Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Elizabeth C. Heaton Date

Striatal melanocortin-4 receptor control of action flexibility

By

Elizabeth C. Heaton
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Neuroscience

_____

Shannon L. Gourley, Ph.D.
Advisor

_____

Adriana Galvan, Ph.D.
Committee Member

_____

Vasiliki Michopoulos, Ph.D.
Committee Member

_____

Larry Young, Ph.D.
Committee Member

Accepted:

_____

Kimberly Jacob Arriola, Ph.D., MPH
Dean of the James T. Laney School of Graduate Studies

_____

Date

Striatal melanocortin-4 receptor control of action flexibility

By

Elizabeth C. Heaton
B.S. Psychology, Haverford College, 2018

Advisor: Shannon L. Gourley, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Neuroscience
2023

Abstract

Striatal melanocortin-4 receptor control of action flexibility

By Elizabeth C. Heaton

More than half of all individuals in treatment for substance use disorder (SUD) will relapse. Inflexibility in selecting between familiar, habitual behaviors that have been rewarded in the past (drug seeking) and novel strategies that might be more advantageous (rehabilitation) may be a factor that preserves SUD. The dorsomedial striatum (DMS) is a brain region that receives and integrates glutamatergic input from cortical and subcortical regions required for goal-directed action selection. However, the factors in the DMS responsible for coordinating this incoming information remain incompletely understood. This dissertation begins by describing the extra-hypothalamic functions of the melanocortin-4 receptor (MC4R), a receptor that is well-positioned in the DMS to control flexible, goal-directed action. Next, I report that MC4R in the DMS appears to propel familiar reward-seeking behavior (habit), even when it is not fruitful, and moderating MC4R presence improves the capacity for goal-directed behavior. I then demonstrate that this process requires inputs from the orbitofrontal cortex, a brain region canonically associated with response strategy switching. Then, I further investigate how striatal melanocortin systems propel familiar behaviors, particularly via interaction with the central nucleus of the amygdala (CeA). I demonstrate that MC4R-expressing cells in the DMS are 1) predominantly expressed on dopamine D1-type receptor-expressing medium spiny neurons and 2) are necessary and sufficient for controlling the capacity of mice to arbitrate between actions and habits. I next use site-selective gene silencing and pharmacological techniques to reveal that MC4R presence suppresses goal seeking. I also find that MC4R-expressing neurons are functionally integrated into an amygdalo-striatal circuit that suppresses action flexibility in favor of routinized behaviors. Additionally, I use publicly available spatial transcriptomics datasets to reveal differences in the gene transcript correlates of *Mc4r* across the striatum, with considerable co-variation in dorsal structures. Guided by these results, I lastly discovered that MC4R function in the dorsolateral striatum complements that in the DMS, here *suppressing* habitual behavior. Together, these findings provide insight into the molecular and circuit-level mechanisms by which MC4R in the DMS propels habitual behavior. This dissertation thus illuminates mechanistic factors that support the development of automatized routines when flexible decision making is no longer adaptive, which may provide insight into therapeutic targets for neuropsychiatric disorders in which decision making is impaired.

Striatal melanocortin-4 receptor control of action flexibility

By

Elizabeth C. Heaton
B.S. Psychology, Haverford College, 2018

Advisor: Shannon L. Gourley, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Neuroscience
2023

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1: Extra-hypothalamic functions of the melanocortin-4 receptor

CHAPTER 2: Selective breeding reveals control of reward-related action selection by the melanocortin-4 receptor

CHAPTER 3: Striatal melanocortin systems propel familiar actions via interaction with the central nucleus of the amygdala

CHAPTER 4: Concluding remarks

Chapter 1

Extra-hypothalamic functions of the melanocortin-4 receptor

1.1 CONTEXT AND AUTHOR'S CONTRIBUTION

The following chapter reviews the extra-hypothalamic functions of the melanocortin-4 receptor. We focus on receptor function within regions necessary for goal-directed decision making. The dissertation author contributed by researching and writing the manuscript, with editorial feedback from Dr. Shannon Gourley.

1.2 ABSTRACT

The melanocortin-4 receptor (MC4R) is well-known for its role in regulating energy homeostasis and feeding behavior via hypothalamic systems. Despite nearly ubiquitous expression throughout the brain, research into extra-hypothalamic functions of MC4R is comparatively thin. Here, we review the mechanisms and functions of MC4R signaling in brain regions outside of the hypothalamus, with the goal of building a more nuanced perspective on MC4R, as well as informing future clinical studies of the melanocortin system. We first provide a broad overview of the central melanocortin system, emphasizing its ubiquity. Next, we discuss the anatomy, mechanism, and function of MC4R in four brain regions necessary for decision-making behavior: the cortex, hippocampus, amygdala, and striatum. We conclude by outlining the framework of this dissertation, revealing a novel role of striatal MC4R in reward-related decision making.

1.3 INTRODUCTION

The melanocortin-4 receptor (MC4R) is intensively studied in the hypothalamus for its regulation of feeding behavior and energy homeostasis. Though MC4R is expressed ubiquitously in the brain, comparatively little research has been dedicated to MC4R function in other cortical and subcortical regions, such as the hippocampus and striatum. In this review, we describe the

mechanistic and functional outcomes of MC4R activation in regions throughout the brain. We believe that a global perspective on central MC4R not only contributes to a more complete understanding of MC4R function, but, moreover, informs current and future clinical studies of the melanocortin system.

1.3.1 Overview of the Central Melanocortin System

The melanocortin system consists of two neuronal populations in the arcuate nucleus of the hypothalamus (Arc) that express proopiomelanocortin (POMC) and agouti-related protein (AgRP), respectively. Posttranslational processing of the prohormone POMC produces four primary melanocortin agonists: α-, β, and γ-melanocyte stimulating hormone (MSH), and adrenocorticotropic hormone (ACTH) (Bertagna, 1994; Mountjoy, 2015; Smith & Funder, 1988). In contrast, AgRP and agouti-signaling protein (ASIP), a peripheral paracrine signaling molecule, act as endogenous melanocortin antagonists/competitive agonists (Gantz & Fong, 2003). Together, these peptides are responsible for the diverse functions of the melanocortin system.

Five melanocortin receptors mediate the activity of melanocortin peptides (Tao, 2010). Of these five receptors, two are primarily expressed in the CNS: the melanocortin-3 receptor (MC3R) and MC4R (Kishi et al., 2003; H. Liu et al., 2003; Mountjoy et al., 1994). The central melanocortin system is classically studied for its role in regulating food intake and energy homeostasis. MC4R, in particular, seems to be important for maintaining homeostatic balance: *Mc4r*-null mice are chronically obese (Huszar et al., 1997), and re-expressing MC4R on cells in the paraventricular nucleus of the hypothalamus (PVH) is sufficient to reduce food intake and rescue body weight (Balthasar et al., 2005; Shah et al., 2014). Furthermore, mutations in the *MC4R* gene remain the only known monogenic cause of obesity in humans (Mutch & Clément, 2006; Yeo et al., 1998),

and systemic administration of the MC4R agonist setmelanotide successfully reverses hyperphagia and weight gain in humans (Sweeney et al., 2023). Thus, MC4R remains an important target for treating dysfunctions in energy homeostasis.

Hypothalamic MC4R is also involved in a diverse range of behaviors outside of food consumption. MC4R activation in the hypothalamus decreases ethanol intake (Lerma-Cabrera et al., 2013, 2020), and increases sexual drive in rodents and humans (Argiolas et al., 2000; Martin & MacIntyre, 2004). Bremelanotide, a synthetic MC4R agonist, is currently in clinical trials for the treatment of hypoactive sexual disorder in female humans (Dhillon & Keam, 2019; Simon et al., 2019). And hypothalamic MC4R activation regulates the central oxytocinergic system. Acute administration of MC4R agonists activates hypothalamic oxytocin neurons and potentiates oxytocin release in the PVH, resulting in enhanced social behavior (Barrett et al., 2014; Modi et al., 2015). MC4R could therefore be a viable target for enhancing social function in psychiatric disorders. Due, in part, to this broad translational relevance, MC4R function in the hypothalamus remains a primary focus of melanocortin research.

Research over the past decade has resolved much of MC4R signaling in the hypothalamus. MC4R is a G protein-coupled receptor (GPCR) with seven transmembrane domains (Tao, 2010). MC4R activation in excitatory single minded 1 (SIM1)-expressing neurons in the PVH triggers the activation of adenyl cyclase, which leads to increases in levels of cyclic AMP (cAMP) and activation of protein kinase A (PKA) signaling through mitogen-activated protein kinases (MAPKs) (Daniels et al., 2003; Li et al., 2019; Münzberg et al., 2020; Shah et al., 2014; Tao, 2010). This signaling pathway is necessary for MC4R function within the hypothalamus and is hypothesized to be dysregulated in humans with monogenic *MC4R* mutations (He & Tao, 2014). Notably, activation of hypothalamic MC4R impacts measures of postsynaptic, but not presynaptic,

neuronal plasticity (Fenselau et al., 2017). It is therefore possible, though not proven, that MC4R-dependent changes in synaptic plasticity could play an important role in MC4R function within the hypothalamus.

Hypothalamic circuits containing melanocortins have been thoroughly investigated, resulting in a greater understanding of homeostatic mechanisms and meaningful translational outcomes. Not because, but rather in spite of this success, the functions of the central melanocortin system outside of the hypothalamus are comparatively understudied. Hypothalamic POMC- and AgRP-expressing neurons project throughout the brain, including cortical and subcortical regions involved in learning and memory, decision making, and motor control (Eskay et al., 1979; Saper et al., 1986; Shen et al., 2013; Wang et al., 2015). Concurrently, MC4R is expressed almost ubiquitously in the CNS, with particularly dense expression in the cortex, hippocampus, striatum, and amygdala (Kishi et al., 2003; Lim et al., 2012; Liu et al., 2003; Shen et al., 2013). This review will focus on the mechanisms and functions of MC4R binding in brain regions outside of the hypothalamus, with the goal of building a more nuanced perspective on MC4R, as well as informing future clinical studies of the melanocortin system. The contents of this review are summarized in part in Tables 1.1 and 1.2.

1.4 FUNCTIONS OF MC4R OUTSIDE OF THE HYPOTHALAMUS

1.4.1 Cortex

The cortex is innervated by α-MSH-immunoreactive fiber systems originating in the arcuate nucleus of the hypothalamus and the zona incerta (Saper et al., 1986; Shiosaka et al., 1984, 1984; Wang et al., 2015). A particularly dense set of projections is directed to the medial prefrontal cortex (mPFC) (Ross et al., 2023). The presence of α-MSH is matched by an abundant expression

of MC4R throughout the cortex. Regions of note include the ventral and lateral orbital frontal cortices (layer 5/6), prelimbic prefrontal cortex (layer 5/6), infralimbic cortex, and olfactory tubercle (Kishi et al., 2003; H. Liu et al., 2003).

What research there is to date on MC4R function in the cortex has focused primarily on the function of this receptor within the mPFC. The mPFC is involved in numerous aspects of complex decision making, including outcome-related learning, consolidation of memories, and forming associations between contexts and responses (Sato et al., 2023). The mPFC can be divided into subregions, including the prelimbic prefrontal cortex (PL) and the infralimbic cortex (IL), which have distinctive functions in these processes (Gourley & Taylor, 2016).

Very little is known about the mechanism of MC4R signaling in the cortex, and the evidence that exists is focused primarily on the mPFC. Activating MC4R in the mPFC of mice increases excitatory postsynaptic potentials (EPSPs) in excitatory pyramidal neurons, suggesting that cortical MC4R activation increases neuronal excitability (Ross et al., 2023). Evidence from studies in amphibians suggests that cortical melanocortin receptor activation may be involved in a negative feedback loop between the cortex and the hypothalamus, which regulates α-MSH secretion (Bercu & Brinkley, 1967). More recent studies show that intraventricular administration of AgRP increases dopaminergic turnover in the mPFC (Davis et al., 2011). These results, in conjunction with thorough work on melanocortin regulation of the mesocortical dopaminergic system, link MC4R to dopamine signaling in the cortex (Roseberry et al., 2015).

As in the hypothalamus, MC4R in the cortex is also linked to feeding behavior. MC4R-expressing glutamatergic cells in the IL project to multiple regions that coordinate responses to food-related stimuli, such as the lateral hypothalamus and nucleus accumbens shell. *Mc4r*

knockdown in the IL, but not PL, increases nighttime feeding and gross body weight (Ross et al., 2023) – highlighting subregional distinctions in MC4R function.

The activity of MC4R-expressing cells in the IL is additionally implicated in impulsive- and anxiety-like behavior. Specifically, cortical MC4R-expressing cells are necessary for prevention of unwanted actions, such that inhibiting MC4R-expressing cells in the IL causes perseverative errors during task reversal training (Ross et al., 2023). Inhibition of MC4R-expressing cells in the IL also impacts novelty-suppressed feeding, a procedure in which mice are placed in open fields containing unfamiliar food. An increased latency by the mouse to approach the strange food is considered anxiety-like behavior. Although all mice consumed the same amount of food, mice in the inhibition condition had increased latency to begin eating (Ross et al., 2023). This finding suggests that MC4R binding in the mPFC is anxiolytic. Indeed, knocking down *Mc4r* expression induces higher retention of learned fear during an extinction learning task, suggesting these mice have trouble shifting their learning when placed in a new context, potentially due to heightened anxiety.

Cortical MC4R is also implicated in the mechanisms of addictive drugs. Chronic administration of nicotine concurrently increases *Pomc* mRNA in the hypothalamus and *Mc4r* mRNA in the mPFC, thus boosting both pre- and post-synaptic melanocortin signaling (Tapinc et al., 2017). This effect is seen systemically in humans: Peripheral POMC levels are elevated in individuals who regularly smoke nicotine (Muschler et al., 2018). On the other hand, unlike nicotine, cocaine does not change *Mc4r* expression across the cerebral cortex (Alvaro et al., 2003). This is likely due to differences in the cellular targets of nicotine *versus* cocaine (Radua et al., 2023).

1.4.2 Hippocampus

α-MSH has long been recognized for its effects on learning and memory. As early as 1966, de Wied and colleagues discovered that rats treated systemically with an "ACTH-like peptide" displayed increased retention of conditioned avoidance responses during an extinction task (de Wied, 1966). This peptide, now known to be α-MSH, also facilitated acquisition and retention of a learned appetitive response (Kastin et al., 1974; Sandman et al., 1969). Since the publication of these studies, α-MSH has also been shown to regulate memory consolidation, particularly when context is in flux (Gonzalez et al., 2009), displaying rather unambiguously that stimulation of the melanocortin system improves performance on learning and memory tasks (McLay et al., 2001).

Given the role of α-MSH in learning and memory, it should therefore come as no surprise that *Mc4r* is present in the hippocampus. The hippocampus is divided into subsections, including: the dentate gyrus, cornu ammonis 3 (CA3), and cornu ammonis 1 (CA1). Each of these regions are a part of the trisynaptic circuit, a relay loop within the hippocampus that supports adaptive decision-making behavior (Mehrotra & Dubé, 2023). *Mc4r* is expressed throughout the hippocampus, particularly within CA1 (Gantz et al., 1993; H. Liu et al., 2003; Siljee-Wong, 2011), and at levels similar to those seen in the hypothalamus (Shen et al., 2013), suggesting that MC4R plays a notable role in hippocampal function. In addition to receiving POMC+ projections from the arcuate nucleus of the hypothalamus (Saper et al., 1986), the hippocampus may also be a site of POMC and AgRP production: α-MSH, POMC, and AgRP mRNA are all found in CA3 (Shen et al., 2013, 2016). POMC-expressing neurons in CA3 project to MC4R-expressing cells in CA1, thereby forming a functional melanocortin circuit within the hippocampus (Shen et al., 2016).

Seminal work from Shen and colleagues illuminates the role of MC4R in regulating hippocampal synaptic plasticity: POMC-expressing pyramidal neurons in CA3 project to CA1 where they secrete α-MSH, thereby activating MC4R (Shen et al., 2016). Within the postsynaptic

compartment, MC4R activation triggers Gs signaling and increases intracellular levels of cyclic adenosine monophosphate (AMP) (Shen et al., 2013; Tao, 2010). This amplifies protein kinase A (PKA) signaling, resulting in increased cyclic AMP response element-binding protein (CREB) activation and changes in gene transcription (Shen et al., 2013, 2016). Functionally, this MC4R/PKA/CREB signaling pathway regulates synaptic plasticity, allowing for MC4R activation to increase densities of mature dendritic spines, elevate long-term potentiation (LTP) induction, and increase synaptic levels of GluA1-containing AMPARs (Shen et al., 2013, 2016). These effects can be abolished via *Mc4r* knockdown or inhibiting PKA signaling in MC4R-expressing neurons (Shen et al., 2016), supporting previous findings that cAMP/PKA/MAPK signaling is vital for hippocampal synaptic plasticity (Waltereit & Weller, 2003). Taken together, these results suggest that MC4R is at least partially responsible for the regulation of synaptic plasticity within the hippocampus.

MC4R function in the hippocampus has implications for human health. MC4R is implicated in Alzheimer's Disease (AD), a disorder characterized by the build-up of tau and beta-amyloid proteins in the central nervous system, resulting in dysregulation of synaptic plasticity in the hippocampus. Plasma and cerebrospinal fluid levels of α-MSH are reduced in patients with AD (Rainero et al., 1988). Administering α-MSH in rodents is protective against symptoms of AD: α-MSH administration decreases beta-amyloid deposits in the hippocampus and cortex (Giuliani et al., 2014); protects the hippocampus from IL1-beta-induced memory impairment (Ma & McLaurin, 2014); and prevents GABAergic neuronal loss and improves cognitive function with age (Gonzalez et al., 2009). Therefore, MC4R may play a role in regulating AD-related pathology in the hippocampus.

Perturbed MC4R signaling in the hippocampus exacerbates LTP impairment in mouse models of AD (Shen et al., 2016). Further, activating the hippocampal melanocortin circuit rescues this LTP impairment, along with dysregulated synaptic morphology and neurotransmission (Shen et al., 2016). This is in congruence with previous findings that plasma and cerebrospinal fluid levels of α-MSH are reduced in individuals with AD (Rainero et al., 1988). Decreased α-MSH levels could reflect a downregulation of melanocortin activity throughout the brain, and including in the hippocampus where MC4R activation is partially responsible for maintaining synaptic plasticity required for cognitive function.

MC4R may have other neuroprotective functions within the hippocampus. Recent evidence reveals that intracellular tau accumulation induces synaptic dysregulating in individuals with AD via inhibition of PKA/GluA1 signaling in the hippocampus (Ye et al., 2020), a pathway up-regulated by MC4R activation (Shen et al., 2013). Activating MC4R, and therefore increasing PKA activation and synaptic GluA1-continaing AMPARs, may ameliorate the effects of intracellular tau accumulation. Intracellular tau build-up also inhibits the function of brain-derived neurotrophic factor (BDNF) and its receptor tropomyosin receptor kinase B (TrkB) (Ye et al., 2020). TrkB signaling mediates dendritic spine remodeling in the hippocampus, which influences learning and memory (Numakawa & Odaka, 2022) – dysregulation of which results in hippocampal synaptic impairments (Ye et al., 2020). MC4R activation increases BDNF release and TrkB signaling in the hypothalamus (Flores-Bastías et al., 2020) and amygdala (Boghossian et al., 2010), likely via stimulation of the transcription factor CREB. Activating MC4R is also found to increase BDNF release by astrocytes (Ramírez et al., 2015). Furthermore, MC4R activation has a central anti-inflammatory effect (Giuliani et al., 2017; Lasaga et al., 2008), a

function that may be mediated through regulation of TrkB, the activation of which is considered to be broadly neuroprotective (Shkundin & Halaris, 2023).

Hippocampal function is also impacted by ethanol, which results in dysfunctional neuronal communication due to oxidative stress and mitochondrial impairment (Tapia-Rojas et al., 2019). MC4R stimulation in cultured hippocampal neurons activates the Nrf-2 pathway, which protects against damage by ethanol exposure (Goodfellow et al., 2020; Quintanilla et al., 2020). Stimulating MC4R also reduces ethanol intake (Boghossian et al., 2010; York et al., 2011). Thus, MC4R both influences the effects of ethanol and is affected by ethanol administration.

1.4.3 Amygdala

The amygdala contains some of the highest expression of *Mc4r* outside of the hypothalamus, including in the medial, central, and basal nuclei (Liu et al., 2003). Moreover, POMC neurons in the arcuate nucleus of the hypothalamus project to the amygdala, also innervating medial, central, and basal nuclei (Eskay et al., 1979).

As in other brain regions, MC4R in the amygdala regulates food intake. The amygdala, and especially the central amygdala (CeA), regulates food intake in response to insulin signaling (Boghossian et al., 2009). Infusion of the MC4R agonist Melanotan II (MTII) into the central amygdala reduces high-fat food intake in rats but has no effect on consumption of typical vivarium chow. Meanwhile, inhibiting *Mc4r* in the CeA increased consumption of high-fat food (Boghossian et al., 2010). This is consistent with previous findings showing that re-expression of *Mc4r* in the CeA *Mc4r*-null transgenic mice is sufficient to rescue some, but not all, typical feeding behavior (Balthasar et al., 2005). Further research is needed to determine if the CeA serves as a melanocortin-driven fat-specific taste-aversion center.

MC4R in both the CeA and basolateral amygdala (BLA) alters preference for water and ethanol, independent of its effects on food intake. Stimulating MC4R in the amygdala reduces food intake but increases water intake (Boghossian et al., 2010; York et al., 2011). Ethanol consumption is reduced following both stimulating and inhibition of amygdalar MC4R activity (Boghossian et al., 2010; York et al., 2011). Intra-CeA infusion of a MC4R antagonist further attenuates ethanol withdrawal-induced hyperalgesia in ethanol-dependent rats. Meanwhile, intra-CeA infusions of α-MSH induce hyperalgesia (Avegno et al., 2018). Taken together, this research suggests that MC4R in the amygdala plays an important role in regulating the appetiveness/aversiveness of stimuli.

1.4.4 Striatum

The striatum is divided into two subcompartments: the dorsal striatum, which includes the dorsomedial striatum (DMS) and dorsolateral striatum (DLS), and the ventral striatum, which includes the nucleus accumbens (NAc). The striatum is predominantly composed of medium spiny neurons (MSNs), large projection neurons that serve as the primary output for the region. MSNs are categorized by the expression of dopamine D1-type receptors (D1R) or D2-type receptors (D2R). Anatomically, D1R-MSNs make up the "direct pathway" while D2R-MSNs form the "indirect pathway," which together sum the primary output of the striatum (for a review of striatal anatomy, see Rocha et al., 2023). MC4R is preferentially expressed on D1R-MSNs in the dorsal and ventral striatum, with some expression on D2R-MSNs (Oude Ophuis et al., 2014). Striatal MC4R signaling regulates the localization of glutamate receptor subunits in the postsynaptic compartment of D1R-MSNs: for instance, activating MC4R prompts movement of AMPA receptors containing subunit GluA2 (AMPAR-GluA2) to the synapse, thus increasing cell excitability (Lim et al., 2012). Whether and how striatal MC4R regulates the localization of other glutamate receptor subunits is not yet fully understood.

MC4R-related changes in cell excitability appear to have lasting impacts on behavior. Typical mice will avoid environments associated with aversive stimuli (e.g. foot shock). In contrast, global deletion of *Mc4r* biases mice towards either indifference or preference for environments associated with aversive stimuli. Such a drastic change in behavior was associated with increases activity in the dopaminergic system within the striatum, and re-expression of *Mc4r* in D1R-MSNs in the striatum normalized responses to aversive stimuli (Klawonn et al., 2018). The value of reward is also impacted by striatal MC4R. Chronic stress induces anhedonic-like behavior (losing preference for otherwise preferred stimuli like sweetened sucrose solutions), a process that requires MC4R signaling in D1R-MSNs in the ventral striatum (Lim et al., 2012).

Difficulties in determining or responding to the motivational valence of rewards may contribute to impairments in procedural learning, the ability to acquire motor skills (such as seeking out and obtaining a reward) by practice. Ventral striatal MC4R is important for procedural learning: Global *Mc4r* knockout mice are significantly slower at acquiring operant responding and are slow to learn the Morris water maze task compared to control mice. Restoring *Mc4r* in D1R-MSNs in the nucleus accumbens core and shell (as well as the paraventricular nucleus of the hypothalamus (PVN) and lateral olfactory tract) rescued procedural learning in both cases (H. Cui et al., 2012).

The first research into melanocortin function in the 1960s described compulsive-like grooming following intracerebral infusions or intraperitoneal injections of MC4R agonists (Ferrari et al., 1963; Gessa et al., 1967). Later investigations hypothesized that this compulsive-like phenotype is dependent on MC4R regulation of dopaminergic activity in the VTA and striatum (Bertolini & Gessa, 1981). In conjunction with these early findings, Xu and colleagues (2013)

found that *Mc4r* deletion in the striatum corrects compulsive-like grooming in *SAPAP3* knockout mice (a common model for studying compulsive-like behavior).

Compulsive-like behaviors may contribute to the initiation and/or maintenance of substance use disorder. Indeed, in addition to promoting compulsive-like grooming, striatal MC4R is required for the motivational, rewarding, and locomotor effects of cocaine in mice (Hsu et al., 2005a). Further, striatal *Mc4r* expression is upregulated by chronic cocaine administration (Alvaro et al., 2003; Hsu et al., 2005a), creating a closed-loop system by which cocaine administration increases MC4R activity, which in turn promotes the compulsivity that might maintain routine substance use behavior.

1.5 DISCUSSION

In this review, I focused on the functions of MC4R in brain regions outside of the hypothalamus, with the goal of building a more nuanced perspective on MC4R. MC4R is classically studied for its role in maintaining energy homeostasis by regulating metabolism and feeding behavior. Here, we identify additional attributes of MC4R across regions: controlling procedural learning, defining motivational valence, gating responses to addictive drugs, and impacting anxiety-like behavior, to name a few. A more complete understanding of MC4R will inform current and future clinical investigations of the melanocortin system.

The intersection of MC4R function and human health is relevant given the FDA approval of the MC4R agonists bremelanotide (Vyleesi) and setmelanotide (Imcivree) for the treatment of hyposexuality and monogenic obesity, respectively, in humans. Both bremelanotide and setmelanotide are given as subcutaneous injections with the goal of targeting hypothalamic nuclei relevant for reproductive behavior and food intake. However, systemic injections of any compound

are expected to act across the central nervous system, regardless of initial intent. The research summarized in this review would suggest that the activation of MC4R by an exogenous agonist could illicit a range of behavioral phenotypes, including a susceptibility to addictive drugs. Dysregulation of the melanocortin system in humans is also linked to both intellectual disability (Beleford et al., 2020) and ADHD (Caruso et al., 2014), perhaps due to MC4R activity in the mPFC.

That being said, MC4R binding has a profoundly seemingly positive effect on hippocampal plasticity: MC4R activation is sufficient to increase synaptic AMPAR expression, change dendritic spine morphology, and influence LTP. Hippocampal MC4R has the greatest relevance to human health in cases where synaptic plasticity is dysregulated: substance use disorder and Alzheimer's disease. Although researchers have begun to characterize the impacts of MC4R signaling in the hippocampus, comparatively little information is known about the impact of hippocampal MC4R on behavior. Shen and colleagues (2016) conclude that MC4R activation could rescue neuronal symptomatology in rodent models of AD by increasing neuronal excitability. These findings, in conjunction with work revealing that MC4R activity improves learning and memory in rodents (Cui et al., 2012), beg the question of whether MC4R activation could similarly alleviate behavioral AD symptoms, such as cognitive impairment.

1.6 FRAMEWORK AND OVERVIEW OF THE DISSERTATION.

More than half of all individuals in treatment for substance use disorder (SUD) will relapse. Inflexibility in selecting between familiar, habitual behaviors that have been rewarded in the past (drug seeking) and novel strategies that might be more advantageous (rehabilitation) may be a factor that preserves SUD. The goal of my thesis research is to identify neural factors supporting

goal-directed action selection, which could provide insight into therapeutic targets for disorders in which goal-oriented action selection is impaired.

The dorsomedial striatum (DMS) is a brain region that receives and integrates glutamatergic input from cortical and subcortical regions required for goal-directed action selection. However, the factors in the DMS responsible for coordinating this incoming information remain incompletely understood. One candidate factor is MC4R. MC4R regulates glutamatergic excitability of D1R-MSNs in the striatum. Inhibiting MC4R reduces the expression of repetitive, familiar behaviors and improves flexible action selection in mice (Allen et al., 2022; Xu et al., 2013). Thus, MC4R seems well-positioned in the DMS to integrate incoming glutamatergic signals and control flexible, goal-directed action.

This dissertation attempts to determine the molecular- and circuit-level mechanisms by which striatal MC4R regulates decision-making behavior in mice. In Chapter 2, I report that MC4R in the DMS appears to propel reward-seeking behavior, even when it is not fruitful, and moderating MC4R presence increases the capacity of mice to inhibit such behaviors. I then demonstrate that this process requires inputs from the orbitofrontal cortex, a brain region canonically associated with response strategy switching.

In Chapter 3, I further investigate how striatal melanocortin systems propel familiar behaviors, particularly via interaction with the CeA. I demonstrate that MC4R-expressing cells in the DMS are 1) predominantly expressed on dopamine D1R-MSNs and 2) are necessary and sufficient for controlling the capacity of mice to arbitrate between actions and habits. I next use site-selective gene silencing and pharmacological techniques to reveal that MC4R presence suppresses goal seeking. I also find that MC4R-expressing neurons are functionally integrated into an amygdalo-striatal circuit that suppresses action flexibility in favor of routinized behaviors.

Additionally, I use publicly available spatial transcriptomics datasets to reveal differences in the gene transcript correlates of *Mc4r* across the striatum, with considerable co-variation in dorsal structures. Guided by these results, I discovered that MC4R function in the dorsolateral striatum complements that in the DMS, here *suppressing* habitual behavior.

In Chapter 4, I summarize the findings of the dissertation, incorporate my findings into existing literature, and propose future directions to further our understanding of how homeostatic systems influence complex decision-making behavior.

Region	Manipulation	Agent	Method	Result	Citation
Cortex					
medial PFC	↑ MC4R	selective agonist THIQ		↑ magnitude of EPSPs in pyramidal cells	Ross 2023
Hippocampus	↑ MC4R	selective agonist D-Tyr MTII	in vitro hippocampal slice	↑ mature dendritic spines ↑ surface expression of AMPARs containing GluA1 ↑ phosphorylation of GluA1 at Ser845 ↑ amplitude and frequency of mEPSCs ↑ magnitude of SC-CA1 LTP	Shen 2013
	↑ MC4R	selective agonist D-Tyr MTII	i.p. injection; i.c.v. infusion	↑ magnitude of SC-CA1 LTP ↑ mature dendritic spines	Shen 2013; Shen 2016
	↑ MC4R	selective agonist RO27-3225	in vitro hippocampal neurons	↓ ethanol-induced oxidative damage and mitochondrial stress ↑ activation of Nrf-2 pathway	Quintanilla 2020
	↑ MC4R	ligand α-MSH	viral-mediated α-MSH overexpression in hippocampal POMC cells	↑ magnitude of SC-CA1 LTP	Shen 2016
	↓ MC4R	MC4R shRNA	in vitro hippocampal slice	↓ dendritic spine density and volume in CA1 ↓ mature (mushroom) dendritic spines ↑ thin dendritic spines	Shen 2013
	↓ MC4R	MC4R shRNA	viral-mediated shRNA expression	↓ magnitude of SC-CA1 LTP	Shen 2016
	↓ MC4R	selective antagonists HS024 and MCL-0020	i.c.v. infusion	↓ magnitude of SC-CA1 LTP	Shen 2016
	↓ MC4R	antagonist AgRP	virus-mediated overexpression	↓ magnitude of SC-CA1 LTP	Shen 2016
Striatum	↑ MC4R	selective re-expression in D1-MSNs	*Mc4r-STOP-flox* mice	normalized responses to aversive stimuli	Klawonn 2018
	↑ MC4R	selective re-expression in D1-MSNs	*Mc4r-STOP-flox* mice	↓ effect of cocaine on phosphorylation of DARPP-32 and GluA1-continaing AMPARs	Cui and Lutter 2013
N. accumbens	↑ MC4R	selective agonist cyclo(NH-CH2-CH2-CO-His-d-Phe-Arg-Trp-Glu)-NH2	intra-accumbens infusion	↓ voluntary ethanol consumption ↓ ethanol palatability	Lerma-Cabrera 2012, 2013; Carvajal 2017
N. accumbens	↑ MC4R	agonist MTII	chronic i.c.v. infusion	↑ D1-like receptor binding	Lindblom 2001
	↑ MC4R	ligand α-MSH	intra-accumbens infusion	↑ cocaine-induced locomotor sensitization	Hsu 2005
	↑ MC4R	ligand α-MSH	in vitro accumbens slice	↓ synaptic AMPARs containing GluA2 in D1-MSNs ↓ NMDAR-dependent LTD in D1-MSNs no effect on NMDAR-mediated synaptic transmission	Lim 2012
	↓ MC4R	global *Mc4r* knockout	*Mc4r-STOP-flox* mice	↑ dopaminergic transmission following aversive stimuli	Klawonn 2018
	↓ MC4R	global *MC4R* knockdown	endogenous *MC4R* mutations in humans	no difference in BOLD response to food-related stimuli from lean controls	van der Klaauw 2014
	↓ MC4R	MC4R shRNA	viral-mediated shRNA expression	↓ effect of chronic stress on food intake ↓ effect of chronic stress on sucrose preference ↓ effect of chronic stress on cocaine-induced conditioned place preference	Lim 2012

				no effect on forced swim test no effect on tail suspension test	
	↓ MC4R	antagonist SHU9119	intra-accumbens infusion	↓ cocaine-conditioned place preference ↓ cocaine self-administration ↓ cocaine locomotor sensitization	Hsu 2005
Dorsal striatum	↑ MC4R	selective agonist [Nle^4,D-Phe^7]α-MSH	i.c.v. infusion	↑ dopaminergic transmission	Florijn 1993
	↑ MC4R	agonist MTII	chronic i.c.v. infusion	↑ D1-like receptor binding ↓ D2-like receptor binding	Lindblom 2001
	↓ MC4R	global *MC4R* knockdown	endogenous *MC4R* mutations in humans	no difference in BOLD response to food-related stimuli from lean controls	van der Klaauw 2014
Amygdala	↑ MC4R	agonist MTII	intra-amygdalar infusion	↓ food and ethanol intake ↑ water intake	York 2011
	↑ MC4R	agonist MTII	intra-amygdalar infusion	↓ food intake	Boghossian 2010
	↓ MC4R	selective antagonist SHU-9119	intra-amygdalar infusion	↑ food and water intake ↓ ethanol intake	York 2011; Boghossian 2010
	↓ MC4R	antagonist AgRP	intra-amygdalar infusion	↑ food intake	Boghossian 2010
	↓ MC4R	global *MC4R* knockdown	endogenous *MC4R* mutations in humans	no difference in BOLD response to food-related stimuli from lean controls	van der Klaauw 2014
Central nucleus	↓ MC4R	antagonist HS014	i.p. injection	rescued thermal hyperalgesia during ethanol withdrawal	Avegno 2018
	↑ MC4R	agonist α-MSH	i.c.v. infusion	↑ hyperalgesia	Avegno 2018

Table 1.1 Manipulations of MC4R have neurobehavioral consequences. In the second column, the up arrow refers to stimulation of the receptor or genetic re-expression strategies, while the down arrow refers to receptor antagonism, gene silencing, or genetic mutation. **Abbreviations:** Agouti-related protein (AgRP), AMPA receptor (AMPAR), α-melanocyte stimulating hormone (α-MSH), blood-oxygen-level-dependent response (BOLD response), cornu Ammonis 1 (CA1), dopamine D1-type receptor (D1), dopamine D2-type receptor (D2), dopamine- and cAMP-regulated phosphoprotein (DARPP-32), Designer Receptors Exclusively Activated by Designer Drugs (DREADDs), excitatory postsynaptic potentials (EPSP), AMPA receptor subunit GluA1 (GluA1), AMPA receptor subunit GluA2 (GluA2), intracerebroventricular injection (i.c.v.), long-term potentiation (LTP), melanocortin-4 receptor (MC4R), medium spiny neuron (MSN), melanotan II (MTII), prefrontal cortex (PFC), proopiomelanocortin (POMC), subiculum (SC), 1,2,3,4-tetrahydroisoquinoline (THIQ).

Region	Manipulation	Method	Result	Citation
Cortex	ligand α-MSH, acute	0.5 mg/kg i.p.	no effect	Saba 2019
	ligand α-MSH	0.5 mg/kg i.p. daily for 2 days	↑ MC4R	Saba 2019
	cocaine, chronic	15 mg/kg i.p. twice daily for 14 days	no effect	Alvaro 2003; Hsu 2005
medial PFC	nicotine, chronic	0.6 mg/kg s.c. daily for 5 days	↑ MC4R	Tapinc 2017
Hippocampus	cocaine, chronic	15 mg/kg i.p. twice daily for 14 days	↑ MC4R	Alvaro 2003
Striatum	ligand α-MSH, acute	0.5 mg/kg i.p.	↑ MC4R	Saba 2019
	ligand α-MSH, chronic	0.5 mg/kg i.p. daily for 2 days	↑ MC4R	Saba 2019
	cocaine, acute	15 mg/kg i.p.	no effect	Alvaro 2003
	cocaine, chronic	15 mg/kg i.p. twice daily for 14 days	↑ MC4R	Alvaro 2003; Hsu 2005
N. accumbens	morphine, acute	75 mg tablet implanted s.c.	↓ MC4R	Alvaro 1996
	morphine, chronic	75 mg tablet implanted s.c. daily for 5 days	no effect ↓ MC4R	Alvaro 1996; Alvaro 2003
	morphine, chronic	2 mg/kg i.p. every other day for 10 days	↑ MC4R	Alvaro 2003
	stress, acute	3-4 h restraint stress	no effect	Lim 2012
	stress, chronic	3-4 h restraint stress daily for 7-8 days	↑ MC4R	Lim 2012
	intracranial self-stimulation	7 days	↑ MC4R	Upadhya 2020
Amygdala	ethanol, chronic	2 g/kg i.p. daily for 14 days (P28-41)	↑ MC4R	Kokare 2017
Central nucleus	ethanol, chronic	BAL 150-250 mg/dl for at least 4 weeks	↓ MC4R	Avegno 2018

Table 1.2 Experimental manipulations that regulate MC4R. α-melanocyte stimulating hormone (α-MSH), blood alcohol level (BAL), intraperitoneal (i.p.), melanocortin-4 receptor (MC4R), postnatal day (P), prefrontal cortex (PFC), subcutaneous (s.c.).

Chapter 2

Selective breeding reveals control of reward-related action strategies by the melanocortin-4 receptor

2.1 CONTEXT, AUTHOR'S CONTRIBUTION, AND ACKNOWLEDGEMENT OF REPRODUCTION

The following chapter describes alterations in protein levels and decision-making behavior in a line of mice that are selectively bred for the expression of habit-like behavior. We then detail the role of striatal MC4R in the control of reward-related action strategies. This chapter is excerpted from: Allen AT, Heaton EC, Shapiro LP, Butkovich LM, Yount ST, Davies RA, Li DC, Swanson AM, and Gourley SL. (2022) Inter-individual variability amplified through breeding reveals control of reward-related action strategies by melanocortin-4 receptor in the dorsomedial striatum. *Communications Biology*, 5, 116; doi: 10.1038/s42003-022-03043-2. Edits to this chapter highlight the contributions of the dissertation author, who contributed by designing and conducting experiments, analyzing data, and editing the manuscript.

2.2 ABSTRACT

In day-to-day life, we often must choose between pursuing familiar behaviors or adjusting behaviors when new strategies might be more fruitful. The dorsomedial striatum (DMS) is indispensable for arbitrating between old and new action strategies. To uncover molecular mechanisms, we trained mice to generate nose poke responses for food, then uncoupled the predictive relationship between one action and its outcome. We then bred the mice that failed to rapidly modify responding. This breeding created offspring with the same tendencies, failing to inhibit behaviors that were not reinforced. These mice had less post-synaptic density protein 95 in the DMS. Also, densities of the melanocortin-4 receptor (MC4R), a high-affinity receptor for α-melanocyte-stimulating hormone, predicted individuals' response strategies. Specifically, high MC4R levels were associated with poor response inhibition. We next found that reducing *Mc4r* in

the DMS in otherwise typical mice expedited response inhibition, allowing mice to modify behavior when rewards were unavailable or lost value. This process required inputs from the orbitofrontal cortex, a brain region canonically associated with response strategy switching. Thus, MC4R in the DMS appears to propel reward-seeking behavior, even when it is not fruitful, while moderating MC4R presence increases the capacity of mice to inhibit such behaviors.

2.3 INTRODUCTION

In day-to-day life, we often pursue familiar behavioral sequences that have been reinforced in the past – e.g., driving a familiar route home from work – or inhibit behaviors when they fail to be reinforced – like avoiding that route when construction blocks our path. The dorsomedial, or associative, striatum (DMS), roughly analogous to the primate caudate, is indispensable for arbitrating between familiar and new action strategies. For instance, damage to the DMS causes rats to pursue familiar behavioral sequences even when they cease to be rewarded (Bradfield et al., 2013; Braun & Hauber, 2012; Lex & Hauber, 2010; Pauli, Clark, et al., 2012; Yin, Ostlund, et al., 2005). Motor task learning recruits neural ensembles in the DMS that decline in activity with task proficiency (Yin et al., 2009). Further, instrumental conditioning – learning to perform a behavior for reward – triggers immediate-early gene expression and transcriptional activity in the DMS (Hernandez et al., 2006; Maroteaux et al., 2014; Matamales et al., 2020; Peak et al., 2020) and requires direct spiny projection neurons in the DMS (Peak et al., 2020). Nevertheless, the molecular mechanisms by which the DMS coordinates the flexible modification of behavior are still emerging.

A strategy by which to identify molecular factors regulating a given behavior is to manipulate the levels or activities of proteins that are predicted to control that behavior. A

limitation of this approach is that unpredicted factors – those that we might not anticipate – remain obscure. Here, we instead used a discovery-driven strategy. We first bred mice that displayed a particular behavioral trait – resistance to inhibiting behaviors when they failed to be rewarded. Their offspring displayed the same behavioral patterns, providing a tool to investigate mechanistic factors. We measured proteins associated with synaptic presence and function, these efforts ultimately leading us to the hypothesis that melanocortin-4 receptor (MC4R) in the DMS controls response flexibility – defined here as the ability to inhibit instrumental behaviors when they are not fruitful.

Melanocortins are peptide hormones including adrenocorticotropic and melanocyte-stimulating hormones. Of the five melanocortin receptors, two are primarily expressed in the central nervous system – MC3R and MC4R. MC4R is a high-affinity receptor for α-melanocyte-stimulating hormone (α-MSH) and has been intensively studied in the hypothalamus, where its role in energy homeostasis is now well-understood (Anderson et al., 2016; Tao, 2010). Striatal MC4R function has also been investigated for >4 decades, but overwhelmingly focused on the ventral striatum. For instance, melanocortins trigger excessive grooming (Gispen et al., 1975), which is attributable to activity at MC4R in the ventral striatum (reviewed Alvaro et al., 2003). Further, cocaine increases *Mc4r* and synaptic MC4R content in the ventral striatum, where its activity masks the aversive properties of cocaine, and also potentiates drug seeking, sensitization, cocaine-elicited grooming, and compulsive-like behaviors (Alvaro et al., 2003; Gawliński et al., 2020; Hsu et al., 2005a; Xu et al., 2013).

Despite this historical focus on ventral striatal melanocortin function, dorsal striatal levels of MC4R are rich (Alvaro et al., 1996; Kishi et al., 2003; Mountjoy & Wild, 1998), and their function remains incompletely understood. We found that MC4R in the DMS propels reward-

seeking behavior. Meanwhile, moderating MC4R presence via site-selective gene silencing increased the capacity of mice to inhibit nonreinforced responses; this occurs at least in part via interactions with the orbitofrontal cortex (OFC), a cortical brain region canonically involved in modifying action strategies.

2.4 RESULTS

2.4.1 Individual differences in reward-related response strategies in mice

Here we bred mice that displayed particular behavioral traits, with the ultimate goal of creating a tool by which to identify molecular factors controlling animals' propensity to inhibit behaviors that are unlikely to be reinforced with desired outcomes. Fifty-two mice were initially screened. Testing occurred in three stages: training, when mice were trained in operant conditioning chambers to respond on two nose poke ports for food. A third, "inactive" port was never reinforced. Next occurred noncontingent pellet delivery, when pellets associated with one familiar response were delivered regardless of the animals' behaviors (and responding was not reinforced); and then a brief probe test the next day, conducted in extinction, when mice could choose between the intact vs. now-defunct contingencies (Fig. 2.1a). The mice selected for breeding fulfilled two or three of the following criteria: 1) >20% of responses were directed to the inactive nose poke port during training; 2) they failed to reduce responding when pellets were delivered noncontingently (meaning, they generated the same or more responses relative to a session when pellets were delivered contingently); or 3) they failed to prefer the reinforced behavior during the probe test (meaning, they generated the same or more responses on the aperture associated with noncontingent vs. contingent pellet delivery).

In this and all other experiments, mice did not develop side biases during training that could impact later response patterns; thus, response rates on both active nose poke ports are collapsed for simplicity. Means and SEMs of all 52 mice are represented in black in Fig. 2.1b-d, with the individual mice that were bred in symbols at right. Mice could differentiate between active and inactive nose pokes ports during training (Fig. 2.1b). The inset in Fig. 2.1b represents total responses on the inactive port over the entire course of training. Individual points represent mice that generated >20% of all responses on the inactive port and also fulfilled another breeding criterion and thus were bred. The mice selected for breeding were not ultimately distinguishable based on this singular criterion. Thus, it seems unlikely that this behavioral characteristic contributed to later response patterns; it is included merely for transparency.

Next, one response ceased to be reinforced, and pellets associated with that response were provided noncontingently. As a group, mice inhibited responding (Fig. 2.1c); however, not all individuals inhibited the nonreinforced response. Those mice selected for breeding based on this criterion are represented by individual lines, highlighting their marked divergence from the group means. Similarly, in a subsequent probe test, mice as a group preferred the response associated with reinforcement (Fig. 2.1d), but again, some individual mice failed to demonstrate this preference. The mice selected for breeding based on this criterion are represented by individual lines, again highlighting their divergence from the group mean.

Ultimately, 15 mice were selected for breeding, and they generated 6 litters (the F1 generation), which were trained and tested identically, as were their offspring (F2). They were compared to same-age control counterparts (mice of the same strain bred in the laboratory) whose parents had also undergone identical testing. Two mice from each litter were tested, and each litter was considered a single, independent sample (the mean of mice in that litter).

Response rates during training of filial generations did not differ between groups or generations (Fig. 2.1e). Next, one port was occluded, and responses on the remaining port ceased to be reinforced; instead, pellets were delivered noncontingently. Control mice overwhelmingly inhibited responding during this session, relative to a session when the other port was available and responding was reinforced. Meanwhile, response patterns in the experimentally bred mice were less flexible, as can be appreciated in Fig. 2.1f. As an additional example of this phenomenon: Response rates in the control mice in Fig. 2.1f were 4.1-fold higher, on average, when responding was explicitly reinforced than when it was not. Meanwhile, experimental offspring in Fig. 2.1f responded only twice as much on average when responding was reinforced, and they were sufficiently variable such that the contingent vs. noncontingent conditions did not statistically differ (Fig. 2.1f).

Interestingly, experimental offspring throughout consistently favored the reinforced behavior during probe tests conducted a day later. Therefore, our breeding strategy spared contingency memory formation. Our studies thus focus on striatal factors controlling rapid, "in-the-moment" response inhibition, occurring when mice first encounter violated response-reward contingencies.

Next, we tested all progeny of the F3 generation (78 mice) and calculated the proportion of each litter that inhibited nonreinforced responses. The majority of typical offspring inhibited nonreinforced behaviors, as expected, but only about half of animals in each experimental litter inhibited responding when it was not reinforced (Fig. 2.1g).

2.4.2 Individual differences in instrumental response strategies are associated with striatal protein composition

Instrumental response flexibility requires synaptic signaling in the DMS (see Introduction). Thus, we next quantified PSD-95, synaptophysin, and CNPase in the DMS and ventral striatum, for comparison. These proteins are commonly considered markers of the excitatory postsynaptic compartment, the presynaptic compartment, and mature oligodendrocytes, respectively. PSD-95 was lower in mice with poor response flexibility across both regions (Fig. 2.2a), while synaptophysin was unaffected (Fig. 2.2b). CNPase was qualitatively lower in mice with poor response flexibility (Fig. 2.2c-d), but this comparison did not reach significance following Benjamini-Hochberg correction for multiple comparisons.

One additional protein, MC4R, was measured based on the results of an exploratory transcriptomic analysis of the DMS from the F3 generation. MC4R levels did not differ between groups (all *p*s>0.2, not shown). Interestingly, however, protein levels correlated with behavioral response strategies: Specifically, we distilled response strategies down to a single value by dividing response rates generated during the contingent pellet delivery/noncontingent pellet delivery sessions. Scores >1 indicate that response rates were higher when responding was explicitly reinforced than when it was not, while scores ~1 indicate that mice responded equivalently in both conditions. MC4R levels negatively correlated with response ratios (Fig. 2.2e), suggesting that mice with high MC4R fail to inhibit responding that is not reinforced, while mice with low MC4R modify response strategies. Meanwhile, ventral striatal MC4R did not correlate with response patterns (Fig. 2.2f).

2.4.3 MC4R control of action strategies

Our findings predict that inhibiting MC4R presence might facilitate response inhibition. To test this hypothesis, we obtained ‘floxed’ *Mc4r* mice, a well-established tool in MC4R research,

in which the single coding exon is flanked by loxP sites, and the introduction of Cre-recombinase (Cre) obstructs MC4R production (Sohn et al., 2013). Cre was delivered selectively to the DMS via CaMKII-driven adeno-associated viral vectors (Fig. 2.3a). *Mc4r* status did not affect response rates during training (Fig. 2.3b), important given that global knockout can reduce operant response rates for food (Cui et al., 2012), and suggesting that gross locomotor activity did not differ between groups.

Next, one nose poke behavior failed to be reinforced, and instead, pellets were delivered noncontingently. We extracted response rates in bins to compare groups across time. Response rates increased as animals first experienced the contingency violation, resembling a so-called "extinction burst," as previously reported in mice performing the same task (Zimmermann et al., 2016). All mice ultimately inhibited responding with time, though, importantly with *Mc4r* knockdown mice responding less overall (Fig. 2.3b).

To further solidify our interpretation that site-selective *Mc4r* knockdown facilitates response inhibition, we reinstated responding in *Mc4r*-deficient mice, then tested their behavioral sensitivity to reinforcer devaluation. In this case, mice will inhibit responding for a devalued outcome. Mice were given free access to one of the two reinforcer pellets in a clean cage, followed by an injection of LiCl, inducing transient malaise and decreasing the value of that pellet via conditioned taste aversion (CTA). The other pellet was paired with NaCl. With repeated pairings, typical mice will inhibit the behavior that leads to the LiCl-paired, devalued outcome, while responding for the NaCl-paired pellet will remain intact – reflecting response plasticity based on reward value [for discussion of reinforcer devaluation, see Balleine & O'Doherty, 2010]. We hypothesized that *Mc4r* knockdown mice would more readily inhibit responding than control mice. To generate the resolution to detect such an effect, we tested response strategies at two time points:

after only a few LiCl pairings, before pellet aversion was strong, and following more pairings, when it was robust (arrows, Fig. 2.3c). We envisioned that this approach might allow for the resolution to detect enhancements in response inhibition, if they existed.

Upon CTA, mice decreased ad libitum consumption of the LiCl-associated pellet, but not NaCl-paired pellet, as expected (Fig. 2.3c). When returned to the conditioning chambers at the early time point, control mice showed no evidence yet of changing response strategies, indicated by equivalent responding on the ports associated with the valued vs. devalued outcomes. Meanwhile, a majority of knockdown mice (73%) favored the response associated with the valued outcome (Fig. 2.3d). Thus, knockdown enriched response plasticity, triggering mice to inhibit a behavior associated with devalued food.

Group differences can be further appreciated by converting response rates to ratios: valued/devalued. Scores >1 reflect preference for the port associated with the valued pellet and neglect of the devalued pellet, while scores of ~1 indicate no change in behavior based on outcome value. As expected, knockdown mice generated higher ratios early in conditioning, while control mice required more CTA to generate response preferences (Fig. 2.3e). Thus, reducing striatal *Mc4r* expedites the ability of mice to inhibit actions when appropriate.

Importantly, following both probe tests, we assessed the propensity of mice to consume freely available pellets placed in their cages. At both time points, both groups consumed far more of the pellet that had been paired with NaCl, relative to the pellet that had been paired with LiCl (Fig. 2.3f). Thus, instrumental response strategies could not be attributable to differences in CTA.

Given that hypothalamic *Mc4r* controls feeding, and our tasks are food-reinforced, it was also important to measure general food intake following DMS-specific knockdown. Ad libitum chow intake and body weights did not differ between groups (Fig. 2.3g-h).

2.4.4 MC4R control of action strategies via the OFC

MC4R presence controls the localization of GluA2-containing AMPA receptors (AMPARs) at the cell membrane of striatal medium spiny neurons (MSNs). Specifically, MC4R binding triggers internalization of these receptors (Lim et al., 2012), leading to the hypothesis that MC4R presence may control response strategies by gating sensitivity to excitatory inputs. Implicit in this model is that behavioral effects of *Mc4r* silencing are dependent on glutamatergic afferents to the DMS.

To begin to identify projections that might be important for MC4R-controlled behavior, we returned to our original population of experimentally bred response-inflexible mice and quantified dendritic spine densities on distal dendritic segments – considered highly labile (McEwen & Morrison, 2013) – as a general measure of neural plasticity, akin to measuring immediate-early gene expression. Densities on excitatory layer V OFC neurons (ventrolateral subregion) were higher in response-inflexible mice vs. age-matched controls (Fig. 2.4a), but not in prelimbic, infralimbic, or hippocampal CA1 regions (Fig. 2.4a).

Next, we classified dendritic spines into their primary subtypes, including mushroom-shaped spines, which are considered mature, stable, and synapse-containing, compared to thin- or stubby-shaped spines, which by contrast are immature and functionally variable (Berry & Nedivi, 2017). Mice that failed to inhibit responding when pellets were delivered noncontingently (contingent/noncontingent scores <1) had more immature, thin-type spines. Meanwhile, mice that did inhibit responding (scores >1) were considered resilient (Fig. 2.4b) and had more mature, mushroom-shaped spines on OFC neurons (Fig. 2.4c). Thin-type spine densities also correlated with response strategies in 2 independent cohorts of mice (Fig. 2.4d). Thus, poor response

inhibition is associated with immature spine types, while successful strategy shifting is associated with mature spine types in the OFC, leading to the hypothesis that the OFC is part of a network controlling response inhibition.

OFC-to-DMS inputs are organized largely ipsilaterally in the brain, including in mice (Zimmermann et al., 2017). We took advantage of these segregated projections to use a "disconnection" design to test the possibility that connections with the OFC were necessary for the behavioral flexibility conferred by silencing *Mc4r* in the DMS. Here, we reduced *Mc4r* unilaterally in one DMS and placed Gi-coupled Designer Receptors Exclusively Activated by Designer Drugs (DREADDs) unilaterally in one OFC (Fig. 2.4e). When infusions are ipsilateral and the DREADDs ligand Clozapine N-oxide (CNO) is delivered, one DMS lacks *Mc4r*, which should improve response inhibition, but it is devoid of the typical OFC signal. We thus anticipated that this group would resemble mice with control viral vectors. Meanwhile, in the contralateral ("asymmetric") group, mice also experience unilateral OFC inactivation, but the healthy OFC is projecting to an *Mc4r* knockdown DMS. If these OFC-to-DMS connections can account for response inhibition following *Mc4r* knockdown, we reasoned that this group should be better able to inhibit responding when food is delivered noncontingently, relative to the control groups.

OFC-targeted infusions were largely contained within the ventrolateral region, and terminals were detected in the DMS, overlapping with areas in which *Mc4r* was reduced (Fig. 2.4f). In the control group, some spread into the ventral striatum was noted (Fig. 2.4f) but did not have obvious consequences. Groups did not differ during response training, conducted in the absence of CNO (Fig. 2.4g).

When one familiar behavior failed to be reinforced, and instead, pellets were delivered noncontingently, the contralateral group generated the lowest response rates (Fig. 2.4g), differing

from mice bearing control viral vectors in the final three time bins. Importantly, while the ipsilateral mice responded less than control mice during the third time bin, this difference was transient and they ultimately were not as adept at inhibiting nonreinforced behaviors as the contralateral group (Fig. 2.4g). These patterns together suggest that response inhibition conferred by *Mc4r* silencing in the DMS requires input from the ventrolateral OFC.

2.5 DISCUSSION

Here we trained mice to generate two responses in operant conditioning chambers for food reinforcers. We then uncoupled the predictive relationship between one response and its outcome by providing food pellets noncontingently and responding was not reinforced. Typically, mice inhibit that response and favor the other, but individual differences exist, such that a minority of mice here failed to readily inhibit familiar behaviors, even when those behaviors were not explicitly reinforced. We bred these mice, generating offspring with the same tendencies. By thereby generating large numbers of mice that failed to readily inhibit reward-seeking behaviors, we were able to resolve correlations between MC4R in the DMS and response strategies. These patterns led to experiments revealing that MC4R presence in the DMS propels reward-seeking behavior, while reducing MC4R expedites response inhibition, an effect that relies, at least in part, on OFC input.

What might account for transgenerational response biases? We used transgenic mice expressing YFP and bred on an inbred C57BL/6 background, which makes genetic variation unlikely. Experimental mice were compared to the offspring of other C57BL/6 mice bred in our lab that had also been behaviorally tested; thus, epigenetic effects of behavioral testing, writ large, are also unlikely. Conceivably, other epigenetic effects and/or familial factors could play a role.

We did not observe gross differences in maternal behavior when quantified during the light cycle (Allen et al., 2022), but potentially, maternal care differed between groups during the dark cycle, which could propel behavioral differences in adulthood. These and other possibilities could be investigated in the future. Our present goal was to amplify individual differences in response inhibition capacity by breeding response-inflexible mice and thereby creating a tool by which to better understand the neurobiology of instrumental behavior.

Several independent investigations indicate that the DMS is necessary for rodents to modify familiar reward-seeking behaviors (Bradfield et al., 2013; Braun & Hauber, 2012; Lex & Hauber, 2010; Pauli, Clark, et al., 2012; Yin, Ostlund, et al., 2005). These observations motivated us to measure synaptic markers in the DMS of experimentally bred, response-inflexible mice. PSD-95, a post-synaptic marker associated with synaptic strength (Béïque & Andrade, 2003), was lower than in typical mice. Meanwhile, synaptophysin, a presynaptic marker associated with synapse density (Navone et al., 1986), was unaffected. Less PSD-95 thus likely reflects weaker excitatory synapses in the DMS, rather than the loss of inputs from extra-striatal regions, per se.

Striatal CNPase, a marker of mature oligodendrocytes, was also quantified. Once considered merely an insulator of neurons, oligodendrocytes are dynamic, sensitive to stressors, alcohol, motor skill learning, and electrical and synaptic activity (McKenzie et al., 2014; Mensch et al., 2015; Nickel & Gu, 2018; Wake et al., 2015). It appeared that experimental breeding reduced CNPase, but this effect did not survive correction for multiple comparisons.

Next, we quantified MC4R, the high-affinity receptor for α-MSH, a peptide produced by proopiomelanocortin (POMC)-expressing neurons in the arcuate nucleus of the hypothalamus. Levels of MC4R in the DMS correlated with response strategies, such that high levels were associated with pursuit of familiar response strategies. Meanwhile, mice with low levels

demonstrated response flexibility, reminiscent of evidence that low *Mc4r* confers resilience to compulsive-like behavior (Xu et al., 2013).

These patterns led us to test MC4R function in the DMS using viral-mediated site-selective *Mc4r* gene silencing. Reducing MC4R expediated response inhibition, enriching the capacity of mice to restrain behaviors that were not reinforced. We also tested the capacity of mice to modify behavior based on reward value. We reasoned that if silencing *Mc4r* enriches response plasticity, then *Mc4r*-deficient mice would more rapidly inhibit responding when a reward lost value. Indeed, inhibiting MC4R in the DMS conferred response flexibility, since *Mc4r*-deficient mice more rapidly inhibited responding when foods were devalued than control mice.

Why might melanocortin-MC4R action in the DMS propel familiar reward-seeking behaviors? In the striatum, MC4R preferentially expresses on dopamine D1 receptor (D1R)-containing medium spiny neurons (MSNs) (H. Cui et al., 2012; Hsu et al., 2005a; Oude Ophuis et al., 2014). MC4Rs, like D1Rs, are positively coupled to the cAMP second messenger cascade (Gantz et al., 1993; Mountjoy & Wild, 1998), and thus can enhance D1R function (Lezcano et al., 1995). D1R stimulation is necessary for learning new skills (Yin et al., 2009), and D1R+ MSNs in the DMS are involved in the development of goal-directed action strategies (Peak et al., 2020) – a process that requires inhibiting unproductive behaviors – and recalling memories linking actions and outcomes (Renteria et al., 2021). *Mc4r*-null mice are delayed in learning to nose poke for food, and restoration of MC4R in D1R-containing cells reinstates this capacity (H. Cui et al., 2012). Possibly, MC4R+D1R stimulation synergistically attunes mice to actions predictive of reward, particularly when learning new tasks, thus propelling those actions. Conceivably, high levels of MC4R (as in inflexible mice) could overly drive reward-seeking behaviors at the expense of adaptive response plasticity.

Why might reducing *Mc4r* facilitate response inhibition? MC4Rs regulate GluA2 AMPAR subunit availability at the membrane. α-MSH-MC4R binding triggers GluA2 internalization (Lim et al., 2012). Meanwhile, decreasing MC4R enhances glutamatergic signaling in the striatum (Xu et al., 2013). Given that dopamine agonists increase POMC, the precursor for α-MSH (Tong & Pelletier, 1992), and cocaine increases striatal α-MSH content (Sarnyai et al., 1992), rewarding events may result in α-MSH-MC4R binding. This binding would cause GluA2-AMPAR internalization, decreasing the synaptic sensitivity of DMS MSNs to cortico-striatal glutamatergic afferents, which otherwise trigger response plasticity and suppression in many contexts (Gremel et al., 2016; Gremel & Costa, 2013; Hart et al., 2018). Thus, reducing MC4R levels or activity would increase sensitivity to cortico-striatal projections that might trigger response inhibition when adaptive.

Implicit in this model is that the apparent "pro-flexibility" effects of *Mc4r* silencing depend on glutamatergic input to the DMS. We attempted to identify likely sources of inputs, first returning to our original experimentally bred response-inflexible mice. We quantified dendritic spines on terminal dendrites in multiple brain regions, because terminal dendrites are highly plastic and can be viewed as a general proxy of neural plasticity – conceptually similar to measuring immediate-early gene expression (McEwen & Morrison, 2013). Response-inflexible mice had higher densities of thin-type dendritic spines on excitatory neurons in the OFC, which are unstable and typically pruned with instrumental conditioning (Whyte et al., 2019). Meanwhile, dendrites from response-flexible mice hosted more mature, mushroom-shaped spines. Notably, we found no obvious group differences on dendrites in the PL, IL, or hippocampal CA1, even while neuronal structural plasticity in the PL, for example, has been associated with instrumental response strategies in the same task (Swanson et al., 2017). Further, stress-induced failures in response

flexibility in a very similar task are associated with dendritic spine loss on proximal branches of apical PL dendrites (and also loss of terminal branches (Dias-Ferreira et al., 2009)). A key difference, though, is that the majority of investigations into dendritic spine densities, particularly ex vivo investigations, focus on dendritic segments at some fixed distance from the soma, while we instead imaged distal, terminal tufts, which are considered more plastic and subject to in-the-moment events and stimuli. Putting the pieces together, then, we might imagine that previously reported modifications in the PL could reflect long-term changes (for instance, associated with initially learning action-reward contingencies), rather than acute effects (for instance, of detecting the violation of learned rules).

We next hypothesized that excitatory plasticity in the OFC may be involved in response flexibility conferred by moderating MC4R tone in the DMS. The OFC and DMS are connected by unidirectional projections organized largely ipsilaterally in the brain (Zimmermann et al., 2017). We capitalized on this anatomical organization and infused into the OFC of one hemisphere inhibitory Gi-coupled DREADDs. In the ipsilateral or contralateral DMS, *Mc4r* was reduced. In the ipsilateral condition, one DMS had less *Mc4r*, but was deprived of typical OFC input – we anticipated that these mice would resemble control mice (those bearing control viral vectors). Meanwhile, in the contralateral condition, mice had the same manipulations, but the DMS that had less *Mc4r* received input from the OFC. If OFC input on striatal neurons with low MC4R optimizes adaptive response inhibition – as we predicted – we expected that this group would be best able to inhibit responding. This was indeed the case. Thus, reducing *Mc4r* appears to facilitate response plasticity at least in part via OFC input.

A final note is that MC4R levels in the ventral striatum did not correlate with response patterns here. This outcome was interesting, given that the ventral striatum is more strongly

innervated by α-MSH-containing projections from the arcuate nucleus than the DMS (Lim et al., 2012). MC4R antagonism and gene silencing in the ventral striatum mitigate cocaine-seeking, anhedonic-like, and compulsive-like behaviors (Gawliński et al., 2020; Hsu et al., 2005a; Lim et al., 2012; Xu et al., 2013), and ventral striatal MC4R controls approach and avoidance of both appetitive and aversive stimuli (Klawonn et al., 2018). Altogether, then, it appears that ventral striatal MC4R stimulation promotes drug seeking and compulsion, while MC4R activity in the DMS appears to propel reward-seeking behaviors. Meanwhile, inhibiting MC4R appears to combat drug seeking and anhedonic-like behavior and promote the capacity for behavioral inhibition – qualities that could be favorable in treating addictions and other illnesses.

2.6 METHODS

2.6.1 Subjects

Initial experiments bred mice with particular behavioral traits and tested their offspring. These mice were maintained on a C57BL/6 background and expressed Thy1-driven YFP (Feng et al., 2000, H line, Jackson Labs), allowing us to visualize neurons and enumerate dendritic spines in some experiments. In experiments in which we manipulated *Mc4r*, mice were homozygous for a 'floxed' *Mc4r* gene (Sohn et al., 2013, Jackson Labs). These mice were maintained on a mixed C57BL/6J-129S1/SvlmJ background.

Mice were weaned from the dam at or soon after postnatal day (P) 21 and housed in single-sex cages with siblings or unrelated mice of the same age. Mice were maintained on a 12-hour light cycle (0700 on) and provided food and water ad libitum except during food-reinforced behavioral testing when food was restricted to motivate responding. Experiments used both sexes. Procedures were approved by the Emory University IACUC.

2.6.2 Ages of mice at testing

Behavioral testing used to identify mice for breeding was initiated between postnatal day (P) 27-30. Once identified, mice were paired with opposite-sex counterparts at or soon after P56. In other experiments, mice were >P56 at the time of testing and behaviorally naïve.

2.6.3 Test of action strategies

Mice were food restricted to motivate food-reinforced responding. In young mice, body weights were maintained at 100% of the expected growth curves for C57BL/6 mice (Jackson Labs) to maintain animals' health. In mature (>P56) mice, body weights dropped to ~93% of their free-feeding weight. Operant conditioning chambers (Med-Associates) were equipped with 3 nose poke ports, as well as a separate food magazine. Responding on 2 of the ports was reinforced with food pellets (20 mg, Bio-serv) using a fixed ratio 1 (FR1) schedule of reinforcement. Up to 30 pellets were available for responding on each port, resulting in 60 pellets/session. Sessions ended when 60 pellets were delivered or at 70 min., whichever came first. Mice did not develop side or pellet preferences, and response acquisition curves represent both nose poke responses/min. Nose poke training occurred over 7-9 days, with 1 session/day.

Next, one port was occluded, and responding on the other had no programmed consequences. Instead, pellets were delivered into the magazine at a rate matched to each animal's reinforcement rate from the previous day (i.e., pellets were delivered "for free"). Thus, the response-reward relationship linking this nose poke and reward was violated, which typically causes mice to cease responding at this port. This session is referred to as the "noncontingent" session. A 25-min. "contingent" session served as a control; here, the other nose poke port was

available, and responding remained reinforced according to an FR1 schedule of reinforcement. The location of the “noncontingent” port within the chamber was counter-balanced.

In mice screened for breeding, we also assessed responding the next day, during a brief probe test, in which both ports were available for 10-min. Responses were recorded but not reinforced. Groups did not differ during this phase, so responding by two cohorts is shown, but not for others.

2.6.4 Breeding strategy

Mice were paired for breeding if they fulfilled 2/3 of the following criteria: 1) >20% of total responses occurred on the inactive port during response training; 2) they failed to inhibit responding during the “noncontingent” session relative to “contingent” session; or 3) they failed to prefer the “contingent” nose poke during the probe test. We first behaviorally characterized 52 mice. Fifteen mice created the parental generation, and their offspring created the F1 generation, which was then tested as its parents were. They were compared to same-age, same-strain mice whose parents had also been behaviorally tested. Mice were again selected for breeding based on the above-described criteria, and their offspring created the F2 generation, which was tested as its parents were. Following the F1 generation, care was taken to ensure that siblings were not bred. These mice are represented in Fig. 2.1. For subsequent studies, experimental mice were the offspring of mice that had been selected for breeding as described above. Control mice were age- and strain-matched mice bred in our colony.

2.6.5 Reinforcer devaluation

One group of mice tested in the above-described behavioral assay was next used in a devaluation experiment. Mice had one re-training session according to an FR1 schedule of reinforcement for 70 min. to reinstate responding on both nose poke ports. As above, responding on two ports was reinforced with either a grain-based or chocolate-flavored pellet (20 mg, Bio-serv). Mice did not display systematic pellet preferences, as can be seen in the associated figure.

Conditioned taste aversion (CTA) was then used to decrease the value of one of the pellets. Mice were placed individually in clean cages with free access to one of the two pellets. After 60 min., mice were injected with lithium chloride (LiCl; 0.15 M in saline, 4 ml/100 g, i.p., Sigma), which induces temporary gastric malaise. The following day, mice were given ad libitum access to the other pellet for 60 min., followed by a vehicle injection (NaCl). Mice experienced 6 pairing sessions/pellet across 12 days. Pellet intake was measured and compared between groups and conditions.

Our hypothesis was that DMS-selective *Mc4r* knockdown would enhance the ability of mice to inhibit responding. To test this possibility, we placed mice in the conditioning chambers for a probe test after only 3 CTA pairings (15 min., conducted in extinction), before mice developed robust CTA. The idea was that this timing would allow us the resolution to detect enhanced performance, if it indeed existed. The probe test was then repeated following all 6 pairings to confirm that CTA would, with sufficient training, reduce responding for the LiCl-paired pellet as expected.

After both probe tests, mice were placed individually in a clean cage with an abundant, equivalent supply of both pellets, allowing them to freely consume pellets. Remaining pellets were measured after 60 min. to quantify ad libitum intake. The point of this measure is to confirm that

CTA is effective, and thus, behavioral responding in the probe test reflects the propensity (or not) of mice to modify behaviors based on goal features.

2.6.6 Intracranial surgery and viral vectors

Mc4r-flox mice were anesthetized via ketamine (100 mg/kg, i.p.) and dexmedetomidine (0.5 mg/kg, i.p.). Mice were administered the analgesic meloxicam (5 mg/kg, s.c.) and revived using atipamezole (1 mg/kg, i.p.). Drugs were dissolved in saline and administered in a volume of 1 ml/100 g.

For DMS infusions, adeno-associated viral vectors (AAV8) expressing Green Fluorescence Protein (GFP) ± Cre-Recombinase (Cre) with a CamKIIα promotor were supplied by the UNC Viral Vector Core. Viral vectors were infused at a rate of 0.1 µl/min., with a total volume of 0.5 µl, at +0.5 mm anteroposterior (AP), -4.5 mm dorsoventral (DV), +/-1.6 mm mediolateral (ML) relative to Bregma. The micro-syringe was left in place for 5 min. following infusion.

In some experiments, viral vectors were also delivered to the OFC. For OFC infusions, mice received unilateral infusions of AAV5-CaMKIIα-mCherry ± hM4D(Gi) (UNC Viral Vector Core) in the ventrolateral OFC (0.5µl/infusion over 5 min. at AP+2.6, ML±1.2, DV-2.8). Simultaneously, they received unilateral infusions of AAV ± Cre into the DMS as above. Infusions were either ipsilateral or contralateral. The micro-syringes were left in place for 5 additional min. prior to withdrawal and suture. The ipsilateral and contralateral control groups (i.e., mice that received the control viral vector in the OFC and DMS) did not differ and were combined for statistical and graphical purposes. For general description of DREADDs, see Urban & Roth (2015). Mice were allowed >3 weeks for recovery and viral vector expression.

2.6.7 CNO administration and timing in DREADDs experiments

Mice with DREADDs were trained to nose poke as described, and they received injections of saline 30 min. before the instrumental training sessions to habituate them to injection stress. Then, CNO (Sigma) was delivered at 1 mg/kg, i.p., dissolved in 2% DMSO and saline (1 ml/100 g) 30 min. before the "noncontingent" session of our procedure. All mice received CNO, regardless of condition, to equally expose animals to any unintended consequences of CNO (Gomez et al., 2017).

2.6.8 Assessments of food intake

To determine whether reducing *Mc4r* in the DMS impacted free-feeding behaviors, we reduced *Mc4r* in the DMS bilaterally, and we then assessed food intake using established methods (Ellacott et al., 2010; Huszar et al., 1997; M. M. Li et al., 2019). Mice were single-housed for 2 weeks prior to the experiment. Mice were given ad libitum standard chow and water. Baseline body weight was collected, and then body weight and food intake were subsequently measured daily for 7 days, 3 hours after lights on.

2.6.9 Histology

Following testing, mice with viral vectors were euthanized either by decapitation following brief anesthesia with isoflurane or more commonly, by deep anesthesia with ketamine/xylazine (100 and 10 mg/kg, i.p.), followed by intracardiac perfusion with chilled saline and 4% paraformaldehyde. Brains were soaked in 4% paraformaldehyde for 48 hours, then transferred to 30% w/v sucrose, and sectioned into 40-50 μm-thick sections on a freezing microtome. Tissues

were plated, then imaged using a fluorescence microscope. If infusions were not contained within the DMS or OFC, mice were excluded.

2.6.10 Immunoblotting

Mice had been trained and tested in the first behavioral task described above. They were returned to free-feeding and left undisturbed for roughly 1 week. Then, they were briefly anaesthetized with isoflurane and euthanized by rapid decapitation, and brains were extracted and frozen at -80ºC. Brains were sectioned into 1 mm coronal sections using a chilled brain matrix, and punches aimed at the DMS and ventral striatum were extracted using tissue corers. Ventral striatal tissue extractions took care to avoid the anterior commissure, and some were unintentionally lost. Tissues were homogenized by sonication in lysis buffer [200 µl: 137 mM NaCl, 20 mM tris-Hcl (pH=8), 1% NP-40, 10% glycerol, 1:100 Phosphatase Inhibitor Cocktails 2 and 3, 1:1000 Protease Inhibitor Cocktail (Sigma)] and stored at -80ºC. Protein concentrations were determined using a Bradford colorimetric assay (Pierce).

Equal amounts of protein (15 µg) were separated by SDS-PAGE on 7.5% or 4-20% gradient Tris-glycine gels (Bio-rad). Following PVDF membrane transfer, blots were blocked with 5% nonfat milk or 5% BSA for 1 hour. Membranes were incubated with primary antibodies at 4ºC overnight and then incubated in horseradish peroxidase secondary antibodies for 1 hour. Primary antibodies were PSD-95 (Ms, Cell Signaling #3450, 1:1000), Synaptophysin (Rb, Abcam #32127, 1:20,000), CNPase [Ms, Millipore (multiple tested), 1:1000], or MC4R (Rb, Abcam #150419, 1:1000).

Immunoreactivity was assessed using a chemiluminescence substrate (Pierce) and measured using a ChemiDoc MP Imaging System (Bio-rad). Densitometry values were

individually normalized to the corresponding loading control (HSP-70; Ms, Santa Cruz Biotechnology #7298, 1:5000), which did not change as a function of breeding, and then normalized to the control sample mean from the same membrane in order to control for fluorescence variance between gels.

2.6.11 Dendritic spine imaging and reconstruction

Mice had been trained and tested in the first behavioral task described above. Roughly 24 hours later, mice were briefly anaesthetized by isoflurane and euthanized by rapid decapitation. Brains were submerged in chilled 4% paraformaldehyde for 48 h, then transferred to 30% w/v sucrose, and sectioned into 40-50 µm-thick sections on a freezing microtome. Mice carried Thy1-driven YFP, resulting in YFP expression in layer V cortical neurons and hippocampal CA1. Z-stacks were collected with a 100X 1.4 numerical port objective using a 0.1 µm step size on a spinning disk confocal (VisiTech International) on a Leica microscope. 6-10 segments/mouse were imaged. They ranged from 19-31 µm in length. Experimenters were blind to group in all experiments.

Experiment 1. Multi-site quantification of dendritic spine densities. Dendritic segments in the prelimbic prefrontal cortex (PL), infralimbic prefrontal cortex (IL), ventrolateral OFC, and dorsal hippocampal CA1 were imaged, with The Mouse Brain in Stereotaxic Coordinates (Franklin & Paxinos, 2001) as reference. We endeavored to image terminal segments, which are considered highly plastic. Dendritic spines were manually counted, normalized to the length of the dendrite (spines/µm), and each mouse contributed a single value (its mean density) to comparisons.

Experiment 2. Defining individual differences in dendritic spine densities and morphologies. Response inhibition in our decision-making task triggers the elimination of thin-

type dendritic spines in the ventrolateral OFC, increasing the proportion of mushroom-shaped spines (Whyte et al., 2019). We thus characterized dendritic spine morphologies in several mice that had been behaviorally characterized. We separated mice by those that failed to inhibit responding when pellets were delivered noncontingently (contingent/noncontingent scores <1) vs. mice that did inhibit responding (scores >1) for comparisons. Using ImageJ, dendritic spines were enumerated. Also, dendritic spine heads were traced at the widest point, and the length of each spine was collected, allowing us to classify spines into their primary subtypes. Dendritic spines with heads >0.35 μm in diameter and >0.45 μm in length were considered mushroom-like, while dendritic spines that were >0.45 μm in length with heads smaller than 0.35 μm in diameter were considered thin-type. Spines <0.45 μm in length were considered stubby. Again, each mouse, rather than each dendrite, was considered an independent sample.

2.6.12 Statistics and reproducibility

Our initial experiment contained 52 mice, each considered an independent sample. The parental generation from this experiment created the F1 generation. A male and female from each F1 litter were tested. In the rare instances that the litters contained only 1 sex, then 2 mice of the same sex were tested. Here, each litter was considered an independent sample, reflecting the mean of the 2 mice tested from that litter. The same approach was taken with the F2 generation. With the F3 generation, we tested all mice in a litter, then calculated the proportion of mice that were able inhibit a nonreinforced response (that is, they generated at least one fewer response when pellets were delivered noncontingently relative to a session of the same duration when pellets were delivered contingently). Each litter contributed one proportion value to the comparison. In

subsequent experiments, experimentally-bred mice were derived from independent litters and treated as independent samples.

In our initial experiment (Fig. 2.1), response rates during training were compared by ANOVA with repeated measures, then response rates between the contingent vs. noncontingent response conditions were compared by paired t-tests. Proportions in Fig. 2.1g were compared by unpaired t-test. In subsequent experiments, response rates, body weights, food intake, and maternal care counts were compared by multi-factor ANOVA, with repeated measures when appropriate. In the case of interactions or main effects between >2 groups, post-hoc comparisons used Tukey's tests; all possible comparisons were made, and any significant differences are reported. Alpha was set at 0.05.

Western blot values were compared by or 1- or 2-factor ANOVA. Dendritic spine densities were compared by unpaired t-tests or 2-factor ANOVA. These exploratory comparisons (Fig. 2.2, 2.4) were subject to the Benjamini-Hochberg Procedure for correcting for multiple comparisons, with a false discovery rate of 5%.

Western blot values and dendritic spine densities were also compared by linear regression against response preference scores – the response rates in the contingent/noncontingent conditions. Scores >1 reflect inhibition of the nonreinforced behavior, while scores at <1 reflect no change in response strategies.

Exclusions: Values >2 standard deviations outside of the mean were considered outliers; thus, one mouse from each group in the "disconnection" experiment in the final figure generated multiple outlying values during training and was excluded. Histological analyses of viral vector placements resulted in the additional exclusion of 1 control, 3 ipsilateral, and 2 contralateral mice from the same experiment. Finally, 1 mouse in the delay discounting procedure did not nose poke

and was excluded. Final *n*'s are reported in the figure captions. SPSS v.28 and SigmaPlot v.11 were used to analyze data.

2.7 FUNDING

This work was supported by NIH R01DA044297, R01MH117103, T32NS96050, T32GM008602, NSF 1937971, the Brain and Behavior Foundation (NARSAD), the Emory University Research Council, and Children's Healthcare of Atlanta. The Emory National Primate Research Center is supported by NIH P51OD011132.

2.8 ACKNOWLEDGEMENTS

We thank Ms. Courtni Andrews and Dr. Elizabeth Hinton for their assistance. We thank Dr. Hasse Walum for critical insights and feedback.

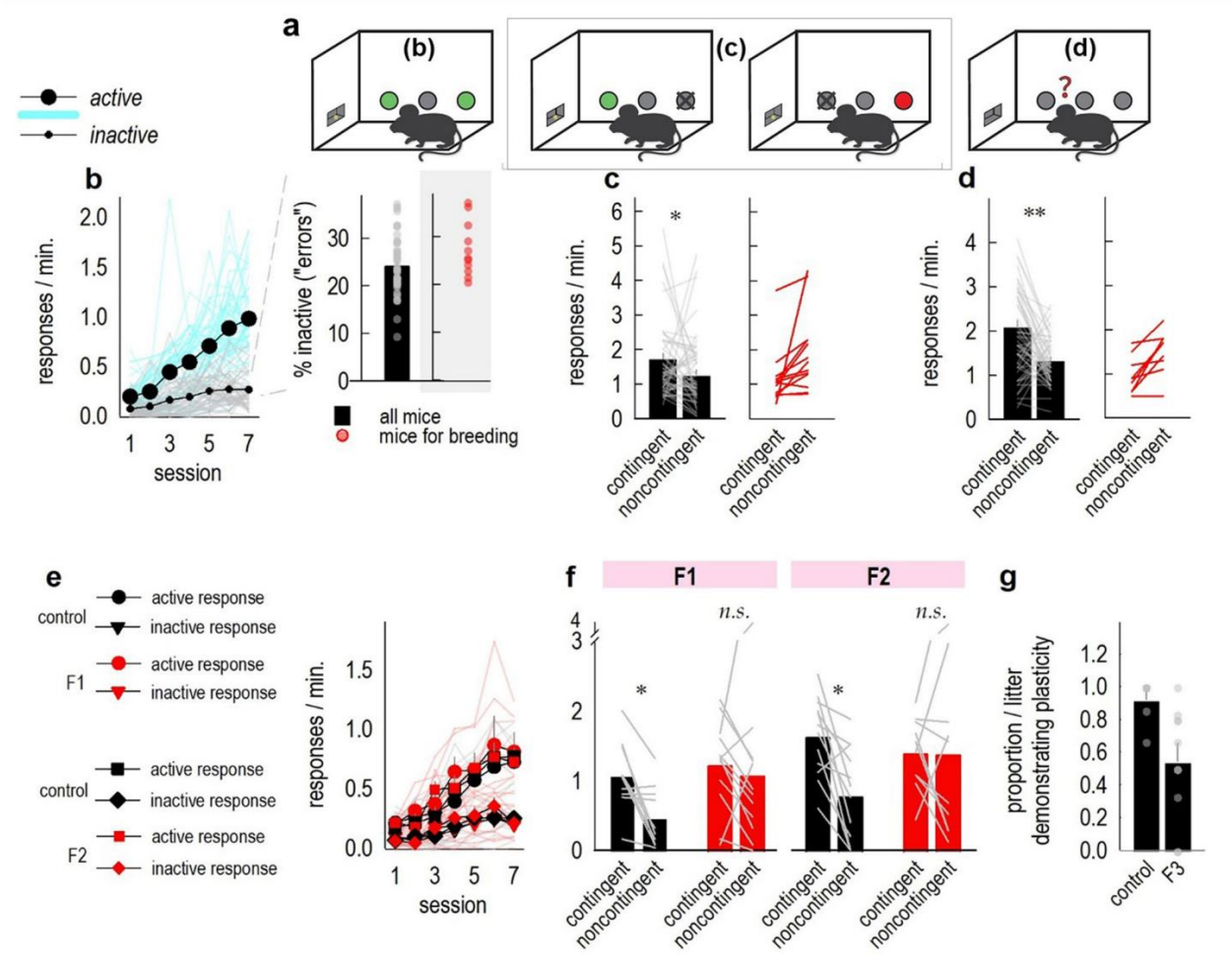

Figure 2.1 Multi-generational biases in reward-related response strategies. a, Task schematic. Responding at two ports resulted in food pellet delivery. Next, we provided pellets associated with one response noncontingently and responding was not reinforced, while responding at the other port remained reinforced. Finally, response preference was assessed during a probe test (right). Small letters in the boxes correspond with the figures below. **b**, In an initial screen of 52 mice, all mice acquired the reinforced ("active") nose poke responses, relative to responding that was not reinforced ("inactive"; active × day $F_{6,300}$=61.98, p<0.001). Inset: Inactive responses as a percentage of all responses. Mice selected for breeding based on this criterion are shown in individual points at right (n=12 selected). **c**, Next, one response was no longer reinforced, and food pellets were delivered randomly. Mice as a group generated lower response rates during

this session, relative to a session when the other response remained reinforced (t_{51}=-2.53, p=0.02). However, some individual mice did not follow this pattern; those selected for breeding are shown at right (n=9 selected). **d**, During a probe test, mice as a group again generated higher response rates on the port associated with reinforcement (t_{51}=4.04, p<0.001]. But again, not all mice followed this pattern; those selected for breeding are shown at right (n=14 selected). Mice deviating from the typical pattern of responding on 2/3 measures were bred. **e**, The F1 and F2 generations acquired reinforced nose poke responses (main effect of day $F_{6,108}$=14.4, p<0.001; no main effects of group, generation, or interactions Fs<1). A port × day interaction indicated that mice differentiated between the active and inactive ports with time ($F_{6,108}$=30.2, p<0.001; main effect of port $F_{1,18}$=129.7, p<0.001). **f**, When responding was not reinforced and pellets were delivered noncontingently, experimental offspring as a group did not modify their behavior relative to a session when responding was reinforced (interaction $F_{1,18}$=4.3, p=0.05; main effect of contingency $F_{1,18}$=6.7, p=0.02). We detected no effects of generation (no main effect $F_{1,18}$=2.4, p=0.18; no generation × contingency or generation × group interaction Fs<1), suggesting that the behavioral phenotype was stable. Control F1 litter n=5, F1 litter n=6, control F2 litter n=6, F2 litter n=5. **g**, Finally, in the F3 generation, we tested all possible progeny and calculated the proportion of each litter than inhibited responding when it was not explicitly reinforced. The majority of control mice displayed this capacity, while only roughly half of experimental offspring did. Control F3 litter n=6, F3 litter n=10. Individual symbols and gray lines represent individual mice. Bars and closed symbols represent means (+SEMs). *p<0.05, **p<0.001, *n.s.* non-significant. F1, F2, and F3 refer to filial 1, 2, 3 generations. Illustration by Aylet Allen.

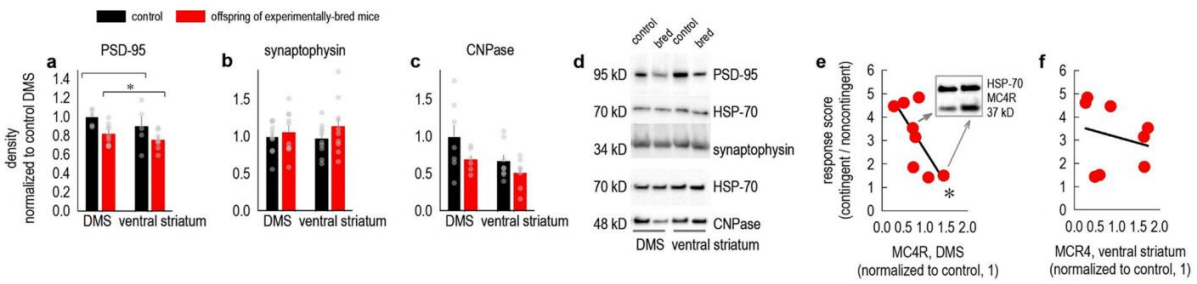

Figure 2.2 Individual differences in response strategies associate with striatal protein content. a, PSD-95 was measured in the dorsomedial and ventral striatum, revealing lower levels in the offspring of experimentally bred mice (main effect $F_{1,23}$=9.6, p=0.005; no other effects Fs<1). n=5-9/group. **b**, Meanwhile, levels of the presynaptic protein synaptophysin did not differ between groups or regions (all Fs<1). n=10-11/group. **c**, The oligodendrocyte marker CNPase appeared lower in the offspring of experimentally-bred mice (effect of group $F_{1,24}$=4.1, p=0.05; effect of brain region $F_{1,24}$=4.9, p=0.04; no interaction F<1), but the effect did not survive correction for multiple comparisons. n=6-8/group. **d**, Representative blots loaded in the order indicated, including corresponding HSP-70 loading controls. **e**, MC4R levels correlated with response scores, such that higher MC4R was associated with response inflexibility (scores ~1), and lower levels were associated with response plasticity (scores >1; r=0.71, p=0.047). Representative blots are inset, with arrows linking each lane to the respective mouse. **f**, MC4R in the ventral striatum of the same mice did not correlate with response strategies (r=0.23, p=0.59). n=8. Bars represent means + SEMs. Symbols represent individual mice. *p<0.05. DMS refers to the dorsomedial striatum. All gels were run at least twice, with concordant results.

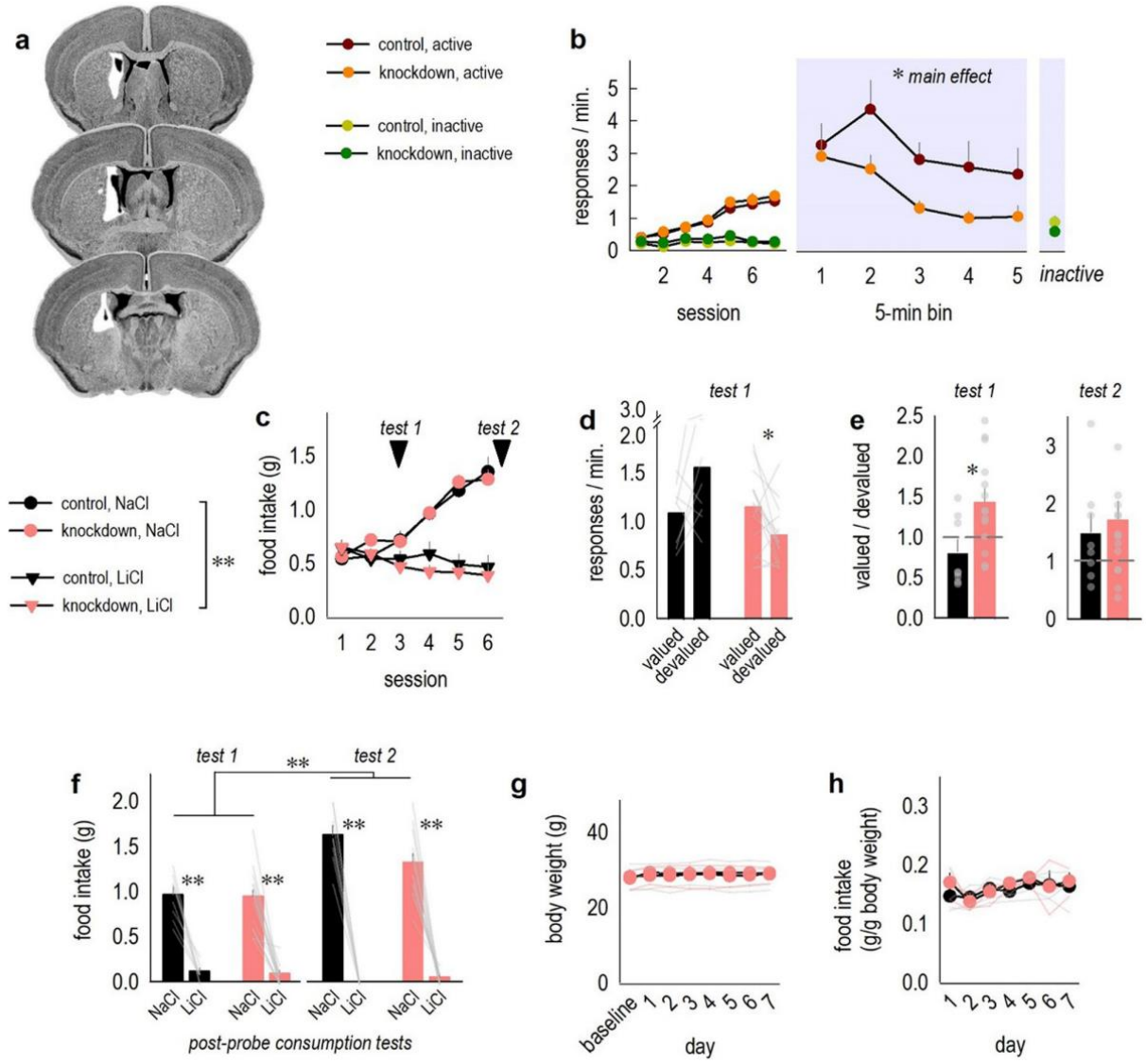

Figure 2.3 ***Mc4r*** **knockdown in the DMS expedites response inhibition. a**, Viral vector infusion sites on coronal sections represent areas of *Mc4r* knockdown. White represents large spread, with black the smallest. **b**, Response acquisition (left) was unaffected, despite *Mc4r* knockdown in the DMS (no main effect of group $F<1$; main effect of day $F_{6,60}=9.8$, $p<0.001$; no group × day interaction $F<1$). Groups differentiated between active and inactive ports, as expected (day × port interaction $F_{6,60}=21.36$, $p<0.001$; no day × port × group interaction $F<1$). When pellets were delivered noncontingently and responding was not reinforced (right), responding decreased

across time (main effect of time bin $F_{4,40}$=9.4, p<0.001), and knockdown mice responded less (main effect of group $F_{1,10}$=13.6, p=0.004; no time × group interaction F<1). Rates on the inactive port did not differ and are collapsed for simplicity. n=7 control, 5 knockdown. **c**, Next, one of the reinforcer pellets was paired with LiCl (decreasing its value), while the other pellet was paired with NaCl (a control). Ad libitum consumption of the LiCl-paired pellet decreased over repeated pairings (effect of day $F_{5,100}$=24.16, p<0.001; effect of pellet $F_{1,20}$=49.09, p<0.001; pellet × day interaction $F_{5,100}$=35.91, p<0.001; no effect or interactions with group Fs<1). **d**, Mice were returned to the testing chambers. Early in conditioning ("test 1"), only *Mc4r* knockdown mice inhibited responding for the devalued reinforcer (interaction $F_{1,20}$=6.64, p=0.02; no main effect of port F<1; no main effect of group $F_{1,20}$=3.19, p=0.09]. **e**, The same data were converted to preference scores (valued/devalued), in which case, scores >1 reflect response preference. The dashed line at 1 represents no change in behavior based on outcome value. Knockdown mice generated higher scores in the initial test, again indicating that they inhibited one behavior over another (t_{20}=2.61, p=0.02). Following more conditioning ("test 2"), both groups inhibited responding for the devalued pellet as expected (t_{20}=0.47, p=0.64). **f**, In post-probe consumption tests, mice overwhelmingly preferred the NaCl-paired pellet and avoided the LiCl-paired pellet, an effect that intensified with time, indicating that the CTA procedure was successful (main effect of pellet $F_{1,20}$=253.37, p<0.001; main effect of test 1 vs. 2 F(1,20)=37.17, p<0.001; pellet x test interaction F(1,20)=63.63, p<0.001). No main effect of group was detected ($F_{1,20}$=2.01, p=0.17). n=8 control, 14 knockdown. **g**, *Mc4r* knockdown did not affect free-feeding body weights (Fs<1) or (**h**) chow intake (no effect of group F<1; no effect of day $F_{6,30}$=2.2, p=0.1; no interactions F<1). n=4 control, 3 knockdown. Bars and closed symbols represent means + SEMs. Gray lines represent

individual mice. *p<0.05, **p<0.001. Instrumental conditioning experiments were conducted twice, with concordant results.

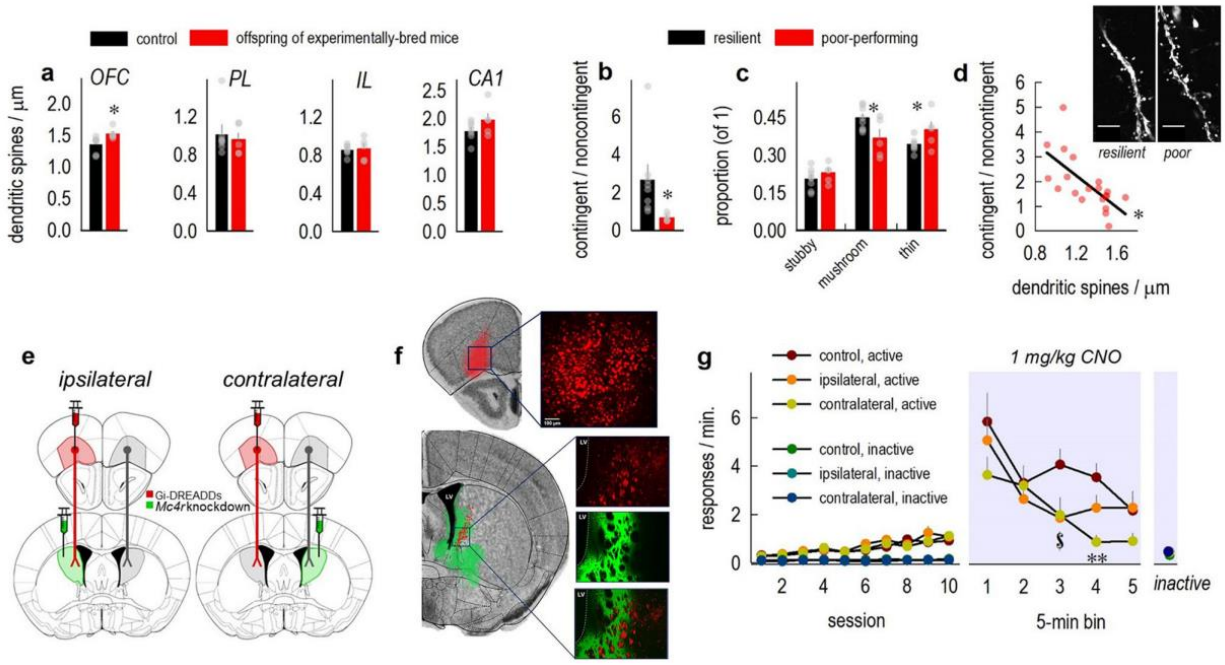

Figure 2.4 Reducing *Mc4r* in the DMS expedites response inhibition in an OFC-dependent manner. a, Terminal dendrites on neurons in multiple brain regions from the progeny of experimentally-bred mice were imaged, and dendritic spines were enumerated. Densities differed relative to typical mice in the ventrolateral OFC (t_{11}=-2.65, p=0.02). Meanwhile, groups did not differ in the prelimbic cortex (PL; t_{11}=0.43, p=0.67) or infralimbic cortex (IL; t_{11}=-0.2, p=0.85) or hippocampal CA1 (t_{11}=-1.44, p=0.18). n=7 control, 5 offspring of experimentally bred mice. **b**, OFC neurons from a separate cohort of experimental progeny were next imaged. Poor-performing progeny (those generating preference scores of ~1) were compared to resilient progeny (preference scores >1, reflecting increased responding in the contingent condition). **c**, Immature, thin-type dendritic spines on OFC dendrites were in excess in poor-performing mice, while resilient mice had more mature, mushroom-shaped spines (group × spine type interaction $F_{2,22}$=4.49, p=0.023; main effect of spine type $F_{2,22}$=33.41, p<0.001; no main effect of group $F_{1,11}$=1.60, p=0.23]. n=8 resilient, 5 poor-performing. **d**, Further, response scores correlated with thin-type dendritic spine densities in 2 independent cohorts (r^2=0.45, p<0.001). Representative

dendrites at right. Scale bar=5 μm. **e**, We next used an asymmetric infusion design to determine whether OFC-to-DMS projections are necessary for the response inhibition capacity conferred by *Mc4r* silencing. In the ipsilateral condition, one DMS lacks *Mc4r* (green), and the upstream OFC has Gi-coupled DREADDs (red). In the contralateral ("asymmetric") condition, mice also experience OFC inactivation, but the healthy OFC projects to an *Mc4r* knockdown DMS. Control mice bear control viral vectors, and thus have intact MC4R levels and no DREADDs. Cartoon adapted from Franklin & Paxinos (2001). **f**, OFC terminals overlapped with transduced regions of the DMS. Histological traces are represented on images from the Mouse Brain Library (Rosen et al., 2000). Scale bar=100 μm. "LV" refers to the lateral ventricle. **g**, Response acquisition did not differ between groups (left) (main effect of day $F_{9,342}$=27.1, p<0.001; no main effect of group $F_{2,38}$=1.65, p=0.21; no group × day interaction $F_{18,342}$=1.34, p=0.27). Groups differentiated between active and inactive ports during training (day × port interaction $F_{9,342}$=12.64, p<0.001; no day × port × group interaction $F_{18,342}$=1.29, p=0.23). When a familiar behavior failed to be reinforced, the contralateral group inhibited responding relative to control mice (right; group × time interaction $F_{8,152}$=2.39, p=0.02; main effect of time bin $F_{4,152}$=16.55, p<0.001; no main effect of group $F_{2,38}$=2.32, p=0.11). The syringe icon indicates that CNO was delivered to all mice prior to this session. n=14 control, 11 ipsilateral, 11 contralateral. Bars and closed symbols represent means + SEMs. Symbols in **d** represent individual mice. *p<0.05. \$$p$<0.05 contralateral and ipsilateral vs. control. **p<0.05 contralateral vs. ipsilateral and control. Instrumental conditioning experiments were conducted twice, with concordant results.

Chapter 3

Striatal melanocortin systems propel familiar actions via interaction with the central nucleus of the amygdala

3.1 CONTEXT AND AUTHOR'S CONTRIBUTION

The following chapter describes the cellular and anatomical mechanisms by which striatal MC4R controls the capacity for action-habit switching. The dissertation author contributed by researching and writing the manuscript, with editorial feedback from Dr. Shannon Gourley, Dr. Laura Butkovich, and Sophie Yount. Dr. Butkovich also conducted the experiment described in Fig. 3.1a with the author, while Sophie Yount conducted the experiment in Fig. 3.6k with the author.

3.2 ABSTRACT

The dorsomedial striatum (DMS) is a striatal subregion long associated with flexible goal seeking, suppressing routinized habits in dynamic environments. Whether local mechanisms brake this function, for instance when habits may be adaptive, is incompletely understood. Here we demonstrate that: melanocortin-4 receptor (MC4R)-expressing cells in the DMS (predominantly dopamine D1 receptor-containing) are necessary and sufficient for controlling the capacity of mice to arbitrate between actions and habits; MC4R presence suppresses goal seeking; and MC4R+ neurons are functionally integrated into an amygdalo-striatal circuit that suppresses action flexibility in favor of routine. Publicly available spatial transcriptomics datasets revealed differences in the gene transcript correlates of *Mc4r* expression across the striatum, with considerable co-variation in dorsal structures. This led to the discovery that MC4R function in the dorsolateral striatum complements that in the DMS, in this case *suppressing* habitual behavior. Altogether, our findings indicate that striatal MC4R controls the capacity for action-habit switching.

3.3 INTRODUCTION

In everyday life, we often must modify familiar behaviors in response to new information: When driving, road construction may require us to discard familiar routines (driving the typical route home) in favor of novel strategies (changing the route). The dorsomedial striatum (DMS) is essential for this kind of behavioral flexibility. The DMS is engaged when familiar behaviors are flexibly updated using new information regarding goal features or how to obtain them (Cruz et al., 2023). DMS lesions ablate flexible behavior, for instance rendering rats trained to press levers for food incapable of modifying response strategies when a given behavior is no longer rewarded (Yin, Knowlton, et al., 2005; Yin, Ostlund, et al., 2005; Yin & Knowlton, 2004). Over time, as behaviors become more automatized, neural activity in the DMS is quieted, coinciding with the ascendancy of habit-promoting dorsolateral striatal pathways (Gremel & Costa, 2013; Thorn et al., 2010; Yin et al., 2009; Yu et al., 2021). Whether neuromodulatory systems within the DMS serve to actively *inhibit* action flexibility, for example when habits may be advantageous, remains unclear.

One candidate factor is the melanocortin-4 receptor (MC4R), a high-affinity receptor for α-melanocyte-stimulating hormone (α-MSH). α-MSH is released in response to satiety from neurons originating in the arcuate nucleus of the hypothalamus and projecting throughout the brain, corresponding with the nearly ubiquitous expression of MC4R. MC4R is classically studied for its role in suppressing feeding and increasing metabolism (Quarta et al., 2021), which likely accounts for the notable success of MC4R agonists in combatting obesity (Sweeney et al., 2023). However, MC4R in the ventral striatum also contributes to non-feeding-related phenomena: anhedonic- and compulsive-like behavior (Alvaro et al., 2003; Lim et al., 2012), as well as cocaine seeking and sensitization (Hsu et al., 2005b). MC4R is also highly expressed in the dorsal subregions (Liu et al., 2003), but its functions there are not fully understood (Roseberry et al.,

2015). One insight came from a recent report in which MC4R levels in the DMS, but not the ventral striatum, correlated with the capacity of mice to flexibly modify familiar behaviors – suggesting that MC4R impacts the ability to adjust or alter routines (Allen et al., 2022).

Here we investigated the hypothesis that MC4R-expressing cells in the DMS are necessary and sufficient for controlling reward-related decision making and are functionally integrated within an amygdalo-striatal circuit that influences the capacity of organisms to arbitrate between flexible and more automatized action strategies. Given the widespread (and growing) use of pharmacological agents that curb appetite and suppress feeding, including in non-clinical populations, understanding the intersections between drug targets and non-feeing-related sequelae is imperative.

3.4 RESULTS

3.4.1 MC4R-expressing cells in the DMS bidirectionally control action flexibility

We first quantified *Mc4r* content in the DMS, revealing expression in ~20% of cells (Fig. 3.1a-b). Medium spiny neurons (MSNs) make up the vast majority of cells of the striatum and are largely segregated based on dopamine D1 or D2 receptor occupancy. We next found that *Mc4r* overwhelmingly co-localizes with *Drd1,* the gene encoding D1 receptors (Fig. 3.1c). This pattern is consistent with a great degree of MC4R co-localization with prodynorphin, which also overwhelmingly co-localizes with dopamine D1 receptors (Hsu et al., 2005b). Thus, MC4R is predominantly expressed on D1 receptor-containing MSNs.

We next measured the ability of mice to flexibly update their behavior when a familiar action is no longer rewarded (Fig. 3.1d). Mice were first trained to respond at two apertures for sweetened grain or chocolate-flavored pellets delivered into a separate food magazine. Next, mice

were given isolated access to one aperture at which responding remained reinforced. The following day, the opposite aperture was available – here, responding was unexpectedly no longer reinforced, and instead pellets were delivered "for free" at a rate equal to that of the prior session. The number of pellets delivered during these reinforced *vs.* non-reinforced sessions was thus equivalent; the distinction was that one action was no longer reinforced. The following day, mice were tested in a brief choice test conducted in extinction, during which they had access to both apertures. Mice that are sensitive to changes in response contingencies will inhibit behaviors that failed to be reinforced and favor the reinforced response, deviating from equivalent responding during training.

We first induced inhibitory (Gi-coupled) chemogenetic receptors in *Mc4r*-expressing cells in the DMS (Fig. 3.1e-f), thus inhibiting their activity in the presence of clozapine *N*-oxide (CNO) (Supplementary Fig. 3.S1a). Throughout training, conducted in the absence of CNO, mice responded equivalently on both reinforced apertures without side preference and thus, responses on both ports are collapsed (Fig. 3.1g). Next, one response was not reinforced, and CNO was delivered. No immediate effects were apparent (Fig. 3.1h), but when choice was assessed later, drug free, only control mice favored the rewarded behavior. Meanwhile, mice with inhibited *Mc4r*-expressing cells did not distinguish between trained behaviors, utilizing habit-like response strategies (Fig. 3.1i). These data can also be converted to response preferences: response rates at the reinforced / non-reinforced apertures, with values greater than 1 indicating preference for the reinforced behavior (*i.e.*, more flexibility). Again, only control mice flexibly favored the reinforced behavior, while this flexibility was blocked by silencing *Mc4r*+ cells (Fig. 3.1i right).

We next induced excitatory (Gq-coupled) chemogenetic receptors to *stimulate Mc4r*-expressing cells in the DMS (Fig. 3.1e-f, j; Supplementary Fig. 3.S1b), predicting that stimulation would prompt behavioral flexibility. Stimulating *Mc4r*+ cell populations indeed caused mice to

more strongly favor reinforced behaviors during the initial experience with non-reinforcement (Fig. 3.1k). We view this pattern as expedited response inhibition, given that both groups ultimately inhibited the non-reinforced behavior during the choice test (Fig. 3.1l).

Next, mice were trained further using random interval schedules of reinforcement, in which case, intervals of time are randomly inserted when responding is not reinforced (Fig. 3.1f, j). Interval schedules weaken the association between a behavior and its reward, resulting in automatized, habitual responses, as opposed to goal-driven actions (Adams & Dickinson, 1981; de Wit et al., 2009). Interval training allowed us the resolution to further detect *improvements* in behavioral flexibility, if any, and will be referred to as "extended training" throughout this manuscript. Stimulating *Mc4r*+ cells again caused mice to more strongly favor a reinforced behavior than unstimulated control mice during the initial experience with non-reinforcement (Fig. 3.1m). During the choice test, control mice failed to differentiate between reinforced and non-reinforced behaviors, as expected, while stimulated mice remained flexible, favoring the reinforced behavior (Fig. 3.1n). Importantly, response patterns here and throughout are not obviously attributable to unintended effects on locomotor activity (Supplementary Fig. 3.S1c). Thus, *Mc4r*-expressing cells in the DMS are both necessary and sufficient for behavioral flexibility according to reward availability.

3.4.2 Reducing *Mc4r* in the DMS prompts flexible behavior

We next began to ask why and how MC4R in the DMS impacts reward-seeking strategies. Viral vectors ± Cre were infused into the DMS of *Mc4r*-flox mice, thereby reducing MC4R in the Cre condition. *Mc4r*-deficient mice and control counterparts were trained and euthanized at two time points: following the non-reinforced session, when they first have the opportunity to update expectations based on new information, or following the choice test, when mice can use previously

learned information to guide choice (Fig. 3.2a). cFos levels in the DMS were highest when mice first encountered unexpected non-reinforcement, but regardless of timepoint, DMS tissue lacking *Mc4r* had more cFos, indicating more immediate-early gene expression in the absence of MC4R (Fig. 3.2b-c). These patterns are consistent with the notion that MC4R acts as a brake on cell activity in the striatum (Pandit et al., 2013).

In the ventral striatum, this "braking" influence appears attributable, at least in part, to MC4R control of AMPA-GluA2 receptor localization at the synaptic membrane (Lim et al., 2012). To determine whether the same principles apply to the DMS, we generated mice with viral-mediated *Mc4r* silencing in one DMS and a control viral vector in the opposite DMS. Dissected hemispheres were used to create and isolate synaptoneurosomes, composite particles that contain both pre- and post-synaptic cell membranes (Fig. 3.2d). The pre-synaptic marker synaptophysin and post-synaptic marker PSD-95 were detected only in the synaptic and not extra-synaptic fraction, as expected (Fig. 3.2e). Levels of α-tubulin remained constant between fractions, also as expected (Fig. 3.2e). MC4R protein content was also lower in the *Mc4r*-deficient hemisphere, again as expected (Fig. 3.2f-g). Unexpectedly, however, synaptic GluA2 content did not differ based on *Mc4r* status, nor did GluA1. Instead, we found that synaptic fractions with reduced *Mc4r* contained elevated levels of NMDA receptor subunit GluN2B (Fig. 3.2h-j).

These patterns lead to the prediction that *decreasing* MC4R availability in the DMS would *increase* behavioral flexibility, given that the DMS is a key hub controlling action flexibility. *Mc4r* was depleted selectively from MSNs in the DMS, and mice were trained and tested as before (Fig. 3.3a). During choice test 1, both groups inhibited behaviors when they were no longer reinforced, as expected (Fig. 3.3b-d). After extended interval training, the control group lost this preference, instead utilizing habit-based response strategies, also as expected. Meanwhile, mice with reduced

Mc4r in the DMS maintained the capacity to distinguish between reinforced and non-reinforced behaviors (Fig. 3.3e). Thus, MC4R in the DMS obstructs action flexibility. Consistent with this notion, the MC4R agonist setmelanotide also suppressed action flexibility in the same task and importantly, at doses that did not impact food-reinforced responding in general (Supplementary Fig. 3.S2). Altogether, these patterns suggest that availability and binding of MC4R control the strategies by which organism seek reward, with MC4R biasing mice towards habit-like actions.

To further solidify this discovery, we next tested whether silencing *Mc4r* could reinstate flexible behavior in cases when it is impaired, causing mice to defer to habit-like behaviors. Here, we exposed mice to cocaine, which occludes behavioral flexibility in a number of assays, including the present one (Li et al., 2022). *Mc4r* knockdown ± cocaine did not impact response training (Fig. 3.3f-g). Cocaine nevertheless obstructed flexible behavior when one response was unexpectedly not reinforced, as expected. Reducing *Mc4r* in the DMS of cocaine-exposed mice reinstated flexible behavior (Fig. 3.3h). Thus, silencing *Mc4r* in the DMS promotes behavioral flexibility, combatting inflexible, routinized responding caused by extensive training and pharmacological insult alike.

3.4.3 MC4R in amygdalo-DMS circuits controls action selection

We next questioned which projections terminate on MC4R+ cells in the DMS. We used a rabies virus-mediated retrograde trans-synaptic tracing technique to selectively label cell populations that synapse directly onto MC4R+ cells in the DMS, focusing on afferent regions involved in behavioral flexibility or habit-based behavior (Fig. 3.4a). The complete list of projections to MC4R+ cells in the DMS is reported in Supplementary Table 1.S1. Importantly,

labeling was found in the arcuate nucleus of the hypothalamus, reflecting monosynaptic input from the primary site of α-MSH production, as expected (Supplementary Table 1.S1).

The highest density of projections to MC4R+ cells originated in the central nucleus of the amygdala (CeA; Fig. 3.4b), consistent with evidence that CeA projections terminate in the DMS, in general (Pan et al., 2010; Wall et al., 2013). Whether CeA-to-DMS projections regulate instrumental behavior is not known. First, we chemogenetically stimulated inhibitory CeA-to-DMS projections (Fig. 3.4c-d). Control mice lacking the chemogenetic receptor favored a rewarded behavior over one that was not, as expected, while stimulating CeA-to-DMS projections occluded flexible response updating, such that mice demonstrated no preference for either behavior (Fig. 3.4e-g). Thus, CeA projections to the DMS regulate behavioral flexibility, providing a brake on action flexibility.

We next asked if MC4R and plasticity in the CeA synchronize to provide a brake on DMS-mediated action flexibility. We reduced *Mc4r* unilaterally in one DMS and placed inhibitory (Gi-coupled) chemogenetic receptors unilaterally in one CeA (Fig. 3.4h-i). Mice in this condition should be adept in inhibiting responding when reinforcement likelihood declines due to the combination of disinhibition of DMS MSNs in one hemisphere (by virtue of *Mc4r* silencing) and inhibited CeA cell activity in one hemisphere, which would similarly disinhibit MSNs in the DMS. Existing literature indicates that unilateral *Mc4r* silencing in the absence of CeA manipulation has no impact (Allen et al., 2022), nor would we expect DREADDs in the unilateral CeA alone to have impact (Lingawi & Balleine, 2012), though this has not been explicitly confirmed. Accordingly, control mice had Gi-DREADDs in one CeA and a control viral vector in the DMS, or they had control viral vectors in both regions. Following extended training, both groups of control mice demonstrated no response preference, as expected (Fig. 3.4j-l). Meanwhile, mice with dual *Mc4r*

gene silencing and inhibitory DREADDs in the CeA preferred the reinforced response, resilient to habit-like behavior. These patterns suggest that MC4R in the DMS and plasticity in the CeA synchronistically brake DMS-mediated action flexibility, for instance, when actions are highly familiar and habitual behavior may be adaptive.

3.4.4 Spatial transcriptomics reveals a diversity of *Mc4r* expression co-variates across the striatum

Given that MC4R in the DMS has a profound effect on behavior, we hypothesized that, within the DMS, *Mc4r* would strongly co-vary with an abundance of genes. How gene transcripts co-vary in a given biological system may shed light onto gene function within that context (de Torrenté et al., 2020). Roman and colleagues recently used Digital Spatial Profiling to create transcriptomic profiles of D1-expressing cells in the DMS, dorsolateral striatum (DLS), and ventrolateral striatum (VLS) (Fig. 3.5a). We analyzed transcript reads (counts of transcripts of a given gene in each region of interest), seeking genes whose transcription levels correlated with that of *Mc4r*. On the whole, gene profiles and numbers varied widely between the dorsal (DMS and DLS) and ventrolateral striatum. In the dorsal striatum, transcription levels of 9,691 out of 12,306 genes (~79%) correlated with *Mc4r*, a measure of interdependence of transcript quantities, compared to 749 (~1%) genes in the ventral striatum (Fig. 3.5b). Within the dorsal striatum, *Mc4r* co-variates were most robust in the DMS: 7,300 out of 8,906 correlations (~82%) in the DMS had a r^2 value > 0.8, compared to 5,094 out of 7,470 correlations (~68%) in the DLS (Fig. 5c). The r^2 frequency distribution varied much less in the VLS: 100% of correlations had a r^2 value > 0.8 despite representing only ~1% of the gene set (Fig. 3.5c).

Mc4r transcript signal correlated with 616 genes across all three regions (Fig. 3.5b). Across the DMS, DLS, and VLS, the greatest variability in correlation strength was within patterns of positive correlations (Fig. 3.5d). Several groups of genes were consistently associated with *Mc4r*, including those involved in GABAergic and glutamatergic signaling (*Gabrd, Gria3, Grina, Grip2)*, potassium channels (*Kcnj6, Kcnq2, Kcnt1, Kctd2, Kctd4*), GPCR function (*Adgrl4, Gpr150, Rgs12, Rgs16*), and Ras signaling (*Rap2a, Rassf5, Rassf8, Rhob, Rhoh, Rin2*) (Fig. 3.5e).

The strength and number of correlations with *Mc4r* transcript was most similar between the DMS and DLS relative to VLS. Correlations with *Mc4r* transcript in the DMS and DLS overlapped significantly, sharing 6,685 (~54% of total) genes (Fig. 3.5b). We conducted gene ontology analyses on *Mc4r* co-variates in the DMS and the DLS to provide insight into any potential similarities in MC4R function. This shared gene set contained several enriched pathways involved in homeostasis and synaptic signaling, as well as intracellular signaling and general cell function (Fig. 5f). The enrichment of each pathway did not vary between the DMS and DLS (Fig. 3.5f). Within each enriched pathway, there were genes whose expression was correlated with *Mc4r* transcription in the DMS and DLS. Visualizing the Pearson's *r* values of these correlations revealed that, for instance, *Mc4r* transcription is positively correlated with transcripts of genes in the metabolic pathway and negatively correlated with transcripts of genes that control long-term potentiation (Fig. 3.5f). Overall, the strength and directionality of correlates with *Mc4r* transcription in each pathway did not differ across the dorsal striatum. The original report of Roman and our subsequent analyses highlight that the transcriptomic profiles of the DMS, DLS, and VLS are distinct, but significant correlations between *Mc4r* and several other transcripts exist in both the DMS and DLS (again, Fig. 3.5b).

3.4.5 MC4R acts as a molecular brake on DLS function in prompting habit-like behavior

Spatial transcriptomics suggests that MC4R functions comparatively in the DMS and DLS. Given that MC4R in the DMS robustly impacts behavior, we next hypothesized that MC4R+ cells may also control DLS function, in particular in executing habitual routines. We induced excitatory (Gq-coupled) chemogenetic receptors in MC4R+ cells in the DLS, then tested mice as before (Fig. 3.6a-c). When one trained response was unexpectedly not reinforced, control mice favored the reinforced behavior, as expected (Fig. 3.6d). Meanwhile, stimulating *Mc4r*+ cells in the DLS prompted habit-like behavior, with mice generating reinforced and non-reinforced behaviors equivalently. Thus, *Mc4r*+ cells in the DLS can induce habit-like insensitivity to new contingencies.

We next investigated whether the availability of MC4R itself in the DLS impacts behavioral flexibility. We again used viral-mediated gene silencing to reduce *Mc4r*, predicting that decreasing MC4R in the DLS would *promote* habit-like behavior (Fig. 3.6e). To address this question, an additional choice test was conducted midway through extended training (Fig. 3.6f-g). This "intermediate" choice test allows for the resolution to detect any early expression of habitual behavior as it develops across repeated training sessions. At the first choice test following short training, both groups inhibited behaviors when they were no longer reinforced, as expected (Fig. 3.6h). At the intermediate test, control mice again inhibited behaviors that were no longer reinforced, as expected (Fig. 3.6i). However, mice with reduced *Mc4r* in the DLS responded equally across contingency conditions. Thus, *Mc4r* gene silencing expedited habit formation. As a positive control, both groups then completed extended training, which induced non-preferential responding in both conditions, as expected (Fig. 3.6j). Altogether, this pattern suggests that MC4R in the DLS brakes habitual behavior.

Like the DLS, the ventral striatum (VS) can support the execution of habitual routines. However, very few gene transcripts correlated with *Mc4r* expression in the VS compared to dorsal striatal subregions (Fig. 3.5b). Moreover, MC4R in the VS does not correlate with the capacity of mice to modify familiar behaviors (Allen et al., 2022), suggesting that MC4R availability in the VS would not impact behavioral flexibility. Again, we used viral-mediated gene silencing to reduce *Mc4r* in the VS (Fig. 3.6k). If, as in the DLS, MC4R in the VS brakes habitual behavior, we predicted that decreasing MC4R in the VS would *promote* habit-like behavior. Mice were tested once following initial training and again at an "intermediate" training point (Fig. 3.6l-m). At choice test 1, both groups inhibited behaviors when they were no longer reinforced, as expected (Fig. 3.6n). At the intermediate test, control mice continued to favor the reinforced response (Fig. 3.6o). Likewise, responding at the reinforced aperture remained stable in mice with reduced *Mc4r* in the VS. Thus, site-selective reduction of *Mc4r* in the NAc did not impact habit formation.

Altogether, MC4R in the DMS dampens action flexibility, while in the DLS, it obstructs habit formation. Striatal MC4R thus serves as an endogenous brake on action flexibility and habitual behavior alike, potentially allowing organisms to arbitrate between strategies when adaptive.

3.5 DISCUSSION

Flexible behavior requires knowledge of the relationship between actions and their specific outcomes, while habitual behavior is triggered by environmental stimuli, independent of goal features or availability. Current models argue that the DMS supports flexible goal seeking, and as behaviors become more familiar, the DLS coordinates the formation and adherence to habits (Guida et al., 2022). Less clear is whether systems (molecular- or circuit-level) within the DMS

mitigate action flexibility as habits form. We found that ~20% of cells in the DMS are MC4R+, they are overwhelmingly D1 receptor-expressing, and despite their relative paucity, they control the capacity of mice to modify familiar routines. MC4R presence brakes action flexibility, reducing immediate-early gene levels in the DMS and conferring bias towards habit-like behaviors. Long-range inhibitory projections from the CeA – a brain region necessary for habit formation (Lingawi & Balleine, 2012) – terminate on MC4R+ neurons in the DMS. Their activity coordinates with MC4R to *moderate* action flexibility. We posit that ligand binding to MC4R in the DMS serves as a neuromodulator suppressing activity of D1-MSNs, which in coordination with inputs from regions such as the CeA, mitigates flexible goal seeking when habits might be more advantageous.

3.5.1 Striatal MC4R controls action flexibility

The striatum is composed predominantly of MSNs, inhibitory GABAergic projection neurons characterized by dopamine D1 or D2 receptor presence (Gerfen, 2022). D1-MSNs form the "direct pathway," a circuit that promotes deliberate movement (Kravitz et al., 2010) and action flexibility (Matamales et al., 2020). As familiar behaviors are flexibly updated, activity in D1-MSNs increases, inducing lasting changes in synaptic signaling that are necessary for strategy switching (Cui et al., 2013; Fisher et al., 2020). And as inflexible habits form, postsynaptic depression of excitatory synapses on D1-MSNs in the DMS suppress action flexibility (Yu et al., 2021). We found that *Mc4r* is present on ~20% of neurons in the DMS, overwhelmingly D1-MSNs, and accordingly, inhibiting *Mc4r*-expressing cells ablated action flexibility, as with silencing D1-MSNs (Kwak & Jung, 2019) and MSNs in general (*e.g.*, Shan et al., 2023). Further, *stimulating Mc4r*-expressing cells *prompted* response switching.

These experiments confirmed that MC4R+ cell populations bidirectionally control action flexibility but did not reveal MC4R function. For this, we genetically silenced *Mc4r* in MSNs and first measured immediate-early expression, important because ligand binding can stimulate Gs, Gq, or Gi/o-mediated signaling (Liu & Hruby, 2022). As such, stimulating MC4R can increase (Shen et al., 2013, 2016) or decrease (Bruschetta et al., 2020; Lim et al., 2012) molecular markers of cell excitability. This effect may be cell-type dependent; for instance, MC4R-mediated suppression of cell excitability has so far been found only in GABAergic cells (Bruschetta et al., 2020; Lim et al., 2012). Accordingly, we found that *Mc4r* silencing *increased* cFos in the DMS, accompanied by elevated synaptic GluN2B, an NMDAR subunit that can form triheteromeric NMDARs promoting long-term potentiation and cell plasticity (France et al., 2017; Joo et al., 2015; Wong & Gray, 2018).

Our results suggest that MC4R dampens neuronal plasticity in the DMS, which could have profound consequences for DMS-mediated action flexibility. Indeed, we next found via site selective *Mc4r* gene silencing that MC4R in the DMS suppressed flexible goal seeking, causing mice to defer to familiar routines. This discovery was notable because addictive drugs, such as cocaine, increase striatal MC4R content (Alvaro et al., 2003; Hsu et al., 2005b) and modify decision-making behavior – for instance, by narrowing individuals' capacity to engage in long-term goal planning (Ognibene et al., 2019). We thus next reduced *Mc4r* in cocaine-exposed mice, which was sufficient to reinstate action flexibility. Altogether, MC4R in the DMS appears to suppress flexible action strategies, for instance when habits may be more advantageous, and may be hijacked by addictive drugs to induce inflexible reward seeking.

Interestingly, cFos levels were higher overall (regardless of *Mc4r*) following the session when mice first encountered unexpected non-reinforcement, compared to the choice test, when

they act on that information. Cell activity at this time may subserve new long-term learning, because chemogenetically silencing MC4R+ cells during this time obstructed the capacity of mice to later use new information to guide later choice.

3.5.2 MC4R and amygdalo-striatal interactions suppress action flexibility

The striatum is a site of converging synaptic inputs, integrating information from cortical and limbic regions into action strategies. We next questioned which projections terminate on MC4R+ DMS neurons. Among several brain regions identified, abundant inputs originated in the CeA, a subregion of the amygdala primarily composed of GABAergic projection neurons (Yang et al., 2023). The CeA is *necessary* for the development of habits, such that lesions render rats perpetually sensitive to change in outcome value, despite extended training (Lingawi & Balleine, 2012). Further, CeA output appears to modulate striatal dopamine systems following the completion of an action, providing feedback on the success of that action to propel rewarded behaviors (Pauli, Hazy, et al., 2012). We first chemogenetically stimulated CeA-to-DMS projections, which induced habit-like choice, presumably via the GABAergic inhibition of MSNs in the DMS. We next chemogenetically suppressed activity of CeA neurons in one hemisphere and silenced *Mc4r* in one hemisphere. If these two systems (plasticity in the CeA and MC4R in the DMS) act synergistically, these mice should be highly adept in inhibiting responding when reward is withheld. This was indeed the case, suggesting that MC4R in the DMS and plasticity in the CeA synchronistically brake DMS-mediated action flexibility, for instance, when actions are highly familiar and habitual behavior may be adaptive.

Interestingly, CeA neurons project to both the DMS and DLS (Pan et al., 2010). As behaviors become more routine, the activity of amygdalar projections to the DLS shifts from being

predominantly from the BLA to predominantly from the CeA (Murray et al., 2015), consistent with evidence that CeA-to-DLS projections are necessary for habitual behavior (Lingawi & Balleine, 2012). It is plausible that the same phenomena occur in the DMS, with BLA input to the DMS promoting flexible behavior (Wassum, 2022), and inputs from the CeA suppressing DMS cell activity as habits form. Another interesting possibility is that CeA-to-DMS interactions may suppress action flexibility when reward *values* (as opposed to contingencies) change, given that persistent responding for devalued outcomes is often used to identify habitual behavior in rodents.

We next used publicly available spatial transcriptomics datasets to investigate the degree to which *Mc4r* transcript quantity correlates with gross cell transcription, and thus, overall cell function. We found that *Mc4r* is highly correlated with the DMS transcriptome compared to other striatal subregions, a phenomenon that may be explained by differential gene expression across the medial-lateral and dorsal-ventral axes of the striatum (Roman et al., 2023). By comparison, apparent coordination between *Mc4r* and other genes is impoverished in the ventral striatum, even though most striatal MC4R research has overwhelmingly focused on the ventral striatum (Roseberry et al., 2015). Future experiments evaluating the impact of *Mc4r* on striatal neurobiology could be enormously fruitful in fully understanding the mechanisms by which MC4R presence and activity has behavioral consequences. For instance, hyperphagia is induced by reducing *Mc4r* in the ventral striatum (Cui & Lutter, 2013) but not the dorsal striatum (Allen et al., 2022).

The transcriptomic profile of DLS samples resembled in many ways that of the DMS, which led us to next question whether MC4R+ cells or MC4R itself would impact DLS function in propelling habitual behavior. Indeed, stimulating MC4R+ cells and reducing *Mc4r* prompted habit-like behavior. MC4R thus plays a dual role in behavioral flexibility, providing a functional

brake on both DMS- and DLS-dependent action. How might this be? As flexible actions become more routinized, task-related cell activity in the striatum shifts from predominating in the DMS to the DLS. By acting on the "dominant" neural pattern at any given time, MC4R systems could impact action flexibility and habit alike.

3.5.3 Conclusions

Pro-opiomelanocortin+ cells in the arcuate nucleus of the hypothalamus produce the MC4R ligand, α-MSH, in response to satiety. As such, the control of decision-making behavior by melanocortin systems may be framed, at least as a starting point, within the context of food seeking. Hunger is a salient motivator and can override other bodily demands in the pursuit of food (Sutton & Krashes, 2020). Animals that are hungry often display flexible behavior: for instance, a lack of food can prompt animals to abandon familiar regions and forage in new areas (Stephens, 2018). This phenomenon, termed optimal foraging theory, similarly drives humans to abandon their empty pantries and travel to grocery stores. In laboratory rodents, hunger confers resistance to habits: over-hungry mice that weigh <80% of their baseline body weight will resist routines and maintain vigilance to changes in environmental cues that may signal food availability (Rossi & Yin, 2012). When food is obtained and satiety achieved, α-MSH activation of MC4R in the DMS would be expected to reduce the excitability of DMS neurons, potentially allowing organisms to abandon flexible food seeking strategies that are no longer in needed, in favor of routinized habits. Habits are advantageous because they free executive resources as organisms to attend to other events while performing behaviors that have been reliably reinforced in the past.

Of course, this model cannot account for why MC4R in the DLS would *moderate* habit-based behavior. Perhaps, α-MSH release when mice are engaged in habitual food intake serves as

a signal to *inhibit* intake as organisms become sated. This phenomenon could contribute to the notable success of MC4R agonists in curbing obesity (Sweeney et al., 2023). A final question is how and why MC4R systems control decision-making behavior *unrelated to food*, for instance promoting cocaine self-administration and drug seeking (Hsu et al., 2005b). Because of these discoveries, we think that the phenomena we report here may not be restricted to food seeking. The development of novel tasks that can reliably disentangle goal- *vs.* habit-based behavior without food reinforcement in model organisms could help to resolve this question. Further, tools by which to measure α-MSH would help to reveal conditions – related *and unrelated* to feeding – when α-MSH is released.

3.6 METHODS

3.6.1 Animals

Experiments were conducted in adult (≥ postnatal day 56) male and female mice. Sex differences were not detected. Several strains of mice were used and bred in-house from Jackson Laboratories stock: For chemogenetic manipulation and rabies tracing, mice were *Mc4r-2a-Cre* knock-in mice (#030759) maintained on a C57BL/6 background (Garfield et al., 2015). For manipulation of *Mc4r*, mice were homozygous for a "floxed" *Mc4r* gene (*Mc4r-flox*, #023720) and maintained on a mixed C57BL/6-129S1/SvlmJ background (Sohn et al., 2013). Otherwise, mice were wildtype C57BL/6 (#000664).

Mice were maintained on a 14-h light cycle (0700 on). Mice were provided food and water *ad libitum* except during instrumental conditioning when body weights were reduced to 90% of baseline to motivate responding. All procedures were performed in accordance with NIH

Guidelines for the Care and Use of Laboratory Animals and were approved by the Emory University IACUC.

3.6.2 RNAScope

In situ RNA analysis was performed with the RNAScope Multiplex Fluorescent v2 kit (ACD #323100) and conducted following the manufacturer's protocol (ACD #323100-USM) using probes for *Drd1* (ACD #316671-C2, Lot 22284E) and *Mc4r* (ACD #319181-C3, Lot 23054A) in Probe Diluent (ACD #300041). Mice were deeply anesthetized (120 mg/kg, *intraperitoneal* [*i.p.*]) and xylazine (10 mg/kg, *i.p.*) and trans-cardially perfused with ice-cold PBS and 4% paraformaldehyde. Brain tissue was extracted and incubated in 4% paraformaldehyde overnight at 4°C. Over the following 3 days, brains went through overnight incubations in solutions of increasing sucrose concentrations (10%, 20%, 30%), then flash frozen and stored at -80°C. Sections were collected at 12µm on a CryoStar NX70 cryostat on Superfrost Plus Slides and stored at -80°C. To prevent tissue detachment, steps described in the ACD Technical Note (ACD #320535-TN) were conducted immediately prior to the Multiplex Fluorescent v2 assay.

Images were acquired on a Keyence BZ-X710 microscope (Keyence Corporation) and analyzed with the open access image analysis software, CellProfiler (Stirling et al., 2021). Cells expressing *Drd1* and/or *Mc4r* were determined by co-localization of DAPI with transcript puncta within a region of interest (ROI) of uniform size across samples. 2-9 images from each mouse were analyzed, and each mouse contributed a single value (its mean) to the analyses.

3.6.3 Surgery and viral vectors

Intracranial surgeries were performed 2 weeks before behavioral testing to allow time for expression of viral vectors. Mice were anesthetized with ketamine (100 mg/kg, *i.p.*) and dexmedetomidine (0.5 mg/kg, *i.p.*) and placed in a digitized stereotaxic frame (Stoelting). Surgeries were performed under aseptic conditions. A mid-sagittal incision exposed the skull, and a craniotomy was performed to allow for intracranial viral vector infusion. Infusion volumes and coordinates relative to bregma were (unless otherwise specified): DMS (0.5 µl per hemisphere; AP: +0.5 mm, ML: ±1.6 mm, DV: -4.25 mm), DLS (0.5 µl per hemisphere; AP: +0.5 mm, ML: ±2.7 mm, DV: -3.5 mm), VS (0.5 µl per hemisphere; AP: +1.5 mm, ML: ±0.95 mm, DV: -4.5 mm), and CeA (0.25 µl per hemisphere; AP: -1.4 mm, ML: ±3.0 mm, DV: -4.9 mm). Viral vectors were infused over 5 minutes using a microliter syringe (Hamilton) and left in place for an additional 5-10 minutes before retraction to restrict off-target viral vector spread.

Chemogenetic manipulations of Mc4r+ neurons. To achieve chemogenetic receptor expression in *Mc4r*+ neurons, a Cre-dependent chemogenetic receptor construct (AAV5-hSyn-DIO-hM3D(Gq)-mCherry or AAV5-hSyn-DIO-hM4D(Gi)-mCherry) or control construct (AAV5-hSyn-DIO-mCherry) was infused bilaterally into the DMS of *Mc4r-2a-Cre* transgenic mice. Complete viral vector product information is listed in Table 1. After testing, mice were euthanized, and brains were fixed and prepared as described below. mCherry expression was examined to confirm viral vector placement. Infusing the Cre-dependent viral vectors into the brains of non-Cre-expressing mice resulted in no mCherry expression, confirming their specificity (Supplementary Figure 1d).

Site-selective Mc4r knockdown. AAV8-CamKII-HI-GFP-Cre-WPRE-SV40 or AAV8-CamKII-eGFP was infused into the DMS, DLS, or VS of *Mc4r-flox* mice. After testing, mice were

euthanized, and brains were fixed and prepared as described below. GFP expression was examined to confirm viral vector placement.

For synaptoneurosome preparations, *Mc4r-flox* mice received a unilateral infusion of AAV8-CamKII-HI-GFP-Cre-WPRE-SV40 into one DMS and a unilateral infusion of AAV8-CamKII-eGFP into the opposite DMS, allowing for cross-hemisphere comparisons within mice.

Trans-synaptic retrograde tracing. Tracing experiments were performed in *Mc4r-2a-Cre* transgenic mice. A Cre-dependent helper construct (AAV8-CMV-FLEX-TVAmCherry2A-oG) was infused unilaterally into the DMS (0.25 µL per hemisphere). Mice were allowed two weeks for recovery. Then, in a second surgery, we unilaterally infused pseudotyped rabies virus (EnvA G-Deleted Rabies-eGFP) at the same DMS coordinates as the helper construct (0.5 µL per hemisphere). The pseudotyped rabies virus was allowed to replicate and spread for 7 days, at which point mice were euthanized and brains were fixed and prepared as described below.

Chemogenetic stimulation of CeA-to-DMS projections. To achieve projection-selective chemogenetic receptor expression, wildtype C57BL/6 mice underwent surgery to receive a retrogradely transported Cre-recombinase construct (rgAAV-hSyn-GFP-Cre) infused bilaterally into the DMS and an anterogradely transported Cre-dependent chemogenetic receptor construct (AAV5-hSyn-DIO-hM3D(Gq)-mCherry) infused bilaterally into the CeA. After testing, mice were euthanized, and brains were fixed and prepared as described below. GFP and mCherry expression were examined to confirm viral vector placement.

Chemogenetic manipulation of CeA neurons, concurrent with Mc4r silencing. Mice received a unilateral infusion of an inhibitory chemogenetic receptor construct (AAV5-CamKII-hM4D(Gi)-mCherry or AAV5-CamKII-mCherry) in the CeA. These *Mc4r-flox* transgenic mice also received unilateral infusions of Cre-expressing viral vectors in the DMS as described above.

Infusions were ipsilateral or contralateral and counterbalanced between hemispheres. After testing, mice were euthanized, and brains were fixed and prepared as described below. mCherry was examined to confirm viral vector placement.

3.6.4 Instrumental response training

Behavioral testing was conducted using operant conditioning chambers equipped with two distinct nose-poke apertures (left versus right), with a separate magazine for food pellet delivery (Med Associates). For all experiments, each mouse performed all tasks in the same chamber. Mice were trained to nose-poke for 20 mg pellets (chocolate or purified grain; one associated with each nose-poke; Bio-Serv). Training commenced using a fixed ratio 1 (FR1) schedule of reinforcement whereby each response resulted in a single reinforcer. Mice underwent daily training sessions, which lasted either 70 minutes or until mice earned 30 reinforcers from responding at each of the two apertures (60 total reinforced responses), whichever occurred first. Response training was considered complete when mice acquired 60 pellets within one session.

For experiments with multiple training conditions, we next transitioned to a 30-second and then 60-second random interval (RI30 and RI60) schedule of reinforcement. During RI training sessions, random intervals of time (30 or 60 seconds on average) were introduced, during which responses produced no effect. These sessions ended at either 70 minutes or when 60 reinforcers were earned, whichever occurred first. The number of sessions is indicated graphically.

3.6.5 Test for response flexibility

We assess the capacity of mice to modify response strategies based on reinforcer likelihood. This procedure manipulated response-outcome contingencies. It consisted of two 25-

minute sessions occurring on consecutive days. On the first day, one nose-poke aperture was occluded, while responding on the remaining available aperture remained reinforced. On the second day, the opposite aperture was occluded, but responding on the available aperture was no longer reinforced. Instead, pellets were delivered into the magazine at a rate equal to what each mouse experienced during their most recent day of training, independent of responding. Thus, for one aperture, the association between responding and reinforcer delivery was maintained, whereas this association was violated at the other aperture. Which aperture (left versus right) was designated to be reinforced versus non-reinforced was counterbalanced within and between groups.

Mice were returned to the same chambers on a third day for a 15-minute choice test conducted in extinction. Both apertures were available. Flexibly favoring the action that remained reinforced requires the DMS (Pitts et al., 2018), while persisting in the familiar strategy of responding on both ports following extended training requires the DLS (Zimmermann et al., 2018). For experiments when mice underwent multiple tests for response flexibility, the assignment of which aperture was to be reinforced versus non-reinforced was reversed so that mice could not use information from prior tests to inform their subsequent response strategies.

3.6.6 CNO administration

The chemogenetic receptor ligand CNO (1 mg/kg, *i.p.*, in 2% DMSO and saline; RTI International) was administered *i.p.* (1 mL/100 g) 30 minutes before the non-reinforced session. All mice received CNO to equally expose mice to any unintended consequences of the drug. Importantly, this dose does not result in detectable reverse-metabolism to clozapine (Manvich et al., 2018), nor does it impact responding in this task (D. C. Li et al., 2022; Whyte et al., 2019).

3.6.7 Cocaine administration

Cocaine (30 mg/kg in saline; Sigma-Aldrich) was administered *i.p.* daily for 14 days in a volume of 1 mL/100 g. This was followed by a two-week washout period when mice remained undisturbed in their home cages. Instrumental response training then commenced.

3.6.8 Setmelanotide administration

(*Supplementary Fig. 3.S2*)**.** Setmelanotide hydrochloride (0.2-1.0 mg/kg in 0.1% BSA and saline; Creative Peptides Inc.) was administered *i.p.* one or two times per mouse, as described in the figures, in a volume of 1 mL/100 g. Setmelanotide was administered immediately prior to sessions measuring food intake and locomotor activity or 30 minutes prior to both the reinforced and non-reinforced sessions of the test for response flexibility.

3.6.9 Locomotion and *ad libitum* feeding

(*Supplementary Fig. 3.S1-S2*). Mice were placed in large clean cages equipped with 16 photobeams in dimly lit testing rooms. Locomotor activity was monitored by photobeam for 1 hour, then CNO or setmelanotide was administered as described. CNO-treated mice were monitored for another hour. Setmelanotide-treated mice were monitored for two more hours in the presence of vivarium chow, allowing for simultaneous assessment of food consumption.

CNO-treated mice were lastly injected with D-amphetamine (3 mg/kg in saline; Sigma-Aldrich), *i.p.,* ahead of a third hour of locomotor monitoring. The purpose was to stimulate locomotor activity and evaluate consequences of chemogenetic constructs, if any.

3.6.10 Histology

Mice were deeply anesthetized with ketamine (120 mg/kg, *i.p.*) and xylazine (10 mg/kg, *i.p.*) and trans-cardially perfused with ice-cold PBS and 4% paraformaldehyde (PFA). Brains completed fixation in 4% PFA over 48 hours, then were transferred to 30% w/v sucrose. Fixed brains were sectioned into 50 µm coronal sections at -20 ºC using a freezing microtome. The fluorescent tags on the viral vectors were visualized using a Keyence BZ-X710 microscope by a blinded investigator.

3.6.11 Immunofluorescence imaging

Fixed sections were generated as above and blocked in a solution containing 2% NGS, 1% BSA, and 0.03% Triton X-100 (Sigma) for 90 minutes at room temperature. Then, sections were incubated with anti-cFos (1:1000; Abcam, ab190289, lot GR3433493-1), 2% NGS, and 0.03% Triton X-100 at 4ºC overnight. Sections were then incubated in a secondary antibody solution containing Alexa Fluor 647 (1:500; Life Technologies, lot GR113912-1), 2% NGS, and 0.03% Triton X-100 at room temperature for 1 hour. Sections were mounted and cover slipped.

Immunostained sections were imaged using a Keyence BZ-X710 microscope. Images were obtained at 20x magnification. Uniform exposure parameters were used. For cFos quantification, analyses were performed using ImageJ software. The analysis pipeline included standardizing a region of interest (ROI), background subtracting, intensity thresholding (Otsu method), and automated cell counting within the defined ROI (Woon et al., 2022). Sections from each mouse were considered independent samples, and each mouse contributed 3-5 values to the analyses.

3.6.12 Synaptoneurosome preparation

Mice were deeply anesthetized with isoflurane and rapidly decapitated. Brains were immediately extracted, flash-frozen, and stored at -80 ºC. Frozen brains were sectioned into 1 mm coronal slices and the DMS from each hemisphere (control versus *Mc4r* knockdown) was dissected using a fluorescent dissecting microscope to visualize and confirm viral vector placement. Tissue was homogenized on ice in Syn-PER Reagent (10 mL per g tissue; Thermo Fisher Scientific) using a Dounce tissue grinder. Homogenate was centrifuged at 1,200 x *g* for 10 minutes at 4 ºC. The resulting pellet was discarded. The supernatant was transferred to a clean centrifuge tube and centrifuged at 15,000 x *g* for 20 minutes at 4 ºC. The supernatant (cytosolic fraction) and pellet (synaptoneurosome fraction) were separated, and the pellet was resuspended in Syn-PER Reagent. Samples were stored at -80 ºC prior to protein quantification by western blot.

3.6.13 Western blotting

Protein concentrations were determined using a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Then, 15 µg of total protein for each sample was separated by SDS-PAGE on 4-20% gradient Tris-glycine gels (Bio-Rad). Protein was transferred to a PVDF membrane (Bio-Rad), blocked with 5% non-fat dry milk and incubated overnight at 4 ºC in the primary antibodies listed in Table 2. Membranes were incubated in horseradish peroxidase-conjugated secondary antibody (goat anti-rabbit-HRP or horse anti-mouse-HRP; Table 2) for 1 hour at room temperature. Immunoreactivity was assessed using Pierce ECL chemiluminescence substrate (Thermo Fisher Scientific) and measured using a ChemiDoc XRS+ Imaging System (Bio-Rad). Densitometry values were obtained using Image Lab Software (Bio-Rad, version 5.0). All samples were processed and quantified in duplicate by experimenters blinded to group. Protein levels were

expressed as a fold change compared to the mean of the control samples from each respective membrane. Each mouse contributed a single value per hemisphere.

3.6.14 Trans-synaptic retrograde tracing

Mice were euthanized via trans-cardial perfusion with 4% PFA as described above. Each fixed whole brain was sectioned serially into 50 µm coronal sections using a freezing microtome. Sections were imaged to visualize GFP (the rabies virus) and mCherry (the helper construct) using a Keyence BZ-X710 microscope. Brain structures were registered using the Allen Mouse Brain Atlas (Allen Institute for Brain Science, 2004). Qualitative estimates of GFP expression were made by considering the signal strength and number of labeled cells. The following four-point density scale was used: +++, high density; ++, moderate density; + low density above background; –, only background.

3.6.15 Statistical analysis

Behavioral measures, immunohistochemistry puncta, and western blot measures were compared by analysis of variance (with repeating measures when necessary) or *t*-test, as appropriate, using IBM SPSS (v.28.0.0.0). Post-hoc tests (paired or unpaired *t*-tests, or Tukey's HSD, as appropriate) were applied following interactions or main effects between >2 groups, and significant comparisons are denoted in the figures. Comparisons were 2-tailed except for western blot measures of MC4R, conducted with the *a priori* hypothesis that silencing *Mc4r* gene expression would lead to a reduction in MC4R protein content. Planned comparisons were applied to response rates in Supplementary Fig. 3.S2g, based the data reported in main text Fig. 3.2-3.3 concerning the impact of *Mc4r* knockdown. *p*-values of ≤ 0.05 were considered significant. All

experiments were replicated at least once. Group sizes and the number of replications are denoted in the figure legends. Graphs were made using GraphPad Prism (v.10.0.2).

Exclusions. Values ±2 standard deviations from the mean were considered outliers. Outlying values were excluded if the associated mouse was tested only in a single test. If a mouse was tested repeatedly, the offending datapoint was replaced with the group mean. No group mean replacement affected the outcome of the experiment, relative to simply excluding the mouse. Across all experiments and multiple replications, 21 nose poke counts during choice tests were affected.

Mice were also excluded if histological analyses revealed that viral vectors were mis-located, being detected outside of the target region (DMS, CeA, DLS, VS, or combination, as appropriate). A total of 41 mice across all experiments were excluded for this reason, primarily due to the small size of the CeA, making localization at times challenging. Importantly, histological analyses on all mice were completed while blind to animal ID.

Analysis of spatial transcriptomics data. We analyzed the publicly available dataset collected by Roman et al., 2023, which reports cell-type specific transcriptional signatures of D1-MSNs within anatomical domains of the dorsal striatum. *Mc4r* transcript counts were correlated with transcript counts of each of 12,306 genes. Benjamini-Hochberg correction (Q = 0.2) was used to determine significant co-variates. The ShinyGO v0.77 application (http://bioinformatics.sdstate.edu/go/) was used for Gene Ontology (GO) enrichment analysis of differentially correlated genes.

3.7 FUNDING

This work was supported by the National Institutes of Health [DA055447, DA044297, and MH117103] and the Marcus Foundation. The Emory National Primate Research Center is supported by NIH OD P51 011132.

3.8 ACKNOWLEDGEMENTS

We thank Trinity Pruitt, Travis Fulton, and Dr. Sumeet Sharma for valuable assistance.

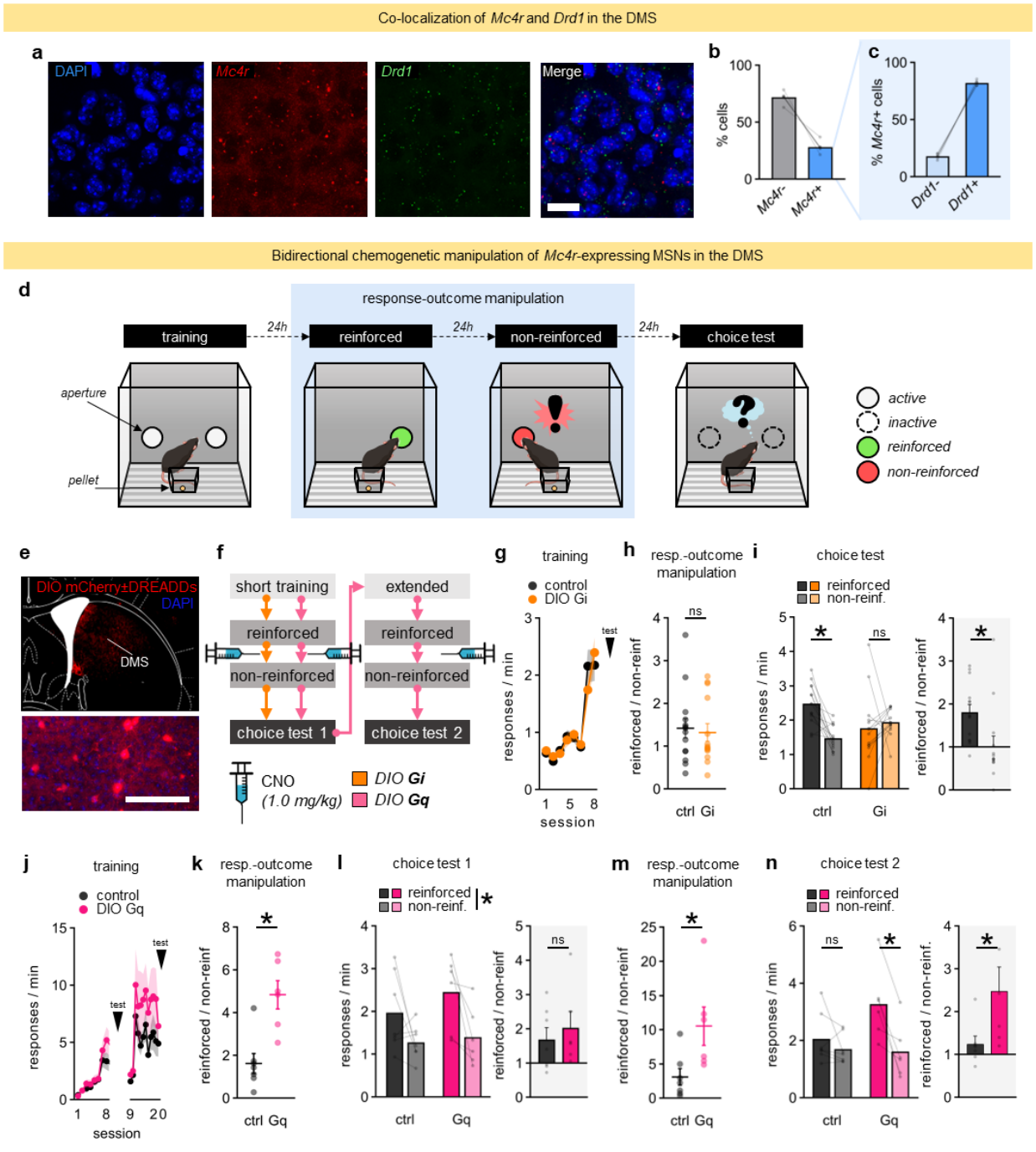

Figure 3.1 MC4R-expressing cells in the DMS bidirectionally control behavioral flexibility. a, Localization of *Mc4r* and *Drd1* mRNA in the DMS via RNAscope. Scale bar, 20 µm. **b-c**, A minority of cells in the DMS express *Mc4r* (**b**), which is preferentially colocalized with *Drd1* (**c**). **d**, Mice were trained to generate two food-reinforced nose-poke responses, and then the

food associated with one response was delivered independent of responding. Choice was assessed the next day. **e**, Viral vector expression in *Mc4r*-expressing cells. Scale bar, 50 µm. **f**, Experiment timeline. **g**, Responses across training (session: $F_{7,182}$=47.0, p<0.001; session × group: $F_{7,182}$=0.94, p=0.48). **h**, Responses during the response-outcome manipulation (Gi *vs.* control: t_{26}=0.33, p=0.75). **i**, Responses during the subsequent choice test differed between groups (reinforcement × group: $F_{1,26}$=9.74, p=0.004). **Inset**: ratios of responding for reinforced / non-reinforced condition (Gi *vs.* control: t_{26}=2.66, p=0.01). **j**, Responses across training (session: $F_{19,209}$=17.3, p<0.001; session × group: $F_{19,209}$=1.15, p=0.31). **k**, Responses during the response-outcome manipulation following short training differed between groups (Gq *vs.* control: t_{11}=-4.06, p=0.002). **l**, Choice test responses following short training (reinforcement × group: $F_{1,11}$=0.42, p=0.53). **Inset**: ratios of responding for reinforced / non-reinforced condition (Gq *vs.* control: t_{11}=-0.59, p=0.57). **m**, Responses during the response-outcome manipulation following extended training differed between groups (Gq *vs.* control: t_{11}=-2.57, p=0.03). **n**, Choice test responses following extended training also differed between groups (reinforcement × group: $F_{1,11}$=6.83, p=0.02). **Inset**: ratios of responding for reinforced / non-reinforced condition (Gq *vs.* control: t_{11}=-2.23, p=0.05). Experiments were replicated at least once with concordant results. See Supplementary Table 2 for complete statistics. *p<0.05, *ns* non-significant.

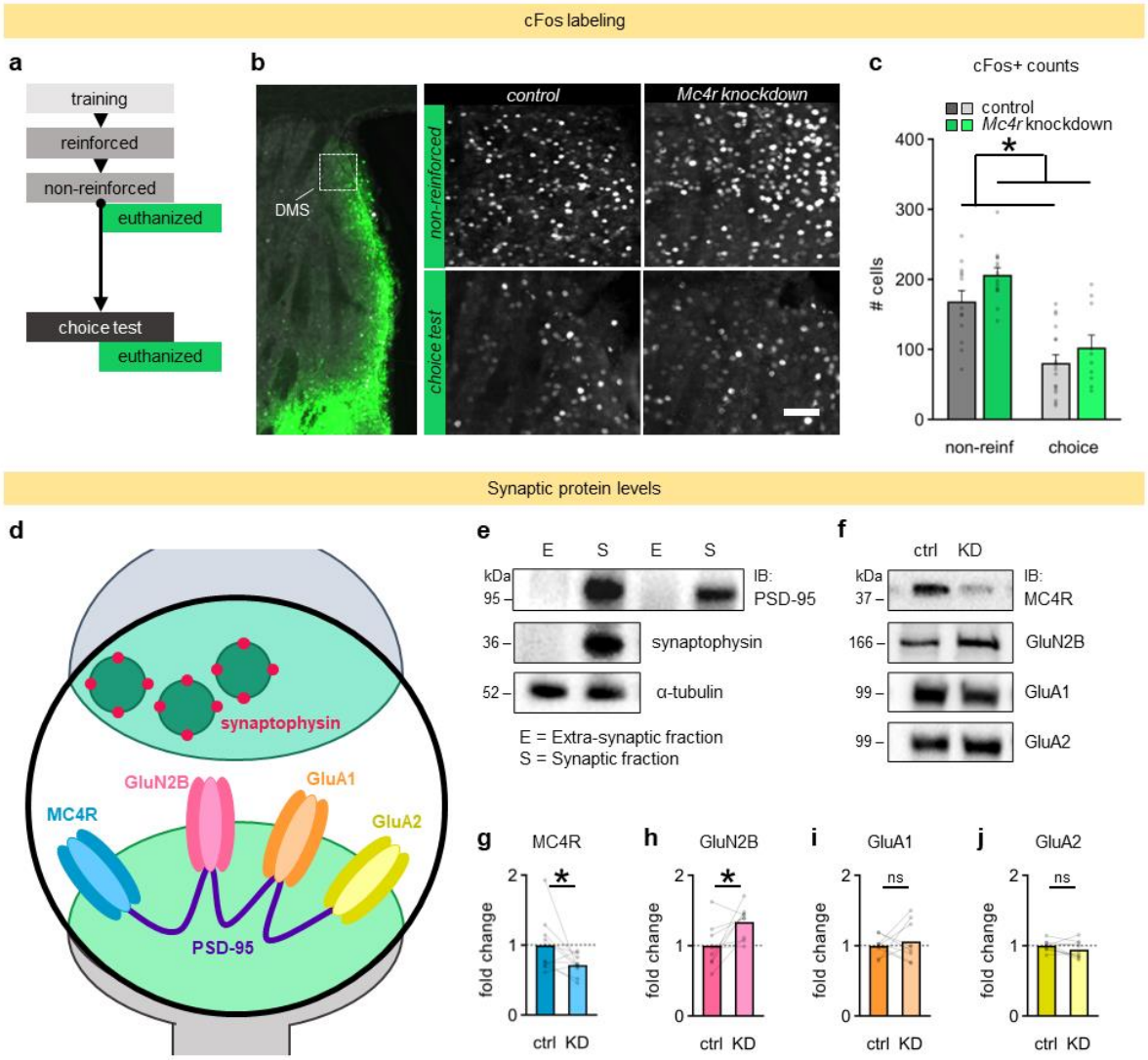

Figure 3.2 MC4R suppresses molecular markers of cell activity and neuronal plasticity in the DMS. a, Timing of euthanasia following behavioral testing. **b**, Representative cFos (white) at the viral vector infusion site (green). Scale bar, 50 µm. **c**, cFos was elevated by *Mc4r* silencing and in the non-reinforcement condition overall (group: $F_{1,51}$=5.32, p=0.03; timepoint: $F_{1,51}$=51.1, p<0.001). **d**, Diagram of protein localization within a synaptoneurosome. **e-f**, Representative immunoblots (**e**) from extra-synaptic (E) and synaptic (S) fractions, validating synaptoneurosome preparation, and (**f**) measuring protein levels in the S fraction. **g-j**, Fold change from control mean

set at 1 of synaptic levels of (**g**) MC4R (t_9=2.2, p=0.03), (**h**) GluN2B (t_8=-2.77, p=0.02), (**i**) GluA1 (t_8=-0.54, p=0.61), and (**j**) GluA2 (t_8=1.48, p=0.18). Experiments were replicated at least once with concordant results. See Supplementary Table 3 for complete statistics. *p<0.05, *ns* non-significant.

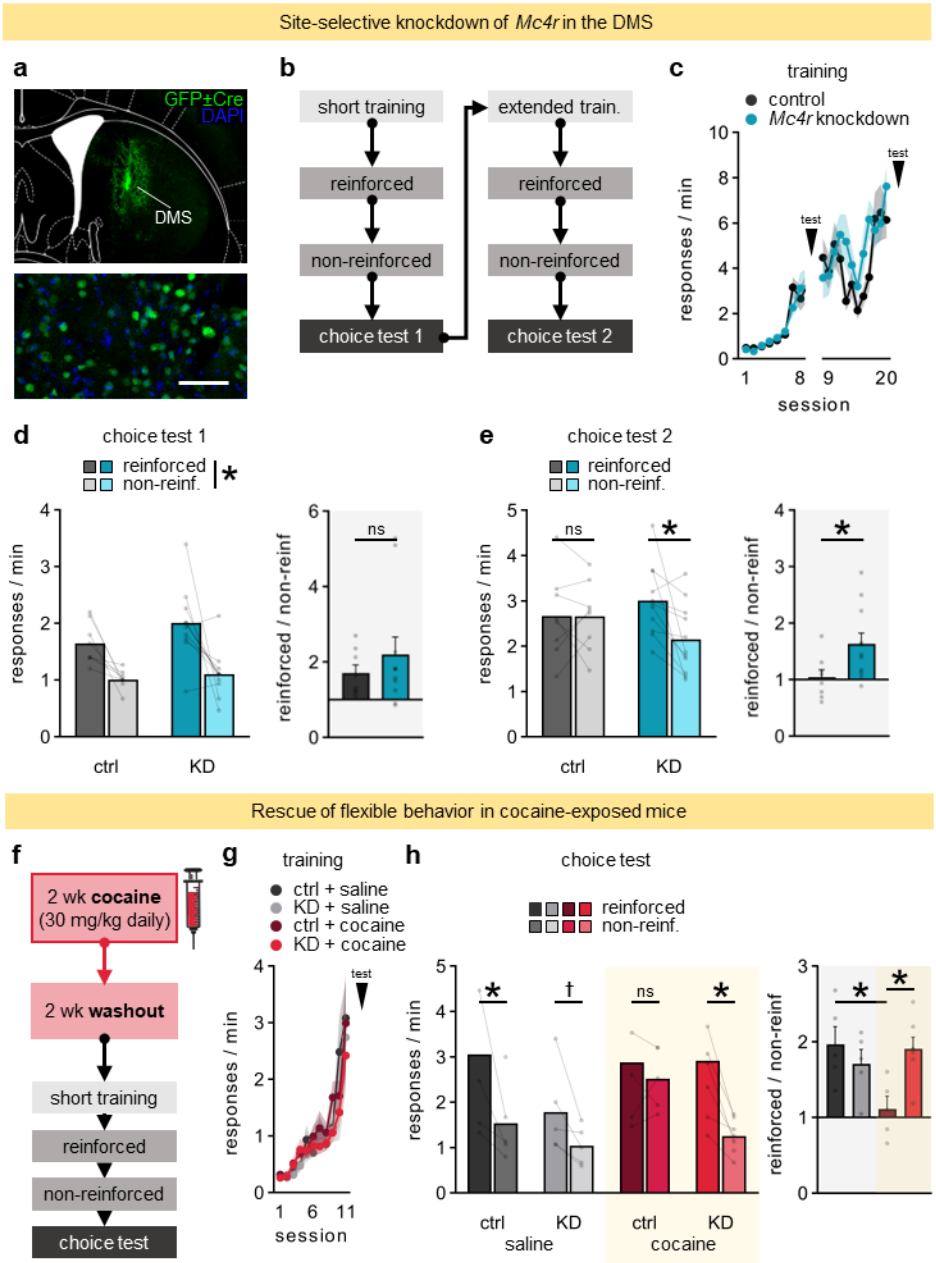

Figure 3.3 **MC4R brakes flexible action selection. a**, Representative viral vector infusion in the DMS. Scale bar, 50 µm. **b**, Experiment timeline. **c**, Responses across training (session: $F_{19,323}$=27.2, p<0.001; session × group: $F_{19,323}$=1.39, p=0.131). **d**, Choice test responses following short training (reinforcement × group: $F_{1,17}$<1). **Inset**: ratios of responding for reinforced / non-reinforced condition (knockdown [kd] *vs.* control [ctrl]: t_{17}=-0.85, p=0.41). **e**, Choice test

responses following extended training differ between groups (reinforcement × group: $F_{1,17}$=5.51, p=0.03). **Inset**: ratios of responding for reinforced / non-reinforced condition (knockdown *vs.* control: t_{17}=-2.31, p=0.03). **f**, Experiment timeline. **g**, Responses across training (session: $F_{10,180}$=30.2, p<0.001; all interactions non-significant). **h**, Choice test responses following cocaine and *Mc4r* gene silencing differed between groups (reinforcement × knockdown × cocaine: $F_{1,18}$=4.46, p=0.049). **Inset**: ratios of responding for reinforced / non-reinforced condition (knockdown × cocaine: $F_{1,18}$=7.89, p=0.01; ctrl+sal *vs.* ctrl+coc: p=0.03; ctrl+coc *vs.* kd+coc: p=0.03). Experiments were replicated at least once with concordant results. See Supplementary Table 4 for complete statistics. *p<0.05, †p<0.1, *ns* non-significant.

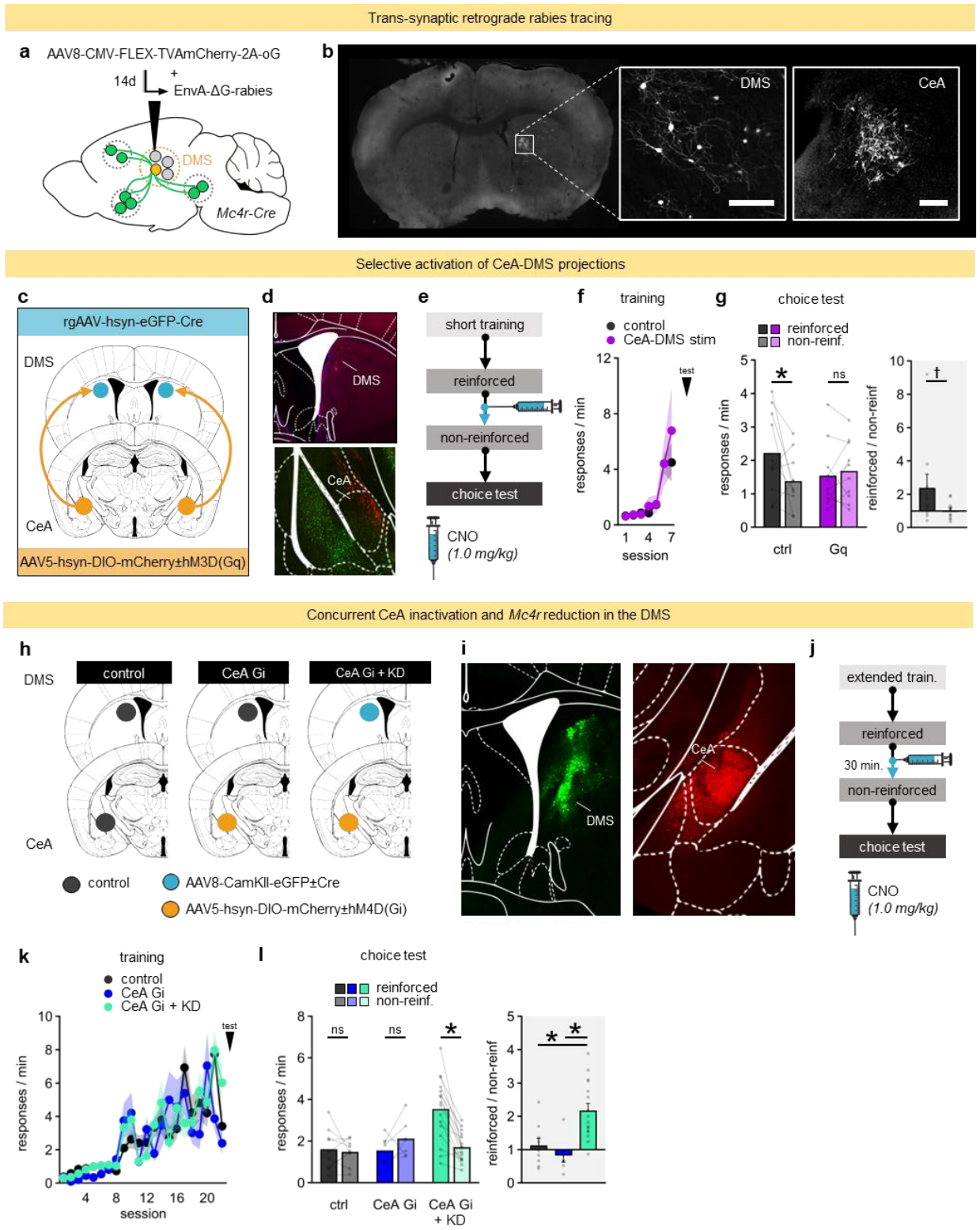

Figure 3.4 Direct CeA-DMS projections onto MC4R+ cells impede DMS-dependent behavioral flexibility. a, Schematic of modified rabies virus expression in MC4R-expressing cells

in the DMS. **b**, Representative images of rabies virus in the DMS and CeA. Scale bar, 50 µm. **c-d**, Viral vector to achieve stimulation of CeA-to-DMS projections. Here, the CeA expresses mCherry, with terminals visible in the DMS. **e**, Experiment timeline. **f**, Responses across training (session: $F_{8,80}$=15.97, p<0.001; session × group: $F_{8,80}$<1). **g**, Activating CeA-to-DMS projections obstructed flexible choice (reinforcement × group: $F_{1,10}$=5.40, p=0.012). **Inset**: ratios of responding for reinforced / non-reinforced condition (stimulation *vs.* control: t_{17}=3.11, p=0.01). **h-i**, Schematic of unilateral inhibitory chemogenetic viral vector in the CeA and site-selective *Mc4r* knockdown in the DMS. **j**, Experiment timeline. **k**, Responses across training (session: $F_{21,609}$=21.21, p<0.001; session × group: $F_{42,609}$=2.16, p<0.001). **l**, Choice test responses following extended training (reinforcement × group: $F_{2,29}$=13.67, p<0.001). Experiments were replicated at least once with concordant results. See Supplementary Table 5 for complete statistics. *p<0.05, *ns* non-significant.

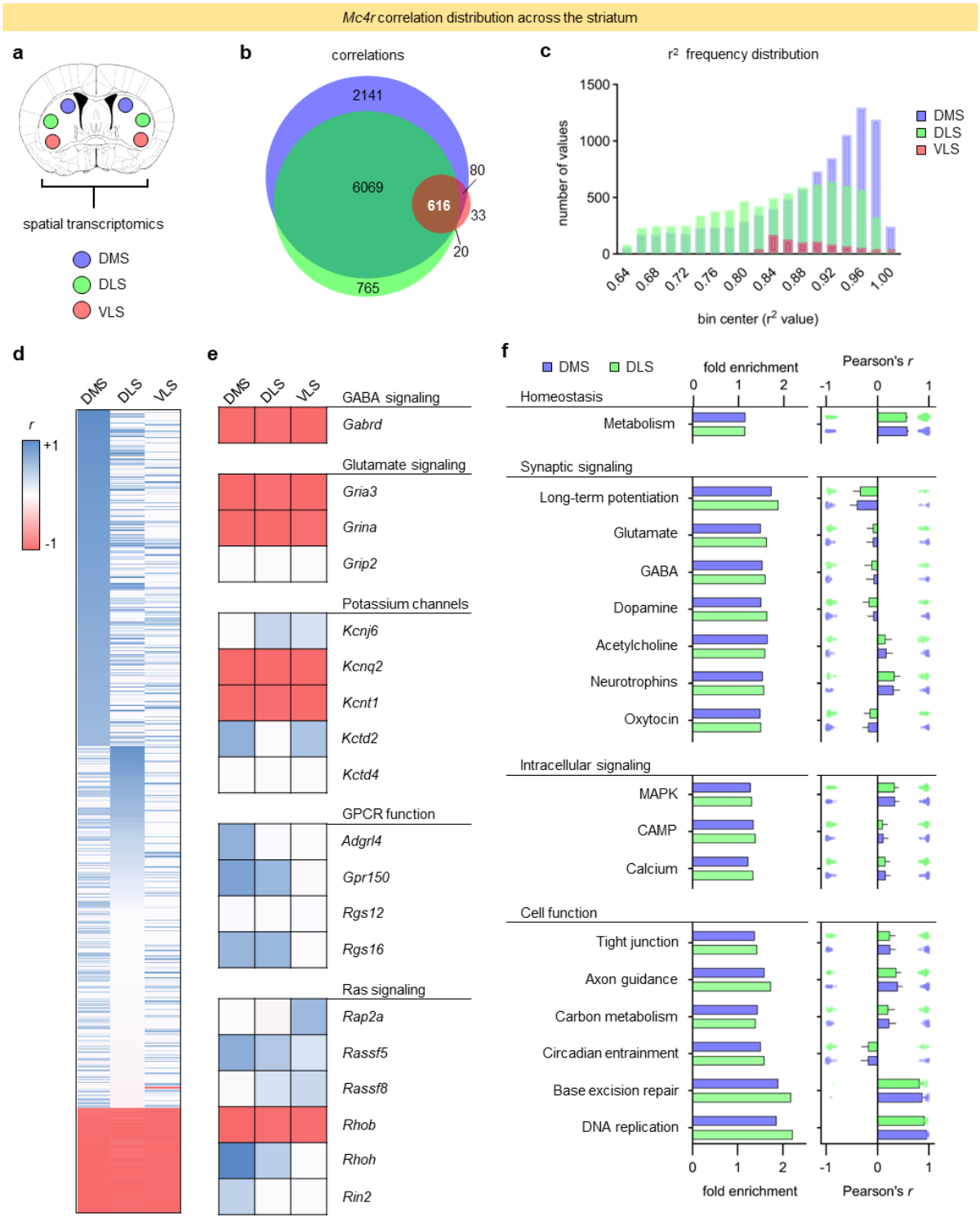

Figure 3.5 Genetic co-variates of *Mc4r* transcript vary in number, strength, and directionality across the dorsal and ventral striatum. a, Digital Spatial Profiling was used to

create transcriptomic profiles of D1 receptor-expressing cells in the DMS, DLS, and VLS (publicly available in Roman et al., 2023). **b**, A Venn diagram of the numbers of co-variates with *Mc4r* transcription in the DMS (blue), DLS (green), and VLS (pink). **c**, The frequency distribution of r^2 values measuring correlations with *Mc4r* transcription by region. Following Benjamini-Hochberg corrections, r^2 values ≥ 0.64 were considered significant. **d**, Heat map of Pearson's *r* values for the 616 genes correlated with *Mc4r* transcription in the DMS, DLS, *and* VLS. **e**, Comparison of Pearson's *r* values across regions for highlighted genes. **f**, Pathways enriched in the genetic co-variates shared between the DMS and DLS.

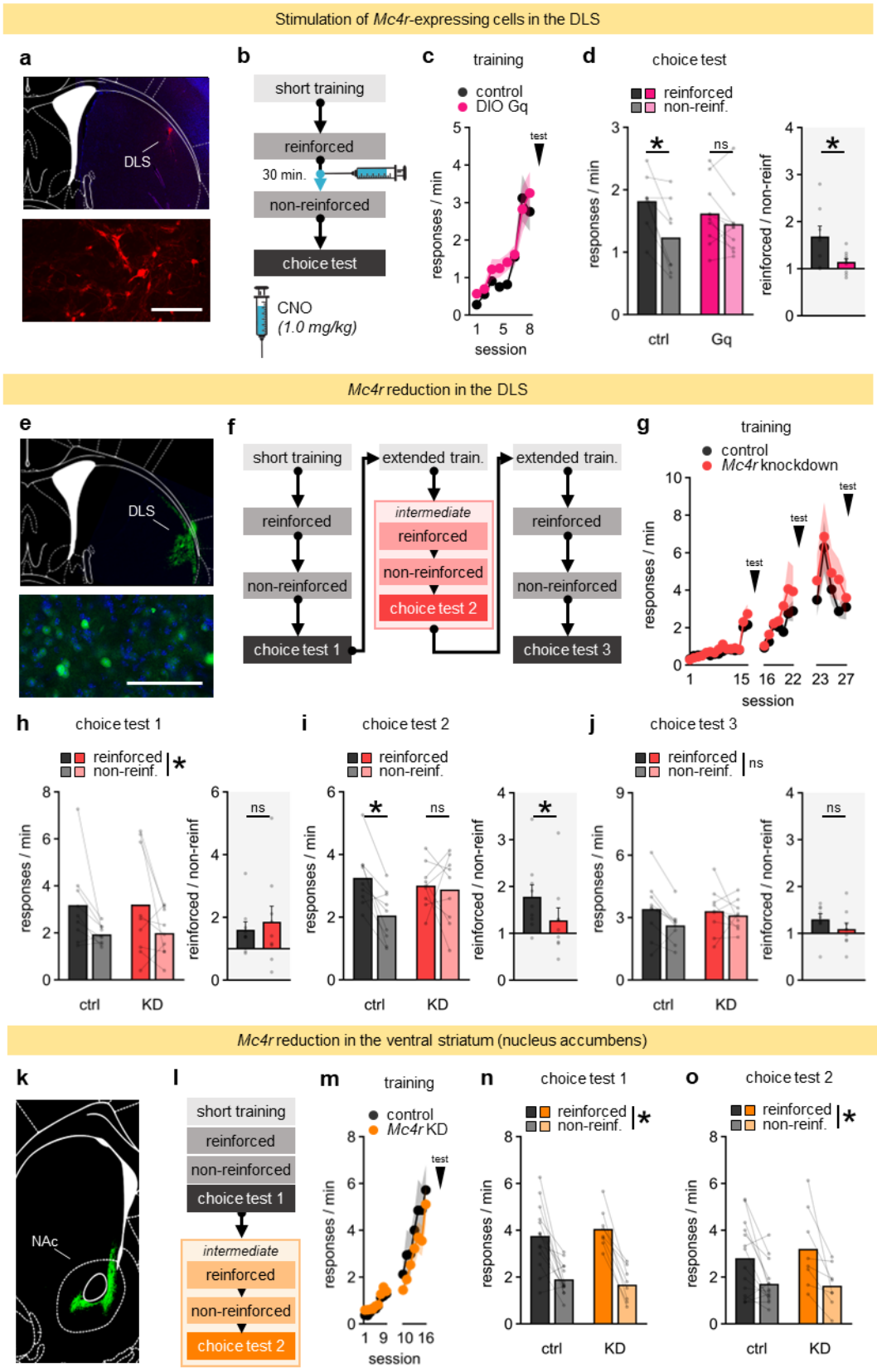
Stimulation of Mc4r-expressing cells in the DLS
a
DLS
b
short training
reinforced
30 min.
non-reinforced
choice test
CNO
(1.0 mg/kg)
c
training
control
DIO Gq
test
responses / min
session
d
choice test
reinforced
non-reinf.
responses / min
ctrl
Gq
reinforced / non-reinf
Mc4r reduction in the DLS
e
DLS
f
short training
reinforced
non-reinforced
choice test 1
extended train.
intermediate
reinforced
non-reinforced
choice test 2
extended train.
reinforced
non-reinforced
choice test 3
g
training
control
Mc4r knockdown
test
responses / min
session
h
choice test 1
reinforced
non-reinf.
responses / min
ctrl
KD
reinforced / non-reinf
i
choice test 2
j
choice test 3
Mc4r reduction in the ventral striatum (nucleus accumbens)
k
NAc
l
short training
reinforced
non-reinforced
choice test 1
intermediate
reinforced
non-reinforced
choice test 2
m
training
control
Mc4r KD
n
choice test 1
o
choice test 2

Figure 3.6 Both chemogenetic stimulation of MC4R+ cells and site-selective reduction of Mc4r in the DLS promote habitual behavior. **a**, Representative infusion of an excitatory chemogenetic construct in the DLS. Scale bar, 50 µm. **b**, Experiment timeline. **c**, Responses across training (session: $F_{7,112}$=26.4, p<0.001; session × group: $F_{7,112}$<1). **d**, Choice test responses following short training (reinforcement × group: $F_{1,16}$=5.46, p=0.03). **Inset:** ratios of responding for reinforced / non-reinforced condition (Gq *vs.* control: t_{16}=2.60, p=0.02). **e**, Representative infusion resulting in site-selective reduction of *Mc4r* in the DLS. Scale bar, 50 µm. **f,** Experiment timeline. **g**, Responses across training (session: $F_{26,416}$=17.1, p<0.001; session × group: $F_{26,416}$<1). **h**, Choice test responses following short training (reinforcement × group: $F_{1,16}$<1). **Inset:** ratios of responding for reinforced / non-reinforced condition (knockdown [kd] *vs.* control: t_{16}=-0.46, p=0.65). **i**, Choice test responses following intermediate training (reinforcement × group: $F_{1,16}$=4.3, p=0.055). **Inset:** ratios of responding for reinforced / non-reinforced condition (kd *vs.* control: t_{16}=2.78, p=0.014). **j**, Choice test responses following extended training (reinforcement: $F_{1,16}$=4.09, p=0.06; reinforcement × group: $F_{1,16}$=1.44, p=0.25). **Inset:** ratios of responding for reinforced / non-reinforced condition (kd *vs.* control: t_{16}=1.19, p=0.25). **k**, Representative infusion resulting in site-selective reduction of *Mc4r* in the NAc. **l**, Experiment timeline. **m**, Responses across training (session: $F_{15,300}$=26.5, p<0.001; session × group: $F_{15,300}$<1). **n**, Choice test responses following short training (reinforcement × group: $F_{1,20}$<1). **o**, Choice test responses following intermediate training (reinforcement × group: $F_{1,20}$<1). See Supplementary Table 6 for complete statistics. *p<0.05, *ns* non-significant.

Vector	Manufacturer	Product Number	Lot	Depositor
AAV5-hSyn-DIO-hM3D(Gq)-mCherry	Addgene	44361	v112948	B. Roth
AAV5-hSyn-DIO-hM4D(Gi)-mCherry	Addgene	44362	v80579	B. Roth
AAV5-hSyn-DIO-mCherry	Addgene	50459	v168135	B. Roth
AAV8-CamKII-HI-GFP-Cre-WPRE-SV40	UNC Core	n/a	AV5055B/C	
AAV8-CamKII-eGFP	Addgene	50469	v105998	B. Roth
AAV8-CMV-FLEX-TVAmCherry2A-oG	Salk Institute	102985	n/a	
EnvA G-Deleted Rabies-eGFP	Salk Institute	32635	n/a	
AAV5-CamKII-hM4D(Gi)-mCherry	Addgene	50477	v60658	B. Roth
AAV5-CamKII-mCherry	UNC Core	n/a	AV4809E	
rgAAV-hSyn-GFP-Cre	Addgene	105540	v133141	J. Wilson

Table 3.1 Viral vector information.

Antigen	Host	Manufacturer	Product	Lot	Dilution
PSD-95	Rabbit	Cell Signaling	3450S	5	1:1000
α-tubulin	Mouse	Cell Signaling	3873S	15	1:1000
Synaptophysin	Rabbit	Abcam	ab32127	GR223336-10	1:10000
MC4R	Rabbit	Abcam	ab150419	GR105886-8	1:500
GluA1	Rabbit	Abcam	ab31232	GR152884-2	1:1000
GluA2	Rabbit	Abcam	ab206293	GR3266953-1	1:1000
GluN2B	Mouse	Novus	NB100-74475	UB295336	1:500
Rabbit-HRP	Goat	Cell Signaling	7074S	28	1:20000
Mouse-HRP	Horse	Cell Signaling	7076S	35	1:2000

Table 3.2 Western blot antibody information.

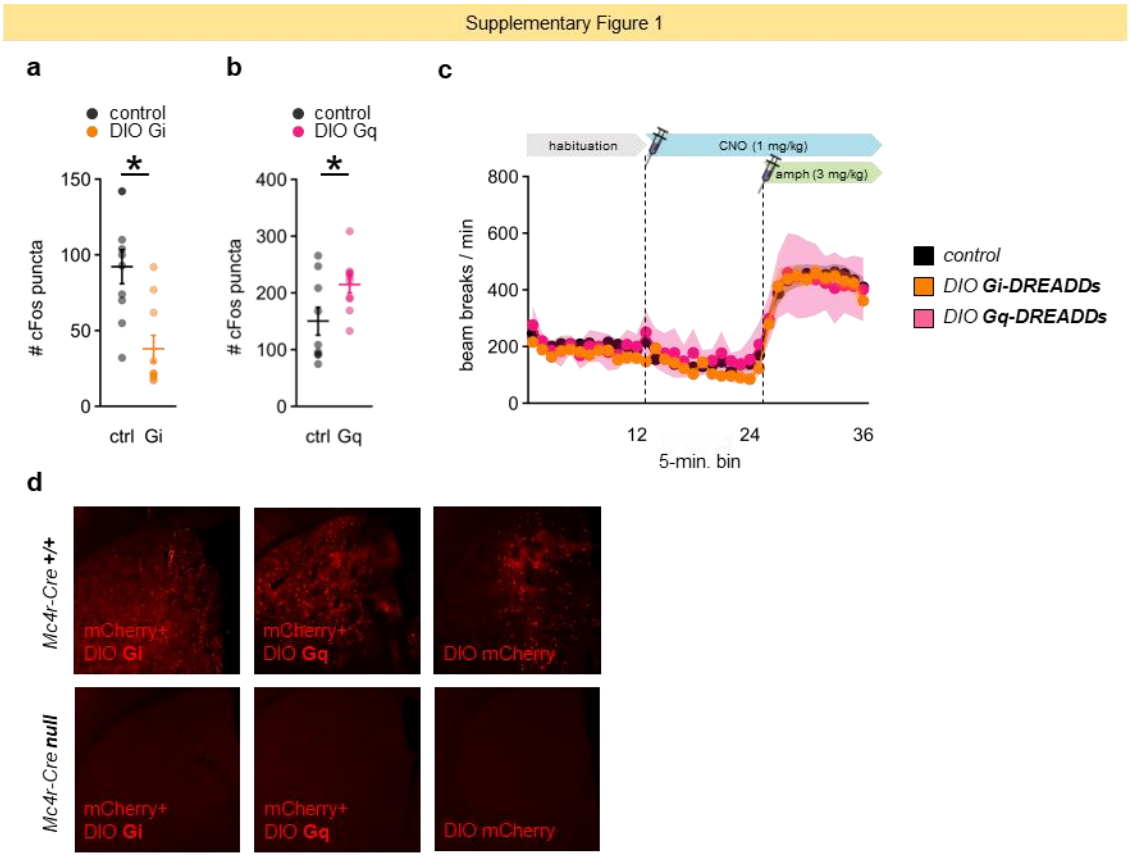

Figure 3.S1 Validation of Cre-recombinase-dependent chemogenetic viral vector constructs. a, Activating inhibitory Gi-DREADD receptors in the DMS decreased c-Fos puncta in the DMS, as expected (ctrl *vs.* Gi: t_{18}=3.79, p<0.001). **b**, Activating excitatory Gq-DREADD receptors in the DMS increase c-Fos puncta in the DMS, also as expected (ctrl *vs.* Gq: t_{17}=-2.45, p=0.047). **c**, Mice were placed in open chambers equipped with photobeams, with the rate of beam breaks representing locomotor activity. Following 1 hour of habituation, the DREADDs ligand CNO was administered. Chemogenetic receptor presence did not impact locomotion, including the locomotor response to amphetamine in hour 3 (group: $F_{2,32}$<1; time × group: $F_{70,1120}$<1). **d**, Viral vectors containing Cre-recombinase-dependent chemogenetic constructs were infused into the DMS. Cre-dependent expression was found only in Cre-expressing transgenic mouse lines. *p<0.05, *ns* non-significant.

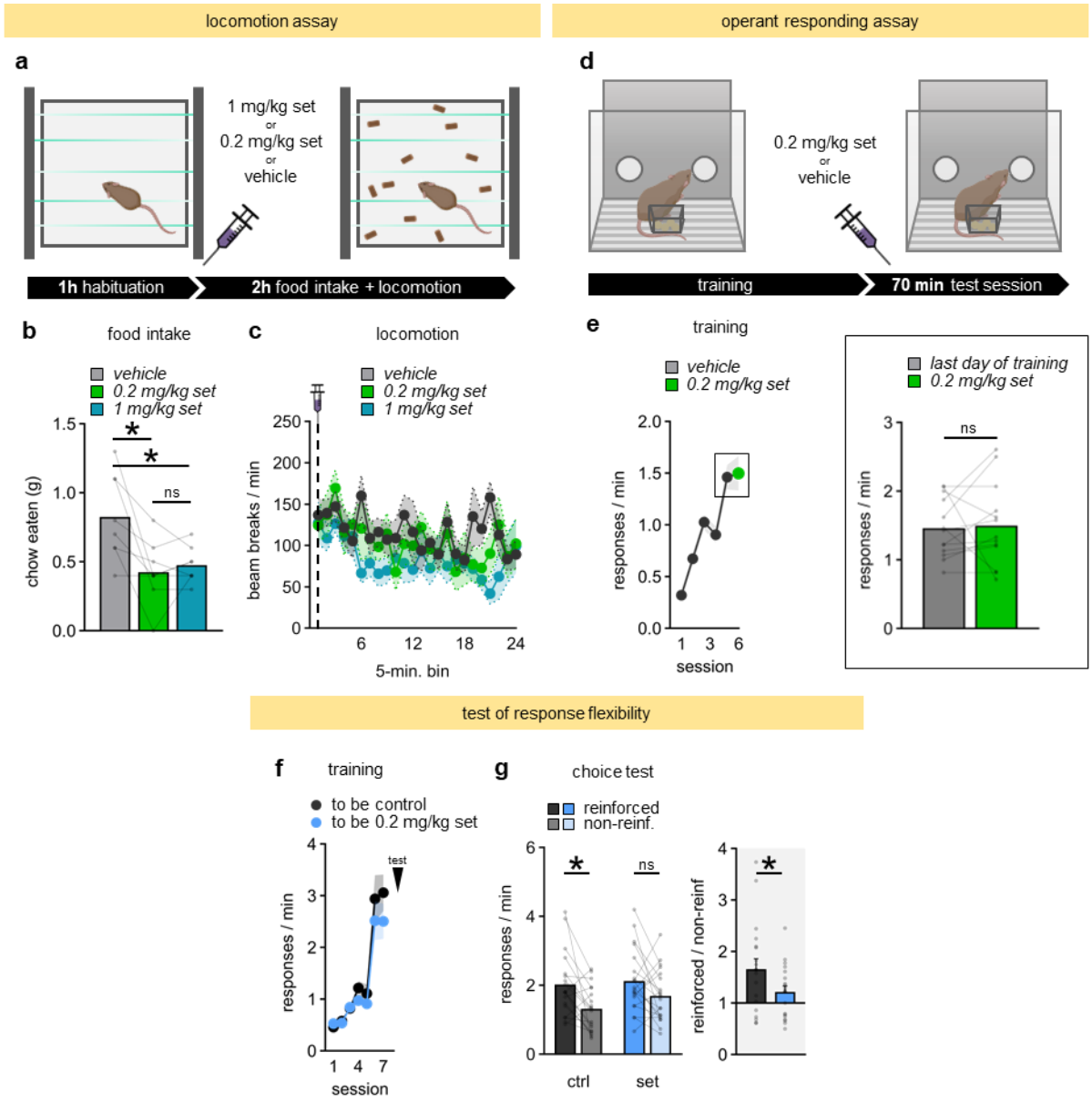

Figure 3.S2 Stimulating MC4R promotes inflexible behavior at doses that do not impact instrumental responding for food. a, Mice were habituated for one hour in open chambers equipped with photobeams, with the rate of beam breaks representing locomotor activity. Following habituation, mice were injected i.p. with one of three drug conditions: 1 mg/kg setmelanotide (set), 0.2 mg/kg set, or vehicle control. Typical vivarium chow was then evenly distributed across the floor of the cages. Mice were placed back in the cages for monitoring. Three

tests were conducted over three days with each animal experiencing each condition in a counterbalanced fashion. Food intake was measured. **b**, Both 1 and 0.2 mg/kg setmelanotide reduced food intake (dose: $F_{2,14}$=9.61, p=0.002; vehicle *vs.* 0.2 mg/kg: t_7=3.43, p=0.011; vehicle *vs.* 1 mg/kg: t_7=3.18, p=0.016; 0.2 mg/kg *vs*. 1 mg/kg: t_7=-0.80, p=0.451). **c**, Setmelanotide did not disrupt gross locomotion (dose: $F_{2,14}$=3.61, p=0.055; group × timepoint: $F_{70,490}$<1). **d**, A separate cohort of mice was trained to respond at two apertures for the delivery of a food reward. Once responding was stable, mice were injected with 0.2 mg/kg setmelanotide or vehicle control and tested in a 70 min training session. Two tests were conducted over two days with each animal experiencing both conditions in a counterbalanced fashion. **e**, Responding increased as training progressed (black line; $F_{4,48}$=41.2, p<0.001). **Inset**: Setmelanotide did not disrupt food-reinforced responding when compared to the last day of training (t_{12}=0.23, p=0.819). Thus, low-dose setmelanotide did not impact instrumental responding for food (despite suppressing *ad libitum* food intake). **f**, A separate group was trained to respond for two distinct food reinforcers. As in the main text, response rates are collapsed across both apertures. Mice were drug-naïve. Groups designated to later receive either 0.2 mg/kg setmelanotide or vehicle control did not differ (session: $F_{6,222}$=53.0, p<0.001; group: $F_{1,37}$=1.01, p=0.302; session × group: $F_{6,222}$<1). **g**, Setmelanotide was administered, and one response ceased to be reinforced, while the other remained reinforced. Control mice reduced responding when responding was not reinforced (reinforced *vs.* non-reinforced: t_{18}=3.08, p=0.007), while responding in the setmelanotide-treated mice did not significantly differ between conditions (reinforced *vs.* non-reinforced: t_{19}=1.90, p=0.073). To further substantiate this perspective, we generates response ratios, which significantly differed between groups (t_{37}=1.93, p=0.03), indicating that blocking MC4R reduced response flexibility. Experiments were replicated at least once with concordant results. *p<0.05, *ns* non-significant.

Region	Density
Cortex	
Medial prefrontal cortex	–
Orbitofrontal cortex	+
Primary motor cortex	++
Primary sensory cortex	++
Amygdala	
Central amygdala	+++
Basal amygdala	–
Thalamus	
Central thalamic nucleus	+
Parafasicular thalamic nucleus	++
Ventromedial thalamic nucleus	+
Subthalamic nucleus	+++
Hypothalamus	
Paraventricular nucleus	–
Arcuate nucleus	+
Lateral hypothalamus	+
Basal Ganglia	
Nucleus accumbens shell	–
Nucleus accumbens core	–
Globus pallidus	+++

Table 3.S1 List of projections to MC4R+ cells in the DMS. Qualitative estimates of GFP-tagged rabies tracing were made by considering both signal strength and the number of labeled cells. The following four-point density scale was used: +++, high density; ++, moderate density; + low density above background; –, only background.

Figure	Measure	Groups	Analysis	Comparison	Statistic	p-value	Effect size
1b	% cells	*n*=6	paired t-test	*Mc4r+ vs. Mc4r-*	t_5=-10.7	p<0.001	d=-4.379
1c	% *Mc4r+* cells			*Drd1+ vs. Drd1-*	t_5=36.4	p<0.001	d=14.876
1g	response rate	Gi inhibition *(n=14)* Control *(n=14)*	RM ANOVA	session	$F_{7,182}$=47.0	p<0.001	η^2=0.644
				group	$F_{1,26}$=0.01	p=0.913	η^2=0.000
				session × group	$F_{7,182}$=0.94	p=0.476	η^2=0.035
1h	response ratio		unpaired t-test	Gi *vs.* control	t_{26}=0.325	p=0.747	d=0.123
1i left	response rate		RM ANOVA	aperture	$F_{1,26}$=4.76	p=0.038	η^2=0.155
				group	$F_{1,26}$=0.75	p=0.396	η^2=0.028
				aperture × group	$F_{1,26}$=9.74	p=0.004	η^2=0.273
			post-hoc: paired t-test	reinf. *vs.* non-reinf. (ctrl)	t_{13}=4.93	p<0.001	d=1.318
				reinf. *vs.* non-reinf. (Gi)	t_{13}=-0.56	p=0.587	d=-0.149
1i right	response ratio		unpaired t-test	Gi *vs.* control	t_{26}=2.66	p=0.013	d=1.006
1j	response rate	Gq excitation *(n=6)* Control *(n=7)*	RM ANOVA	session	$F_{19,209}$=17.3	p<0.001	η^2=0.612
				group	$F_{1,11}$=1.51	p=0.245	η^2=0.121
				session × group	$F_{19,209}$=1.15	p=0.309	η^2=0.094
1k	response ratio		unpaired t-test	Gq *vs.* control	t_{11}=-4.06	p=0.002	d=-2.258
1l left	response rate		RM ANOVA	aperture	$F_{1,11}$=9.79	p=0.010	η^2=0.471
				group	$F_{1,11}$=1.00	p=0.339	η^2=0.083
				aperture × group	$F_{1,11}$=0.42	p=0.533	η^2=0.036
1l right	response ratio		unpaired t-test	Gq *vs.* control	t_{11}=-0.59	p=0.565	d=-0.330
1m	response ratio		unpaired t-test	Gq *vs.* control	t_{11}=-2.57	p=0.026	d=-1.428
1n left	response rate		RM ANOVA	aperture	$F_{1,11}$=16.5	p=0.002	η^2=0.600
				group	$F_{1,11}$=1.58	p=0.235	η^2=0.126
				aperture × group	$F_{1,11}$=6.83	p=0.024	η^2=0.383
			post-hoc: paired t-test	reinf. *vs.* non-reinf. (ctrl)	t_6=1.14	p=0.297	d=0.431
				reinf. *vs.* non-reinf. (Gq)	t_5=4.24	p=0.008	d=1.730
1n right	response ratio		unpaired t-test	Gq *vs.* control	t_{11}=-2.23	p=0.048	d=-1.238

Table 3.S2 Complete statistics for experiments reported in Figure 3.1.

Figure	Measure	Groups	Analysis	Comparison	Statistic	p-value	Effect size
2c	cFos puncta	non-reinf ctrl *(n=13)* non-reinf kd *(n=15)* choice ctrl *(n=17)* choice kd *(n=10)*	ANOVA	timepoint	$F_{1,51}=51.1$	$p<0.001$	$\eta^2=0.501$
				group	$F_{1,51}=5.32$	$p=0.025$	$\eta^2=0.094$
				timepoint × group	$F_{1,51}=0.49$	$p=0.487$	$\eta^2=0.010$
2g	fold change	*n=10*	1-tail paired t-test	control *vs.* knockdown	$t_9=2.20$	$p=0.028$	$d=0.695$
2h		*n=9*	2-tail paired t-test		$t_8=-2.77$	$p=0.024$	$d=-0.924$
2i					$t_8=-0.54$	$p=0.608$	$d=-0.178$
2j					$t_8=1.48$	$p=0.176$	$d=0.495$

Table 3.S3 Complete statistics for experiments reported in Figure 3.2.

Figure	Measure	Groups	Analysis	Comparison	Statistic	p-value	Effect size
3c	response rate	control *(n=8)* *Mc4r* kd (*n=11)*	RM ANOVA	session	$F_{19,323}=27.2$	$p<0.001$	$\eta^2=0.616$
				group	$F_{1,17}=0.42$	$p=0.525$	$\eta^2=0.024$
				session × group	$F_{19,323}=1.39$	$p=0.131$	$\eta^2=0.075$
3d left			RM ANOVA	aperture	$F_{1,17}=20.7$	$p<0.001$	$\eta^2=0.549$
				group	$F_{1,17}=3.54$	$p=0.077$	$\eta^2=0.172$
				aperture × group	$F_{1,17}=0.614$	$p=0.444$	$\eta^2=0.035$
3d right	response ratio		unpaired t-test	control *vs. Mc4r* kd	$t_{17}=-0.85$	$p=0.408$	$d=-0.394$
3e left	response rate		RM ANOVA	aperture	$F_{1,17}=5.67$	$p=0.029$	$\eta^2=0.250$
				group	$F_{1,17}=0.02$	$p=0.883$	$\eta^2=0.001$
				aperture × group	$F_{1,17}=5.51$	$p=0.031$	$\eta^2=0.245$
			post-hoc: paired t-test	reinf. *vs.* non-reinf. (ctrl)	$t_7=0.03$	$p=0.976$	$d=0.011$
				reinf. *vs.* non-reinf. (kd)	$t_{10}=3.16$	$p=0.010$	$d=0.954$
3e right	response ratio		unpaired t-test	control *vs. Mc4r* kd	$t_{17}=-2.31$	$p=0.034$	$d=-1.073$
3g	response rate	ctrl+saline *(n=5)* kd+saline *(n=5)* ctrl+cocaine *(n=5)* kd+cocaine *(n=7)*	RM ANOVA	session	$F_{10,180}=30.2$	$p<0.001$	$\eta^2=0.627$
				knockdown	$F_{1,18}=0.63$	$p=0.438$	$\eta^2=0.034$
				cocaine	$F_{1,18}=0.37$	$p=0.551$	$\eta^2=0.020$
				session × knockdown	$F_{10,180}=0.91$	$p=0.523$	$\eta^2=0.048$
				session × cocaine	$F_{10,180}=1.24$	$p=0.268$	$\eta^2=0.064$
				knockdown × cocaine	$F_{1,18}=0.49$	$p=0.491$	$\eta^2=0.027$
				session × kd × cocaine	$F_{10,180}=0.82$	$p=0.613$	$\eta^2=0.043$
3h left			RM ANOVA	aperture	$F_{1,18}=19.1$	$p<0.001$	$\eta^2=0.514$
				knockdown	$F_{1,18}=3.25$	$p=0.088$	$\eta^2=0.153$
				cocaine	$F_{1,18}=1.69$	$p=0.210$	$\eta^2=0.086$
				aperture × knockdown	$F_{1,18}=0.29$	$p=0.599$	$\eta^2=0.016$
				aperture × cocaine	$F_{1,18}=0.06$	$p=0.803$	$\eta^2=0.004$
				knockdown × cocaine	$F_{1,18}=0.10$	$p=0.753$	$\eta^2=0.006$
				aperture × kd × cocaine	$F_{1,18}=4.46$	$p=0.049$	$\eta^2=0.198$
			post-hoc: paired t-test	re. *vs.* non-re. (ctrl+sal)	$t_4=3.11$	$p=0.036$	$d=1.391$
				re. *vs.* non-re. (kd+sal)	$t_4=2.47$	$p=0.069$	$d=1.102$
				re. *vs.* non-re. (ctrl+coc)	$t_4=0.76$	$p=0.489$	$d=0.341$

				re. *vs.* non-re. (kd+coc)	t_6=3.07	p=0.022	d=1.160
3h right	response ratio		ANOVA	knockdown	$F_{1,18}$=2.08	p=0.167	η^2=0.103
				cocaine	$F_{1,18}$=3.08	p=0.096	η^2=0.146
				knockdown × cocaine	$F_{1,18}$=7.89	p=0.012	η^2=0.305

Table 3.S4 Complete statistics for experiments reported in Figure 3.3.

Figure	Measure	Groups	Analysis	Comparison	Statistic	p-value	Effect size
1f	response rate	control *(n=9)* stimulation *(n=10)*	RM ANOVA	session	$F_{5,90}=15.3$	$p<0.001$	$\eta^2=0.473$
				group	$F_{1,18}=0.48$	$p=0.498$	$\eta^2=0.027$
				session × group	$F_{5,90}=0.66$	$p=0.653$	$\eta^2=0.037$
1g left			RM ANOVA	aperture	$F_{1,17}=3.50$	$p=0.079$	$\eta^2=0.171$
				group	$F_{1,17}=0.49$	$p=0.494$	$\eta^2=0.028$
				aperture × group	$F_{1,17}=6.21$	$p=0.023$	$\eta^2=0.268$
			post-hoc: paired t-test	reinf. *vs.* non-reinf. (ctrl)	$t_8=2.38$	$p=0.045$	$d=0.793$
				reinf. *vs.* non-reinf. (stim)	$t_9=-0.66$	$p=0.526$	$d=-0.209$
1g right	response ratio		unpaired t-test	control *vs.* stimulation	$t_{17}=1.84$	$p=0.083$	$d=0.845$
1k	response rate	control *(n=9)* CeA Gi *(n=7)* CeA Gi+KD *(n=16)*	RM ANOVA	session	$F_{21,609}=21.2$	$p<0.001$	$\eta^2=0.422$
				group	$F_{2,29}=0.12$	$p=0.886$	$\eta^2=0.008$
				session × group	$F_{42,609}=2.16$	$p<0.001$	$\eta^2=0.130$
			post-hoc: one-way ANOVA	session 2: group	$F_{2,31}=7.83$	$p=0.002$	$\eta^2=0.351$
				session 3: group	$F_{2,31}=13.4$	$p<0.001$	$\eta^2=0.481$
				session 5: group	$F_{2,31}=4.37$	$p=0.022$	$\eta^2=0.232$
				session 11: group	$F_{2,31}=3.75$	$p=0.036$	$\eta^2=0.205$
				session 21: group	$F_{2,31}=3.47$	$p=0.044$	$\eta^2=0.193$
				session 22: group	$F_{2,31}=3.88$	$p=0.032$	$\eta^2=0.211$
1l left			RM ANOVA	aperture	$F_{1,29}=4.87$	$p=0.035$	$\eta^2=0.144$
				group	$F_{2,29}=5.14$	$p=0.012$	$\eta^2=0.262$
				aperture × group	$F_{2,29}=13.7$	$p<0.001$	$\eta^2=0.486$
			post-hoc: paired t-test	reinf. *vs.* non-reinf. (ctrl)	$t_8=0.39$	$p=0.704$	$d=0.131$
				reinf. *vs.* non-reinf. (Gi)	$t_6=-1.48$	$p=0.189$	$d=-0.560$
				reinf. *vs.* non-reinf. (Gi+KD)	$t_{15}=5.84$	$p<0.001$	$d=1.460$
1l right	response ratio		ANOVA	group	$F_{2,31}=10.3$	$p<0.001$	$\eta^2=0.416$
			post-hoc: Tukey HSD	control *vs.* CeA Gi	$\Delta=\pm0.300$	$p=0.705$	n/a
				control *vs.* CeA Gi + KD	$\Delta=\pm1.045$	$p=0.006$	
				CeA Gi *vs.* CeA Gi + KD	$\Delta=\pm1.345$	$p=0.001$	

Table 3.S5 Complete statistics for experiments reported in Figure 3.4.

Figure	Measure	Groups	Analysis	Comparison	Statistic	p-value	Effect size
6c	response rate	control *(n=8)* Gq *(n=10)*	RM ANOVA	session	$F_{7,112}=26.4$	$p<0.001$	$\eta^2=0.622$
				group	$F_{1,16}=0.74$	$p=0.403$	$\eta^2=0.044$
				session × group	$F_{7,112}=0.56$	$p=0.785$	$\eta^2=0.034$
6d left			RM ANOVA	aperture	$F_{1,16}=18.1$	$p<0.001$	$\eta^2=0.531$
				group	$F_{1,16}=0.002$	$p=0.963$	$\eta^2=0.000$
				aperture × group	$F_{1,16}=5.46$	$p=0.033$	$\eta^2=0.254$
			paired t-test	reinf. *vs.* non-reinf. (ctrl)	$t_7=3.74$	$p=0.007$	$d=1.322$
				reinf. *vs.* non-reinf. (Gq)	$t_9=1.73$	$p=0.118$	$d=0.547$
6d right	response ratio		unpaired t-test	control *vs.* Gq	$t_{16}=2.60$	$p=0.019$	$d=1.231$
6g	response rate	control *(n=9)* knockdown *(n=9)*	RM ANOVA	session	$F_{26,416}=17.7$	$p<0.001$	$\eta^2=0.525$
				group	$F_{1,16}=0.99$	$p=0.334$	$\eta^2=0.058$
				session × group	$F_{26,416}=0.44$	$p=0.993$	$\eta^2=0.027$
6h left			RM ANOVA	aperture	$F_{1,16}=7.95$	$p=0.012$	$\eta^2=0.332$
				group	$F_{1,16}=0.005$	$p=0.946$	$\eta^2=0.000$
				aperture × group	$F_{1,16}=0.001$	$p=0.980$	$\eta^2=0.000$
6h right	response ratio		unpaired t-test	control *vs.* knockdown	$t_{16}=-0.46$	$p=0.653$	$d=-0.216$
6i left	response rate		RM ANOVA	aperture	$F_{1,16}=6.71$	$p=0.020$	$\eta^2=0.296$
				group	$F_{1,16}=0.71$	$p=0.413$	$\eta^2=0.042$
				aperture × group	$F_{1,16}=4.30$	$p=0.055$	$\eta^2=0.212$
			paired t-test	reinf. *vs.* non-reinf. (ctrl)	$t_8=4.13$	$p=0.003$	$d=1.378$
				reinf. *vs.* non-reinf. (KD)	$t_8=0.31$	$p=0.762$	$d=0.105$
6i right	response ratio		unpaired t-test	control *vs.* knockdown	$t_{16}=2.78$	$p=0.014$	$d=1.308$
6j left	response rate		RM ANOVA	aperture	$F_{1.16}=4.09$	$p=0.060$	$\eta^2=0.203$
				group	$F_{1,16}=0.20$	$p=0.663$	$\eta^2=0.012$
				aperture × group	$F_{1,16}=1.44$	$p=0.247$	$\eta^2=0.083$
6j right	response ratio		unpaired t-test	control *vs.* knockdown	$t_{16}=1.19$	$p=0.252$	$d=0.561$
6m	response rate	control *(n=14)* knockdown *(n=8)*	RM ANOVA	session	$F_{15,300}=26.5$	$p<0.001$	$\eta^2=0.570$
				group	$F_{1,20}=0.51$	$p=0.484$	$\eta^2=0.025$
				session × group	$F_{15,300}=0.82$	$p=0.655$	$\eta^2=0.039$
6n			RM ANOVA	aperture	$F_{1,20}=18.3$	$p<0.001$	$\eta^2=0.478$
				group	$F_{1,20}=0.10$	$p=0.752$	$\eta^2=0.005$

				aperture × group	$F_{1,20}$=0.58	p=0.454	η^2=0.028
6o			RM ANOVA	aperture	$F_{1,20}$=46.0	p<0.001	η^2=0.725
				group	$F_{1,20}$=0.02	p=0.895	η^2=0.001
				aperture × group	$F_{1,20}$=0.83	p=0.372	η^2=0.040

Table 3.S6 Complete statistics for experiments reported in Figure 3.6.

Chapter 4

Concluding remarks

4.1 ABSTRACT

The findings described in previous chapters provide molecular- and circuit-level insights into how MC4R controls flexible decision-making strategies. Here, I summarize those findings and discuss 1) how considerations of etiological motivations provide context for decision-making mechanisms, and 2) potential future directions.

4.2 ETILOGICAL MOTIVATIONS OF DECISION-MAKING BEHAVIOR

In Chapter 1, I describe the functions of MC4R in extra-hypothalamic brain regions that are involved in decision-making behavior, highlighting the need for a greater understanding of MC4R function across the brain. In Chapter 2, I report that MC4R presence in the DMS biases animals towards inflexible, habitual routines. In Chapter 3, I show that striatal melanocortin systems propel familiar actions via interactions with the central nucleus of the amygdala. In sum, this dissertation describes how MC4R is molecularly and anatomically integrated into striatal-dependent mechanisms of behavioral flexibility.

Motivated behaviors have fascinated neuroscientists and ethologists for decades due to their necessity for survival of the organism. Motivations guide behavioral choice through an intricate synthesis of internal state detection, external stimulus exposure, and learned associations. However, neuroscientists commonly focus research efforts on neural circuits *underlying* individual motivations, sacrificing ethological relevance for tight experimental control. My own work is an example, focusing on two different learning processes in decision-making behavior: one encoding the relationship between actions and their consequences and the second involved in the formation of stimulus-response associations (Balleine & O'Doherty, 2010). But these flexible or inflexible action strategies are not derived in a vacuum. It is thus necessary to consider how the

neurobiological underpinnings of ethological motivations are integrated into decision-making processes, and how integration of competing motivations may influence action selection.

In 1943, Abraham Maslow first describe his theory of hierarchical needs, a pyramid of five motivational categories organized from most to least necessity for survival: physiological, safety, social, esteem, and self-actualization (Maslow, 1943). Meeting each need is not an all-or-none phenomenon, and one behavior may address two or more needs (Maslow, 1971). Maslow's hierarchy has been used by behavioral neuroscientists for more than 80 years as a framework for understanding the motivations of the organism. In mice, particular focus is given to the first three levels of the hierarchy: physiological, safety, and social needs (Sutton & Krashes, 2020). Here, I will walk through an example motivation from each of these three levels – hunger, stress, and reproduction – and describe how the related neurobiological mechanisms are integrated into and influence decision-making behavior.

4.3 PHYSIOLOGICAL MOTIVATION: HUNGER

Hunger is one of the most salient physiological needs, the foundational category of Maslow's hierarchy. Hunger, along with other physiological needs such as sleep, water intake, and thermoregulation, are critical for survival: failure to meet any one of these individual needs leads to a rapid breakdown in molecular, cellular, and whole-body functions can only be restored by capitulating to the need. As such, the peripheral and central mechanisms regulating hunger and our subsequent behavioral responses are numerous and well-studied.

4.3.1 The cycle of hunger

The term “hunger” describes the hormonal and neural cues informing the body and brain that calories are imminently required to maintain homeostasis. Physically, hunger is reported to be a desire to eat and an uncomfortable emptiness in the abdomen. One of the first peripheral hunger cues is increased gastrointestinal motility: a pattern of motility and secretion termed the migrating motor complex (MMC) empties the gastrointestinal system to remove the sense of fullness caused by gastric distention (Sanger et al., 2011). At the same time, a dearth in dietary nutrients disinhibits the release of ghrelin from the mucosal lining of the stomach and small intestine (Mace et al., 2015). As plasma levels of ghrelin increase, levels of leptin released from fatty tissue decrease (Chin-Chance et al., 2000), resulting in a characteristic high-ghrelin, low-leptin hunger state.

Leptin and ghrelin act on discrete subsets of agouti-related protein (AgRP) neurons and proopiomelanocortin (POMC) neurons in the arcuate nucleus of the hypothalamus, the primary hub for the central regulation of hunger (Andermann & Lowell, 2017). AgRP neurons, which are activated by ghrelin and inhibited by leptin, produce the peptide AgRP, which antagonizes melanocortin-4 receptors (MC4R) in the paraventricular nucleus of the hypothalamus (PVH). POMC neurons, on the other hand, which are inhibited by ghrelin and activated by leptin, produce α-melanocyte stimulating hormone (α-MSH), a MC4R agonist. In the high-ghrelin, low-leptin hunger state, POMC neurons are inhibited and neuropeptide release from AgRP neurons dominate (Heisler & Lam, 2017). This cascade decreases MC4R activity in the PVH, resulting in a heightened drive to seek food and lower rates of metabolism. This signaling cascade beginning with ghrelin and leptin in the gastrointestinal track and ending with MC4R in the hypothalamus will continue with heightened urgency until the organism begins to consume food. As nutritional content fills the stomach, serotonin (5HT), cholecystokinin (CCK), and peptide tyrosine tyrosine (PYY) are released peripherally to inhibit hypothalamic AgRP neurons and activate neurons

expressing POMC (Marciani et al., 2015). POMC neurons release α-MSH, which activates MC4R to reduce food seeking and consumption, increase metabolism, and increase satiety. The organism will remain in this sated state until stores of nutrients run low, wherein the cycle of hunger will begin again.

4.3.2 Hunger and behavioral flexibility

Hunger is a particularly salient motivator and can override other bodily demands to initiate a series of stereotypic behaviors: food seeking, detection, and consumption (Sutton & Krashes, 2020). This transition from everyday behavior to goal-directed action necessitates behavioral flexibility: new homeostatic information prompts organisms to update their behavior. This phenomenon, termed optimal foraging theory, drives organisms to abandon familiar regions and forage in new areas (Stephens, 2018). It follows that the neuropeptidergic cues that cycle hunger signaling also innervate the corticostriatal and amygdalo-striatal systems that direct flexible decision-making behavior.

Ghrelin is not only the first sign of hunger, but it also biases organisms towards flexible, goal-directed food seeking. Ghrelin binding in the CA1 region of the hippocampus promotes dendritic spine synapse formation and generation of long-term potentiation, and enhances spatial learning and memory in rodents (Diano et al., 2006). Ghrelin delivery to the ventral hippocampus also increases willingness to work for sucrose and increases spontaneous meal initiation in non-deprived rats following presentation of a food stimulus (Kanoski et al., 2013). Furthermore, infusing ghrelin into the medial prefrontal cortex (mPFC) increases food seeking and consumption (Parent et al., 2015). The mPFC is necessary for determining the value of a given reward; here, it

seems likely that ghrelin acting in the mPFC increases the rewarding value of food, such that an organism is more likely to seek food at the cost of other resources.

As the organism eats and hunger pains decline, PYY is released from the gastrointestinal tract to inhibit AgRP neurons, stimulate POMC neurons, and increase satiety (Adewale et al., 2007). During increased periods of PYY circulation, the strength of blood-oxygen-level-dependent (BOLD) response in the orbitofrontal cortex (OFC), a region necessary for flexible behavior, is negatively correlated with meal pleasantness (Batterham et al., 2007; De Silva et al., 2011). This finding suggests that PYY increases cell activity in the OFC to decrease the reward value of food, thus opposing the function of ghrelin in the mPFC and decreasing feeding when it is no longer adaptive. Nutritional intake also leads to heightened levels of leptin, which acts in the Arc to increase production of the MC4R agonist α-MSH. How the satiety-dependent melanocortin system biases animals back towards their routines following food consumption is described in the prior chapter and summarized in Figure 4.1. In sum, neuropeptides that function as hunger and satiety cues such as ghrelin and PYY are fundamentally integrated into the circuits that govern behavioral flexibility.

4.4 SECURITY MOTIVATION: STRESS

The second tier of motivation in Maslow's hierarchy of needs is concerned with security. The world is full of dangers that require constant vigilance to avoid: pain, sickness, aggression from conspecifics, and threats of predation from other species. In any of these cases, the organism runs the risk of irreversible bodily harm, and thus is extremely motivated to avoid these safety concerns.

4.4.1 The hypothalamic-pituitary-adrenal axis

One of the primary systems by which the organism maintains vigilance of its surroundings is the hypothalamic-pituitary-adrenal (HPA) axis. Following an alarming stimulus, corticotropin-releasing hormone (CRH) and vasopressin (AVP) are secreted from the paraventricular nucleus of the hypothalamus. Both CRH and AVP act on corticotrope cells in the anterior lobe of the pituitary gland to upregulate transcription of *Pomc*, which, in addition to producing the MC4R agonist α-MSH, is also derived into adrenocorticotropic hormone (ACTH). ACTH is transported in the blood to the adrenal cortex of the adrenal gland where it acts on the melanocortin-2 receptor (MC2R) to rapidly stimulate the biosynthesis of corticosteroids from cholesterol (cortisol in humans, corticosterone in rodents). Corticosteroids simultaneously alert target organs and immune systems to the potential threat and provide a negative feedback loop onto the HPA axis to dampen further signaling. This cascade, termed the "stress response", then resets to baseline levels to await the next alarming stimulus.

4.4.2 Stress and behavioral flexibility

The stress response is rapid and reliable; as such, organisms exposed to a stressor are often inferred to have heightened corticosteroid levels. Acute instances of stress and subsequent bursts in corticosteroid levels can in some instances enable organisms to flexibly make in-the-moment updates to behavior (Sapolsky, 2021). Following acute stress, humans with high corticosteroid levels are more likely to forgo immediate rewards in favor of larger, delayed rewards, particularly when learning from experience (Byrne et al., 2019; Cooper et al., 2013). In risky situations, acute stress biases participants towards "high-risk, high reward" decisions, even if they result in less optimal performance on the task compared to controls (Pabst et al., 2013; Starcke et al., 2008;

Starcke & Brand, 2016). “Risky” behavior is often flexible in nature: instead of relying on more cautious routines, individuals wildly change their behavior to gain future safety in an increasingly stressful environment, even if it means gambling with immediate security. This increase in flexible behavior following an acute stressor is concurrent with increased glutamatergic and dopaminergic signaling in the PFC and striatum (Arnsten, 2009; Popoli et al., 2011), two regions canonically associated with response strategy switching. Corticostriatal glutamate and dopamine release, which is partially mediated by corticosteroids (Holloway et al., 2023; Musazzi et al., 2015), is thought to increase reward salience and value and, in turn, enhance learning from positive outcomes (Mather & Lighthall, 2012). Thus, stress hormones like corticosteroids help the brain store learned information for future use when behaviors must be flexibly modified to meet an organism’s needs.

The neurobiological response to chronic stress is fundamentally different from acute stress. While acute stress can promote certain forms of flexible behavior, chronic stress exposure impedes goal-directed action. Extended periods of both stress and exogenous corticosteroid exposure impair one’s ability to select actions based on their outcomes (Dias-Ferreira et al., 2009; Gourley et al., 2012). This behavioral impairment is likely due, at least in part, to stress-dependent modifications of neuronal dendritic architecture and neurotransmission in regions necessary for goal-sensitive decision-making behavior, such as the OFC. For instance, excitatory neurons in the OFC are particularly susceptible to chronic stress during the adolescent period (reviewed in Sequeira & Gourley, 2021). Atypical dendritic spine plasticity in the OFC could hinder excitatory corticostriatal signaling necessary for flexible decision making. Indeed, mice exposed to excess corticosterone or that are socially isolated during adolescence are consistently impaired in the ability to select actions based on reward likelihood (Barfield et al., 2017; Hinton et al., 2019; Li et

al., 2021). Thus, the relationship between stress and behavioral flexibility can be conceptualized as an inverted parabola. Too little stress, and organisms are not motivated to update their behaviors when routine choices may endanger the security. Too much stress, and, over time, organisms also lose the ability to flexibly act in stressful situations. This balance between stress exposure and behavioral flexibility is thus both delicate in nature and vital for continued survival.

4.5 SOCIAL MOTIVATION: PAIR BONDING

Physiological needs increase an organism's rate of survival by supporting homeostasis; security needs increase survival rates by defending the organism from external threat. Social motivations, on the other hand, increase an organism's chance at survival via species propagation. Mating leads to offspring, which in many cases necessitates rearing those offspring until they are ready to reproduce on their own. In some species, including humans, producing offspring is a method of ensuring one's survival and comfort in later periods of life. In all of these cases, the act of choosing a mate determines the viability of those offspring, any potential co-parenting, and thus impacts the survival chances of the parent organism. How systems that drive reproduction are integrated into decision-making systems is thus a topic of great interest.

4.5.1 The neurobiology of pair bonding

The phenomenon of pair bonding is inextricably linked to the neuropeptide oxytocin (OT). The OT system is classically studied for its role in maternal behaviors, such as parturition, uterine contractions, and suckling (Gimpl & Fahrenholz, 2001). Pair bonding likely evolved from maternal attachment, creating a lasting social connection between two conspecifics rather than a mother and child. It follows that OT also promotes pair bonding. Monogamous prairie voles that form pair

bonds with sexual partners have significantly greater OT receptor density in the mPFC and dorsal and ventral striatum, compared to polygamous meadow voles (Insel & Shapiro, 1992). OT receptor activation in the nucleus accumbens (NAc) and mPFC, but not the dorsal striatum, is necessary for the formation of pair bonds in female monogamous prairie voles (Young et al., 2001).

Pair bonding also relies on dopaminergic signaling in the striatum (Burkett & Young, 2012). Mating, which promotes the formation of pair bonds, increases dopamine release in the dorsal and ventral striatum, including the NAc (Damsma et al., 1992). Pair bonding in prairie voles is prevented if both D1- and D2-type receptors are blocked in the NAc shell, and non-specific activation of dopamine receptors in the NAc shell is sufficient to induce pair bonding (Aragona et al., 2003).

4.5.2 Pair bonding and behavioral flexibility

How does pair bonding intersect with behavioral flexibility? Choosing a partner takes effort and comes with no small amount of risk. An organism must flexibly adapt its behavior in response to interactions with conspecifics and environmental information – even traveling to a new location when partner prospects are looking thin. Once a partner is chosen, much of the stress and risk of seeking a mate is done away with. Instead, an organism routinely returns to its familiar partner, freeing cognitive resources for the tasks of parenting offspring, defending the nest from threats, and gathering physiological necessities like food and water. OT, which is imperative for forming monogamous pair bonds in mammals, is critical for the processing of stimuli, particularly sensory information extracted from social contexts (Coccia et al., 2022). Initially, intimate behaviors between new reproductive partners can be considered *goal*-directed: each conspecific wants to please their partner to retain relationship benefits (safety, resources, offspring). OT is

released centrally and peripherally during such periods of intimacy between partners. Over time, the association between partner stimulus (exaggerated by OT) and relationship-typical behavior would be expected to grow stronger. Eventually, partners would be expected to participate in their relationship out of habit – the presence of a familiar stimulus elicits a familiar action – rather than a drive to gain resources.

This monogamy-as-habit framework is supported neurobiologically. OT promotes pair bonding via binding in the NAc shell, a region that supports habitual behavior (Di Chiara, 2002). Selective pharmacological activation of NAc D1R-MSNs, which promote flexible decision making (Peak et al., 2020), but not D2R-MSNs, prevents the formation of a pair bond (Aragona et al., 2006; Aragona & Wang, 2009). A proportion of central OT is produced in MC4R-expressing cells in the PVH. Here, MC4R activation via α-MSH increases OT production (Barrett et al., 2014; Modi et al., 2015). It is conceivable that during periods of growing familiarity between two partners, changes in homeostatic state induce the release of α-MSH, which activates MC4R in the DMS and separately in the PVH to increase OT release to the NAc shell, thus playing a dual role in promoting the formation of habitual behavior.

4.6 CONCLUSIONS

Here I use the term "behavioral flexibility" to describe the ability to integrate new learning into familiar routines. This conceit is deceptively simple, given the multitude of physical, mental, and environmental factors that contribute to a single "flexible" decision. In laboratory settings, one is able to boil a concept like behavioral flexibility down to its simplest pieces. Indeed, my dissertation work has focused almost exclusively on how mice make decisions in response to changes in reward likelihood. In reality, the decision-making process is confounded by an

innumerable complexity of features. These other features – reward value, hunger level, context cues – may also be modulated equally or to a greater extent by melanocortin signaling in the striatum. Given the many different complex processes that are likely involved in the consideration of everyday decisions, it is perhaps unsurprising that deficits in decision making are a central feature of nearly every neuropsychiatric disorder.

To conclude, these findings provide insight into the molecular and circuit-level mechanisms by which MC4R in the DMS propels habitual behavior. This dissertation thus illuminates mechanistic factors that support the development of automatized routines when flexible decision making is no longer adaptive, which may provide insight into therapeutic targets for neuropsychiatric disorders in which decision making is impaired.

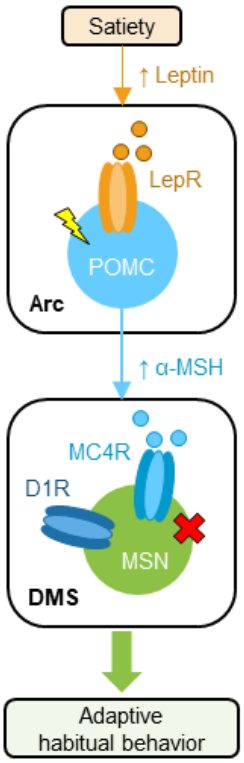

Figure 4.1 Satiety-dependent melanocortin systems drive a transition to adaptive habitual behavior following food consumption. Animals that are hungry often display flexible behavior (Stephens, 2018). As such, the control of decision-making behavior by melanocortin systems may be framed, at least as a starting point, within the context of hunger. When food is obtained and satiety achieved, leptin is released from peripheral adipose deposits and activates leptin receptors (LepR) on pro-opiomelanocortin (POMC)+ cells in the arcuate nucleus of the hypothalamus (Arc). POMC+ cells produce the melanocortin-4 receptor (MC4R) ligand α-melanocyte stimulating hormone (α-MSH). α-MSH activation of MC4R on dopamine D1-type receptor (D1R)+ neurons in the dorsomedial striatum (DMS) would be expected to reduce the excitability of these neurons, potentially allowing organisms to abandon flexible food seeking strategies that are no longer needed, in favor of adaptive habitual behavior (that is, pursuing familiar routines that have been reliably reinforced in the past).

APPENDIX A: PUBLICATIONS TO WHICH THE AUTHOR HAS CONTRIBUTED

Compton RJ, **Heaton EC**, & Ozer E. (2017). Inter-trial interval duration affects error monitoring. *Psychophysiology*, 54(8), 1151-1162.

Compton RJ, **Heaton EC**, & Gaines A. (2018). Is attention enhanced following performance errors? Testing the adaptive control hypothesis. *Psychophysiology*, 55(4), e13022.

Compton RJ, Gearinger D, Wild H, Rette D, **Heaton EC**, Histon S, Thiel P, & Jaskir M. (2020). Simultaneous EEG and pupillary evidence for post-error arousal during a speeded performance task. *European Journal of Neuroscience*, 53(2), 543-555.

Hedges VL, **Heaton EC**, Amaral C, Benedetto LE, Bodie CL, D'Antonio BI, Davila Portillo DR, Lee RH, Levine MT, O'Sullivan EC, Pisch NP, Taveras S, Wild HR, Grieb ZA, Ross AP, Albers HE, & Been LE. (2021). Estrogen withdrawal increases postpartum anxiety via oxytocin plasticity in the paraventricular hypothalamus and dorsal raphe nucleus. *Biological Psychiatry*, 89(9), 929-938.

Wong JC, Shapiro L, Thelin JT, **Heaton EC**, Zaman RU, D'Souza MJ, Murnane KS, & Escayg A. (2021). Nanoparticle-encapsulated oxytocin increases resistance to induced seizures and restores social behavior in *Scn1a*-derived epilepsy. *Neurobiology of Disease*, 147, 105147.

Allen AG, ***Heaton EC**, *Shapiro LP, Butkovich LM, Davies RA, Yount ST, Li DC, Swanson AM, & Gourley SL. (2022). Inter-individual variability amplified through breeding reveals control of reward-related action strategies by Melanocortin-4 Receptor in the dorsomedial striatum. *Communications Biology*, 5(1), 1-12.

Michopoulos V, Turkson S, Dyer S, Howell P, **Heaton EC**, Hart J, Powers A, Mekawi Y, Carter S, Ofotokun I, Jovanovic T, & Neigh G. A history of childhood trauma influences PTSD symptom severity in Black women living with HIV. Submitted.

Heaton EC, Seo E, Yount S, Butkovich LM, & Gourley SL. Striatal melanocortin-4 receptor controls behavioral flexibility in mice. Submitted.

REFERENCES

Adams CD & Dickinson A. (1981). Instrumental responding following reinforcer devaluation. *Experimental Psychology*, 33(2b), 109–121.

Adewale AS, Macarthur H, & Westfall TC. (2007). Neuropeptide Y-induced enhancement of the evoked release of newly synthesized dopamine in rat striatum: Mediation by Y2 receptors. *Neuropharmacology*, 52(6), 1396–1402.

Allen AT, Heaton, EC, Shapiro LP, Butkovich, LM, Yount ST, Davies RA, Li, D. C., Swanson AM, & Gourley SL. (2022). Inter-individual variability amplified through breeding reveals control of reward-related action strategies by melanocortin-4 receptor in the dorsomedial striatum. *Communications Biology*, 5(1), 116.

Allen Institute for Brain Science. (2004). *Allen Mouse Brain Atlas*. https://mouse.brain-map.org/

Alvaro JD, Tatro JB, Quillan JM, Fogliano M, Eisenhard M, Lerner MR, Nestler EJ, & Duman RS. (1996). Morphine down-regulates melanocortin-4 receptor expression in brain regions that mediate opiate addiction. *Molecular Pharmacology*, 50(3), 583–591.

Alvaro JD, Taylor JR, & Duman RS. (2003). Molecular and behavioral interactions between central melanocortins and cocaine. *The Journal of Pharmacology and Experimental Therapeutics*, 304(1), 391–399.

Andermann ML, & Lowell BB. (2017). Toward a wiring diagram understanding of appetite control. *Neuron*, 95(4), 757–778.

Anderson EJP, Çakir I, Carrington SJ, Cone RD, Ghamari-Langroudi M, Gillyard T, Gimenez LE, & Litt MJ. (2016). 60 YEARS OF POMC: Regulation of feeding and energy homeostasis by α-MSH. *Journal of Molecular Endocrinology*, 56(4), T157-174.

Aragona BJ, Liu Y, Curtis JT, Stephan FK, & Wang Z. (2003). A critical role for nucleus accumbens dopamine in partner-preference formation in male prairie voles. *Journal of Neuroscience*, 23(8), 3483–3490.

Aragona BJ, Liu Y, Yu YJ, Curtis JT, Detwiler JM, Insel TR, & Wang Z. (2006). Nucleus accumbens dopamine differentially mediates the formation and maintenance of monogamous pair bonds. *Nature Neuroscience*, 9(1), 133–139.

Aragona BJ, & Wang Z. (2009). Dopamine regulation of social choice in a monogamous rodent species. *Frontiers in Behavioral Neuroscience*, 3, 15.

Argiolas A, Melis MR, Murgia S, & Schiöth HB. (2000). ACTH- and alpha-MSH-induced grooming, stretching, yawning and penile erection in male rats: Site of action in the brain and role of melanocortin receptors. *Brain Research Bulletin*, 51(5), 425–431.

Arnsten AFT. (2009). Stress signaling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, 10(6).

Avegno EM, Lobell TD, Itoga CA, Baynes BB, Whitaker AM, Weera MM, Edwards S, Middleton JW & Gilpin NW. (2018). Central amygdala circuits mediate hyperalgesia in alcohol-dependent rats. *Journal of Neuroscience*, 38(36), 7761–7773.

Balleine BW & O'Doherty JP. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35(1), 48–69.

Balthasar N, Dalgaard LT, Lee CE, Yu J, Funahashi H, Williams T, Ferreira M, Tang V, McGovern RA, Kenny CD, Christiansen LM, Edelstein E, Choi B, Boss O, Aschkenasi C, Zhang C, Mountjoy K, Kishi T, Elmquist JK & Lowell BB. (2005). Divergence of melanocortin pathways in the control of food intake and energy expenditure. *Cell*, 123(3), 493–505.

Barfield ET, Gerber KJ, Zimmermann KS, Ressler KL, Parsons RG, & Gourley SL. (2017). Regulation of actions and habits by ventral hippocampal trkB and adolescent corticosteroid exposure. *PLoS Biology*, 15(11), e2003000.

Barrett CE, Modi ME, Zhang BC, Walum H, Inuoe K, & Young LJ. (2014). Neonatal melanocortin receptor agonist treatment reduces play fighting and promotes adult attachment in prairie voles in a sex-dependent manner. *Neuropharmacology*, 85, 357–366.

Batterham RL, Ffytche DH, Rosenthal JM, Zelaya FO, Barker GJ, Withers DJ, & Williams SCR. (2007). PYY modulation of cortical and hypothalamic brain areas predicts feeding behavior in humans. *Nature*, 450(7166).

Béïque JC & Andrade R. (2003). PSD-95 regulates synaptic transmission and plasticity in rat cerebral cortex. *The Journal of Physiology*, 546(3), 859–867.

Beleford DT, Van Ziffle J, Hodoglugil U, & Slavotinek AM. (2020). A missense variant, p.(Ile269Asn), in MC4R as a secondary finding in a child with BCL11A-related intellectual disability. *European Journal of Medical Genetics*, 103969.

Bercu BB & Brinkley HJ. (1967). Hypothalamic and cerebral cortical inhibition of melanocyte-stimulating hormone secretion in the frog, *Rana pipiens*. *Endocrinology*, 80(3), 399–403.

Berry KP & Nedivi E. (2017). Spine dynamics: Are they all the same? *Neuron*, 96(1), 43–55.

Bertagna X. (1994). Proopiomelanocortin-derived peptides. *Endocrinology and Metabolism Clinics of North America*, 23(3), 467–485.

Bertolini A & Gessa GL. (1981). Behavioral effects of ACTH and MSH peptides. *Journal of Endocrinological Investigation*, 4(2), 241–251.

Boghossian S, Lemmon K, Park M, & York DA. (2009). High-fat diets induce a rapid loss of the insulin anorectic response in the amygdala. *American Journal of Physiology*, 297(5), R1302-1311.

Boghossian S, Park M, & York DA. (2010). Melanocortin activity in the amygdala controls appetite for dietary fat. *American Journal of Physiology*, 298(2), R385-393.

Bradfield LA, Bertran-Gonzalez J, Chieng B, & Balleine BW. (2013). The thalamostriatal pathway and cholinergic control of goal-directed action: Interlacing new with existing learning in the striatum. *Neuron*, 79(1), 153–166.

Braun S, & Hauber W. (2012). Striatal dopamine depletion in rats produces variable effects on contingency detection: Task-related influences. *European Journal of Neuroscience*, 35(3), 486–495.

Bruschetta G, Jin S, Liu ZW, Kim JD, & Diano S. (2020). MC4R signaling in dorsal raphe nucleus controls feeding, anxiety, and depression. *Cell Reports*, 33(2).

Burkett JP & Young LJ. (2012). The behavioral, anatomical and pharmacological parallels between social attachment, love and addiction. *Psychopharmacology*, 224(1), 1–26.

Byrne KS, Cornwall AC, & Worthy DA. (2019). Acute stress improves long-term reward maximization in decision-making under uncertainty. *Brain and Cognition*, 133, 84–93.

Caruso V, Lagerström MC, Olszewski PK, Fredriksson R, & Schiöth HB. (2014). Synaptic changes induced by melanocortin signaling. *Nature Reviews Neuroscience*, 15(2), 98–110.

Chin-Chance C, Polonsky KS, & Schoeller DA. (2000). Twenty-four-hour leptin levels respond to cumulative short-term energy imbalance and predict subsequent intake. *Journal of Clinical Endocrinology and Metabolism*, 85(8), 2685–2691.

Coccia G, La Greca F, Di Luca M, & Scheggia D. (2022). Dissecting social decision-making: A spotlight on oxytocinergic transmission. *Frontiers in Molecular Neuroscience*, 15.

Cooper JA, Worthy DA, Gorlick MA, & Maddox WT. (2013). Scaffolding across the lifespan in history-dependent decision-making. *Psychology and Aging*, 28(2), 505–514.

Cruz KG, Leow YN, Le NM, Adam E, Huda R, & Sur M. (2023). Cortical-subcortical interactions in goal-directed behavior. *Physiological Reviews*, 103(1), 347–389.

Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM, & Costa RM. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*, 494(7436).

Cui H & Lutter M. (2013). The expression of MC4Rs in D1R neurons regulates food intake and locomotor sensitization to cocaine. *Genes, Brain, and Behavior*, 12(6), 658–665.

Cui H, Mason BL, Lee C, Nishi A, Elmquist JK, & Lutter M. (2012). Melanocortin 4 receptor signaling in dopamine 1 receptor neurons is required for procedural memory learning. *Physiology & Behavior*, 106(2), 201–210.

Damsma G, Pfaus JG, Wenkstern D, Phillips AG, & Fibiger HC. (1992). Sexual behavior increases dopamine transmission in the nucleus accumbens and striatum of male rats: Comparison with novelty and locomotion. *Behavioral Neuroscience*, 106(1), 181–191.

Daniels D, Patten CS, Roth JD, Yee DK, & Fluharty SJ. (2003). Melanocortin receptor signaling through mitogen-activated protein kinase in vitro and in rat hypothalamus. *Brain Research*, 986(1), 1–11.

Davis JF, Choi DL, Shurdak JD, Krause EG, Fitzgerald MF, Lipton JW, Sakai RR, & Benoit SC. (2011). Central melanocortins modulate mesocorticolimbic activity and food seeking behavior in the rat. *Physiology & Behavior*, 102(5), 491–495.

De Silva A, Salem V, Long CJ, Makwana A, Newbould RD, Rabiner EA, Ghatei MA, Bloom SR, Matthews PM, Beaver JD, & Dhillo WS. (2011). The gut hormones PYY3-36 and GLP-17-36 amide reduce food intake and modulate brain activity in appetite centers in humans. *Cell Metabolism*, 14(5), 700–706.

de Torrenté L, Zimmerman S, Suzuki M, Christopeit M, Greally JM, & Mar JC. (2020). The shape of gene expression distributions matter: How incorporating distribution shape improves the interpretation of cancer transcriptomic data. *BMC Bioinformatics*, 21(21), 562.

de Wied D. (1966). Inhibitory effect of ACTH and related peptides on extinction of conditioned avoidance behavior in rats. *Experimental Biology and Medicine*, 122(1), 28–32.

de Wit S, Ostlund SB, Balleine BW, & Dickinson A. (2009). Resolution of conflict between goal-directed actions: Outcome encoding and neural control processes. *Journal of Experimental Psychology*, 35, 382–393.

Dhillon S & Keam SJ. (2019). Bremelanotide: First approval. *Drugs*, 79(14), 1599–1606.

Di Chiara G. (2002). Nucleus accumbens shell and co re dopamine: Differential role in behavior and addiction. *Behavioural Brain Research*, 137(1), 75–114.

Diano S, Farr SA, Benoit SC, McNay EC, da Silva I, Horvath B, Gaskin FS, Nonaka N, Jaeger LB, Banks WA, Morley JE, Pinto S, Sherwin RS, Xu L, Yamada KA, Sleeman MW, Tschöp MH, & Horvath TL. (2006). Ghrelin controls hippocampal spine synapse density and memory performance. *Nature Neuroscience*, 9(3).

Dias-Ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Costa RM, & Sousa N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, 325(5940), 621–625.

Ellacott KLJ, Morton GJ, Woods SC, Tso P, & Schwartz MW. (2010). Assessment of feeding behavior in laboratory mice. *Cell Metabolism*, 12(1), 10–17.

Eskay RL, Giraud P, Oliver C, & Brown-Stein MJ. (1979). Distribution of alpha-melanocyte-stimulating hormone in the rat brain: Evidence that alpha-MSH-containing cells in the arcuate region send projections to extrahypothalamic areas. *Brain Research*, 178(1), 55–67.

Feng G, Mellor RH, Bernstein M, Keller-Peck C, Nguyen QT, Wallace M, Nerbonne JM, Lichtman JW, & Sanes JR. (2000). Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron*, 28(1), 41–51.

Fenselau H, Campbell JN, Verstegen AMJ, Madara JC, Xu J, Shah BP, Resch JM, Yang Z, Mandelblat-Cerf Y, Livneh Y, & Lowell BB. (2017). A rapidly-acting glutamatergic ARC→PVH satiety circuit postsynaptically regulated by α-MSH. *Nature Neuroscience*, 20(1), 42–51.

Ferrari W, Gessa GL, & Vargiu L. (1963). Behavioral effects induced by intracisternally injected ACTH and MSH. *Annals of the New York Academy of Sciences*, 104, 330–345.

Fisher SD, Ferguson LA, Bertran-Gonzalez J, & Balleine BW. (2020). Amygdala-cortical control of striatal plasticity drives the acquisition of goal-directed action. *Current Biology*, 30(22).

Flores-Bastías O, Adriasola-Carrasco A, & Karahanian E. (2020). Activation of melanocortin-4 receptor inhibits both neuroinflammation induced by early exposure to ethanol and subsequent voluntary alcohol intake in adulthood in animal models: Is BDNF the key mediator? *Frontiers in Cellular Neuroscience*, 14, 5.

France G, Fernández-Fernández D, Burnell ES, Irvine MW, Monaghan DT, Jane DE, Bortolotto ZA, Collingridge GL, & Volianskis A. (2017). Multiple roles of GluN2B-containing NMDA receptors in synaptic plasticity in juvenile hippocampus. *Neuropharmacology*, 112, 76–83.

Franklin K & Paxinos G. (2001). The mouse brain in stereotaxic coordinates (2nd ed.). *Academic Press*.

Gantz I & Fong TM. (2003). The melanocortin system. *American Journal of Physiology*, 284(3), E468-474.

Gantz I, Miwa H, Konda Y, Shimoto Y, Tashiro T, Watson SJ, DelValle J, & Yamada T. (1993). Molecular cloning, expression, and gene localization of a fourth melanocortin receptor. *Journal of Biological Chemistry*, 268(20), 15174–15179.

Garfield AS, Li C, Madara JC, Shah BP, Webber E, Steger JS, Campbell JN, Gavrilova O, Lee CE, Olson DP, Elmquist JK, Tannous BA, Krashes MJ, & Lowell BB. (2015). A neural basis for melanocortin-4 receptor-regulated appetite. *Nature Neuroscience*, 18(6), 863–871.

Gawliński D, Gawlińska K, Frankowska M, & Filip M. (2020). Maternal high-sugar diet changes offspring vulnerability to reinstatement of cocaine-seeking behavior: Role of melanocortin-4 receptors. *FASEB Journal*, 34(7), 9192-9206.

Gerfen CR. (2022). Segregation of D1 and D2 dopamine receptors in the striatal direct and indirect pathways: An historical perspective. Frontiers in Synaptic Neuroscience, 14, 1002960.

Gessa GL, Pisano M, Vargiu L, Crabai F, & Ferrari W. (1967). Stretching and yawning movements after intracerebral injection of ACTH. *Revue Canadienne De Biologie*, 26(3), 229–236.

Gimpl G & Fahrenholz F. (2001). The oxytocin receptor system: Structure, function, and regulation. *Physiological Reviews*, 81(2), 629–683.

Gispen WH, Wiegant VM, Greven HM, & de Wied D. (1975). The induction of excessive grooming in the rat by intraventricular application of peptides derived from ACTH: Structure-activity studies. *Life Sciences*, 17(4), 645–652.

Giuliani D, Galantucci M, Neri L, Canalini F, Calevro A, Bitto A, Ottani A, Vandini E, Sena P, Sandrini M, Squadrito F, Zaffe D, & Guarini S. (2014). Melanocortins protect against brain damage and counteract cognitive decline in a transgenic mouse model of moderate Alzheimer′s disease. *European Journal of Pharmacology*, 740, 144–150.

Giuliani D, Ottani A, Neri L, Zaffe D, Grieco P, Jochem J, Cavallini GM, Catania A, & Guarini S. (2017). Multiple beneficial effects of melanocortin MC4 receptor agonists in experimental neurodegenerative disorders: Therapeutic perspectives. *Progress in Neurobiology*, 148, 40–56.

Gomez JL, Bonaventura J, Lesniak W, Mathews WB, Sysa-Shah P, Rodriguez LA, Ellis RJ, Richie CT, Harvey BK, Dannals RF, Pomper MG, Bonci A, & Michaelides M. (2017). Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. *Science*, 357(6350), 503–507.

Gonzalez PV, Schiöth HB, Lasaga M, & Scimonelli TN. (2009). Memory impairment induced by IL-1beta is reversed by alpha-MSH through central melanocortin-4 receptors. *Brain, Behavior, and Immunity*, 23(6), 817–822.

Goodfellow JM, Borcar A, Proctor JL, Greco T, Rosenthal RE, & Fiskum G. (2020). Transcriptional activation of antioxidant gene expression by Nrf2 protects against mitochondrial dysfunction and neuronal death associated with acute and chronic neurodegeneration. *Experimental Neurology*, 328, 113247.

Gourley SL, Swanson AM, Jacobs AM, Howell JL, Mo M, Dileone RJ, Koleske AJ, & Taylor JR. (2012). Action control is mediated by prefrontal BDNF and glucocorticoid receptor binding. *PNAS*, 109(50), 20714–20719.

Gourley SL, & Taylor JR. (2016). Going and stopping: Dichotomies in behavioral control by the prefrontal cortex. *Nature Neuroscience*, 19(6), 656–664.

Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K, Lovinger DM, & Costa RM. (2016). Endocannabinoid modulation of orbitostriatal circuits gates habit formation. *Neuron*, 90(6), 1312–1324.

Gremel CM & Costa RM. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, 4, 2264.

Guida P, Michiels M, Redgrave P, Luque D, & Obeso I. (2022). An fMRI meta-analysis of the role of the striatum in everyday-life vs laboratory-developed habits. *Neuroscience and Biobehavioral Reviews*, 141, 104826.

Hart G, Bradfield LA, Fok SY, Chieng B, & Balleine BW. (2018). The bilateral prefronto-striatal pathway is necessary for learning new goal-directed actions. *Current Biology*, 28(14), 2218-2229.e7.

He S & Tao YX. (2014). Defect in MAPK signaling as a cause for monogenic obesity caused by inactivating mutations in the melanocortin-4 receptor gene. *International Journal of Biological Sciences*, 10(10), 1128–1137.

Heisler LK & Lam DD. (2017). An appetite for life: Brain regulation of hunger and satiety. *Current Opinion in Pharmacology*, 37, 100–106.

Hernandez PJ, Schiltz CA, & Kelley AE. (2006). Dynamic shifts in corticostriatal expression patterns of the immediate early genes Homer 1a and Zif268 during early and late phases of instrumental training. *Learning & Memory*, 13(5), 599–608.

Hinton EA, Li DC, Allen AG, & Gourley SL. (2019). Social isolation in adolescence disrupts cortical development and goal-dependent decision-making in adulthood, despite social reintegration. *eNeuro*, 6(5).

Holloway AL, Schaid MD, & Lerner TN. (2023). Chronically dysregulated corticosterone impairs dopaminergic transmission in the dorsomedial striatum by sex-divergent mechanisms. *Neuropsychopharmacology*, 48(9), 1328–1337.

Hsu R, Taylor JR, Newton SS, Alvaro JD, Haile C, Han G, Hruby VJ, Nestler EJ, & Duman RS. (2005). Blockade of melanocortin transmission inhibits cocaine reward. *European Journal of Neuroscience*, 21(8), 2233–2242.

Huszar D, Lynch CA, Fairchild-Huntress V, Dunmore JH, Fang Q, Berkemeier LR, Gu W, Kesterson RA, Boston BA, Cone RD, Smith FJ, Campfield LA, Burn P, & Lee F. (1997). Targeted disruption of the melanocortin-4 receptor results in obesity in mice. *Cell*, 88(1), 131–141.

Insel TR & Shapiro LE. (1992). Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *PNAS*, 89(13), 5981–5985.

Joo K, Rhie DJ, & Jang HJ. (2015). Enhancement of GluN2B subunit-containing NMDA receptor underlies serotonergic regulation of long-term potentiation after critical period in the rat visual cortex. *Korean Journal of Physiology & Pharmacology*, 19(6), 523–531.

Kanoski SE, Fortin SM, Ricks KM, & Grill HJ. (2013). Ghrelin signaling in the ventral hippocampus stimulates learned and motivational aspects of feeding via PI3K-Akt signaling. *Biological Psychiatry*, 73(9), 915–923.

Kastin AJ, Dempsey GL, LeBlanc B, Dyster-Aas K, & Schally AV. (1974). Extinction of an appetitive operant response after administration of MSH. *Hormones and Behavior*, 5(2), 135–139.

Kishi T, Aschkenasi CJ, Lee CE, Mountjoy KG, Saper CB, & Elmquist JK. (2003). Expression of melanocortin 4 receptor mRNA in the central nervous system of the rat. *Journal of Comparative Neurology*, 457(3), 213–235.

Klawonn AM, Fritz M, Nilsson A, Bonaventura J, Shionoya K, Mirrasekhian E, Karlsson U, Jaarola M, Granseth B, Blomqvist A, Michaelides M, & Engblom D. (2018). Motivational valence is determined by striatal melanocortin 4 receptors. *Journal of Clinical Investigation*, 128(7), 3160–3170.

Kravitz AV, Freeze BS, Parker PRL, Kay K, Thwin MT, Deisseroth K, & Kreitzer AC. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature*, 466(7306), 622–626.

Kwak S & Jung MW. (2019). Distinct roles of striatal direct and indirect pathways in value-based decision making. *eLife*, 8, e46050.

Lasaga M, Debeljuk L, Durand D, Scimonelli TN, & Caruso C. (2008). Role of alpha-melanocyte stimulating hormone and melanocortin 4 receptor in brain inflammation. *Peptides*, 29(10), 1825–1835.

Lerma-Cabrera JM, Carvajal F, Alcaraz-Iborra M, de la Fuente L, Navarro M, Thiele TE, & Cubero I. (2013). Adolescent binge-like ethanol exposure reduces basal α-MSH expression in the hypothalamus and the amygdala of adult rats. *Pharmacology, Biochemistry, and Behavior*, 110, 66–74.

Lerma-Cabrera JM, Carvajal F, Garbutt JC, Navarro M, & Thiele TE. (2020). The melanocortin system as a potential target for treating alcohol use disorders: A review of pre-clinical data. *Brain Research*, 1730, 146628.

Lex B & Hauber W. (2010). The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. *Cerebral Cortex*, 20(4), 873–883.

Lezcano NE, De Barioglio SR, & Celis ME. (1995). α-MSH changes cyclic AMP levels in rat brain slices by an interaction with the D1 dopamine receptor. *Peptides*, 16(1), 133–137.

Li DC, Dighe NM, Barbee BR, Pitts EG, Kochoian B, Blumenthal SA, Figueroa J, Leong T, & Gourley SL. (2022). A molecularly integrated amygdalo-fronto-striatal network coordinates flexible learning and memory. *Nature Neuroscience*, 25(9), 1213–1224.

Li DC, Hinton EA, & Gourley SL. (2021). Persistent behavioral and neurobiological consequences of social isolation during adolescence. *Seminars in Cell & Developmental Biology*, 118, 73–82.

Li MM, Madara JC, Steger JS, Krashes MJ, Balthasar N, Campbell JN, Resch JM, Conley NJ, Garfield AS, & Lowell BB. (2019). The paraventricular hypothalamus regulates satiety and prevents obesity via two genetically distinct circuits. *Neuron*, 102(3), 653-667.e6.

Lim BK, Huang KW, Grueter BA, Rothwell PE, & Malenka RC. (2012). Anhedonia requires MC4R-mediated synaptic adaptations in nucleus accumbens. *Nature*, 487(7406), 183–189.

Lingawi NW & Balleine BW. (2012). Amygdala central nucleus interacts with dorsolateral striatum to regulate the acquisition of habits. *Journal of Neuroscience*, 32(3), 1073–1081.

Liu H, Kishi T, Roseberry AG, Cai X, Lee CE, Montez JM, Friedman JM, & Elmquist JK. (2003). Transgenic mice expressing green fluorescent protein under the control of the melanocortin-4 receptor promoter. *Journal of Neuroscience*, 23(18), 7143–7154.

Liu Z & Hruby VJ. (2022). MC4R biased signalling and the conformational basis of biological function selections. *Journal of Cellular and Molecular Medicine*, 26(15), 4125–4136.

Ma K, & McLaurin J. (2014). α-melanocyte stimulating hormone prevents GABAergic neuronal loss and improves cognitive function in Alzheimer's disease. *Journal of Neuroscience*, 34(20), 6736–6745.

Mace OJ, Tehan B, & Marshall F. (2015). Pharmacology and physiology of gastrointestinal enteroendocrine cells. *Pharmacology Research & Perspectives*, 3(4), e00155.

Manvich DF, Webster KA, Foster SL, Farrell MS, Ritchie JC, Porter JH, & Weinshenker D. (2018). The DREADD agonist clozapine N-oxide (CNO) is reverse-metabolized to clozapine and produces clozapine-like interoceptive stimulus effects in rats and mice. *Scientific Reports*, 8(1), 3840.

Marciani L, Cox EF, Pritchard SE, Major G, Hoad CL, Mellows M, Hussein MO, Costigan C, Fox M, Gowland PA, & Spiller RC. (2015). Additive effects of gastric volumes and macronutrient composition on the sensation of postprandial fullness in humans. *European Journal of Clinical Nutrition*, 69(3), 380–384.

Maroteaux M, Valjent E, Longueville S, Topilko P, Girault JA, & Hervé D. (2014). Role of the plasticity-associated transcription factor zif268 in the early phase of instrumental learning. *PloS One*, 9(1), e81868.

Martin WJ & MacIntyre DE. (2004). Melanocortin receptors and erectile function. *European Urology*, 45(6), 706–713.

Maslow AH. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396.

Maslow AH. (1971). The farther reaches of human nature. *Arkana/Penguin Books*.

Matamales M, McGovern AE, Mi JD, Mazzone SB, Balleine BW, & Bertran-Gonzalez J. (2020). Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. *Science*, 367(6477), 549–555.

Mather M & Lighthall NR. (2012). Risk and reward are processed differently in decisions made under stress. *Current Directions in Psychological Science*, 21(1), 36–41.

McEwen BS & Morrison JH. (2013). Brain on stress: Vulnerability and plasticity of the prefrontal cortex over the life course. *Neuron*, 79(1), 16–29.

McKenzie IA, Ohayon D, Li H, de Faria JP, Emery B, Tohyama K, & Richardson WD. (2014). Motor skill learning requires active central myelination. *Science*, 346(6207), 318–322.

McLay RN, Pan W, & Kastin AJ. (2001). Effects of peptides on animal and human behavior: A review of studies published in the first twenty years of the journal *Peptides*. *Peptides*, 22(12), 2181–2255.

Mehrotra D & Dubé L. (2023). Accounting for multiscale processing in adaptive real-world decision-making via the hippocampus. *Frontiers in Neuroscience*, 17, 1200842.

Mensch S, Baraban M, Almeida R, Czopka T, Ausborn J, El Manira A, & Lyons DA. (2015). Synaptic vesicle release regulates myelin sheath number of individual oligodendrocytes in vivo. *Nature Neuroscience*, 18(5), 628–630.

Modi ME, Inoue K, Barrett CE, Kittelberger KA, Smith DG, Landgraf R, & Young LJ. (2015). Melanocortin receptor agonists facilitate oxytocin-dependent partner preference formation in the prairie vole. *Neuropsychopharmacology*, 40(8), 1856–1865.

Mountjoy KG. (2015). Pro-opiomelanocortin (POMC) neurons, POMC-derived peptides, melanocortin receptors and obesity: How understanding of this system has changed over the last decade. *Journal of Neuroendocrinology*, 27(6), 406–418.

Mountjoy KG, Mortrud MT, Low MJ, Simerly RB, & Cone RD. (1994). Localization of the melanocortin-4 receptor (MC4-R) in neuroendocrine and autonomic control circuits in the brain. *Molecular Endocrinology*, 8(10), 1298–1308.

Mountjoy KG & Wild JM. (1998). Melanocortin-4 receptor mRNA expression in the developing autonomic and central nervous systems. *Developmental Brain Research*, 107(2), 309–314.

Münzberg H, Singh P, Heymsfield SB, Yu S, & Morrison CD. (2020). Recent advances in understanding the role of leptin in energy homeostasis. *F1000Research*, 9.

Murray JE, Belin-Rauscent A, Simon M, Giuliano C, Benoit-Marand M, Everitt BJ, & Belin D. (2015). Basolateral and central amygdala differentially recruit and maintain dorsolateral striatum-dependent cocaine-seeking habits. *Nature Communications*, 6(1).

Musazzi L, Treccani G, & Popoli M. (2015). Functional and structural remodeling of glutamate synapses in prefrontal and frontal cortex induced by behavioral stress. *Frontiers in Psychiatry*, 6, 60.

Muschler M, Rhein M, Ritter A, Hillemacher T, Frieling H, Bleich S, & Glahn A. (2018). Epigenetic alterations of the POMC promoter in tobacco dependence. *European Neuropsychopharmacology*, 28(7), 875–879.

Mutch DM & Clément K. (2006). Unraveling the genetics of human obesity. *PLoS Genetics*, 2(12), e188.

Navone F, Jahn R, Di Gioia G, Stukenbrok H, Greengard P, & De Camilli P. (1986). Protein p38: An integral membrane protein specific for small vesicles of neurons and neuroendocrine cells. *Journal of Cell Biology*, 103(6), 2511–2527.

Nickel M & Gu C. (2018). Regulation of central nervous system myelination in higher brain functions. *Neural Plasticity*, 2018, 6436453.

Numakawa T & Odaka H. (2022). The role of neurotrophin signaling in age-related cognitive decline and cognitive diseases. *International Journal of Molecular Sciences*, 23(14), 7726.

Ognibene D, Fiore VG, & Gu X. (2019). Addiction beyond pharmacological effects: The role of environment complexity and bounded rationality. *Neural Networks*, 116, 269–278.

Oude Ophuis RJA, Boender AJ, van Rozen AJ, & Adan RAH. (2014). Cannabinoid, melanocortin and opioid receptor expression on DRD1 and DRD2 subpopulations in rat striatum. *Frontiers in Neuroanatomy*, 8, 14.

Pabst S, Brand M, & Wolf OT. (2013). Stress and decision making: A few minutes make all the difference. *Behavioural Brain Research*, 250, 39–45.

Pan WX, Mao T, & Dudman JT. (2010). Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. *Frontiers in Neuroanatomy*, 4, 147.

Pandit R, la Fleur SE, & Adan RAH. (2013). The role of melanocortins and Neuropeptide Y in food reward. *European Journal of Pharmacology*, 719(1–3), 208–214.

Parent MA, Amarante LM, Swanson K, & Laubach M. (2015). Cholinergic and ghrelinergic receptors and KCNQ channels in the medial PFC regulate the expression of palatability. *Frontiers in Behavioral Neuroscience*, 26(9), 284.

Pauli WM, Clark AD, Guenther HJ, O'Reilly RC, & Rudy JW. (2012). Inhibiting PKMζ reveals dorsal lateral and dorsal medial striatum store the different memories needed to support adaptive behavior. *Learning & Memory*, 19(7), 307–314.

Pauli WM, Hazy TE, & O'Reilly RC. (2012). Expectancy, ambiguity, and behavioral flexibility: Separable and complementary roles of the orbital frontal cortex and amygdala in processing reward expectancies. *Journal of Cognitive Neuroscience*, 24(2), 351–366.

Peak J, Chieng B, Hart G, & Balleine BW. (2020). Striatal direct and indirect pathway neurons differentially control the encoding and updating of goal-directed learning. *eLife*, 9, e58544.

Pitts EG, Li DC, & Gourley SL. (2018). Bidirectional coordination of actions and habits by TrkB in mice. *Scientific Reports*, 8(1), 4495.

Popoli M, Yan Z, McEwen BS, & Sanacora G. (2011). The stressed synapse: The impact of stress and glucocorticoids on glutamate transmission. *Nature Reviews Neuroscience*, 13(1), 22–37.

Quarta C, Claret M, Zeltser LM, Williams KW, Yeo GSH, Tschöp MH, Diano S, Brüning JC, & Cota D. (2021). POMC neuronal heterogeneity in energy balance and beyond: An integrated view. *Nature Metabolism*, 3(3), 299–308.

Quintanilla RA, Pérez MJ, Aranguiz A, Tapia-Monsalves C, & Mendez G. (2020). Activation of the melanocortin-4 receptor prevents oxidative damage and mitochondrial dysfunction in cultured hippocampal neurons exposed to ethanol. *Neurotoxicity Research*, 38(2), 421-433.

Radua J, Fortea L, Goikolea JM, Zorrilla I, Bernardo M, Arrojo M, Cunill R, Castells X, Becoña E, López-Durán A, Torrens M, Tirado-Muñoz J, Fonseca F, Arranz B, Garriga M, Sáiz PA, Flórez G, San L, & González-Pinto A. (2023). Meta-analysis of the effects of adjuvant drugs in co-occurring bipolar and substance use disorder. *Revista De Psiquiatria Y Salud Mental*, S1888-9891(23)00006-X.

Rainero I, May C, Kaye JA, Friedland RP, & Rapoport SI. (1988). CSF alpha-MSH in dementia of the Alzheimer type. *Neurology*, 38(8), 1281–1284.

Ramírez D, Saba J, Carniglia L, Durand D, Lasaga M, & Caruso C. (2015). Melanocortin 4 receptor activates ERK-cFos pathway to increase brain-derived neurotrophic factor expression in rat astrocytes and hypothalamus. *Molecular and Cellular Endocrinology*, 411, 28–37.

Renteria R, Cazares C, Baltz ET, Schreiner DC, Yalcinbas EA, Steinkellner T, Hnasko TS, & Gremel CM. (2021). Mechanism for differential recruitment of orbitostriatal transmission during actions and outcomes following chronic alcohol exposure. *eLife*, 10, e67065.

Rocha GS, Freire MAM, Britto AM, Paiva KM, Oliveira RF, Fonseca IAT, Araújo DP, Oliveira LC, Guzen FP, Morais PLAG, & Cavalcanti JRLP. (2023). Basal ganglia for beginners: The basic concepts you need to know and their role in movement control. *Frontiers in Systems Neuroscience*, 17, 1242929.

Roman KM, Briscione MA, Donsante Y, Ingram J, Fan X, Bernhard D, Campbell SA, Downs AM, Gutman D, Sardar TA, Bonno SQ, Sutcliffe DJ, Jinnah HA, & Hess EJ. (2023). Striatal subregion-selective dysregulated dopamine receptor-mediated intracellular signaling in a model of DOPA-responsive dystonia. *Neuroscience*, 517, 37–49.

Roseberry AG, Stuhrman K, & Dunigan AI. (2015). Regulation of the mesocorticolimbic and mesostriatal dopamine systems by α-melanocyte stimulating hormone and agouti-related protein. *Neuroscience & Biobehavioral Reviews*, 56, 15–25.

Rosen G, Williams A, Capra J, Connolly M, Cruz B, Lu L, Airey D, Kulkarn I, & Williams R. (2000). The Mouse Brain Library @ www.mbl.org. *Int Mouse Genome Conference*, 14(166).

Ross RA, Kim A, Das P, Li Y, Choi YK, Thompson AT, Douglas E, Subramanian S, Ramos K, Callahan K, Bolshakov VY, & Ressler KJ. (2023). Prefrontal cortex melanocortin 4 receptors (MC4R) mediate food intake behavior in male mice. *Physiology & Behavior*, 269, 114280.

Rossi MA & Yin HH. (2012). Methods for studying habitual behavior in mice. *Current Protocols in Neuroscience*, 60, 8.29.1-8.29.9.

Sandman CA, Kastin AJ, & Schally AV. (1969). Melanocyte-stimulating hormone and learned appetitive behavior. *Experientia*, 25(9), 1001–1002.

Sanger GJ, Hellström PM, & Näslund E. (2011). The hungry stomach: Physiology, disease, and drug development opportunities. *Frontiers in Pharmacology*, 1, 145.

Saper CB, Akil H, & Watson SJ. (1986). Lateral hypothalamic innervation of the cerebral cortex: Immunoreactive staining for a peptide resembling but immunochemically distinct from

pituitary/arcuate α-melanocyte stimulating hormone. *Brain Research Bulletin*, 16(1), 107–120.

Sapolsky RM. (2021). Glucocorticoids, the evolution of the stress-response, and the primate predicament. *Neurobiology of Stress*, 14, 100320.

Sarnyai Z, Vecsernyés M, Julesz J, Szabó G, & Telegdy G. (1992). Effects of cocaine and pimozide on plasma and brain alpha-melanocyte-stimulating hormone levels in rats. *Neuroendocrinology*, 55(1), 9–13.

Sato M, Nakai N, Fujima S, Choe KY, & Takumi T. (2023). Social circuits and their dysfunction in autism spectrum disorder. *Molecular Psychiatry*.

Sequeira MK & Gourley SL. (2021). The stressed orbitofrontal cortex. *Behavioral Neuroscience*, 135(2), 202–209.

Shah BP, Vong L, Olson DP, Koda S, Krashes MJ, Ye C, Yang Z, Fuller PM, Elmquist JK, & Lowell BB. (2014). MC4R-expressing glutamatergic neurons in the paraventricular hypothalamus regulate feeding and are synaptically connected to the parabrachial nucleus. *PNAS*, 111(36), 13193–13198.

Shan Q, Yu X, & Tian Y. (2023). Adolescent social isolation shifts the balance of decision-making strategy from goal-directed action to habitual response in adulthood via suppressing the excitatory neurotransmission onto the direct pathway of the dorsomedial striatum. *Cerebral Cortex*, 33(5), 1595–1609.

Shen Y, Fu WY, Cheng EYL, Fu AKY, & Ip NY. (2013). Melanocortin-4 receptor regulates hippocampal synaptic plasticity through a protein kinase A-dependent mechanism. *Journal of Neuroscience*, 33(2), 464–472.

Shen Y, Tian M, Zheng Y, Gong F, Fu A, & Ip N. (2016). Stimulation of the hippocampal POMC/MC4R circuit alleviates synaptic plasticity impairment in an Alzheimer's disease model. *Cell Reports*, 17.

Shiosaka S, Shibasaki T, & Tohyama M. (1984). Bilateral α-melanocyte stimulating hormonergic fiber system from zona incerta to cerebral cortex: Combined retrograde axonal transport and immunohistochemical study. *Brain Research*, 309(2), 346–349.

Shkundin A & Halaris A. (2023). Associations of BDNF/BDNF-AS SNPs with depression, schizophrenia, and bipolar disorder. *Journal of Personalized Medicine*, 13(9), 1395.

Siljee-Wong JE. (2011). Melanocortin MC_4 receptor expression sites and local function. *European Journal of Pharmacology*, 660(1), 234–240.

Simon JA, Kingsberg SA, Portman D, Williams LA, Krop J, Jordan R, Lucas J, & Clayton AH. (2019). Long-term safety and efficacy of bremelanotide for hypoactive sexual desire disorder. *Obstetrics and Gynecology*, 134(5), 909–917.

Smith AI & Funder JW. (1988). Proopiomelanocortin processing in the pituitary, central nervous system, and peripheral tissues. *Endocrine Reviews*, 9(1), 159–179.

Sohn JW, Harris LE, Berglund ED, Liu T, Vong L, Lowell BB, Balthasar N, Williams KW, & Elmquist JK. (2013). Melanocortin 4 receptors reciprocally regulate sympathetic and parasympathetic preganglionic neurons. *Cell*, 152(3), 612–619.

Starcke K & Brand M. (2016). Effects of stress on decisions under uncertainty: A meta-analysis. *Psychological Bulletin*, 142(9), 909–933.

Starcke K, Wolf OT, Markowitsch HJ, & Brand M. (2008). Anticipatory stress influences decision making under explicit risk conditions. *Behavioral Neuroscience*, 122(6), 1352–1360.

Stephens DW. (2018). Optimal foraging theory. In *Encyclopedia of Ecology* (pp. 284–289). Elsevier.

Stirling DR, Carpenter AE, & Cimini BA. (2021). CellProfiler Analyst 3.0: Accessible data exploration and machine learning for image analysis. *Bioinformatics*, 37(21), 3992–3994.

Sutton AK, & Krashes MJ. (2020). Integrating hunger with rival motivations. *Trends in Endocrinology and Metabolism*, 31(7), 495–507.

Swanson AM, DePoy LM, & Gourley SL. (2017). Inhibiting Rho kinase promotes goal-directed decision making and blocks habitual responding for cocaine. *Nature Communications*, 8(1), 1861.

Sweeney P, Gimenez LE, Hernandez CC, & Cone RD. (2023). Targeting the central melanocortin system for the treatment of metabolic disorders. *Nature Reviews Endocrinology*, 19(9), 507–519.

Tao YX. (2010). The melanocortin-4 receptor: Physiology, pharmacology, and pathophysiology. *Endocrine Reviews*, 31(4), 506–543.

Tapia-Rojas C, Torres AK, & Quintanilla RA. (2019). Adolescence binge alcohol consumption induces hippocampal mitochondrial impairment that persists during the adulthood. *Neuroscience*, 406, 356–368.

Tapinc DE, Ilgin R, Kaya E, Gozen O, Ugur M, Koylu EO, Kanit L, Keser A, & Balkan B. (2017). Gene expression of pro-opiomelanocortin and melanocortin receptors is regulated in the hypothalamus and mesocorticolimbic system following nicotine administration. *Neuroscience Letters*, 637, 75–79.

Thorn CA, Atallah H, Howe M, & Graybiel AM. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*, 66(5), 781–795.

Tong Y & Pelletier G. (1992). Role of dopamine in the regulation of proopiomelanocortin (POMC) mRNA levels in the arcuate nucleus and pituitary gland of the female rat as studied by in situ hybridization. *Molecular Brain Research*, 15(1–2), 27–32.

Urban DJ & Roth BL. (2015). DREADDs (designer receptors exclusively activated by designer drugs): Chemogenetic tools with therapeutic utility. *Annual Review of Pharmacology and Toxicology*, 55, 399–417.

Wake H, Ortiz FC, Woo DH, Lee PR, Angulo CM, & Fields RD. (2015). Nonsynaptic junctions on myelinating glia promote preferential myelination of electrically active axons. *Nature Communications*, 6, 7844.

Wall NR, De La Parra M, Callaway EM, & Kreitzer AC. (2013). Differential Innervation of direct- and indirect-pathway striatal projection neurons. *Neuron*, 79(2), 347–360.

Waltereit R & Weller M. (2003). Signaling from cAMP/PKA to MAPK and synaptic plasticity. *Molecular Neurobiology*, 27(1), 99–106.

Wang D, He X, Zhao Z, Feng Q, Lin R, Sun Y, Ding T, Xu F, Luo M, & Zhan C. (2015). Whole-brain mapping of the direct inputs and axonal projections of POMC and AgRP neurons. *Frontiers in Neuroanatomy*, 9.

Wassum KM. (2022). Amygdala-cortical collaboration in reward learning and decision making. *eLife*, 11, e80926.

Whyte AJ, Kietzman HW, Swanson AM, Butkovich LM, Barbee BR, Bassell GJ, Gross C, & Gourley SL. (2019). Reward-related expectations trigger dendritic spine plasticity in the mouse ventrolateral orbitofrontal cortex. *Journal of Neuroscience*, 39(23), 4595–4605.

Wong JM & Gray JA. (2018). Long-term depression is independent of GluN2 subunit composition. *Journal of Neuroscience*, 38(19), 4462–4470.

Woon EP, Butkovich LM, Peluso AA, Elbasheir A, Taylor K, & Gourley SL. (2022). Medial orbitofrontal neurotrophin systems integrate hippocampal input into outcome-specific value representations. *Cell Reports*, 40(11), 111334.

Xu P, Grueter BA, Britt JK, McDaniel L, Huntington PJ, Hodge R, Tran S, Mason BL, Lee C, Vong L, Lowell BB, Malenka RC, Lutter M, & Pieper AA. (2013). Double deletion of melanocortin 4 receptors and SAPAP3 corrects compulsive behavior and obesity in mice. *PNAS*, 110(26), 10759–10764.

Yang T, Yu K, Zhang X, Xiao X, Chen X, Fu Y, & Li B. (2023). Plastic and stimulus-specific coding of salient events in the central amygdala. *Nature*, 616(7957), 510–519.

Ye J, Yin Y, Liu H, Fang L, Tao X, Wei L, Zuo Y, Yin Y, Ke D, & Wang J. (2020). Tau inhibits PKA by nuclear proteasome-dependent PKAR2α elevation with suppressed CREB/GluA1 phosphorylation. *Aging Cell*, 19(1).

Yeo GS, Farooqi IS, Aminian S, Halsall DJ, Stanhope RG, & O'Rahilly S. (1998). A frameshift mutation in *MC4R* associated with dominantly inherited human obesity. *Nature Genetics*, 20(2), 111–112.

Yin HH & Knowlton BJ. (2004). Contributions of striatal subregions to place and response learning. *Learning & Memory*, 11(4), 459–463.

Yin HH, Knowlton BJ, & Balleine BW. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience*, 22(2), 505–512.

Yin HH, Mulcare SP, Hilário MRF, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, & Costa RM. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*, 12(3), 333–341.

Yin HH, Ostlund SB, Knowlton BJ, & Balleine BW. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*, 22(2), 513–523.

York DA, Boghossian S, & Park-York M. (2011). Melanocortin activity in the amygdala influences alcohol intake. *Pharmacology Biochemistry and Behavior*, 98(1), 112–119.

Young LJ, Lim MM, Gingrich B, & Insel TR. (2001). Cellular mechanisms of social attachment. *Hormones and Behavior*, 40(2), 133–138.

Yu X, Chen S, & Shan Q. (2021). Depression in the direct pathway of the dorsomedial striatum permits the formation of habitual action. *Cerebral Cortex*, 31(7), 3551–3564.

Zimmermann KS, Hsu CC, & Gourley SL. (2016). Strain commonalities and differences in response-outcome decision making in mice. *Neurobiology of Learning and Memory*, 131, 101–108.

Zimmermann KS, Li CC, Rainnie DG, Ressler KJ, & Gourley SL. (2018). Memory retention involves the ventrolateral orbitofrontal cortex: Comparison with the basolateral amygdala. *Neuropsychopharmacology*, 43(2), 373–383.

Zimmermann KS, Yamin JA, Rainnie DG, Ressler KJ, & Gourley SL. (2017). Connections of the mouse orbitofrontal cortex and regulation of goal-directed action selection by brain-derived neurotrophic factor. *Biological Psychiatry*, 81(4), 366–377.