Title: Morphine withdrawal recruits lateral habenula cytokine signaling to reduce synaptic excitation and sociability.

Authors: Valentinova K, Tchenio A, Trusel M, Clerke JA, Lalive AL, Tzanoulinou S, Matera A, Moutkine I, Maroteaux L, Paolicelli RC, Volterra A, Bellone C, Mameli M

Journal: Nature neuroscience

Year: 2019 Jul

Issue: 22

Volume: 7

Pages: 1053-1056

DOI: 10.1038/s41593-019-0421-4

Morphine withdrawal recruits lateral habenula cytokine signaling to reduce synaptic excitation and sociability

Kristina Valentinova[1,4,6], Anna Tchenio[1,6], Massimo Trusel[1], Joseph A. Clerke[1], Arnaud L. Lalive[1], Stamatina Tzanoulinou[3], Alessandro Matera[5], Imane Moutkine[2], Luc Maroteaux[2], Rosa C. Paolicelli[5], Andrea Volterra[1], Camilla Bellone[3] and Manuel Mameli[1,2]*

[1] The Department of Fundamental Neuroscience, The University of Lausanne 1005 Lausanne, Switzerland.
[2] Inserm, UMR-S 839, 75005 Paris, France.
[3] Department of Basic Neuroscience, The University of Geneva, Switzerland.
[4] Department of Physiology, The University of Bern, Bern, Switzerland.
[5] Department of Physiology, The University of Lausanne 1005 Lausanne, Switzerland.
[6]These authors equally contributed to the work.

To whom correspondence should be addressed:
Manuel Mameli, PhD
ORCID ID: 0000-0002-0570-6964
The Department of Fundamental Neuroscience, The University of Lausanne 1005 Lausanne, Switzerland.
Email manuel.mameli@unil.ch

Abstract

The lateral habenula (LHb) encodes aversive stimuli contributing to negative emotional states during drug withdrawal. Here, we report that morphine withdrawal (MORwd) in mice leads to microglia adaptations and diminishes glutamatergic transmission onto raphe-projecting LHb neurons. Chemogenetic inhibition of this circuit promotes MORwd-like social deficits. MORwd-driven synaptic plasticity and reduced sociability require TNFα release and neuronal TNF-Receptor-1 activation. Hence, habenular cytokines control synaptic and behavioral adaptations during drug withdrawal.

Opiate withdrawal produces negative states including low mood and reduced sociability, contributing to relapse during drug abstinence[1,2]. Dysfunction of the lateral habenula (LHb) – a nucleus controlling monoaminergic systems and processing aversive stimuli – underlies depressive symptoms typical of drug withdrawal[3], yet how opiates affect the LHb remains poorly known[4,5].

We subjected mice to naloxone-precipitated MORwd to examine its repercussions on glutamatergic synapses onto LHb neurons[1]. Indeed, aberrant LHb excitatory transmission underlies negative symptoms in rodent models of depression and addiction[3]. Spontaneous excitatory postsynaptic current (sEPSC) amplitudes, but not frequencies, were reduced only in LHb neurons located in the medial aspect (^{Med}LHb; lateral LHb, ^{Lat}LHb; Supplementary Fig. 1a–b). Accordingly, MORwd diminished AMPAR:NMDAR ratios solely in ^{Med}LHb (Fig. 1a and Supplementary Fig. 1c) without affecting neurotransmitter release assessed by trains of synaptic stimulation (Supplementary Fig. 1d). Recordings obtained 1 hour after the last MOR injection (without naloxone) yielded saline-comparable AMPAR:NMDAR ratios. In contrast, spontaneous MORwd decreased AMPAR:NMDAR ratios in the ^{Med}LHb up to 30 days after the last MOR injection (Supplementary Fig. 1e).

To assess whether MORwd affects AMPAR conductance or number, we used peak-scaled non-stationary fluctuation analysis of ^{Med}LHb-recorded sEPSCs[6]. While single-channel conductance remained unaffected in MORwd slices, the number of channels opened at the peak positively correlated with amplitude values (Supplementary Fig. 1f). MORwd failed to alter AMPAR-EPSC rectification (Supplementary Fig. 1g), whereas it reduced glutamate uncaging-evoked AMPAR:NMDAR ratios, yielding a decrease only in absolute AMPAR currents (Supplementary Fig. 1h). This suggests that MORwd reduces, in a territory-specific fashion, the number of AMPARs without affecting their biophysical properties, NMDARs or presynaptic glutamate release.

MORwd-evoked plasticity occurs onto ^{Med}LHb neurons, which innervate downstream structures including the raphe nucleus and the ventral tegmental area (VTA)[7]. MORwd diminished AMPAR:NMDAR ratios solely in retrobeads-labeled Raphe– but not VTA–projecting LHb neurons (LHb_{Raphe} and LHb_{VTA}) (Fig. 1b, c) pointing to its specificity for discrete habenular circuits.

Which induction mechanism gates MORwd-driven plasticity onto ^{Med}LHb neurons? Inflammatory responses and glial cell activation emerge during drug withdrawal[8]. Indeed, spontaneous MORwd drives microglia adaptations and pro-inflammatory cytokine release (i.e. tumor necrosis factor-α (TNFα))[9]. Notably, cocaine also leads to reduced microglia arborization along with TNFα-dependent AMPAR internalization, partly underlying drug-mediated behavioral adaptations[8]. We found that, within the ^{Med}LHb, MORwd i. reduced microglial markers including Iba1 and CD68 and ii. diminished microglial cell volume (Fig. 2a–d). In parallel, naloxone- and spontaneous MORwd increased habenular TNFα immunolabeling (Fig. 2e and Supplementary Fig. 2a–d). Altogether, these findings support the engagement of inflammatory responses and cytokine signaling within the LHb during MORwd.

We then reasoned that if MORwd promotes TNFα release, artificially increasing its levels should prove sufficient to recapitulate MORwd-driven synaptic plasticity. Incubating LHb-containing slices from saline-injected mice

with exogenous TNFα reduced AMPAR:NMDAR ratios in the ^{Med}LHb. This effect was absent in ^{Lat}LHb, and occluded by naloxone and spontaneous MORwd (Fig. 2f–g and Supplementary Fig. 3a). TNFα release may arise from microglial Toll-Like Receptor 4 (TLR-4) signaling[10]. Systemically activating TLR-4 with the agonist MPLA in MOR-treated mice, instead of naloxone, mimicked MORwd plasticity (Supplementary Fig. 3b). Moreover, MPLA application in slices from MOR-treated animals reduced AMPAR currents in ^{Med}LHb, but not ^{Lat}LHb (Supplementary Fig. 3c–d). MPLA-driven EPSCs reduction did not occur in presence of a dominant negative peptide, which blocks the soluble form of TNFα (XENP1595; Supplementary Fig. 3e)[8]. Furthermore, MORwd occluded MPLA-driven synaptic depression (Supplementary Fig. 3c) and systemic injection of XENP1595 prevented MORwd-induced plasticity (Supplementary Fig. 3f). Altogether, this supports *i.* TLR-4 expression within the LHb (See Allen Brain Atlas), *ii*. its effect on AMPARs via TNFα signaling, and *iii.* the necessity and sufficiency of TNFα for MORwd-driven reduction of LHb glutamatergic transmission.

TNFα triggers its central effects partly through TNF receptor-type-1 (TNF-R1)[11]. We employed TNF-R1$^{fl/fl}$ mice to Cre-dependently knock-down TNF-R1 expression in LHb neurons (Fig. 3a–c). After viral injection, AAV_{Cre}-TNF-R1$^{fl/fl}$ mice failed to show MORwd-driven AMPAR:NMDAR ratio reduction compared to $AAV_{Control}$-infused mice (Fig. 3d). This highlights the necessity of neuronal TNF-R1 for MORwd-driven depression of synaptic AMPARs in LHb.

MORwd drives negative symptoms among which social detachment[1]. Similarly, LHb dysfunction contributes to the negative states emerging in addiction, although its implications for sociability remains poorly addressed. We examined the contribution of the LHb-to-Raphe pathway, the locus of MORwd plasticity, for social behavior. We employed an intersectional chemogenetic approach to reduce the efficiency of the LHb-to-Raphe projection. This combined the retrograde expression of cre-recombinase (HSV-Cre, dorsal raphe) with cre-dependent expression of hM4Di (rAAV-hM4Di-mCherry, DREADDi, LHb; Fig. 3e). Reducing LHb-to-Raphe efficiency

with clozapine-N-oxide diminished social preference (Fig. 3f), supporting LHb contribution to social behaviors.

Next, we recapitulated MORwd-driven reduction in social preference in C57Bl6 mice (Fig. 3g and Supplementary Fig. 4a–d). We then prepared slices from MORwd mice showing low or high sociability scores, and found that ^{Med}LHb AMPAR:NMDAR ratios positively correlated with the social score (Fig. 3h). This indicates that reduced synaptic strength in the LHb predicts opiate-WD-driven sociability deficits.

Notably, microglia and TNFα signaling also contributes to social behaviors[12]. Accordingly, MORwd-driven sociability deficits were absent after cre-dependent LHb TNF-R1 knock-down (Fig. 3i and Supplementary Fig. 4e–i). This genetic intervention did not affect locomotion (Supplementary Fig. 4j).

We found that MORwd-driven TNFα release requires neuronal TNF-R1 to reduce AMPAR transmission onto raphe-projecting, medially-located, LHb neurons. This ultimately gates MORwd-driven social impairment, a negative symptom typical of opiate withdrawal.

Together with sociability deficits, MORwd also leads to anxiety and hyperalgesia[1]. Since the contribution of LHb on these two behavioral aspects remains elusive we cannot rule out that MORwd-driven habenular plasticity is specific for withdrawal-mediated sociability defects.

The TNFα-TNF-R1 engagement within the LHb represents a previously unidentified mechanism underlying precise cellular and behavioural aspects of MORwd. Yet this is consistent with the following: *i*. drugs and drug-withdrawal-mediated modulation of AMPAR transmission partly rely on cytokine signaling[8]; *ii*. inhibition of TLR-4 attenuates MORwd symptoms[13]; *iii.* TNF-Rs contribute to social behaviors[14]. Notably, in pyramidal neurons of the hippocampus and cortex, TNFα regulates AMPAR surface expression[15,16]. This phenomenon is opposite at striatal synapses where, similarly to the LHb, TNFα application results in decreased AMPAR transmission[8]. This divergence may arise from different TNFα release dynamics, TNF receptors expression and signaling, or alternatively AMPAR anchoring properties within the LHb.

MORwd modifies the morphology of microglia in the LHb. This is, at least partly, consistent with previous findings[8], yet it remains correlative with respect to TNFα levels. This heightens the need to fill the gap in understanding microglia function and its relationship with TNFα within the habenula. Overall, while pharmacotherapies targeting pro-inflammatory pathways in substance abuse are missing, our data support cytokine signaling as a cellular pillar for aspects of drug addiction.

MORwd-driven TNFα-dependent depression of AMPAR transmission occurs at LHb_{Raphe} neurons. From a circuit standpoint, this may provide an 'anti-social' signal likely through reduced actions onto raphe neuronal populations. This is consistent with the evidence reported here that chemogenetic manipulation of the LHb-to-Raphe projection diminishes sociability. Alongside, dopamine- and serotonin-containing raphe neurons contribute to social behaviors, and medially-located LHb neurons monosynaptically connect to the latter[17–19]. Understanding the repercussions of LHb activity onto raphe neuronal subtypes during MORwd remains an important aspect for future investigation.

In conclusion, our data support the participation of cytokine-mediated plasticity for opiate-evoked negative symptoms, a mechanism by which LHb ultimately contributes to the addiction spiral.

Acknowledgements

We thank F.J. Meye and the Mameli Laboratory for comments on the manuscript. This work was supported by funds from the ERC StG SalienSy 335333, the SNSF (31003A) to M.M. We thank David Szymkowski for the donation of XENP1595, Georg Kollias and Hiltrud Strubbe for the use and breeding of TNF-R1 mouse line.

Author contributions

K.V., A.T. and M.M. performed and analyzed *ex vivo* recordings and behavior. A.L.L and J.A.C. contributed to *ex-vivo* recordings. M.T. I.M. and L.M.

performed molecular biology experiments. C.B. and S.T. provided support for behavioral experiments. A.M. and R.C.P. analyzed microglia morphology. A.V. provided conceptual and experimental input related to the TNFα signaling and TNFR1$^{fl/fl}$ mice. K.V. and M.M. conceptualized, designed the study and wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References

1. Goeldner, C. *et al.* Impaired emotional-like behavior and serotonergic function during protracted abstinence from chronic morphine. *Biol Psychiatry* **69**, 236-244 (2011).
2. Lutz, P. E. *et al.* Distinct mu, delta, and kappa opioid receptor mechanisms underlie low sociability and depressive-like behaviors during heroin abstinence. *Neuropsychopharmacology* **39**, 2694-2705 (2014).
3. Meye, F. J., Trusel, M., Soiza-Reilly, M. & Mameli, M. Neural circuit adaptations during drug withdrawal - Spotlight on the lateral habenula. *Pharmacol Biochem Behav* **162**, 87-93 (2017).
4. Margolis, E. B. & Fields, H. L. Mu Opioid Receptor Actions in the Lateral Habenula. *PLoS One* **11**, e0159097 (2016).
5. Wang, J. *et al.* Inhibition of the lateral habenular CaMKII abolishes naloxone-precipitated conditioned place aversion in morphine-dependent mice. *Neurosci Lett* **653**, 64-70 (2017).
6. Valentinova, K. & Mameli, M. mGluR-LTD at Excitatory and Inhibitory Synapses in the Lateral Habenula Tunes Neuronal Output. *Cell Rep* **16**, 2298-2307 (2016).
7. Pollak Dorocic, I. *et al.* A whole-brain atlas of inputs to serotonergic neurons of the dorsal and median raphe nuclei. *Neuron* **83**, 663-678 (2014).
8. Lewitus, G. M. *et al.* Microglial TNF-α Suppresses Cocaine-Induced Plasticity and Behavioral Sensitization. *Neuron* **90**, 483-491 (2016).
9. Campbell, L. A., Avdoshina, V., Rozzi, S. & Mocchetti, I. CCL5 and cytokine expression in the rat brain: differential modulation by chronic morphine and morphine withdrawal. *Brain Behav Immun* **34**, 130-140 (2013).
10. Michaud, M. *et al.* Proinflammatory cytokines, aging, and age-related diseases. *J Am Med Dir Assoc* **14**, 877-882 (2013).
11. Probert, L. TNF and its receptors in the CNS: The essential, the desirable and the deleterious effects. *Neuroscience* **302**, 2-22 (2015).
12. Nie, X. *et al.* The Innate Immune Receptors TLR2/4 Mediate Repeated Social Defeat Stress-Induced Social Avoidance through Prefrontal Microglial Activation. *Neuron* **99**, 464-479.e7 (2018).
13. Hutchinson, M. R. *et al.* Proinflammatory cytokines oppose opioid-induced acute and chronic analgesia. *Brain Behav Immun* **22**, 1178-1189 (2008).
14. Patel, A., Siegel, A. & Zalcman, S. S. Lack of aggression and anxiolytic-like behavior in TNF receptor (TNF-R1 and TNF-R2) deficient mice. *Brain Behav Immun* **24**, 1276-1280 (2010).
15. He, P., Liu, Q., Wu, J. & Shen, Y. Genetic deletion of TNF receptor suppresses excitatory synaptic transmission via reducing AMPA receptor synaptic localization in cortical neurons. *FASEB J* **26**, 334-345 (2012).
16. Stellwagen, D., Beattie, E. C., Seo, J. Y. & Malenka, R. C. Differential regulation of AMPA receptor and GABA receptor trafficking by tumor necrosis factor-alpha. *J Neurosci* **25**, 3219-3228 (2005).
17. Kane, M. J. *et al.* Mice genetically depleted of brain serotonin display social impairments, communication deficits and repetitive behaviors:

possible relevance to autism. *PLoS One* **7**, e48975 (2012).
18. Lecca, S. *et al.* Aversive stimuli drive hypothalamus-to-habenula excitation to promote escape behavior. *Elife* **6**, (2017).
19. Matthews, G. A. *et al.* Dorsal Raphe Dopamine Neurons Represent the Experience of Social Isolation. *Cell* **164**, 617-631 (2016).

Figure legends

Figure 1 MORwd-driven projection-specific synaptic depression in LHb.
(a) Naloxone-precipitated MORwd (NP-MORwd) protocol and AMPAR:NMDAR ratios from ^{Lat}LHb (saline+naloxone ($n_{mice/cells}$=7/11; gray) versus NP-MORwd ($n_{mice/cells}$=8/11; orange): two-sided t-test, t_{20}=0.0548, P=0.957) and ^{Med}LHb: (saline+naloxone ($n_{mice/cells}$=7/12; black) versus NP-MORwd ($n_{mice/cells}$=8/13; red): two-sided t-test, t_{23}=2.210, *P=0.037) (b) Retrobeads in Raphe (left) and retrogradely-labeled LHb_{Raphe} neurons (right) images. AMPAR:NMDAR ratios from LHb_{Raphe} neurons (saline+naloxone ($n_{mice/cells}$=5/10; black) versus NP-MORwd ($n_{mice/cells}$=5/11, red), two-sided t-test, t_{19}=3.153, **P=0.005). (c) Same as c but in VTA (saline+naloxone ($n_{mice/cells}$=2/6; black) versus NP-MORwd ($n_{mice/cells}$=4/7; red), two-sided t-test, t_{11}=0.575, P=0.577). Sal, saline; Mor, morphine; Nlx, naloxone; PAG, periaqueductal gray; DG, dentate gyrus; MHb, medial habenula; SNr, substantia nigra pars reticulata. Data are presented as box plots 10-90 percentiles with median and scatter.

Figure 2 Cytokine signaling in the LHb for MORwd plasticity.
(a) Max-projection of confocal acquisition of LHb of Iba1-positive microglia (left). Max-projection of ^{Med}LHb microglia in saline+naloxone and NP-MORwd mice (top right). 3D reconstruction of Iba1 positive microglia, containing CD68 structures in saline+naloxone and NP-MORwd (bottom right). (b) Analysis of Iba1 microglia immunoreactivity in ^{Med}LHb (Saline+naloxone (black) $n_{mice/cells}$=4/366, NP-MORwd (red) $n_{mice/cells}$=4/328, two-sided t-test, t_{692}=3.305, ***P=0.001; three independent acquisition sessions). (c, d) Quantitative analysis of microglial cell volume (based on Iba1-immunoreactivity) and CD68-positive structures. Values are normalized to Saline control (Saline+naloxone (black) $n_{mice/cells}$=4/89, NP-MORwd (red) $n_{mice/cells}$=4/83; Iba1 volume, two-sided t-test, t_{170}=3.05, **P=0.003; CD68/Iba1 volume, two-sided t-test, t_{170}=2.65, **P=0.008). Measurements were obtained from independent samples. (e) TNFα (cyan) and DAPI (magenta) immunostaining and normalized TNFα optical density in the LHb (saline+naloxone (n_{mice}=7; 6 independent acquisitions/mouse; black) versus NP-MORwd (n_{mice}=8; red; 5

independent acquisitions/mouse), two-sided t-test, t_{13}=2.991, *P=0.0104). (f) Experimental protocol and AMPAR:NMDAR ratios in ^{Lat}LHb without (–) or with (+) exogenous TNFα (saline+naloxone (–) TNFα□□$n_{mice/cells}$=2/9; □□□□□ vs saline+naloxone (+)TNFα □$n_{mice/cells}$=2/9; lighter gray□: two-sided t-test, t_{14}=0.37, P=0.717). (g) AMPAR:NMDAR ratios without (–) or with (+) exogenous TNFα□□□□□□□□□ from saline+naloxone (black, $n_{mice/cells}$=3/10 and gray, $n_{mice/cells}$=4/11) and NP-MORwd ^{Med}LHb slices (red, $n_{mice/cells}$=3/10 and pink, $n_{mice/cells}$=5/15) (interaction factor $F_{(1, 42)}$=4.90 two-way ANOVA, *P=0.039). Data are presented as box plots 10-90 percentiles with median and scatter.

Figure 3 TNFR1 requirements for MORwd-driven synaptic and behavioural adaptations.

(a) Experimental protocol. (b) Image and (c) quantification of AAV-Cre-infected LHb neurons (magenta) and total LHb neurons (cyan; 3 mice: M1, M2 and M3). (d) AMPAR:NMDAR ratios from ^{Med}LHb of: $AAV_{Control}$-TNF-$R1^{fl/fl}$ (saline+naloxone ($n_{mice/cells}$=2/9; gray) versus NP-MORwd ($n_{mice/cells}$=3/14; red)) or AAV_{Cre}-TNF-$R1^{fl/fl}$ (saline+naloxone ($n_{mice/cells}$=4/13; open gray) versus NP-MORwd ($n_{mice/cells}$=3/12; open pink)) (interaction factor $F_{(1,44)}$=4.887 two-way ANOVA, *P=0.032). (e) AAV-Control or AAV-Flex-DREADDi virus injections in the LHb and HSV-Cre virus in the Raphe. (f) Tracking, box and scatter plots of social preference test (SPT) in C57Bl6 mice expressing or not DREADDi in LHb-to-raphe neurons (Control virus (black) versus DREADDi (red), n_{mice}=8/group, t_{18}=2.271, *P=0.043). (g) Tracking plots of social preference test (SPT) in C57/Bl6 mice. Box/scatter plots showing social preference score (saline+naloxone (black) versus NP-MORwd (red), n_{mice}=22 /group, t_{42}=2.559, *P=0.014). Full circles indicate mice used for recordings in h. (h) Correlation of AMPAR:NMDAR ratios and social preference score ($n_{mice/cells}$=4/12 ; Pearson's r^2= 0.954; *P=0.023). Shaded circles represent single cell, full circles represent average per mouse (mean and sem). (i) Tracking and box/scatter plots of SPT in TNF-$R1^{fl/fl}$ mice ($AAV_{Control}$: saline+naloxone (n_{mice}=20; black), NP-MORwd (n_{mice}=20; red), AAV_{Cre}: saline+naloxone (open gray), NP-MORwd (open pink), n_{mice}=13 mice/group,

interaction factor $F_{(1, 65)}$= 7.20 two-way ANOVA, **P=0.009). S, social stimulus; O, object. Data are presented as box plots 10-90 percentiles with median and scatter.

Methods

Animals and morphine treatments. C57Bl/6J wild-type (male) and 129-Tnfrsf1atm3Gkl (male and female, referred as TNF-R1$^{fl/f}$) mice of 4–10 weeks were group-housed (three to five per cage) on a 12:12 h light cycle (lights on at 7 a.m.) with food and water ad libitum. All procedures aimed to fulfill the 3R criterion and were approved by the Veterinary Offices of Vaud (Switzerland; License VD3172). Part of the current study was carried out in the Institut du Fer a Moulin, Paris and experiments were in accordance with the guidelines of the French Agriculture and Forestry Ministry. Morphine withdrawal was either precipitated with naloxone or was induced naturally. For naloxone-precipitated morphine withdrawal, mice were subjected to six-day intraperitoneal (i.p.) morphine (20mg/kg, Cantonal Hospital of Lausanne, CHUV, Switzerland) or saline injections (saline and morphine-treated animals were housed together). On day 6, the last morphine/saline injection was given in a separate cage, thirty minutes after which animals received an i.p. injection of naloxone hydrochloride (2mg/kg, Abcam). Morphine withdrawal dependence symptoms were allowed to develop in the following thirty minutes, after what mice were either sacrificed for *ex vivo* electrophysiological recordings or were subjected to behavioral tests.

For spontaneous withdrawal, mice were treated with morphine or saline for 6 days and were sacrificed for recordings 10-13, 20 or 30 days after the last injection. For recordings in morphine-treated animals not in withdrawal, mice were sacrificed one hour after the last morphine injection on day 6. To assess TNFα involvement in morphine withdrawal plasticity, part of the animals were subjected to an i.p. injection of MPLA (Monophosphoryl Lipid A, 10 □g, a Toll-like receptor 4 activator dissolved in DMSO and saline) or saline (containing the same amount of DMSO as control)[8] instead of naloxone thirty minutes after the last morphine injection on day 6. Another portion of the animals received an i.p. injection of a dominant negative peptide blocking the soluble form of TNFα□□XENP1595, 30mg/kg, Xencor, US)[8] one hour prior the last morphine or saline injection on day 6. Thirty minutes after the morphine/saline injection these animals received naloxone and were sacrificed for recordings

as described above. No statistical methods were chose to pre-determine sample sizes but our sample sizes are similar to those reported previously[6].

Surgery. Animals of at least 4 weeks were anesthetized with ketamine (150 mg/kg)/xylazine (100 mg/kg) i.p. (Veterinary office University of Lausanne) and were placed on a stereotactic frame (Kopf, Germany). Bilateral injections of 200-400 nl volume were performed through a glass needle, at a rate of approximately 100 nl min^{-1}. The injection pipette was withdrawn from the brain 10 min after the infusion. Retrobeads (Lumafluor) were infused into the dorsal raphe nucleus (A-P:-3.5; M-L:0;D-V:-3.8 mm) or ventral tegmental area (A-P:-2.4; M-L:±0.65;D-V:-4.9 mm) of C57Bl6 mice. 129-Tnfrsf1atm3Gki mice were injected with either rAAV2-hSyn-eGFP or rAAV2-hSyn or CMV-Cre-eGFP into the LHb (A-P:-1.35; M-L: ±0.45;D-V:-3.00 mm). In another set of experiments C57Bl6 mice were injected with a herpes simplex virus derived hEF1α-cre vector (MGA Gene delivery technology core, Cambridge, MA, USA) in the raphe nucleus and with rAAV-DJ-EF1□-Flex-hM4D(Gi)-mCherry (Gene vector and virus core, Stanford medicine, CA, USA) in the LHb. Animals were allowed to recover for about 5-7 days after retrobeads injections or 5 weeks after viral infusion before being submitted to morphine/saline treatment. The injection sites were carefully examined for all electrophysiology experiments and only animals with correct injections were used for recordings. Similarly, for behavioral studies only animals with correct injection sites were included in the analysis. Brain slices from mice injected with retrobeads or viruses were directly examined under an epifluorescence microscope.

Ex-vivo electrophysiology. Animals of 5 weeks were anesthetized with ketamine/ xylazine; 150 mg/kg/100 mg/kg i.p. for preparation of LHb-containing brain slices. Slicing was done in bubbled ice- cold 95% O2/5% CO2-equilibrated solution containing (in mM): choline chloride 110; glucose 25; NaHCO3 25; MgCl2 7; ascorbic acid 11.6; sodium pyruvate 3.1; KCl 2.5; NaH2PO4 1.25; CaCl2 0.5. Coronal slices (250 µm) were prepared and transferred for 10 min to warmed solution (34 °C) of identical composition, before they were stored at ~22 °C in 95% O2/5% CO2-equilibrated artificial

cerebrospinal fluid (ACSF) containing (in mM): NaCl 124; NaHCO3 26.2; glucose 11; KCl 2.5; CaCl2 2.5; MgCl2 1.3; NaH2PO4 1. Recordings (flow rate of 2.5 ml/min) were made under an Olympus-BX51 microscope (Olympus) at 32 °C. Patch-clamp experiments were performed using borosilicate glass pipettes (2.7–4 MΩ; Phymep, France). Currents were amplified, filtered at 5 kHz and digitized at 20 kHz (Multiclamp 200B; Molecular Devices, USA). Data were acquired using Igor Pro with NIDAQ tools (Wavemetrics, USA). Access resistance was monitored by a step of −4 mV (0.1 Hz). Experiments were discarded if the access resistance increased more than 20%. All recordings were made in voltage-clamp configuration. Spontaneous EPSCs were recorded either in the lateral or in the medial territory of the LHb at −60 mV in presence of picrotoxin (100 µM, Abcam) and APV (50 µM, Abcam). The internal solution contained (in mM): CsCl 130; NaCl 4; MgCl2 2; EGTA 1.1; HEPES 5; Na2ATP 2; sodium creatine-phosphate 5; Na3GTP 0.6; spermine 0.1. The liquid junction potential was −3 mV and was not compensated. For AMPAR:NMDAR ratios EPSCs were evoked through glass electrodes placed ~200 µm from the recording site using AMPI ISO-Flex stimulator. A mixture of AMPA and NMDA currents were evoked at +40 mV (in presence of picrotoxin). The two components were pharmacologically isolated by adding APV in the recording solution and by subsequent identification of the individual currents via digital subtraction. For glutamate uncaging experiments MNI-glutamate (4-methoxy-7-nitroindolinyl-caged L-glutamate 500µM, Tocris) was added to the recording solution. Uncaging was obtained via a single-path photolysis head (Prairie Technologies) connected to a solid-state laser (Rapp Optolectronics, Germany; 405 nm, duration 1 ms, diameter 3–5µm, 250-300µm from soma). AMPAR:NMDAR ratios in uncaging experiments were calculated as follows: AMPA-EPSC at −60 mV/NMDA-EPSCs at +40 mV and the individual components were identified as previously described, using the late component of the EPSC at 30 ms after the onset. Rectification index was computed by recording AMPA-EPSC at −70 and +40 mV and was calculated as follows: (AMPA-EPSC at −70/AMPA-EPSC at +40)/1.75. To assess presynaptic release properties, trains of AMPAR-EPSCs were evoked using

extracellular stimulating electrode (5 pulses at 5Hz, 10Hz and 20Hz). The amplitudes of EPSCs trains were normalized to the amplitude of the first pulse. When indicated recordings were performed from retrogradely labeled and fluorescently identified LHb neurons. Some experiments were performed in LHb-containing slices incubated for minimum one hour with exogenous TNFα (100ng/ml). To test the effect of MPLA on AMPAR transmission, neurons were patched either in the lateral or the medial territory of the LHb and EPSCs were evoked with extracellular stimulation. Following a ten-minute baseline, MPLA (1□g/ml) was added to the recording solution and EPSCs were recorded minimum 40 minutes after. Some experiments were performed in presence of the TNFα□dominant negative peptide □□6mg/1ml; XENP1595, Xencor, US) in the recording solution.

Non stationary fluctuation analysis.

A peak-scaled nonstationary fluctuation analysis (NSFA) was performed on sEPSCs (# of events, 70–250) (Synaptosoft, USA). sEPSCs were selected by: fast rise time alignment, stable baseline holding current, and the absence of spurious fluctuations during the sEPSCs decay. The variance–amplitude relationship of sEPSC decay was plotted and fitted with the equation $\sigma^2 = iI - I^2/N + \sigma_b^2$ (where i is the mean single-channel AMPA current, I is the mean current, N is the number of channels activated at the peak, N = mean amplitude/i; and σ^2 is the baseline variance). i was estimated as the slope of the linear fit of the first portion of the parabola of the fitted sEPSC decay. The goodness-of-fit was assessed with a least-squares algorithm. The unitary current was converted in conductance based on the reversal potential of evoked EPSCs (□0 mV) and the holding potential (–60 mV). Conductance and average EPSC amplitude, mean rise time, mean decay time, access resistance, or background noise variance had no correlation ($p > 0.4$).

Histology and immunofluorescence. Mice were injected daily with saline/morphine (20mg/kg, i.p.) for 6 days. Some mice were left to develop spontaneous withdrawal, while others received naloxone (2mg/kg, i.p.) injection 30 min after the last saline/morphine injection on day 6. After 10-13

days of spontaneous withdrawal or 30 min after naloxone injection mice were anhestetized and perfused with cold 4% paraformaldehyde (PFA) in PBS (phosphate-buffered saline). The brains were extracted, post-fixed in 4% PFA in PBS, and incubated in 30% sucrose in PBS until they sank. 30 µm slices were cut at the cryostat, and stored in PBS containing 0.02% NaN3 for future analysis. For the immunofluorescence, the slices were incubated 2h in blocking buffer (5% NGS, 0.3% Triton-X in PBS) and then 24h at 4°C with the primary antibody solution (mouse anti-TNFα antibody, ab1793, Abcam, 1:100 in blocking buffer[8]). After extensive rinses, the secondary antibody was applied (goat anti-mouse IgG-conjugated Alexa 488, Invitrogen, 1:400 in blocking buffer, 24h at 4°C). The slices were then incubated in a 1:400 DAPI solution in PBS, extensively rinsed, mounted on glass slides with Pro-Long Gold Antifade Reagent (Invitrogen) and coverslipped. Images were acquired with an epifluorescent microscope with a 20x objective (AxioVision, Zeiss) using the same parameters for all the samples. The images were analyzed and processed with ImageJ software. Optical density was measured on the whole LHb area, and normalized on the neighboring thalamus [LHb-Thal/(LHb+Thal)]. 3-6 slices distributed in the rostrocaudal axis were analyzed per each animal (8 morphine, 7 saline).

Microglia analysis

Mice were anhestetized and perfused with cold 4% paraformaldehyde (PFA) in PBS (phosphate-buffered saline). The brains were extracted, post-fixed in 4% PFA in PBS, and incubated in 30% sucrose in PBS until they sank. 30 µm slices were cut at the cryostat, and stored in PBS containing 0.02% NaN3 for future analysis. Brain sections were permeabilized at room temperature (RT) in 0.5% Triton X-100 (Sigma) for 1 hr RT, followed by 1 hr RT blocking in 2% BSA 0.5% Triton X-100 and overnight incubation with primary antibody (Iba1 1:1000, Wako Chemicals, Cat. No. 019-19741 and CD68 1:400, Bio-Rad Cat. No. MCA1957[8]) at 4°C. Upon washing, sections were incubated 2 hr RT with Alexa-fluorophore-conjugated secondary antibodies (Invitrogen), and counterstained with DAPI (Invitrogen).

Confocal microscopy was performed with a TCS-SP5 (Leica) Laser Scanning System, by using a 20X dry objective and images were processed and

analyzed by Fiji Software or Imaris Software (Bitplane, Switzerland), as appropriate. Imaris was used for 3D rendering of confocal images for quantification of volumes.

For density analysis, for each acquisition, the DAPI channel was max-projected and the medial and lateral portions of the lateral habenula were manually drawn as region of interest. Then, stacks ranging from 15 to 20µm in thickness, with z-step size of 1µm, were processed as follows: Iba1 and DAPI channels were thresholded in Fiji and multiplied to each other for each stack, with the image calculator function. The resulting thresholded stack was max-projected and the microglia nuclei were counted with Analyze Particle function.

For cell soma size and Iba1 intensity, each acquisition was max-projected and the contour of cell somata in the medial portion of the lateral habenula were manually drawn based on the Iba1 immunoreactivity, and analyzed per size in μm^2 and intensity.

3D imaging analysis was performed by Imaris applying recorded algorithms (fixed thresholds for signal intensity) to all the images of the same experiment, in order to produce unbiased signal quantification. In each experiment, one brain slice per animal (n=4) per each group was acquired. The microglial cell volume and the volume of phagocytic structures were reconstructed based on the absolute intensity of Iba1 and CD68 signals, respectively. The volume of CD68 was then normalized for the Iba1 volume, to take in account the cell size.

Behavior.

Social preference test. A three-chambered social preference test was used, consisting in a rectangular Plexiglas arena (60 × 40 × 22 cm) (Ugo Basile, Varese, Italy) divided into three chambers. The walls of the center chamber had doors to allow free access to all compartments. The luminosity was around 10 lux. Thirty minutes after naloxone injection each mouse was placed in the arena for a habituation period of 10 min and was allowed to freely explore the whole empty arena. The social preference test was performed immediately after the end of the habituation: two enclosures with vertical bars were placed in the middle of the two lateral compartments, while the central

chamber remained empty. One enclosure was empty (serving as an inanimate object) whereas the other contained a social stimulus (unfamiliar juvenile mouse 25 ± 1 days old). The enclosures allowed visual, auditory, olfactory and tactile contact between the experimental mice and the social stimuli mice. The juvenile mice in the enclosures were habituated to the apparatus and the enclosures for 3 days before the experiment and each one of them served as a social stimulus for no more than 2 experimental mice (at least 6 weeks old). The test lasted 10 minutes where experimental mice were allowed to freely explore the apparatus and the enclosures. The position of the empty and juvenile-containing enclosures alternated and was counterbalanced for each trial to avoid any bias effects. Every session was video-tracked and recorded using Ethovision XT (Noldus, Wageningen, the Netherlands) or AnyMaze (Stoelting, Ireland), which provided an automated recording of the entries and time spent in the compartments, the distance moved and the velocity. The time spent in each chamber was assessed and then used to determine the preference score for the social compartment as compared to the object compartment (social/(social + object)). The arena was cleaned with 1% acetic acid solution and dried between trials.

Analysis and statistics. Animals were randomly assigned to experimental groups. Compiled data are always reported and represented as whisker box plots (whisker top/bottom represent 90/10th percentile, box top/bottom represent 75/25th percentile and median) or mean ± SEM, with single data points plotted (single cell for electrophysiology and single animal for behavioral experiments). Animals or data points were not excluded unless stated and normality test was applied. Data collection and analysis were not performed blind to the conditions of the experiments. When applicable, statistical tests were paired or unpaired t-test and one-way or two-way ANOVA. Significance for correlations was obtained applying Pearson's estimates. Testing was always performed two-tailed with $\alpha = 0.05$. More information on the methods and analysis can be found in the Life Science Reporting Summary.

Data availability statement

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

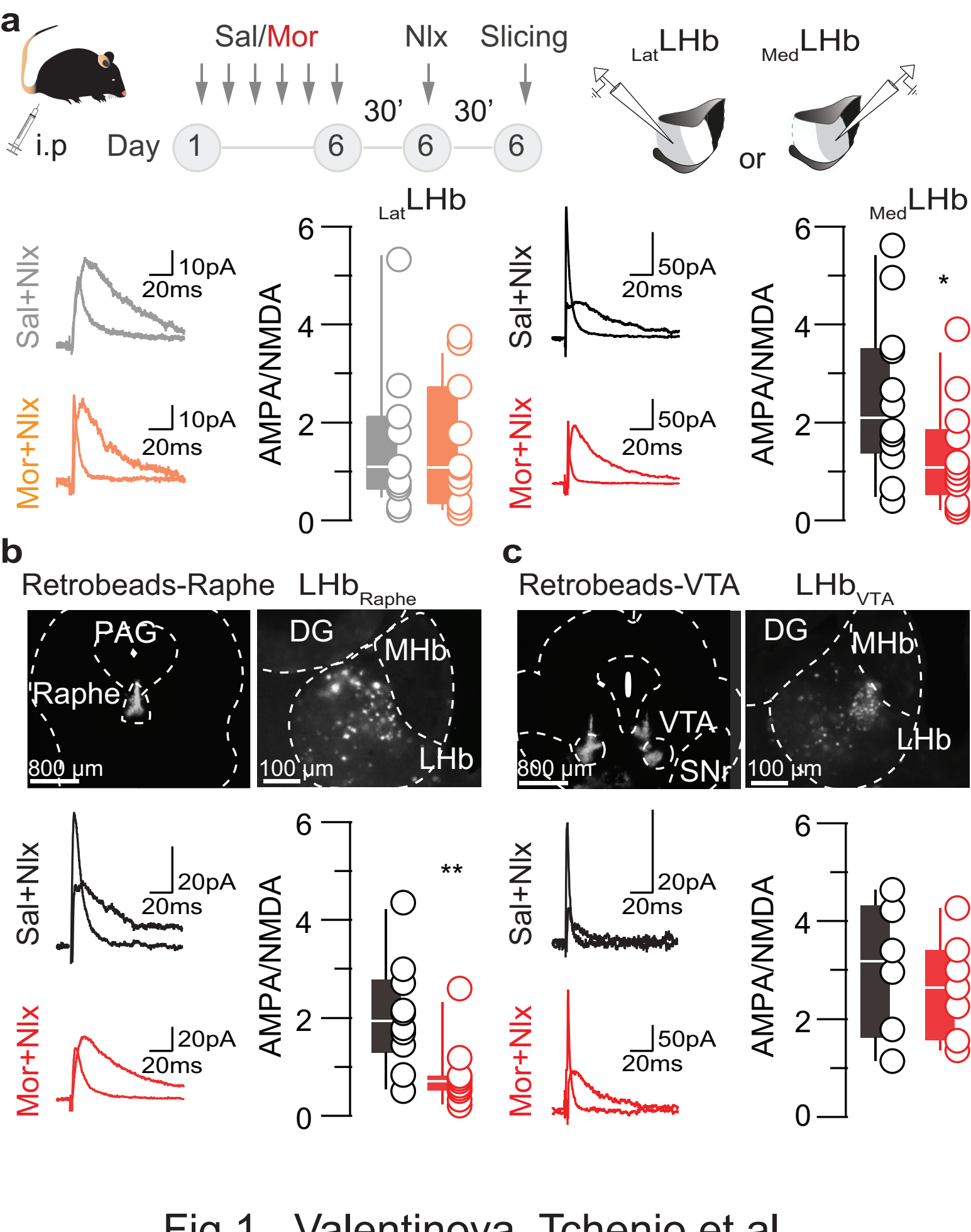

Fig 1., Valentinova, Tchenio et al.

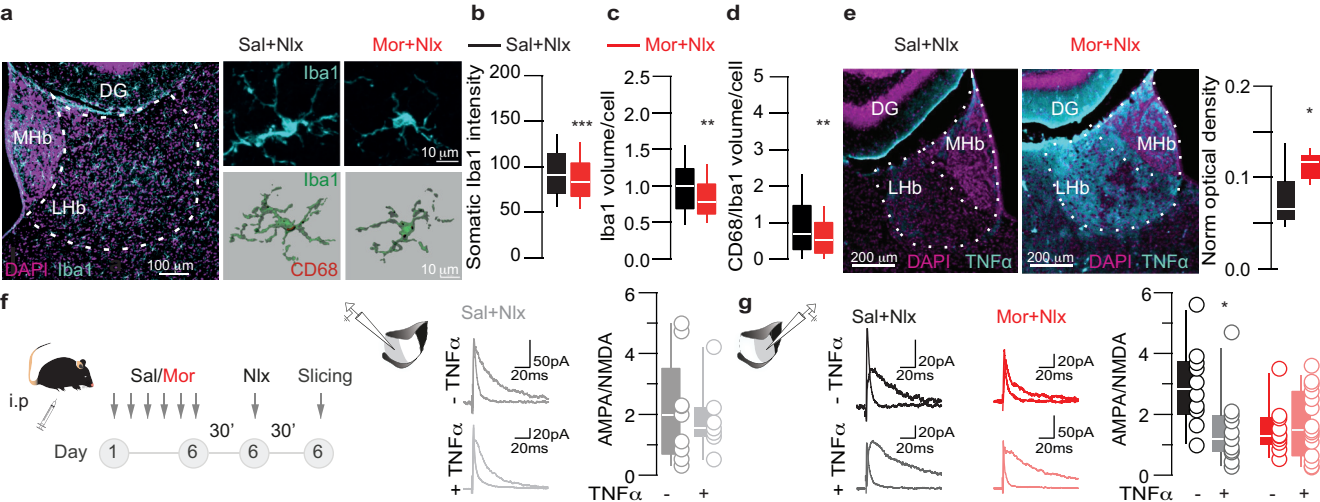

Fig 2., Valentinova, Tchenio et al.

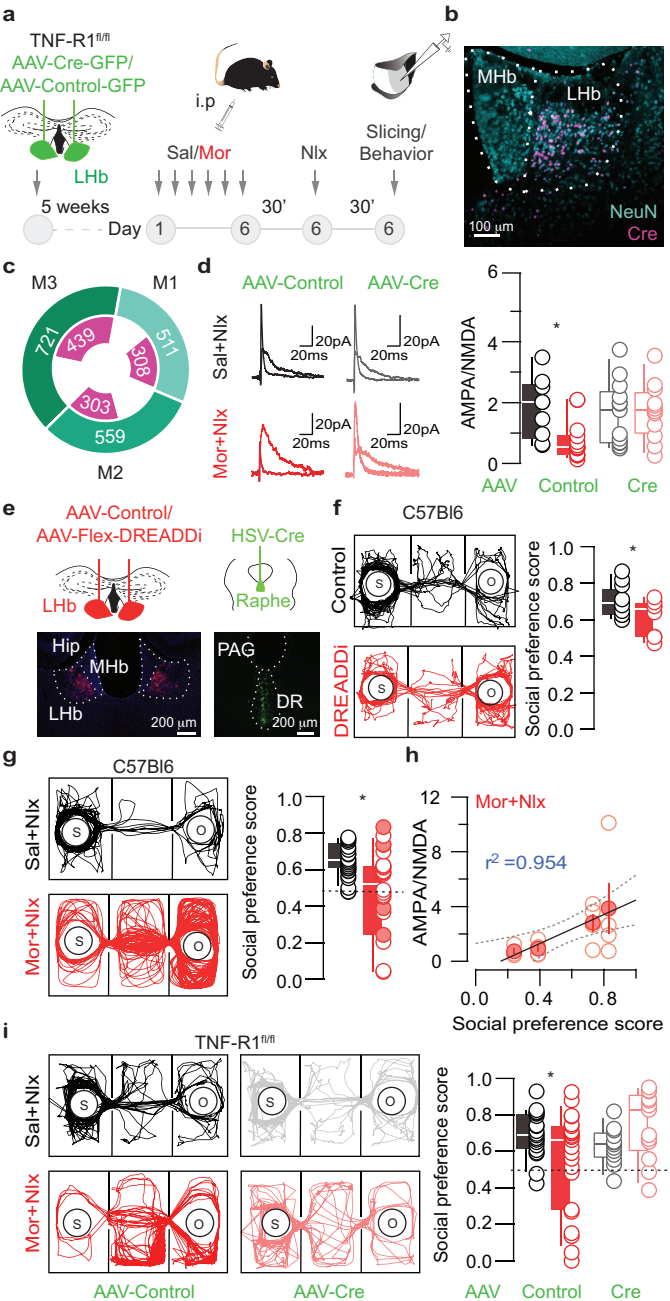

Fig 3., Valentinova, Tchenio et al.