Neuropharmacology 76 (2014) 370–382

Contents lists available at SciVerse ScienceDirect

Neuropharmacology

journal homepage: www.elsevier.com/locate/neuropharm

Invited review

Addiction as a stress surfeit disorder

George F. Koob*, Cara L. Buck, Ami Cohen, Scott Edwards, Paula E. Park, Joel E. Schlosburg, Brooke Schmeichel, Leandro F. Vendruscolo, Carrie L. Wade, Timothy W. Whitfield Jr., Olivier George

Committee on the Neurobiology of Addictive Disorders, The Scripps Research Institute, 10550 North Torrey Pines Road, SP30-2400, La Jolla, CA 92037, USA

ARTICLE INFO

Article history:
Received 25 March 2013
Received in revised form
22 May 2013
Accepted 28 May 2013

Keywords:
Opponent process
Extended amygdala
Corticotropin-releasing factor
Dynorphin
Reward
Compulsive
Impulsive
Sensitization
Abstinence or withdrawal
Prefrontal cortex

ABSTRACT

Drug addiction has been conceptualized as a chronically relapsing disorder of compulsive drug seeking and taking that progresses through three stages: *binge/intoxication*, *withdrawal/negative affect*, and *preoccupation/anticipation*. Drug addiction impacts multiple motivational mechanisms and can be conceptualized as a disorder that progresses from positive reinforcement (*binge/intoxication* stage) to negative reinforcement (*withdrawal/negative affect* stage). The construct of negative reinforcement is defined as drug taking that alleviates a negative emotional state. Our hypothesis is that the negative emotional state that drives such negative reinforcement is derived from dysregulation of key neurochemical elements involved in the brain stress systems within the frontal cortex, ventral striatum, and extended amygdala. Specific neurochemical elements in these structures include not only recruitment of the classic stress axis mediated by corticotropin-releasing factor (CRF) in the extended amygdala as previously hypothesized but also recruitment of dynorphin–κ opioid aversive systems in the ventral striatum and extended amygdala. Additionally, we hypothesized that these brain stress systems may be engaged in the frontal cortex early in the addiction process. Excessive drug taking engages activation of CRF not only in the extended amygdala, accompanied by anxiety-like states, but also in the medial prefrontal cortex, accompanied by deficits in executive function that may facilitate the transition to compulsive-like responding. Excessive activation of the nucleus accumbens via the release of mesocorticolimbic dopamine or activation of opioid receptors has long been hypothesized to subsequently activate the dynorphin–κ opioid system, which in turn can decrease dopaminergic activity in the mesocorticolimbic dopamine system. Blockade of the κ opioid system can also block anxiety-like and reward deficits associated with withdrawal from drugs of abuse and block the development of compulsive-like responding during extended access to drugs of abuse, suggesting another powerful brain stress/anti-reward system that contributes to compulsive drug seeking. Thus, brain stress response systems are hypothesized to be activated by acute excessive drug intake, to be sensitized during repeated withdrawal, to persist into protracted abstinence, and to contribute to the development and persistence of addiction. The recruitment of anti-reward systems provides a powerful neurochemical basis for the negative emotional states that are responsible for the dark side of addiction.

This article is part of a Special Issue entitled 'NIDA 40th Anniversary Issue'.

1. Conceptual framework for the dark side of addiction

Addiction has been conceptualized as a three-stage cycle—*binge/intoxication*, *withdrawal/negative affect*, and *preoccupation/anticipation*—with two primary sources of reinforcement: positive and negative reinforcement. Positive reinforcement is defined as the process by which presentation of a stimulus increases the probability of a response; negative reinforcement is defined as the process by which removal of an aversive stimulus (or aversive state in the case of addiction) increases the probability of a response. Secondary sources of reinforcement include conditioned positive and conditioned negative reinforcement. Different theoretical perspectives from experimental psychology (positive and negative reinforcement frameworks), social psychology (self-regulation failure framework), and neurobiology (counteradaptive and sensitization framework) can be superimposed on the stages of the

* Corresponding author. Tel.: +1 858 784 7062; fax: +1 858 784 7405.
E-mail address: gkoob@scripps.edu (G.F. Koob).

0028-3908/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropharm.2013.05.024

addiction cycle (Koob and Le Moal, 1997). These stages are thought to feed into each other, become more intense, and ultimately lead to the pathological state known as *addiction*.

1.1. Motivation, withdrawal, and opponent process

Motivation is a state that involves arousal, emotion, and expectation, all of which direct behavior. William James, describing the far greater variety of impulses in humans, wrote, "some expectation of consequences must in every case like this be aroused; and this expectation, according as it is, that of something desired or of something disliked, must necessarily either reinforce or inhibit the mere impulse" (James, 1918, p. 390). Such motivational states are not constant but rather vary over time. The concept of motivation was inextricably linked with hedonic, affective, or emotional states in addiction in the context of temporal dynamics by Solomon's opponent process theory of motivation. Solomon and Corbit (1974) postulated that hedonic, affective, or emotional states, once initiated, are automatically modulated by the central nervous system through mechanisms that reduce the intensity of hedonic feelings. The *a-process* includes affective or hedonic activation, and the *b-process* includes affective or hedonic withdrawal (abstinence). The *a-process* in drug use consists of positive hedonic responses, occurs shortly after the presentation of a stimulus, correlates closely with the intensity, quality, and duration of the reinforcer, and shows tolerance. In contrast, the *b-process* in drug use appears after the *a-process* has terminated, consists of negative hedonic responses, and is sluggish in onset, slow to build up to an asymptote, slow to decay, and gets larger with repeated exposure. Indeed, one can argue that in fact the *b-process* begins early in response to the *a-process* and opposes the manifestation of the *a-process*, yielding the phenomenon of "apparent tolerance" (Colpaert, 1996). The thesis we have elaborated is that there is a neurocircuitry change of specific neurochemical systems that account for the *b-process*. Thus, such opponent processes are hypothesized to begin early in drug taking, reflecting not only deficits in brain reward system function but also recruitment of function in the brain stress systems. Furthermore, we hypothesize that recruitment of the brain stress systems forms one of the major sources for the development of negative reinforcement in addiction.

The manifestation of a withdrawal syndrome after the removal of chronic drug administration, is thus defined in terms of *motivational* symptoms, such as the emergence of a negative emotional state (e.g., dysphoria, anxiety, and irritability) when access to the drug is prevented (Koob and Le Moal, 2001), which can be exacerbated or even caused by the *physical* signs of withdrawal. Indeed, some have argued that the development of such a negative affective state can define dependence as it relates to addiction (Russell, 1976; Koob et al., 1989; Baker et al., 2004).

Negative emotional states have been characterized in humans during acute and protracted abstinence from all major drugs of abuse (American Psychiatric Association, 1994; Koob, 2012; Khantzian, 1997). Similar results have been observed in animal models with all major drugs of abuse using intracranial self-stimulation as a measure of hedonic tone. Withdrawal from chronic cocaine (Markou and Koob, 1991), amphetamine (Paterson et al., 2000), opioids (Schulteis et al., 1994), cannabinoids (Gardner and Vorel, 1998), nicotine (Epping-Jordan et al., 1998), and ethanol (Schulteis et al., 1995) leads to increases in reward threshold during acute abstinence, and some of these elevations in threshold can last for up to 1 week (Koob, 2009). These observations lend credence to the hypothesis that opponent processes in the hedonic domain have an identifiable neurobiological basis and provide an impetus for defining the mechanisms involved. Understanding the mechanisms that drive this increase in reward thresholds is key to understanding the mechanisms that drive negative reinforcement in addiction.

Such elevations in reward threshold begin rapidly and can be observed within a single session of self-administration (Kenny et al., 2003), bearing a striking resemblance to human subjective reports of acute withdrawal. Dysphoria-like responses also accompany acute opioid and ethanol withdrawal (Liu and Schulteis, 2004; Schulteis and Liu, 2006). Here, naloxone administration following single injections of morphine increased reward thresholds, measured by intracranial self-stimulation (ICSS), and increased thresholds with repeated morphine and naloxone-induced withdrawal experience (Liu and Schulteis, 2004). Similar results were observed during repeated acute withdrawal from ethanol (Schulteis and Liu, 2006).

The development of the aversive emotional state that drives the negative reinforcement of addiction is defined here as the "dark side" of addiction. We have argued that drug addiction progresses from a source of positive reinforcement that may early on involve a form of sensitization of incentive salience, as argued by Robinson and Berridge (1993), to sensitization of the brain stress and anti-reward systems that sets up a powerful negative reinforcement process. Anti-reward is a concept developed by Koob and Le Moal (2008), based on the hypothesis that brain systems are in place to limit reward (see footnote in Koob and Le Moal, 1997), with an opponent process concept that forms a general feature of biological systems. Our concept of an anti-reward system is derived from the hypothesis of both within- and between-system neuroadaptations to excessive activation of the reward system at the neurocircuitry level. Within-system neuroadaptations are defined as the process by which the primary cellular response element to the drug (circuit A) itself adapts to neutralize the drug's effects. Persistence of the opposing effects after the drug disappears produces adaptation. A between-system neuroadaptation is a circuitry change, in which circuit B (i.e., the anti-reward circuit) is activated by circuit A (i.e., the reward circuit) (Koob and Bloom, 1988). In the present treatise, we hypothesize that a within-system neuroadaptation can also result from a between-system neuroadaptation, in which circuit B (i.e., the anti-reward circuit) is activated either in parallel or in series to suppress the activity of circuit A (see below).

1.2. Animal models of the transition to an addiction-like state as defined by escalation in drug self-administration with prolonged access

A progressive increase in the frequency and intensity of drug use is one of the major behavioral phenomena that characterize the development of addiction and has face validity with the criteria of the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) "The substance is often taken in larger amounts and over a longer period than was intended" (American Psychological Association, 1994). A framework with which to model the transition from drug use to drug addiction can be found in recent animal models of prolonged access to intravenous cocaine self-administration. Historically, animal models of cocaine self-administration involved the establishment of stable behavior from day to day to allow the reliable interpretation of data provided by within-subject designs aimed at exploring the neuropharmacological and neurobiological bases of the reinforcing effects of acute cocaine. Up until 1998, after the acquisition of self-administration, rats were typically allowed access to cocaine for 3 h or less per day to establish highly stable levels of intake and patterns of responding between daily sessions. This was a useful paradigm for exploring the neurobiological substrates for the acute reinforcing effects of drugs of abuse.