Plasticity of reward neurocircuitry and the 'dark side' of drug addiction

George F Koob & Michel Le Moal

Drug seeking is associated with activation of reward neural circuitry. Here we argue that drug addiction also involves a 'dark side'—a decrease in the function of normal reward-related neurocircuitry and persistent recruitment of anti-reward systems. Understanding the neuroplasticity of the dark side of this circuitry is the key to understanding vulnerability to addiction.

Drug addiction has been conceptualized as a progression from impulsive to compulsive behavior, ending in chronic, relapsing drug taking. Patients with impulse control disorders experience an increasing sense of tension or arousal before committing an impulsive act; pleasure, gratification or relief at the time of committing the act; and then regret, self-reproach or guilt after the act[1]. In contrast, patients with compulsive disorders experience anxiety and stress before committing a compulsive repetitive behavior, then relief from the stress by performing the behavior[1]. In addiction, drug-taking behavior progresses from impulsivity to compulsivity in a three-stage cycle: binge/intoxication, withdrawal/negative affect and preoccupation/anticipation[2]. In the impulsive stage, the drive for the drug-taking behavior is positive reinforcement, in which stimuli increase the probability of the response. As individuals move to the compulsive stage, the drive transitions to negative reinforcement, in which removal of the aversive state increases the probability of the response. Different theoretical perspectives from experimental psychology (positive and negative reinforcement framework), social psychology (self-regulation failure framework) and neurobiology (counteradaptive and sensitization framework) can be superimposed on the stages of the addiction cycle[2]. These stages are thought to feed into each other, becoming more intense and ultimately leading to the pathological state known as addiction.

Our thesis is that addiction involves a long-term, persistent plasticity in the activity of neural circuits mediating two different motivational systems: decreased function of brain reward systems driven by natural rewards, and recruitment of anti-reward systems that drive aversive states. The concept of anti-reward is based on the hypothesis that there are brain systems in place to limit reward (see footnote in ref. 2), an 'opponent process' concept that is a general feature of biological systems[3].

From a neurobiological perspective, progression through the three stages of the addiction cycle induces plasticity in neural circuitry that drives compulsive drug taking, narrowing the behavioral repertoire to drug seeking. Animal models have been developed that have face validity (resembles the human condition) and some construct validity (possesses explanatory power) for all three stages of the addiction cycle and the transition to drug addiction. Acute self-administration of drugs (intravenous and oral) has construct validity for drug intoxication and elements of drug binges in humans. Self-stimulation and place conditioning (learning to avoid a location previously paired with an aversive stimulus or state) are sensitive measures of 'motivational' withdrawal. Cue-induced or stress-induced reinstatement has face validity and is currently under test for construct validity. Although more construct validation relative to the human condition is needed, neural substrates for each of the stages have already been identified using these models[4]. Different theoretical positions favor models from each of the three stages, although in our view, models of the transition to dependence have the most heuristic value for the human condition.

For the binge-intoxication stage, studies of the acute reinforcing effects of drugs of abuse *per se* have identified key neurobiological substrates. Important anatomical circuits include the mesocorticolimbic dopamine system originating in the ventral tegmental area and projecting to the nucleus accumbens and the extended amygdala. The extended amygdala comprises the central nucleus of the amygdala, the bed nucleus of the stria terminalis and a transition area in the medial (shell) part of the nucleus accumbens and a major projection to the lateral hypothalamus. Neurotransmitter/neuromodulator systems implicated in the acute reinforcing effects of drugs of abuse in these neuroanatomical sites include dopamine, opioid peptides, γ-aminobutyric acid (GABA), glutamate, neuropeptide Y and glucocorticoids of the hypothalamic-pituitary-adrenal (HPA) axis[5]. There is strong evidence for a role of dopamine in the acute reinforcing actions of psychostimulants, for opioid peptide receptors in the acute reinforcing effects of opioids, and for GABA and opioid peptides in the acute reinforcing actions of alcohol. Although acute drug use is not, in and of itself, addiction, the study of the neuropharmacological mechanisms for the acute reinforcing effects of drugs of abuse has had heuristic value in two major domains. Such studies provide a framework for examining neuroadaptive changes in the reward circuits with the development of addiction, and they also provide a valid model for development of medications to treat excessive drug intake (such as naltrexone for excessive drinking).

George F. Koob is in the Molecular and Integrative Neurosciences Department, The Scripps Research Institute, La Jolla, California, 92037, USA, and Michel Le Moal is at the Laboratoire de Physiopathologie des Comportements, Institut National de la Santé et de la Recherche Médicale, Unite 588, Université Victor Segalen Bordeaux 2, Bordeaux, France.
e-mail: gkoob@scripps.edu

For the purposes of this opinion piece, the withdrawal/negative affect stage can be defined as the presence of motivational signs of withdrawal in humans: chronic irritability, emotional pain, malaise, dysphoria, alexithymia and loss of motivation for natural rewards. It is characterized in animals by increases in reward thresholds during withdrawal from all major drugs of abuse. Significant plasticity occurs in the neurotransmitter circuits identified above as critical for the acute reinforcing effects of drugs of abuse. In animal models of the transition to addiction, similar changes in brain reward threshold occur that temporally precede and highly correlate with escalation in drug intake[6]. During such acute withdrawal, there is decreased activity of the mesocorticolimbic dopamine system as measured by electrophysiological recordings and *in vivo* microdialysis, and there is also decreased activity in opioid peptide, GABA, glutamate and neuropeptide Y in elements of the extended amygdala and/or nucleus accumbens. Human imaging studies of addicts during withdrawal or protracted abstinence give results that are consistent with the animal studies, including decreases in dopamine D_2 receptors (hypothesized to reflect hypodopaminergic functioning) and hypoactivity of the orbitofrontal-infralimbic cortex system[7]. These neurotransmitter/neuromodulator system changes may persist during protracted abstinence and include hypofunctioning of the HPA axis[8].

More importantly for the present thesis, as dependence and withdrawal develop, brain anti-reward systems such as corticotropin-releasing factor (CRF), norepinephrine and dynorphin are recruited. For example, extracellular CRF in the extended amygdala is increased during acute withdrawal from drugs of abuse, and critically, CRF receptor antagonists block excessive drug taking during dependence[8]. These neurotransmitter systems are activated during the development of excessive drug taking, and this activation is manifest when the drug in removed (acute withdrawal and protracted abstinence). The observation that CRF receptor antagonists in the amygdala can block excessive drug intake associated with the development of dependence provides a compelling example of a key player in the plasticity of the extended amygdala in the development of addiction. We hypothesize that anti-reward circuits are recruited as between-system neuroadaptations[9] during the development of addiction, producing aversive or stress-like states[8,10,11]. We further hypothesize that within the motivational circuits of the extended amygdala, the combination of decreases in reward neurotransmitter function and recruitment of anti-reward systems provides a powerful source of negative reinforcement that defines compulsive drug-seeking behavior and addiction. The development of the aversive emotional state that drives the negative reinforcement of addiction is here termed the 'dark side' of addiction. We further hypothesize that this chronic aversive state manifested by motivational signs of withdrawal in humans is produced in part by recruitment of the brain anti-reward systems. We believe that research in this domain has been largely neglected by the field, mainly because of an excessive focus on psychostimulant drugs and reward pathways (largely misattributed to the mesolimbic dopamine system).

A critical problem in drug addiction is chronic relapse, in which addicts return to compulsive drug taking long after acute withdrawal. This corresponds to the preoccupation/anticipation stage of the addiction cycle, outlined above. Both animal and human neuroimaging studies show that the prefrontal cortex system (orbitofrontal, medial prefrontal, prelimbic/cingulate) and the basolateral amygdala are key mediators of drug- and cue-induced reinstatement in animal models and craving and relapse in humans. Neurotransmitter systems implicated in drug- and cue- or context-induced craving again include dopamine, opioid peptides, glutamate and GABA. Neurotransmitter/neuromodulator systems implicated in stress-induced relapse include CRF, glucocorticoids and norepinephrine, suggesting that there is reactivation of both reward and anti-reward systems during relapse[12–14]. Although the relapse models have face validity, there remain serious concerns about construct validity relative to the human condition. Most reinstatement studies to date have been done with nondependent animals and may be of little more relevance to the study of addiction than studies of reinstatement of responding for a nondrug, high-incentive stimulus such as a saccharin solution, a control rarely explored in reinstatement studies. In other words, do the neuropharmacological substrates for the reinstatement of responding for saccharin—or any other nondrug reinforcer of high incentive value—parallel those of the neural substrates for nondependent doses of cocaine or heroin?

We also hypothesize that the dysregulations that constitute the dark side of drug addiction persist during protracted abstinence to set the tone for vulnerability to 'craving' by activation of the drug-, cue- and stress-induced reinstatement neurocircuits now driven by a reorganized and hypofunctioning prefrontal system[15]. A reward allostasis model is proposed to explain how dysregulation of the reward system associated with the development of motivational aspects of withdrawal is a major source of potential allostatic changes that drive and maintain addiction[2]. In this context, 'allostasis' is defined as the process of achieving stability of the reward system through change. An allostatic state is a state of chronic deviation of the reward system from its normal (homeostatic) operating level, which ultimately leads to the pathological state of addiction. More specifically, in drug addiction, allostasis is the process of attempting to maintain apparent reward function stability by changes in reward and anti-reward system neurocircuitry[5].

Neuroplasticity in the natural reward system is highlighted by decreased dopaminergic activity and hypofrontality. Neuroplasticity in the anti-reward system is highlighted by increased CRF function and is hypothesized to be particularly slow to return to homeostasis. This makes the system that drives the dark side potentially more important for driving dependence than decreases in natural reward function. For example, there is evidence of residual dysregulation of the HPA axis[16] and of the brain CRF system weeks after acute withdrawal from alcohol[17]. Similar observations are made in human addicts[18]. In contrast, withdrawal-induced decreases in dopaminergic function are relatively transient[10]. Thus, the drug addict, futilely in the short term, attempts to misregulate these drug-induced neuroplasticities by taking more drug, which only serves in the long-term to dysregulate the system further, leading to a worsening of the condition. The most prominent functional increase of the anti-reward system identified to date involves activation of the CRF-HPA axis and subsequent activation of the CRF–extended amygdala system, but other neuroadaptive processes associated with behavioral responses to stressors also may have potential roles such as neuropeptide Y, dynorphin and norepinephrine. The allostatic dysregulated reward state not only produces the motivational symptoms of acute withdrawal and protracted abstinence, but also provides the background by which drug priming, drug cues and acute stressors acquire even more power to elicit drug-seeking behavior.

Clearly appropriate, construct-validated animal models for the stages of the addiction cycle, the motivational aspects of drug-seeking and genetic vulnerability to addiction are critical to test these hypotheses. Recently, animal models with excellent face validity for the transition to addiction have been developed, which include escalated drug intake driven by dependence[6,15], responding for drug despite adverse consequences[19,20] and a narrowing of the behavioral repertoire for drug[19,20]. These models have strong face validity for the *Diagnostic and Statistical Manual of*

Mental Disorders[1] and *International Statistical Classification of Diseases*[21] criteria for addiction, are currently under test for construct validity and show promise for measuring the genetic and environmental contributions to vulnerability to addiction.

Thus, in our view, a perspective often overlooked in the drug abuse field is that there is a long-term persistent decrease in function of normal motivational systems driven by two sources: decreased function of brain reward systems (mediating natural rewards) and increased anti-reward systems (recruited as an opponent process to excessive activation of the brain reward system)[22]. It is the deficit state for normal reward, produced by excessive drug taking, that provides the core element of the motivation to seek drugs, not a hyperactive or sensitized reward state for drugs *per se*. In our view, understanding the neuroplasticity of the dark side of this circuitry will be the key to understanding individual vulnerability to addiction.

ACKNOWLEDGMENTS

We thank M. Arends for his assistance with the preparation of this manuscript.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders* 4th edn. (American Psychiatric Press, Washington, D.C., 1994).
2. Koob, G.F. & Le Moal, M. *Science* **278**, 52–58 (1997).
3. Martin, W.R. in *The Addictive States* (ed. Wikler, A.) 206–225 (Williams and Wilkins, Baltimore, 1968).
4. Shippenberg, T.S. & Koob, G.F. in *Neuropsychopharmacology: The Fifth Generation of Progress* (eds. Davis, K.L., Charney, D., Coyle, J.T. & Nemeroff, C.) 1381–1397 (Lippincott Williams and Wilkins, Philadelphia, 2002).
5. Koob, G.F. & Le Moal, M. *Neuropsychopharmacology* **24**, 97–129 (2001).
6. Ahmed, S.H., Kenny, P.J., Koob, G.F. & Markou, A. *Nat. Neurosci.* **5**, 625–626 (2002).
7. Volkow, N.D., Fowler, J.S. & Wang, G.J. *J. Clin. Invest.* **111**, 1444–1451 (2003).
8. Koob, G.F. *Alcohol. Clin. Exp. Res.* **27**, 232–243 (2003).
9. Koob, G.F. & Bloom, F.E. *Science* **242**, 715–723 (1988).
10. Nestler, E.J. *Nat. Rev. Neurosci.* **2**, 119–128 (2001).
11. Aston-Jones, G., Delfs, J.M., Druhan, J. & Zhu, Y. *Ann. NY Acad. Sci.* **877**, 486–498 (1999).
12. Piazza, P.V. & Le Moal, M. *Annu. Rev. Pharmacol. Toxicol.* **36**, 359–378 (1996).
13. Shaham, Y., Erb, S. & Stewart, J. *Brain Res. Brain Res. Rev.* **33**, 13–33 (2000).
14. See, R.E., Fuchs, R.A., Ledford, C.C. & McLaughlin, J. *Ann. NY Acad. Sci.* **985**, 294–307 (2003).
15. Le Moal, M. in *Psychopharmacology: The Fourth Generation of Progress* (eds. Bloom, F.E. & Kupfer, D.J.) 283–294 (Raven, New York, 1995).
16. Rasmussen, D.D. *et al. Alcohol. Clin. Exp. Res.* **24**, 1836–1849 (2000).
17. Valdez, G.R. *et al. Alcohol. Clin. Exp. Res.* **26**, 1494–1501 (2002).
18. Kreek, M.J. & Koob, G.F. *Drug Alcohol Depend.* **51**, 23–47 (1998).
19. Deroche-Gamonet, V., Belin, D. & Piazza, P.V. *Science* **305**, 1014–1017 (2004).
20. Vanderschuren, L.J. & Everitt, B.J. *Science* **305**, 1017–1019 (2004).
21. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems*, 10th revision (World Health Organization, Geneva, 1992).
22. Solomon, R.L. & Corbit, J.D. *Psychol. Rev.* **81**, 119–145 (1974).