs D2-P2A,  $p > 0.05$ ). **E.** The average ISI duration was not altered by D2R upregulation ( $F_{(2,46)} = 0.4685$ ,  $p = 0.6289$ ). **F.** In a smaller subset of neurons that received both ACSF and sulpiride (10  $\mu$ M), pause duration was significantly reduced by sulpiride pretreatment. A 2-way ANOVA found a statistically significant difference in pause duration by treatment ( $F_{(1,19)} = 40.82$ ,  $p < 0.001$ ) and by AAV ( $F_{(2,19)} = 9.645$ ,  $p = 0.0013$ ), and a significant treatment x AAV interaction ( $F_{(2,19)} = 9.603$ ,  $p = 0.0013$ ). Bonferroni *post hoc* tests revealed no significant pairwise differences following sulpiride treatment between groups (all  $p$ 's  $> 0.05$ ). **G.** Peristimulus histograms of mean firing from 10 consecutive trials (0.1s bins).



**Figure 3. D2R upregulation in NAc CINs alters ACh levels in a continuous reinforcement (CRF) task.**

**A.** AAV-GRAB<sub>ACh3.0</sub> (ACh3.0) was infused together with either AAV-DIO-D2R-IRES-mCherry (or AAV-DIO-mCherry) into the NAc. An optic fiber was implanted to measure task-evoked NAc ACh3.0 fluorescence signals. *Inset*, representation of expected cell targeting of D2R AAV to NAc CINs (red), with a broader expression of ACh3.0 signal (green). **B.** Normalized mean dLight1.2 fluorescent signals in NAc of a representative mouse, aligned to the lever extension across 3 days of training on CRF schedule. **C.** Press latency across days did not differ between the two groups (virus effect:  $F_{(1,12)} = 0.91$ ,  $p = 0.36$  or virus x day interaction:  $F_{(5,60)} = 1.14$ ,  $p = 0.35$ ). **D.** Normalized mean ACh3.0 fluorescent signals aligned to the lever extension across 6 days of training (Days 2-7; signals were not recorded on the first day of training). **E.** Peak amplitude was decreased in both groups (day effect:  $F_{(5,60)} = 3.06$ , \*  $p = 0.016$ ). Peak amplitude was reduced in

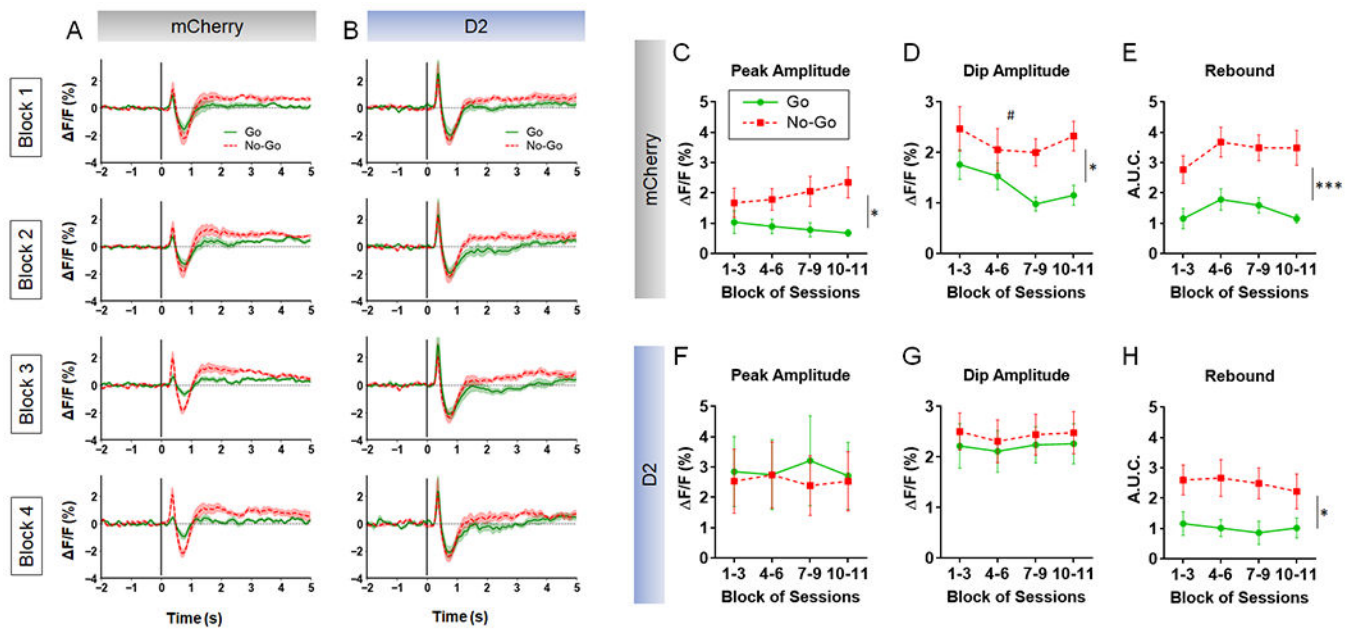
D2R-OE<sub>NacChAT</sub> mice (virus effect:  $F_{(1,12)} = 10.52$ ,  $^{##}p = 0.007$ ). **F.** Dip amplitude increased with training (day effect:  $F_{(5,60)} = 17.74$ ,  $^{***}p = 0.0001$ ). A main effect of virus was also observed ( $F_{(1,12)} = 6.33$ ,  $^{#}p = 0.027$ ). **G.** D2R upregulation lead to a trend towards a longer dip duration ( $F_{(1,12)} = 4.60$ ,  $p = 0.053$ ). **H.** A.U.C. above baseline was significantly reduced by D2R upregulation ( $F_{(1,12)} = 6.76$ ,  $^{#}p = 0.023$ ). **I.** A.U.C. below baseline was increased in D2R-OE<sub>NacChAT</sub> mice ( $F_{(1,12)} = 9.32$ ,  $^{#}p = 0.01$ ).  $n = 7$  mice/group for panels C-I.



**Figure 4. D2R upregulation in NAc CINs impairs No-Go responding.**

**A.** Schematic of first training phase consisting of 60 Go trials. Each Go trial is started with house light illumination and lever presentation, and mice must press the lever within 5 s to receive a reward. New trials begin after a variable ITI. **B.** Go responding was measured across 8 days and expressed as the average percent correct Go trials. This “hit rate” increased similarly in both groups with training (day effect:  $F_{(7,161)} = 21.8$ ,  $p < 0.0001$ ; AAV effect:  $F_{(1,23)} = 0.011$ ,  $p = 0.91$ ; AAV x day interaction:  $F_{(7,161)} = 0.75$ ,  $p = 0.63$ ). EGFP,  $n = 11$ ; D2,  $n = 14$  mice. **C.** In the second phase, which consisted of 30 days, 30 Go trials were intermixed with 30 No-Go trials. Unlike Go trials, No-Go trials were signaled

by the presentation of the lever and LED lights above the lever without a house light. Withholding from pressing for 5 s during No-Go trial led to reward. **D.** D2R upregulation did not alter accuracy of responding during Go trials (AAV effect:  $F_{(1,23)} = 0.52$ ,  $p = 0.48$ ); AAV x day interaction ( $F_{(29,662)} = 1.183$ ,  $p = 0.24$ ). EGFP,  $n = 11$ ; D2,  $n = 14$  mice. **E.** In No-Go trials, premature responding (false alarm rate) decreased with training in both groups (day effect:  $F_{(29,663)} = 88.97$ ,  $p < 0.0001$ ), yet this transition was significantly delayed in D2R-OE<sub>NacChAT</sub> mice (AAV x day interaction:  $F_{(29,663)} = 3.099$ ,  $*p < 0.0001$ ). EGFP,  $n = 11$ ; D2,  $n = 14$  mice.



**Figure 5. CIN D2R upregulation reduces contrast in cue-evoked phasic ACh signaling in Go/No-Go task.**

**A, B.** Mean ACh3.0 signals aligned to the onset of Go (green) and No-Go trials (red) were recorded in NAc of mice expressing mCherry ( $n = 6$ ) or D2-IRES-mCherry ( $n = 5$ ). Fluorescent signals were recorded across 11 sessions (every 3 days in a 31-day period) and analyzed as a function of 4 blocks of recording sessions (1-3, 4-6, 7-9, 10-11). **C-E.** Repeated measures 2-way ANOVA for peak amplitude, dip amplitude and post-dip rebound in the mCherry group showed that all three measures were significantly higher in No-Go relative to Go trials. **C.** Peak amplitude showed a significant main effect of trial type ( $F_{(1,10)} = 5.08$ , \*  $p = 0.048$ ) and a significant session block  $\times$  trial type interaction ( $F_{(3,30)} = 3.86$ ,  $p = 0.019$ ). **D.** Dip amplitude revealed significant main effects of session ( $F_{(3,30)} = 3.54$  #  $p = 0.026$ ) and trial type ( $F_{(1,10)} = 5.61$ , \*  $p = 0.039$ ). **E.** Rebound ACh was also significantly greater in No-Go trials in this group (trial type:  $F_{(1,10)} = 27.8$ , \*\*\*  $p = 0.0004$ ). **F-H.** In contrast, in D2R- $OE_{NacChAT}$  mice, no significant trial-type differences were observed in peak amplitude ( $F_{(1,8)} = 0.043$ ,  $p = 0.84$ ) or dip amplitude ( $F_{(1,8)} = 0.161$ ,  $p = 0.70$ ), but No-Go trials did show a higher rebound than Go trials ( $F_{(1,8)} = 6.35$ , \* $p = 0.036$ ).