Is serotonin an upper or a downer? The evolution of the serotonergic system and its role in depression and the antidepressant response

Paul W. Andrews a,*, Aadil Bharwani a, Kyuwon R. Lee a, Molly Fox b, J. Anderson Thomson Jr. c, d

a Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
b Department of Psychiatry and Human Behavior, University of California Irvine, Orange, CA, USA
c Counseling and Psychological Services, University of Virginia Student Health, Charlottesville, VA, USA
d Institute of Law, Psychiatry, and Public Policy, University of Virginia, Charlottesville, VA, USA

A B S T R A C T

The role of serotonin in depression and antidepressant treatment remains unresolved despite decades of research. In this paper, we make three major claims. First, serotonin transmission is elevated in multiple depressive phenotypes, including melancholia, a subtype associated with sustained cognition. The primary challenge to this first claim is that the direct pharmacological effect of most symptom-reducing medications, such as the selective serotonin reuptake inhibitors (SSRIs), is to increase synaptic serotonin. The second claim, which is crucial to resolving this paradox, is that the serotonergic system evolved to regulate energy. By increasing extracellular serotonin, SSRIs disrupt energy homeostasis and often worsen symptoms during acute treatment. Our third claim is that symptom reduction is not achieved by the direct pharmacological properties of SSRIs, but by the brain’s compensatory responses that attempt to restore energy homeostasis. These responses take several weeks to develop, which explains why SSRIs have a therapeutic delay. We demonstrate the utility of our claims by examining what happens in animal models of melancholia and during acute and chronic SSRI treatment.

© 2015 Elsevier Ltd. All rights reserved.

ARTICLE INFO

Article history:
Received 4 September 2013
Received in revised form 8 January 2015
Accepted 15 January 2015
Available online 24 January 2015

Keywords:
Analysis
Depression
Serotonin
Energy regulation
Learning
Plasticity
Working memory
Distraction
Hippocampus
Prefrontal cortex
Hypothalamus

Contents

1. Introduction ......................................................... 165
2. Serotonin is elevated in multiple depressive phenotypes ......................................................... 168
   2.1. In people ......................................................... 169
       2.1.1. Polymorphism in the SERT gene ......................................................... 169
       2.1.2. 5-HIAA levels in the jugular vein ......................................................... 169
       2.1.3. Tryptophan depletion increases DRN activity in depressed patients taking ADMs ......................................................... 169
       2.1.4. Increased preference for carbohydrates in depression ......................................................... 169
       2.1.5. Tianeptine ......................................................... 170
       2.1.6. Anxiety ......................................................... 170

Abbreviations: 5-HT, 5-hydroxytryptamine (serotonin); DA, dopamine; NE, norepinephrine; ADM, antidepressant medication; SSRIs, selective serotonin reuptake inhibitors; SERT, serotonin transporter; 5-HIAA, 5-hydroxyindoleacetic acid; PFC, prefrontal cortex; mPFC, ventral part of the rodent medial prefrontal cortex; DLPFC, dorsolateral prefrontal cortex; VLPC, ventrolateral prefrontal cortex; DRN, dorsal raphe nucleus; PET, positron emission tomography; ATP, adenosine triphosphate; BDNF, brain-derived neurotrophic factor; NET, norepinephrine transporter; DAT, dopamine transporter.

* Corresponding author. Tel.: +1 905 525 9140x20820; fax: +1 9055296225.
E-mail address: pandrews@mcmaster.ca (P.W. Andrews).

http://dx.doi.org/10.1016/j.neubiorev.2015.01.018
0149-7634/© 2015 Elsevier Ltd. All rights reserved.
1. Introduction

Depression is a heterogeneous suite of states characterized by sad mood and anhedonia (an inability to experience pleasure) (Hyman, 2010; Insel and Charney, 2003). Depressive states share some genes and neurobiology in common, but they otherwise differ in symptom and etiology (Akiskal and Akiskal, 2007; Dantzer et al., 2008; Flint and Kendler, 2014; Lux and Kendler, 2010; Maier and Watkins, 1998; Parker, 2000; Raison and Miller, 2013; Sullivan et al., 2012). For instance, depressive symptoms can occur in response to infection (called sickness behavior) or starvation (Hart, 1988; Keys et al., 1950), though the symptoms are not considered pathological in these contexts (Andrews and Durisko, in press; Dantzer, 2001; Engel and Schmale, 1972). In the fifth edition of the Diagnostic and Statistical Manual for Mental Disorders (DSM-5), the diagnostic category of major depression envelops some of the symptomological heterogeneity by allowing for variability in weight, sleeping, and psychomotor activity (Table 1) (APA, 2013).

Episodes of major depression may be further subdivided into more precise phenotypes. Melancholia (weight loss, insomnia, and agitation/retardation) is considered by many to be the “biological core of depression” (Akiskal and Akiskal, 2007, p. 46). It is the most common and reliably diagnosed subtype, often accounting for 50% or more of clinical episodes (Angst et al., 2007; Taylor and Fink, 2008; Xiang et al., 2012). Melancholia is associated with heightened hypothalamic–pituitary–adrenal (HPA) activity (Taylor and Fink, 2008), which is a physiological indicator of stress (Chrousos, 2009). While it was formerly called endogenous depression, melancholia is in fact associated with stressful life events that are often serious or highly private in nature (Harkness and Monroe, 2002; Leff et al., 1970; Mundt et al., 2000; Willner et al., 1990). Atypical depression (weight gain, hypersomnia, and retardation) is the other major subtype, but it is heterogeneous and not well understood (Stewart et al., 2007).

Despite decades of research, the role serotonin plays in depressive phenotypes has not been conclusively determined. The original clue that monoamines (serotonin, norepinephrine, and dopamine) were involved in depression came from two serendipitous discoveries (Baumeister et al., 2003; Valenstein, 1998). First, during the investigations of iproniazid as a treatment for tuberculosis and imipramine as a treatment for schizophrenia, clinicians reported that these drugs could reduce depressive symptoms. An effort was then made to find a common pharmacological property that could explain their antidepressant effect. Eventually, researchers found that iproniazid inhibits the enzymes that break down the monoamines, while imipramine blocks the serotonin transporter (SERT) and the norepinephrine transporter (NET). Second, clinical observations suggested that reserpine, a drug known to deplete monoamines, increased depressive symptoms. These findings appeared to solve the puzzle. By preventing the breakdown of norepinephrine and serotonin, or preventing their clearance from the synapse, iproniazid and imipramine appeared to increase forebrain monoamine levels. The monoamine-enhancing effect of antidepressant medications (ADMs), coupled with the depression-inducing effects of reserpine, suggested that depression was caused by reduced monoamine neurotransmission (Everett and Toman, 1959; Jacobsen, 1964; Schildkraut, 1965).

Other researchers soon suggested that serotonin was the most important monoamine (Coppen, 1967). Often it is called the ‘monoamine hypothesis’ or the ‘serotonin hypothesis.’ We refer to it as the low serotonin hypothesis because it proposes a particular direction. Researchers then focused on the creation of drugs that could increase synaptic serotonin without perturbing other monoamines by selectively binding to the serotonin transporter (SERT). This research effort was successful, and the selective
serotonin reuptake inhibitors (SSRIs) are now among the most widely prescribed medications (Olfson and Marcus, 2009; Olfson et al., 2002).

However, many problems with the low serotonin hypothesis have prompted a reassessment of serotonin’s role in depression (see Box 1). Although the idea that a single neurochemical is the cause of depression is now considered simplistic, the low serotonin hypothesis still lies at the foundation of most research on depression (Albert et al., 2012). It is generally thought that reduced serotonin transmission is one of the more distal factors in the neurological chain of events that regulate depressive symptoms (Krishnan and Nestler, 2008). The fact that ketamine (which blocks a glutamate receptor) has rapid antidepressant effects lends support to the hypothesis that depressive symptoms are more proximally controlled by glutamate transmission in frontal regions (Mahar et al., 2014; Popoli et al., 2012). Others propose serotonin does not exert any regulatory control over depressive symptoms (Kirsch, 2010; Lacasse and Leo, 2009). Still others have suggested serotonin transmission is elevated in depression (Andrews and Thomson, 2009; Petty et al., 1994; Zangen et al., 1997).

In this paper, we make three major claims. The first claim, discussed in Section 2, is that serotonin transmission is elevated in multiple depressive phenotypes, including melancholia, infection, and starvation. We refer to this as the high serotonin hypothesis. What constitutes evidence of serotonin transmission is the key to the evaluation of this hypothesis. Since depression is a persistent state, reliable indices of stable serotonin transmission are particularly relevant. The 5-HIAA/5-HT ratio is the most reliable index of stable serotonin transmission, although 5-HIAA is also used (Shannon et al., 1986). While the literature on depressed patients is necessarily limited due to the methodological difficulties measuring serotonin and 5-HIAA in the human brain, the most pertinent studies support the high serotonin hypothesis. In non-human animal models of depression—where these indices can be measured more readily—abundant evidence supports the high serotonin hypothesis.

The primary challenge for the high serotonin hypothesis is explaining how ADMS, nearly all of which rapidly increase extracellular serotonin, reduce depressive symptoms. Our second claim, discussed in Section 3, is crucial to resolving this paradox. Specifically, we argue that the evolved function of the serotonergic system is energy regulation—which we define as the coordination of metabolic processes with the storage, mobilization, distribution, production and utilization of energetic resources to meet adaptive demands (Table 2).

In the brain and throughout the body, serotonin is homeostatically regulated (Best et al., 2010; Gershon and Tack, 2007; Mercado and Kilic, 2010). The energy regulation hypothesis predicts that the homeostatic equilibrium level of serotonin transmission is elevated in situations that require limited energetic resources to be reallocated among metabolically expensive processes: growth, reproduction, physical activity, maintenance, immune function, and cognition. Table 3 shows there is a stable increase in serotonin transmission to the hypothalamus in both positive and negative mood states where energy must be reallocated for prolonged periods of time. Similarly, the effects of SSRIs are state-dependent. Depending on the context, SSRIs can increase or decrease anxiety (Robert et al., 2011), motor activity (Altemus et al., 1996; Page et al., 1999), anhedonia (Branchi et al., 2013; Harrison et al., 2001), and neurotrophin signaling (Bland et al., 2007; Freitas et al., 2013; Hellweg et al., 2007; Rasmussen et al., 2002; Van Hoomissen et al., 2003). Thus, serotonin cannot be simply described as an ‘upper’ or a ‘downer’; its symptomatic effects depend on the organism’s state.

### Box 1: Problems with the low serotonin hypothesis

There has been no direct test of the low serotonin hypothesis in humans because it requires invasive techniques (see Section 4). Nevertheless, several findings have cast doubt on the low serotonin hypothesis.

1. Some drugs that block serotonin reuptake (e.g., cocaine and amphetamine) are not effective in treating depression (Charney et al., 1981).
2. Researchers and historians have concluded that reserpine-induced depression is a ‘myth’ (Baumeister et al., 2003), and that it may actually have antidepressant properties (Healy, 2002). The only placebo controlled, randomized, parallel group study of chronic reserpine treatment in anxious or depressed people showed that reserpine had an antidepressant effect (Davies and Shepherd, 1955). Indeed, some researchers argued that reserpine had antidepressant properties (Ayd, 1958), and it was used in the 1970s and 1980s to manage refractory depression (Price et al., 1987).
3. SSRIs and other ADMS decrease extracellular serotonin within minutes to hours of the first dose (Bymaster et al., 2002; Rutter and Auerbach, 1993), but they do not reduce symptoms until after several weeks of continuous treatment (Charney et al., 1981; Oswald et al., 1972). This pattern is called the therapeutic delay.
4. The attempt to reduce serotonin through tryptophan depletion fails to trigger depression in non-depressed participants (Ruhe et al., 2007).
5. Neonatal exposure to SSRIs causes depressive symptoms in adult rodents (Ansorge et al., 2004; Hansen et al., 1997).
6. Genetic downregulation of SERT, which increases synaptic serotonin, is associated with an increase in depressive symptoms (Holmes et al., 2003).
7. Meta-analyses of published and unpublished studies show that ADMS are only modestly more effective than placebo at reducing depressive symptoms (Fournier et al., 2010; Khan et al., 2002, 2005, 2011; Kirsch et al., 2008).

### Table 1

The symptoms of major depression, according to the DSM-5. Episodes of major depression can have melancholic or atypical features.

<table>
<thead>
<tr>
<th>Major depression</th>
<th>Melancholic subtype</th>
<th>Atypical subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sad mood</td>
<td>Sad mood is worse in the morning and not reactive to positive events; different from grief or loss</td>
<td>Sad mood is reactive; brightens in response to positive events</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>Anhedonia</td>
<td></td>
</tr>
<tr>
<td>Weight loss or gain</td>
<td>Weight loss</td>
<td>Weight gain</td>
</tr>
<tr>
<td>Hypersomnia or insomnia</td>
<td>Insomnia with early morning waking</td>
<td>Hypersomnia</td>
</tr>
<tr>
<td>Psychomotor agitation or retardation</td>
<td>Psychomotor agitation or retardation</td>
<td>Leaden paralysis</td>
</tr>
<tr>
<td>Fatigue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excessive feelings of worthlessness or guilt</td>
<td>Excessive guilt</td>
<td></td>
</tr>
<tr>
<td>Difficulty concentrating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suicidal ideation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sensitivity to interpersonal rejection

---

166  
The serotonergic system and energy regulation.

<table>
<thead>
<tr>
<th>Production of adenosine triphosphate (ATP)</th>
<th>Oxidative phosphorylation (slow, efficient)</th>
<th>Aerobic glycolysis (fast, inefficient)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy storage/mobilization</td>
<td>Insulin, glucagon, leptin secretion</td>
<td></td>
</tr>
<tr>
<td>Distribution of energetic resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vasoconstriction/Vasodilation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuronal activity</td>
<td>Activation/inhibition</td>
<td></td>
</tr>
<tr>
<td>Tissue uptake</td>
<td>All major tissues in the body</td>
<td></td>
</tr>
<tr>
<td>Metabolically expensive processes</td>
<td>Growth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reproduction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Immune function</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cognition</td>
<td></td>
</tr>
</tbody>
</table>

Table 2

(i.e., whether it is infected, starving, satiated, physically exhausted, sexually exhausted, etc.).

Table 4 lists the symptoms of three reliably diagnosed depressive states: sickness behavior, starvation depression, and melancholia. Each involves an altered balance between metabolically expensive processes (Fig. 1). In sickness behavior, limited energetic resources are devoted to immune function at the expense of growth and reproduction. In starvation depression, energy is devoted to maintenance functions at the expense of growth, reproduction, and immune function. In melancholia, there is an upregulation in sustained cognition at the expense of growth and reproduction. The energy regulation hypothesis suggests serotonin transmission is elevated in these states to coordinate tradeoffs in energy allocation. In melancholia, this tradeoff is coordinated by serotonin transmission to various regions, including the hypothalamus, amygdala, hippocampus and lateral prefrontal cortex (PFC) (Fig. 2). In the hippocampus and lateral PFC, the processes involved in sustained cognition are energetically expensive and can only be sustained with aerobic glycolysis (the generation of lactate from the metabolism of glucose stored in astrocytes).

Our third major claim, discussed in Section 4, is that the direct pharmacological effects of SSRIs are not responsible for symptom reduction. Rather, by rapidly increasing extracellular serotonin, SSRIs cause a disruption in energy homeostasis (the state-dependent balance between energetically expensive processes that existed prior to medication), and a worsening of symptoms. For instance, in melancholia, serotonin transmission to the PFC causes an increase in energetically expensive glutamatergic activity (Fig. 3B), which is exacerbated during acute SSRI treatment (Fig. 3C). We argue that symptom reduction is due to the compensatory changes made by the brain’s homeostatic mechanisms that attempt to restore energy homeostasis (Fig. 3D). These compensatory changes take several weeks to develop, which explains why symptoms fail to alleviate for several weeks after the initiation of SSRI treatment (the therapeutic delay).

Table 3

<table>
<thead>
<tr>
<th>State</th>
<th>Index</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infection</td>
<td>5-HIAA/5-HT</td>
<td>(Linhorst et al., 1995a)</td>
</tr>
<tr>
<td>Fasting/starvation</td>
<td>5-HIAA, 5-HT</td>
<td>(Brooks et al., 1991; Petisov et al., 2000)</td>
</tr>
<tr>
<td>Satiation</td>
<td>5-HIAA, 5-HT</td>
<td>(Meguid et al., 2000; Orosco and Nicolaidis, 1994)</td>
</tr>
<tr>
<td>Physical exhaustion</td>
<td>5-HIAA, 5-HT</td>
<td>(Blomstrand, 2011)</td>
</tr>
<tr>
<td>Sexual exhaustion</td>
<td>5-HIAA, 5-HT</td>
<td>(Lorrain et al., 1997; Mas et al., 1995)</td>
</tr>
<tr>
<td>Awake &gt; REM</td>
<td>DRN activity</td>
<td>(Monti, 2010)</td>
</tr>
<tr>
<td>Female &gt; male</td>
<td>5-HIAA/5-HT</td>
<td>(Carlson and Carlson, 1988)</td>
</tr>
<tr>
<td>Proestrus</td>
<td>5-HIAA/5-HT</td>
<td>(Kerdéhu et al., 1989)</td>
</tr>
<tr>
<td>Cold &gt; warm</td>
<td>5-HIAA/5-HT</td>
<td>(Ohtani et al., 1999)</td>
</tr>
</tbody>
</table>

Fig. 1. Graphical representation of how depressed organisms make different adaptive trade-offs in allocating limited energetic resources. (The numbers are hypothetical and illustrative.) Relative to normal baseline, infection involves upregulated immune function, while growth and reproduction are downregulated (Dantzer, 2001; Lochmiller and Deerenberg, 2000); in starvation, a higher proportion of energetic reserves are devoted to maintenance (Ruiz-Núñez et al., 2013), while growth, reproduction, and immune function are suppressed (Chandia, 1991; Holliday, 1989); melancholia involves an increase in cognition (Section 5) and possibly immune function (Frank et al., 2013), while growth and reproduction are downregulated (Taylor and Fink, 2008).
Table 4
The symptomatic similarity between sickness behavior, starvation depression, melanocholia, and four commonly studied rat models of depression: inescapable shock, chronic social defeat, chronic mild stress, and the Flinders Sensitive Line. A “-” indicates data are not available. “+” indicates no statistically significant change in the symptom.

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>Sickness behavior</th>
<th>Starvation depression</th>
<th>Melanocholia</th>
<th>Inescapable shock</th>
<th>Chronic social defeat</th>
<th>Chronic mild stress</th>
<th>Flinders Sensitive Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anhedonia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td>+,16,1</td>
<td></td>
<td>30</td>
<td>12</td>
<td>10</td>
<td>32</td>
<td>22</td>
</tr>
<tr>
<td>Sexual behavior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REM sleep</td>
<td>+,16,1</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow wave sleep</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passive coping</td>
<td>Yes⁶,13</td>
<td>Yes¹⁰</td>
<td>Yes¹⁰</td>
<td>Yes¹⁰</td>
<td>Yes¹⁰</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPA axis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference for carbohydrate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altered focus of attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex information processing</td>
<td>No5</td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References: ¹Andrews and Thomson (2009); ²Benca et al. (1996); ³Cheeta et al. (1997); ⁴Christensen and Brooks (2006); ⁵Dantzer (2001); ⁶Deak et al. (1997); ⁷Dess (1992); ⁸Exner et al. (2000); ⁹Ferreira-Nuno et al. (2002); ¹⁰Fuchs and Flugge (2002); ¹¹Hart (1988); ¹²Jackson et al. (1978); ¹³Keeney et al. (2001); ¹⁴Keys et al. (1950); ¹⁵Kramer et al. (2002); ¹⁶Lee and Maier (1988); ¹⁷MacFadyen et al. (1973); ¹⁸Meunier et al. (2007); ¹⁹Minor et al. (1984); ²⁰Moles et al. (2006); ²¹Neumann et al. (2011); ²²O’Malley et al. (2013); ²³Ovendamm (1976); ²⁴Rausch et al. (2003); ²⁵Rising et al. (1992); ²⁶Schwartz and Seeley (1997); ²⁷Shayit et al. (2003); ²⁸Shors (2004); ²⁹Taylor and Fink (2008); ³⁰Ushijima et al. (2006); ³¹Vollmayr and Henn (2003); ³²Wilson et al. (1998); ³³Yan et al. (2010).

In Section 5, we show how these claims help explain what is happening in non-human animal models of melancholia and during acute and chronic treatment with SSRIs. We conclude with implications and suggestions for future research.

2. Serotonin is elevated in multiple depressive phenotypes

It is currently impossible to measure 5-HT in the living human brain because it requires invasive techniques (Marsden, 2010). Moreover, serotonin cannot cross the blood brain barrier (Bouchard, 1972; Genot et al., 1981), so peripheral measures do not accurately reflect brain levels.

Some studies use molecular in vivo neuroimaging techniques to attempt to infer changes in endogenous serotonin levels (Bhagwagar et al., 2007; Savitz et al., 2009; Stockmeier, 2003). These techniques can measure dynamic changes in neurotransmission induced by pharmacological or physiological challenges if radiotracers measuring monoamine receptor or transporter density are sensitive to changes in endogenous monoamine levels (Paterson et al., 2010, 2013). This has been successfully applied...
to the dopaminergic system where such ligands are available (Paterson et al., 2010). However, none of the ligands currently available for the serotonin transporter and its receptors are reliable in imaging endogenous serotonin levels (Paterson et al., 2010, 2013). Thus, current neuroimaging techniques cannot reliably measure endogenous serotonin levels.

In non-human animals, invasive techniques (cyclic voltammetry, microdialysis) can be used, but most only measure extracellular neurotransmitter concentrations (Robinson et al., 2003). Extracellular concentrations are a poor index of serotonin transmission, which ideally requires the ability to measure the rate at which serotonin is released into the synapse. Extracellular concentrations reflect: (1) the rate at which serotonin is released into the synapse (transmission); and (2) the rate at which it is cleared from the synapse. Thus, synaptic serotonin can accrete without an increase in serotonin transmission (e.g., if SERT functioning is downregulated). Conversely, synaptic serotonin concentrations can decline despite elevated transmission if the rate of clearance is faster.

Single-unit recording techniques allow researchers to measure the rate of neuronal firing of individual neurons, which should generally correspond to the rate of synaptic release. But neurons in midbrain nuclei may release several neurotransmitters, so single-unit recordings must be used in conjunction with other techniques (e.g., voltammetry) to determine the rate and type of neurotransmitters that are released (Armstrong-James et al., 1980; Cheer et al., 2005). In short, it is often impractical to directly measure the rate serotonin is released into the synapse.

To deal with these difficulties, researchers have attempted to identify indices of sustained serotonin transmission (Shannon et al., 1986). This research is particularly relevant because depression is more persistent than many other emotional states. Shannon and colleagues (1986) assessed different indices of serotonin transmission to the amygdala, nucleus accumbens, and hypothalamus in response to electrical stimulation of neurons in the dorsal raphe nucleus (DRN), which is the primary source of serotonergic neurons projecting to forebrain regions. The 5-HIAA/5-HT ratio was the only index sensitive to the duration and frequency of electrical stimulation. The effect was driven by an increase in 5-HIAA, although there was a non-significant decrease in serotonin. Consequently, the 5-HIAA/5-HT ratio is the most reliable index of sustained serotonin transmission, although 5-HIAA can also be used (Barton et al., 2008; Dominguez et al., 2003; Kerdelhué et al., 1989; Winberg et al., 1992).

In the absence of information on the 5-HIAA/5-HT ratio or 5-HIAA levels, we rely on the extracellular concentration of serotonin despite its limitations.

2.1. In people

We are unaware of any evidence attempting to assess serotonin transmission in humans during starvation depression or sickness behavior. However, several lines of evidence suggest that serotonin transmission is elevated in patients with major depression.

2.1.1. Polymorphism in the SERT gene

The polymorphism in the promoter region of the SERT gene has two major variants: the short (s) and the long (l) alleles (Munafo et al., 2009). The polymorphism has transcriptional and functional consequences, with the s-allele resulting in lower densities of transporter mRNA and protein, and slower clearance of serotonin from the synaptic cleft (Murphy et al., 2004). By reducing serotonin reuptake, the s-allele keeps extracellular levels of serotonin higher than the l-allele. Consistent with the high serotonin hypothesis, the s-allele is associated with a slightly increased risk of depression in response to stressors (Karg et al., 2011).

2.1.2. 5-HIAA levels in the jugular vein

The level of 5-HIAA in the cerebrospinal fluid is an unreliable indicator of brain serotonin transmission because it is contaminated by peripheral sources (Barton et al., 2008). However, the level of 5-HIAA in the jugular vein is less contaminated because this vein directly drains blood from the brain. In an important study, a group of Australian researchers found that, relative to non-depressed controls, there was a higher overflow of 5-HIAA in the jugular veins of human subjects who met DSM–IV criteria for major depression (Barton et al., 2008). 5-HIAA overflow decreased over 12 weeks of treatment with an SSRI. Finally, among the depressed patients, 5-HIAA overflow was 2.4 times greater for carriers of the s-allele of the serotonin transporter polymorphism than for those who were homozygous for the l-allele. The authors concluded that the pattern of results “appear to run counter to… the conventional view that [major depression] is caused by a relative reduction in brain monoaminergic neuronal activity” (Barton et al., 2008, p. 42). This study provides converging evidence of increased serotonin transmission in the brains of depressed patients.

2.1.3. Tryptophan depletion increases DRN activity in depressed patients taking ADMs

While tryptophan depletion does not trigger depressive symptoms in non-depressed people (Box 1), it does trigger depressive symptoms in remitted patients who have currently or previously used serotonergic ADMs (Rupe et al., 2007). In such patients, it does not suppress DRN activity, as the low serotonin hypothesis predicts. Rather, it activates the DRN (Morris et al., 1999), which is consistent with the high serotonin hypothesis. Perhaps tryptophan depletion causes an initial decrease in serotonin around the DRN, deactivating the 5-HT1A autoreceptor and disinhibiting serotonin transmission to forebrain regions.

2.1.4. Increased preference for carbohydrates in depression

The high serotonin hypothesis is also supported less directly by the increased preference depressed patients have for carbohydrate over fat and protein (Christensen, 2001; Christensen and Brooks, 2006; Christensen and Pettijohn, 2001). This preference for carbohydrate rich food is consistent across depressed patients, regardless of the individual variability in appetite (i.e., increased or decreased appetite). Moreover, the intensity of this preference correlates to the severity of depression (Christensen and Somers, 1996).

The relative increase in carbohydrates intake causes brain serotonin levels to increase (Christensen and Somers, 1996; Fernstrom and Wurtman, 1997). Upon carbohydrate intake, insulin levels increase, stimulating the uptake of large neutral amino acids (LNAA) including valine, leucine, and isoleucine—into skeletal muscle and out of the bloodstream. The exception is tryptophan, which is not taken up into the skeletal muscle along with other LNAA’s because it is the only amino acid that binds to serum albumin. Thus, while most of the other LNAA’s are in the form of free plasma amino acids—and so readily taken up into the muscle tissue—approximately 80–90% of circulating tryptophan is normally bound to serum albumin (Fuller and Roush, 1973; Tricklebank et al., 1979) until tryptophan is released during the perfusion of brain capillaries. All LNAA’s are in competition for transport across the blood brain barrier, and by increasing the tryptophan:LNAA ratio in the blood, carbohydrates enhance the transport of tryptophan into brain tissue (Heine et al., 1995). Since tryptophan is a crucial precursor of serotonin, this can increase serotonin levels in the brain.

The low serotonin hypothesis proposes that individuals are craving carbohydrates to improve mood and seek relief in depressive symptoms by increasing serotonin (Leibenluft et al., 1993). However, if this were true, then a prolonged increase in carbohydrate intake should be an effective treatment for depression by
increasing the available amount of serotonin. Thus, the symptoms of depressed patients on high carbohydrate diets should ameliorate over time relative to depressed patients on low carbohydrate diets. However, high carbohydrate diets appear to increase depressive symptoms rather than reduce them (Cheatham et al., 2009). Moreover, in a 3-week dietary intervention, depressed patients with a restricted intake of sucrose and caffeine, which also increases extracellular serotonin (Nehlig et al., 1992), experienced a persistent amelioration in depressive symptoms (Christensen and Burrows, 1990). Thus, it seems more plausible that “the consumption of sweet carbohydrates may contribute to the development and/or maintenance of emotional distress” (Christensen and Pettijohn, 2001, p. 164).

2.2.1. Stressor models

2.2.1.1. Starvation. Starvation triggers depressive symptoms in humans (Keys et al., 1950). During periods of fasting and starvation, extracellular 5-HIAA and serotonin increase in the hypothalamus (Broocks et al., 1991; Fetissov et al., 2000). During prolonged starvation, the ability to synthesize serotonin could be reduced by a lack of dietary tryptophan. However, the metabolism of muscle tissue could liberate tryptophan to replace declining serotonin levels. In arctic char, serotonin declined in the telencephalon during four weeks of starvation, but the 5-HIAA/5-HT ratio was unaltered (Winberg et al., 1992). Since body weight declined by nearly 20%, we suggest that muscle metabolism during starvation helps maintain serotonin transmission. To help maintain extracellular serotonin levels, the starving brain also appears to downregulate the density of the serotonin transporter (Huether et al., 1997).

2.2.1.2. Infection and immune challenge. Infection also triggers depressive symptoms (Dantzner, 2001; Hart, 1988). During immune challenge, the 5-HIAA/5-HT ratio is elevated in the hypothalamus (Dunn et al., 1989; Linthorst et al., 1995a; Mefford and Heyes, 1990) and remains elevated while the organism is sick (Dunn, 2006). The 5-HIAA/5-HT ratio is elevated in the hippocampus as well (Linthorst et al., 1995b). By themselves, pyrogenic cytokines also increase serotonin transmission. IL-1β has been found to increase 5-HIAA in the PFC, nucleus accumbens and hippocampus (Merali et al., 1997), while IL-6 has been found to increase the 5-HIAA/5-HT ratio in the brain stem, hypothalamus and striatum (Wang and Dunn, 1998; Zhang et al., 2001).

2.2.1.3. Inescapable shock. Inescapable shock is a common rodent model of depression, and it increases extracellular serotonin in the medial PFC (Amat et al., 2005), ventral hippocampus and dorsal periaqueductal gray (Amat et al., 1998b), basolateral amygdala (Amat et al., 1998a), and nucleus accumbens (Bland et al., 2003b). Inescapable shock also increases the activity of serotonergic neurons, as indexed by c-Fos expression (Grahm et al., 1999), suggesting that the increase in extracellular serotonin is caused by an increase in transmission. Since the 5-HIAA/5-HT ratio is our main index of serotonin transmission, it is perhaps more telling that inescapable shock increases this ratio across many regions, including the locus coeruleus, brain stem, thalamus, hypothalamus, striatum, frontal cortex, and hippocampus (Adell et al., 1998).

2.2.1.4. Chronic social defeat. In rats, chronic social defeat has been found to increase extracellular serotonin in the DRN (Amat et al., 2010), 5-HIAA levels in the amygdala and hippocampus, and the 5-HIAA/5-HT ratio in the midbrain and hypothalamus (Blanchard et al., 1993). In mice, chronic social defeat has been found to increase the 5-HIAA/5-HT ratio in the hypothalamus and hippocampus (Beitia et al., 2005; Keeney et al., 2006).

2.2.1.5. Chronic mild stress. In chronic mild stress, serotonin transmission (as indexed by 5-HIAA or the 5-HIAA/5-HT ratio) is elevated in many regions, including the PFC, hypothalamus, hippocampus, and amygdala (Bekris et al., 2005; Gamaro et al., 2003; Li et al., 2012; Tannenbaum and Anisman, 2003; Tannenbaum et al., 2002).

2.2.1.6. Chronic restraint stress. Chronic restraint stress also shows evidence of elevated serotonin transmission in some regions, although there are also many null effects (O’Mahony et al., 2011; Torres et al., 2002). The mixed evidence is probably due to the fact that rodents are more likely to habituate to chronic restraint than other models, thereby lessening its depressogenic impact (Bergström et al., 2008; Marin et al., 2007).

2.2.1.7. Maternal separation and social isolation. Some depression models involve examining how rodents respond to a stressor after...
having been raised apart from their mothers or in social isolation. In a study using this paradigm, there were no differences in serotonin transmission between maternally separated rats and control rats at baseline (Daniels et al., 2004). However, after exposure to a restraint stressor, the maternally separated rats had a higher 5-HIAA/5-HT ratio in the frontal cortex and hypothalamus, and 5-HIAA levels were elevated in the frontal cortex and hippocampus.

Brush-tailed rats (Octodon degus) raised in social isolation show increased innervation of serotonergic fibers to the infralimbic region of the mPFC (Braun et al., 1999). Hooded Lister rats raised in social isolation also showed an increase in serotonin release (as measured by voltammetry and microdialysis) in the frontal cortex in response to KCl and fenfluramine (Crespi et al., 1992), and an increase in extracellular serotonin in the nucleus accumbens in response to a conditioned stress paradigm (Fulford and Marsden, 1997).

2.2.1.8. Neonatal SSRI exposure. Interestingly, neonatal exposure to SSRIs is a model of depression that is also consistent with the high serotonin hypothesis. Adult rats exposed to SSRIs as neonates show increased serotonin transmission (indexed by the 5-HIAA/5-HT ratio) in the hypothalamus (Feenstra et al., 1996; Hilakivi et al., 1987), and exhibit a depressive behavioral profile (Ansorge et al., 2004; Hansen et al., 1997).

2.2.2. Genetic models

2.2.2.1. The Flinders Sensitive Line. In the Flinders Sensitive Line rat, a breed that exhibits many depressive symptoms (Table 4), serotonin and 5-HIAA levels are elevated in the PFC, hippocampus and other regions relative to control rats (Zangen et al., 1997).

2.2.2.2. The congenital learned helplessness breed. We have been unable to find any evidence on serotonin transmission in rats bred for congenital learned helplessness. We predict that the 5-HIAA/5-HT ratio will be elevated in multiple regions, particularly the hypothalamus, PFC and hippocampus.

2.2.2.3. SERT and 5-HT1A knockouts. Rodents that have had the genes for SERT or the 5-HT1A receptor knocked out express higher levels of depressive symptoms (Heisler et al., 1998; Holmes et al., 2003; Ramboz et al., 1998). Consistent with the high serotonin hypothesis, 5-HT1A knockouts were found to have higher 5-HIAA levels in multiple brain regions, including the olfactory bulb, substantia nigra, thalamus, locus coeruleus, and the dorsal and medial raphe nuclei (Ase et al., 2000). While there are differences in the levels of serotonin and 5-HIAA in SERT knockout mice and SERT knockout rats (Olivier et al., 2008), the ratio of 5-HIAA/5-HT is elevated in multiple brain regions in both species (Fabre et al., 2000; Homberg et al., 2007).

2.2.3. Lesion models

2.2.3.1. Olfactory bulbectomy. Olfactory bulbectomy is the only model of depression to show reduced a 5-HIAA/5-HT ratio in multiple brain regions (Song and Leonard, 2005). This is because olfactory bulbectomy causes DRN neurons to degenerate so there is less capacity to transmit serotonin (Song and Leonard, 2005). However, it is possible that the remaining DRN neurons transmit serotonin at a heightened rate, which would be consistent with the high serotonin hypothesis. Indeed, there is an increase in the innervation of serotonin fibers and the synthesis of serotonin in cortical and limbic regions following olfactory bulbectomy (Watanabe et al., 2003).

2.2.3.2. Lesion of the DRN. Lesion of the DRN is not a model of depression, which is problematic for the low serotonin hypothesis. For instance, rats with electrolytic lesion of the DRN were less anhedonic (assessed by intake of a sucrose solution) than sham-operated controls (Wirtshafter and Asin, 1991). Given the state-dependent effects of serotonin, we do not expect DRN lesion to have simple effects on depressive symptoms. But DRN lesion should inhibit the production of depressive symptoms in response to depressogenic stressors. Indeed, DRN lesion inhibits the development of depressive symptoms in the inescapable shock, chronic social defeat, and chronic mild stress models (Chung et al., 1999; Maier et al., 1993; Yalcin et al., 2008).

2.3. Summary

In humans, the strongest evidence that serotonin transmission is elevated in depression and anxiety comes from the jugular sampling studies of 5-HIAA, which is a commonly used index of sustained serotonin transmission. This is strongly supported by the numerous studies in non-human animal models demonstrating elevations in 5-HIAA/5-HT, 5-HIAA, and even extracellular serotonin in many brain regions.

The principle challenge to the high serotonin hypothesis is the fact that the direct pharmacological properties of most antidepressants increase extracellular serotonin, most commonly by SERT blockade. We argue that this puzzle cannot be resolved without understanding the evolved function of the serotonergic system, to which we now turn.

3. The energy regulation function of the serotonergic system

In this section of the paper, we propose a novel hypothesis for the evolved function of the serotonergic system, which includes serotonin, its receptors, SERT, and other components that help regulate serotonin or its effects. Our hypothesis owes much to the research of Efrain Azmitia on the evolution of serotonin (Azmitia, 2001, 2007, 2010). One of our novel contributions is to explicitly identify the evolution of the mitochondrion as the likely point on the tree of life where serotonin evolved. This key fact helped shape our energy regulation hypothesis for the serotonergic system.

3.1. Overview of the serotonergic system

In the brain, the dorsal raphe nucleus (DRN) is the main source of serotonergic neurons that project to forebrain regions (Hornung, 2010). Tryptophan is the crucial precursor used to synthesize serotonin. Animals cannot synthesize tryptophan, so they must acquire it from their diet (Azmitia, 2010), and it goes through three main metabolic pathways: (1) protein synthesis; (2) the kynurenine pathway; and (3) the serotonin pathway. Of the tryptophan not used in protein synthesis, 99% goes down the kynurenine pathway (Stone and Darlington, 2002). The remaining 1% is converted to serotonin in two steps. First, tryptophan is converted to 5-hydroxytryptophan by tryptophan hydroxylase. Second, 5-hydroxytryptophan is converted to serotonin by aromatic L-amino acid decarboxylase (AADC).

There are no enzymes for breaking down serotonin in the extracellular space so it must be transported inside the cell. Most extracellular serotonin is transported into the pre-synaptic neuron by SERT (D’Souza and Craig, 2010). Serotonin is primarily broken down to 5-HIAA by the monoamine oxidase A (MAO-A) enzyme, which is located in mitochondria. SERT is widely expressed throughout the body (Lin et al., 2006). In the periphery, SERT is commonly expressed in many organs that take up serotonin from the bloodstream (Gershon and Tack, 2007; Mercado and Kilic, 2010; Wilson et al., 2002).

Several aspects of the serotonergic system contribute to the ability to produce diverse state-dependent effects. First, the DRN has several anatomically distinct subdivisions (Hale and Lowry, 2011),
which can cause differential transmission to forebrain regions. For instance, activation of the caudal and dorsal DRN has anxiogenic effects, while activation of the ventrolateral DRN/ventrolateral periaqueductal gray has anxiolytic effects (Hale et al., 2012).

Second, the large number of serotonin receptors arguably gives the serotonergic system greater regulatory flexibility than any other neurotransmitter system in the brain. There are 14 serotonin receptors that fall into seven classes (Barnes and Sharp, 1999). The 5-HT1A and 5-HT2A classes are inhibitory, while the 5-HT2, 5-HT3, 5-HT4, 5-HT5, and 5-HT7 classes are excitatory. Multiple serotonin receptor types are commonly co-expressed on a variety of cells throughout the brain and the periphery (Basura et al., 2001; Bickmeyer et al., 2002; Bonsi et al., 2007; Hannon and Hoyer, 2008; Irving et al., 2007; Kellermann et al., 1996; Noh and Han, 1998; Wright et al., 1995). Serotonin receptors can also form homodimers and heterodimers, the functional consequences of which are not fully understood (Albizu et al., 2011; Herrick-Davis, 2013; Renner et al., 2012). The complex control that can be achieved with the diversity of receptor function supports the role of the serotonergic system in energy regulation.

Third, the temporal firing patterns of serotonergic neurons may have different postsynaptic effects. For instance, prolonged exposure to serotonin (but not other neurotransmitters) can cause phasically firing neurons to transition to a repetitive, prolonged (tonic) firing pattern (Garraway and Hochman, 2001a). A sustained increase in serotonin transmission has a similar excitatory effect on cortical networks in the PFC (Puig and Gullidge, 2011), while 5-HT2A receptors mediate the increase in glutamatergic activity (Puig and Gullidge, 2011), while 5-HT1A receptors mediate the tonic increase in motor neuron activity (Harvey et al., 2006a, b; Liu et al., 2011).

3.2. The evolution of serotonin in mitochondria

It is very likely that serotonin evolved in mitochondria or their immediate ancestors. First, serotonin is found in plants, animals, and fungi, so the latest it could have evolved was in the unicellular eukaryotic precursor to multicellular organisms, which is about one billion years ago (Azmitia, 2010). Second, the synthesis of serotonin requires oxygen (Azmitia, 2010), which is also important in mitochondrial function. Third, MAO-A is localized to the inner surface of the outer mitochondrial membrane (Russell et al., 1979; Wang and Edmondson, 2011), which suggests a mitochondrial origin because serotonin must be inside the mitochondrion to be metabolized. Indeed, the mitochondrion may be the most common intracellular location of serotonin (Das and Steinberg, 1985), and at least some mitochondria contain the enzymes for synthesizing serotonin (Basu et al., 2008; Ichiyama et al., 1970). Surprisingly, the genes for the synthesizing enzymes are not located in the mitochondrial genome (Boore, 1999) but in the nuclear genome (Craig et al., 1991; Sumi-Ichinose et al., 1992). How could serotonin evolve in mitochondria if the genes for the synthesizing enzymes are not located in the mitochondrial genome? Of particular importance is AADC, which catalyzes the final step.

AADC belongs to a class of enzymes called pyridoxal phosphate (PLP)-dependent carboxylase enzymes (Jackson, 1990). Mitochondria and PLP-dependent carboxylases have a common phylogenetic origin. Mitochondria evolved approximately 2 billion years ago from an α-proteobacterium that formed an endosymbiotic relationship with an ill-defined larger bacterium (Emelyanov, 2001). Similarly, PLP-dependent carboxylases evolved from α-proteobacteria (Ilyer et al., 2004; Jackson, 1990). Thus, AADC evolved from the PLP-dependent carboxylase precursor, probably in the mitochondrion. As mitochondria evolved and became more integrated with the endosymbiotic host, some mitochondrial genes were lost, and some were transferred to the nuclear genome (Andersson et al., 2003; Emelyanov, 2001). During this process, the AADC gene was transferred to the nuclear genome and deleted from the mitochondrial genome (Ilyer et al., 2004).

3.3. The mitochondrial functions of serotonin

What does serotonin do in mitochondria? Serotonin increases the potential across the inner mitochondrial membrane, although the precise mechanisms by which this is achieved are unknown (Basu et al., 2008). Serotonin may affect mitochondrial function as the precursor to melatonin. Mitochondria have the enzymes that convert serotonin to melatonin, and melatonin increases the efficiency of energy production by accelerating electron transport (Tan et al., 2013). Electron transport generates reactive oxygen and nitrogen species that can damage the mitochondrial and other cellular structures (Tan et al., 2013), and serotonin and melatonin both have powerful antioxidant properties (Park et al., 2002).

3.4. What is the function of the serotonergic system?

The serotonergic system affects so many processes that some researchers despair of ever identifying a unifying function. Based on the foregoing, serotonin probably evolved first to regulate mitochondrial activity. This function could, in principle, affect every major system, organ, and metabolic process in the body. Moreover, it is so important that it is highly likely that any subsequent functions of the serotonergic system were at least consistent with this original function, and probably facilitate it (for a similar point, see Azmitia, 2010).

Mitochondria face adaptive challenges within multicellular organisms, and the serotonergic system could have evolved to solve these problems. Multicellular organisms are composed of specialized cells with different functions that respond to environmental contingencies, and these responses depend on ATP produced by mitochondria (or glycolysis in the cytosol). Multicellular organisms must therefore coordinate the distribution of important energetic resources (glucose, fatty acids, amino acids) throughout the organism with regional mitochondrial activity patterns. We propose that the serotonergic system evolved to promote energy regulation, which we define as the coordination of metabolic processes with the distribution and utilization of limited energetic resources to meet adaptive demands.

Other prominent hypotheses for serotonin propose that it evolved to promote homeostasis (Azmitia, 2007) or phenotypic plasticity (Branchi, 2011; Homberg, 2012). While it is undeniable that serotonin can affect homeostasis and phenotypic plasticity, this is true of all biochemicals: it makes little sense to single out the serotonergic system for these functions. However, the serotonergic system is unique in that it can simultaneously coordinate the production, storage, mobilization, distribution, and utilization of energy. Arguably, no other biochemical system in the body can do this.

3.4.1. Serotonin and energy regulation

3.4.1.1. Glucose metabolism. Serotonin regulates the two major metabolic pathways for generating ATP from glucose. In addition to affecting electron transport in mitochondria (oxidative phosphorylation), serotonin can upregulate or downregulate the production of ATP from glucose in the cytosol from glycolysis (Ashkenazy-Shahar and Beitner, 1997; Assouline-Cohen et al., 1998; Beitner et al., 1983; Coelho et al., 2007, 2012; Lilling and Beitner, 1990; Mansour, 1962). This process is often called aerobic glycolysis because it can take place in the presence of oxygen, even though it does not use oxygen. Oxidative phosphorylation is more efficient because it extracts more molecules of ATP from every molecule of glucose, but aerobic
glycolysis is rapid and generates ATP at a faster rate than oxidative phosphorylation (Pfeiffer et al., 2001). In addition to being faster, glycolysis may produce less reactive oxygen species that can harm the cell or the mitochondrion (Brand and Hermisse, 1997). In the brain, aerobic glycolysis involves the breakdown of glycogen stored in astrocytes, which then transport the end product (lactate) to neurons that preferentially use it as a fuel source (Magistretti and Ransom, 2002). In astrocytes, serotonin regulates aerobic glycolysis through the 5-HT₁Ａ heteroreceptor (Uehara et al., 2006).

3.4.1.2. Blood glucose homeostasis. Serotonin has bidirectional control over glucose homeostasis in the bloodstream by regulating glucagon and insulin secretion from pancreatic cells (Adeghate et al., 1999; Coulie et al., 1998; Sugimoto et al., 1996; Yamada and Sugimoto, 2000; Yamada et al., 1995).

3.4.1.3. Lipid storage and metabolism. Serotonin also has bidirectional control over the homeostatic regulation of stores of body fat through the leptin signaling pathways involved in lipid metabolism (Donovan and Tecott, 2013).

3.4.1.4. The vascular system. Serotonin also exerts control over the vascular system. While mainly known for its vasoconstrictive properties, serotonin is also a vasodilator (Cohen et al., 1996), which gives it bidirectional control over the distribution of energetic resources. Serotonin also regulates vascular networks in plants (Kang et al., 2007, 2009), and future research should test whether serotonin has a similar function in fungal hyphae.

3.4.1.5. Neuronal activity. Neurons are major consumers of energy in the brain, and serotonin exerts complex bidirectional effects on neuronal growth, differentiation, and death (Azmitia, 2001). Moreover, inhibitory and excitatory serotonin receptors are often co-expressed on cholinergic, glutamatergic, GABAergic, dopaminergic, and motor neurons, so serotonin also has bidirectional control over neuronal activity (Fink and Gothert, 2007; Puig and Gulledge, 2011).

3.4.1.6. Organ function. Many organs have large energetic demands, and serotonin is either produced or taken up from the bloodstream by every major organ in the body (Table 5). Indeed, the uptake of serotonin in lung tissue, platelet cells, and chromaffin granules of the adrenal medulla is positively correlated with the level of ATP production in those tissues (Bankston and Guidotti, 1996; Born and Gillson, 1959; Fisher et al., 1974).

3.4.1.7. Metabolically expensive processes. Serotonin also controls the expenditure of energy by regulating metabolically expensive processes—growth, development, reproduction, immune function, and the stress response (Azmitia, 2007), probably by affecting hypothalamic function. The hypothalamus regulates the timing and coordination of these processes (Chrrousos, 2009; Cyr and Eales, 1996; Sower et al., 2009; Tsang et al., 2014; Yang, 2010), and it contains some of the highest concentrations of serotonin in the brain (Bogdanski et al., 1957; Brown et al., 1979; Paasonen et al., 1957).

Important metabolic processes are disturbed when serotonin transmission is disrupted. For instance, monoamine transmission to the hypothalamus is completely inhibited in REM sleep (Parmeggiani, 2011). During this time, important physiological parameters also become less regulated—blood pressure, heart rate, breathing and body temperature (Parmeggiani, 2011). Despite this, the brain’s total energy consumption during REM sleep is nearly the same level as during the awake state (Buchsbaum et al., 1989; Madsen et al., 1991). Similarly, Kanarik and colleagues have found that serotonergic lesions induced by the neurotoxin parachloroamphetamine trigger a compensatory response 28 days later in which cytochrome oxidase expression was increased in multiple regions of the rat brain (Kanarik, 2011; Kanarik et al., 2008). Together, both lines of evidence suggest serotonin increases the energetic efficiency of metabolic processes.

3.4.2. The homeostatic equilibrium level of serotonin transmission is increased in situations requiring a rebalancing of metabolically expensive processes.

Based on the foregoing, we propose that the homeostatic equilibrium level of serotonin transmission increases in situations that require a shift in the balance of metabolically expensive processes to adaptively respond to environmental contingencies. The hypothalamus should be a common site of increased transmission due to its role in coordinating these processes.

In a recent study, muscle glycogen levels were depleted by 82–90% in adult male rats during exhaustive exercise, while brain glycogen levels decreased by 50–64%. During recovery, glycogen reserves were replenished through a supercompensatory response (Matsui et al., 2012). Interestingly, during exercise there is an increase in serotonin transmission to the hypothalamus and other brain regions (Blomstrand, 2011). Another study found that serotonin levels in the lateral hypothalamus increase during exercise and return to baseline during recovery (Smriga et al., 2002), which mirrors what happens to glycogen levels. Indeed, elevated serotonin levels during exercise are associated with fatigue (Blomstrand, 2011), an indicator of energetic stress. We suggest that serotonin is elevated during exercise because the fall in glycogen forces a reprioritization in energy allocation. During recovery, serotonin levels fall as glycogen is replenished and allocation patterns normalize.

The association with energetic stress is not limited to negative situations. Male rats become unresponsive to new mating opportunities for nearly two days after about 3.5 h of ad libitum copulation with successive estrous females (Mas et al., 1995). The most likely reason for the unresponsiveness is the depletion of viable sperm.
Since spermatogenesis is energetically expensive (Dowling and Simmons, 2012; Olsson et al., 1997), sperm depleted males must devote less energy to mating effort and devote more to spermatogenesis. During the period of sexual exhaustion, serotonin is elevated in the hypothalamus and returns to baseline as sexual responsiveness resumes (Hull et al. 2004; Lorrain et al., 1997; Mas et al., 1995). Consistent with the role of serotonin in rebalancing metabolically expensive processes, elevated serotonin levels in the hypothalamus promote spermatogenesis (Aragon et al., 2005; Shishkind and Dygalo, 2000) and inhibit mating behavior (Hull et al., 2004).

In short, the effects of enhanced serotonin transmission are state-dependent. Physical exhaustion, sexual exhaustion, and many other states show evidence of enhanced serotonin transmission (Table 3), yet their symptom profiles differ in important ways. Under the energy regulation hypothesis, state-dependence is expected because situational demands determine how energy should be adaptively reallocated.

State-dependence can explain some inconsistent findings. Homberg and colleagues have shown that the serotonergic system affects rodents’ cognitive flexibility, including reversal learning, attentional set shifting, the ability to form and update representations of stimulus-reward or response-reward contingencies, the inhibition of inappropriate responses, and the ability to postpone immediate reward for a larger delayed reward (Homberg, 2012; Homberg and Lesch, 2011; Nonkes et al., 2012; Nonkes and Homberg, 2013). They argue that the serotonergic system integrates past learning with incoming information from the environment to regulate attention, focusing on the processing of stimuli most relevant to the organism’s survival and reproduction (‘vigilance behavior’). Their hypothesis is consistent with a larger body of evidence implicating the serotonergic system in learning and memory systems (Altman and Normile, 1988; Cassel, 2010). However, the direction of association is unclear, with some studies reporting a positive association between cognitive flexibility and serotonin transmission, and other studies reporting a negative association (Altman and Normile, 1988; Cassel, 2010; Homberg, 2012). The bidirectional findings are explicable by the hypothesis that the serotonergic system is part of the adaptive energy-regulation machinery that balances cognition with other metabolically expensive processes—growth, maintenance, immune function, reproduction—as the situation demands. In Section 5, we discuss how serotonin coordinates the cognitive changes that take place in melancholia.

4. The homeostatic response to SSRIs and symptom reduction

In this section, we argue that depressive symptoms are reduced over several weeks of SSRI treatment, not by their direct pharmacological properties, but due to the compensatory responses of the brain attempting to restore energy homeostasis.

4.1. Acute SSRI treatment disrupts energy homeostasis

Fig. 3D. Extracellular serotonin levels in the hippocampus of BALB/c mice exposed to plain drinking water (control) or fluoxetine (floux) in their drinking water (floux) for 28 days. By 28 days, fluoxetine exposed rats were statistically indistinguishable from control rats. Reprinted with permission from Popa et al. (2010).

The total content of serotonin in the brain is composed of the intracellular pool and the extracellular pool. With acute SSRI treatment, SERT blockade prevents reuptake from the synapse, increasing extracellular serotonin within minutes to hours of administration (Bymaster et al., 2002; Rutter and Auerbach, 1993). Put another way, the distribution of serotonin is rapidly shifted to the extracellular pool, and extracellular levels are perturbed from their homeostatic equilibrium (Fig. 3C).

The increase in extracellular serotonin causes corresponding disruptions to energy homeostasis. In rodents, acute SSRI treatment has been shown to increase glutamatergic activity in the rodent prefrontal cortex (Fu et al., 2012), promote glycolytic activity in the hippocampus (Webhofer et al., 2013), inhibit oxidative phosphorylation in liver and brain mitochondria (Curti et al., 1999; Souza et al., 1994), and inhibit the consumption of blood-borne glucose throughout the brain (Freo et al., 2000).

4.2. The brain’s compensatory responses to SSRI treatment

The brain attempts to restore energy homeostasis through a number of compensatory responses. These compensatory responses take several weeks to develop, which could make them important in the therapeutic delay. One such change is a decline in extracellular serotonin during chronic SSRI treatment that eventually comes back to the premedication equilibrium (Fig. 4) (Popa et al., 2010; Smith et al., 2000). This decline is due to the fact that all ADM classes inhibit the synthesis of serotonin (Bosker et al., 2010; Esteban et al., 1999; Honig et al., 2009; Moret and Briley, 1996; Muck-Seler et al., 1996; Siesser et al., 2013; Yamane et al., 1999, 2001). Over chronic treatment, the cumulative effects of the inhibition of synthesis cause total (intracellular + extracellular) serotonin levels to decline (Fig. 5) (Bosker et al., 2010; Honig et al., 2009; Marsteller et al., 2007; Siesser et al., 2013).

Over several weeks of ADM treatment, the 5-HT1A heteroreceptor also becomes tonically activated in many forebrain regions (Fig. 3D) (Beck et al., 1997; de Bortoli et al., 2006, 2013; Elena Castro et al., 2003; Jongsma et al., 2006; Lopez et al., 1998; Shen et al., 2002; Vicente and Zangrossi Jr, 2014; Welner et al., 1989; Zanoveli et al., 2005, 2007, 2010). This is a postsynaptic effect, so it is not easily explained as an attempt to restore serotonin homeostasis. This is more readily explained as a compensatory response to the disruptions in the allocation of energy caused by acute treatment. Specifically, most cortical neurons are glutamatergic, so activation of the 5-HT1A heteroreceptor, which is inhibitory, counters the stimulatory effect of serotonin on glutamatergic neurons induced by acute SSRI treatment (Fu et al., 2012). The gradual decline in extracellular serotonin from peak value also helps reverse SSRI-stimulated glutamatergic activity in cortical regions (Fig. 3D). These alterations, and possibly others, help restore energy homeostasis after perturbation by SSRI treatment. Indeed, while acute SSRI treatment increases glutamatergic activity in rodent models of depression (Fu et al., 2012), chronic treatment decreases it (Bonanno et al., 2005; Mallei et al., 2011; Musazzi et al., 2010).

This pattern, in which acute and chronic SSRI treatments have opposing phenotypic effects, is a fairly common phenomenon. ADMs of all major classes reduce aggression in rodents during acute treatment, but increase aggression over chronic treatment
(Mitchell, 2005). In healthy volunteers, a single dose of the SSRI citalopram potentiates anxiety, while chronic treatment inhibits it (Grillon et al., 2007, 2009). Similarly, acute and chronic paroxetine treatments exert diametrically opposing effects on the excitability of motor cortex (Gerdelat-Mas et al., 2005; Loubinoux et al., 2002). Acute SSRI treatment stabilizes microtubule structure and potentiates the hippocampal-PFC synapse, while the opposite effects are seen over chronic treatment (Bianchi et al., 2009; Cai et al., 2013). BDNF signaling is decreased with acute SSRI treatment, and chronic treatment increases it (De Foubert et al., 2004; Khundakar and Zetterström, 2006).

The opposing effects are theoretically important because the acute effects are more likely to be due to the direct pharmacological properties of these drugs. That acute SSRI treatment has widespread phenotypic effects is further evidence that they disrupt energy homeostasis. Conversely, the opposing effects that occur over chronic treatment are more likely to be due to the brain’s compensatory responses that attempt to restore homeostasis.

The opposing effects are difficult for the phenotypic plasticity hypothesis to explain. As it is currently described (Branchi, 2011), there is no reason to predict that chronic SSRI treatment should reverse the phenotypic effects of acute treatment. Rather, the most obvious prediction is that chronic treatment will exacerbate the effects of acute treatment, simply because phenotypic changes have more time to develop.

4.3. The mechanisms of symptom reduction

We hypothesize that it is the brain’s compensatory responses to SSRI treatment, rather than the direct pharmacological properties of SSRIs, that are responsible for reducing depressive symptoms. Others have suggested the symptom-reducing effects of SSRIs are attributable to the brain’s attempts to re-establish homeostasis (Hyman and Nestler, 1996). We differ slightly in that we propose that the brain is attempting to restore energy homeostasis rather than serotonin homeostasis. The return of extracellular serotonin to equilibrium conditions is only one component of the homeostatic response to the energy dysregulation caused by SSRI treatment.

If our hypothesis is correct, SSRIs (and perhaps other ADMs) could have opposing effects on depressive symptoms during acute and chronic treatment. Efficacy studies usually do not report the relative effect of ADMs over placebo on depressive symptoms during the early stages of treatment. However, anecdotal evidence suggests that symptoms often worsen before they get better (Haslam et al., 2004). The anecdotal evidence is supported by two pertinent studies. In one placebo-controlled study, imipramine was less effective than placebo during the first week of treatment (Oswald et al., 1972). Imipramine only outperformed placebo over several weeks of treatment. In another study, 30.4% of participants experienced a worsening of depressive symptoms (defined as an increase of five points or more on the Hamilton Depression Research Scale: HDRS) within the first weeks of fluoxetine treatment (Cusin et al., 2007). This is perhaps a surprising finding given the large placebo effect in depression (Kirsch et al., 2008), which could obscure any pharmacological effects that increase symptoms. Moreover, the requirement that the increase be at least five HDRS points is stringent since antidepressant drugs must only reduce symptoms by three HDRS points more than placebo to be deemed clinically significant in the United Kingdom (Excellence, 2004). Indeed, since an increase in depressive symptoms is likely to have a Poisson distribution, the proportion of participants who experienced any increase in symptoms during early treatment is likely to have been much higher.

The initial worsening of symptoms is theoretically important because this is when the largest increases in extracellular serotonin occur (Fig. 4). It is only over several weeks of treatment that depressive symptoms reduce, during which the trajectory of extracellular serotonin is declining from its peak value (Fig. 4). That the therapeutic delay of ADMs might be related to the downward trajectory in serotonin has been noted by other authors. In a study involving Flinders Sensitive Line rats, the symptom-reducing effects of chronic desipramine administration were associated with a reduction in total (intracellular + extracellular) serotonin content in PFC, hippocampus, and nucleus accumbens. The authors suggested that “decreasing 5-HT levels in limbic regions is important for the therapeutic effect of antidepressants” (Zangen et al., 1997, p. 2482). Similarly, in a primate microdialysis study, extracellular serotonin levels in the hippocampus and other brain regions gradually returned to baseline over chronic treatment with fluoxetine. The authors suggested that the brain’s compensatory responses “may contribute to the therapeutic actions of this drug in human depression” (Smith et al., 2000, p. 470).

In short, the upward trajectory in serotonin during initial ADM treatment is often associated with a worsening of symptoms, while the downward trajectory over chronic treatment is associated with symptom reduction. This pattern can be explained by the energy regulation hypothesis. The acute (direct) effects of SSRI treatment disrupt energy homeostasis by exacerbating glutamatergic activity in frontal brain regions, which, according to the glutamate hypothesis (Popoli et al., 2012), should worsen symptoms. The brain develops compensatory responses over chronic treatment that reverse the energy disruptions and reduce symptoms. Specifically, both the reduction in the synthesis of serotonin and the tonic activation of the 5-HT1A heteroreceptor act to reverse the elevated glutamatergic activity induced by the direct effects of SSRI treatment. If the 5-HT1A heteroreceptor is still activated as extracellular serotonin returns to baseline over chronic treatment, glutamatergic activity would fall below equilibrium conditions (Fig. 3D), producing an actual antidepressant effect. We therefore explain the symptom reducing effects of ADMs as due to the brain’s attempts to restore energy homeostasis. Alterations to the serotonergic system are needed to accomplish this, but these alterations cannot all be explained in terms of restoring serotonin homeostasis.

Fig. 5. Total (intracellular + extracellular) serotonin content in different brain tissues declines with chronic citalopram treatment. Gray bars represent 15 days of citalopram treatment (50 mg/ml) plus 2 days of washout. White bars represent 17 days of citalopram treatment (50 mg/ml). Black bars represent chronic saline treatment. Acad = anterior cingulate cortex; NAc = nucleus accumbens; CP = caudate/putamen; dHC = dorsal hippocampus; VHc = ventral hippocampus; Amy = amygdala; PVN = paraventricular nucleus of the hypothalamus; DRN = dorsal raphe nucleus; MRN = median raphe nucleus. Reprinted with permission from Bosker et al. (2010).
4.4. Symptom reduction is a temporary overshoot of the homeostatic equilibrium

When a homeostatic mechanism is perturbed, it often exhibits a dampened oscillation around its equilibrium, as in the case of a spring that is released from a compressed position. We suggest this is what is happening over the course of SSRI treatment. Acute treatment often causes a worsening of symptoms relative to the premedicated state. Over chronic treatment (several weeks) symptoms are alleviated relative to the premedicated state, and symptoms return to the premedicated baseline over more prolonged treatment periods. Indeed, ADM users commonly experience relapses over months or years of treatment (Byrne and Rothschild, 1998). In one study of fluoxetine, the relapse rate was 35% at six months and 46% at 12 months (McGrath et al., 2006). In another study, the relapse rate over two years of continuous ADM treatment was 60% (Bockting et al., 2008).

4.5. The effects of SSRIs during recalibration of serotonin transmission

Homeostasis requires the brain to produce compensatory responses to interventions that perturb serotonin from equilibrium (e.g., SSRIs). However, what happens when SSRIs are initiated while the equilibrium is changing? Under those conditions, the brain may not interpret the synaptic-enhancing effects of SSRIs as a perturbation, but rather as part of the recalibration of equilibrium, and the brain may not produce compensatory responses. If so, then acute and chronic SSRI treatment may produce similar phenotypic effects.

A recent study allows the opportunity to compare the phenotypic effects of SSRIs when they are initiated during environmental change (i.e., a possible recalibration period) or during a continuation of the same environmental conditions (Branchi et al., 2013). Rats were randomly assigned to a sequence of conditions involving chronic mild stress (CMS) and/or an enriched environment (EE) and exposed to the SSRI fluoxetine for three weeks. The authors reported that fluoxetine exerted greater effects on anhedonia (assessed by changes in the preference for saccharine) and hippocampal BDNF signaling when initiated as the rats’ environment changed (i.e., during the transition from EE to CMS or from CMS to EE) than when initiated during a continuation of the same environment (i.e., EE to EE, CMS to CMS). Importantly, acute (1–3 days) and chronic (>3 days) fluoxetine treatment did not affect anhedonia differently when initiated concurrently with environmental change (EE to CMS, or CMS to EE).

Branchi and colleagues interpret their findings in terms of the phenotypic plasticity hypothesis, which is reasonable because the rats were exposed to an altered environment that required a response (i.e., a phenotypic change). However, we have argued that the energy regulation hypothesis more accurately describes serotonin’s unique effects because all biochemistries are involved in phenotypic plasticity or homeostasis. Interestingly, fluoxetine produced significant changes in corticosterone in all four conditions (EE to EE, EE to CMS, CMS to CMS, EE to CMS to CMS) regardless of whether it was administered during a constant or a changing environment. Branchi and colleagues argue that corticosterone “is more sensitive than anhedonia and BDNF to the effects of the combination of the drug and the environment, being altered even after a period of habituation” (p. 6). This finding is more naturally explained by the energy regulation hypothesis since corticosterone is involved in mobilizing energetic resources.

5. What is serotonin doing in melancholia?

Since the effects of serotonin are state-dependent, we demonstrate the utility of our hypotheses in explaining what happens in the melancholic state. In melancholia, the symptoms reflect a trade-off in which energy is reallocated toward cognition at the expense of growth and reproduction. We suggest that the elevation in serotonin transmission coordinates this trade-off and helps explain many of the symptoms of melancholia.

5.1. Energy is reallocated to cognition in melancholia

The fact that melancholia is highly associated with sustained activation of the HPA axis (Taylor and Fink, 2008) indicates that melancholia is energetically expensive. One may wonder what this energy is used for since growth and sexual activity are generally inhibited (Taylor and Fink, 2008). We may gain insight into this question by considering the symptoms of melancholia and other depressive syndromes (Table 4). Of particular interest is the comparison between melancholia and sickness behavior because they share a great many symptoms in common as well as genes and neurobiology. This similarity has led some to suggest depression is a dysregulation in the immune response (Dantzer et al., 2008; Maier and Watkins, 1998; Wager-Smith and Markou, 2011) or an adaptive response to social stressors that predict the risk of infection (Raison and Miller, 2013). However, these hypotheses are difficult to reconcile with the few symptomatic differences that exist between the phenotypes (Andrews and Durisiko, in press). Cognition is generally impaired in sick organisms, and they spend more of their time in slow wave sleep (Dantzer, 2001; Larson and Dunn, 2001). Conversely, melancholia is associated with an increase in ruminination (Jackson, 1983; Nelson and Mazure, 1985) and rapid eye movement (REM) sleep (Steiger and Kimura, 2010; Taylor and Fink, 2008).

These differences suggest that cognition is altered in melancholia (Andrews and Durisiko, in press). Ruminantia refers to persistent, distraction-resistant thoughts about the problems associated with the episode (Andrews and Thomson, 2009). Ruminantia involves an analytical processing style in which complex problems are broken into smaller, more manageable components, which are then sequentially studied (Andrews and Thomson, 2009; Barbic et al., 2014). To keep track of the components of a problem, analysis requires working memory, a memory system in which information is kept active because it is useful in ongoing processing (Baddeley, 2007). As working memory load increases, tasks become more vulnerable to interruption because it is easier for irrelevant stimuli to displace task-relevant information (Kane and Engle, 2002). The distraction-resistant nature of ruminantia may promote analysis by reducing the vulnerability to interruption (Andrews and Thomson, 2009). Finally, the increase in REM sleep also points to complex information processing, because REM sleep promotes consolidation of hippocampal memory representations encoding complex information (Rasch and Born, 2013).

The symptomatic differences between melancholia and sickness behavior suggest that the melancholic brain has been primed by evolution to process complex information (Andrews and Durisiko, in press; Andrews and Thomson, 2009). Indeed, the cognitive aspect of melancholia may be phylogenetically ancien (Andrews and Durisiko, in press). Of the non-human animal models of depression, inescapable shock exhibits the closest symptomatic correspondence with human melancholia (Table 4). This correspondence extends to a complex information processing style that may be analogous to analytical ruminantia. Early research suggested that inescapable shock led to a cognitively helpless state—the learned helplessness hypothesis (Seligman, 1975). However, this hypothesis has been refuted (Shors, 2004). Rather, the uncontrollably stressed organism gives up on tactics it has learned
are futile and searches its environment for alternatives (Lee and Maier, 1988; Minor et al., 1984). When given a task in which cues in its environment are relevant, uncontrollably stressed organisms often outperform control groups (Lee and Maier, 1988; Minor et al., 1984; Rodd et al., 1997). In one study, a complex pattern of behavior led researchers to suggest that inescapably shocked rats may be “more prone to test hypotheses” about their control over the environment (Minor et al., 1984, p. 553).

One indicator that the melancholic cognitive pattern is metabolically expensive comes from the fact that, in melancholic patients and rodent models, glutamatergic activity is elevated and GABAergic activity is reduced in many brain regions (Alcaro et al., 2010; Petty and Sherman, 1981; Sanacora et al., 2004, 2012). Since glutamate is the most excitatory neurotransmitter in the brain, and GABA is the most inhibitory, this pattern suggests increased energy consumption. In neuroimaging studies, depressed patients show higher resting activity in the default mode network (Kühn and Gallinat, 2013), which is related to social cognition (Lieberman, 2013). They also show higher resting activity and increased functional connectivity in the network involved in self-reflection and attentionally demanding tasks (Schilbach et al., 2014; Zhou et al., 2010). These two networks are rarely co-activated (Lieberman, 2013), and this is probably an energetically expensive state.

Another clue that melancholia is energetically expensive is the fact that it is supported by glycolysis. Rodent models of depression show increased expression of glycolytic genes and an increase in glycolytic metabolism (Mallei et al., 2011; Uehara et al., 2006, 2007).

An increased reliance on aerobic glycolysis to support ruminating explains some puzzling findings in neuroimaging studies of melancholic and depressed patients. PET studies using [18F]fluorodeoxyglucose (FDG) tend to show hypoactivity in the DLPCF and other frontal regions (Fitzgerald et al., 2008). This might seem inconsistent with an increase in energy devoted to cognition, since the DLPCF is associated with working memory and executive function (Nee et al., 2013). However, aerobic glycolysis can dramatically alter the interpretation of neuroimaging signals (Dienel, 2012; Pellerin et al., 2007). Studies using [18F]-FDG PET may underestimate energy consumption by 50% or more because astrocytes readily incorporate labeled glucose, and much of the labeled lactate produced by aerobic glycolysis diffuses into the bloodstream and escapes the brain (Dienel, 2012). All brain regions utilize a mixture of aerobic glycolysis and oxidative phosphorylation, but the highest resting rates of aerobic glycolysis occur in the DLPCF and VLPCF, while the amygdala and hippocampus have relatively low resting rates (Goyal et al., 2014; Vaishnavi et al., 2010). Thus, the lower [18F]-FDG PET signal in the DLPCF of depressed patients probably reflects an increase in aerobic glycolysis, and a decrease in oxidative phosphorylation, rather than actual hypoactivity. Indeed, two neuroimaging studies (Drevets et al., 1992; Dunn et al., 2005) show patterns suggesting that aerobic glycolysis is positively associated with depressive symptoms in patients diagnosed with major depression (Figs. 6 and 7).

An increase in aerobic glycolysis also provides an elegant explanation for why c-Fos and cytochrome oxidase c expression are reduced in rodent models of depression, especially in regions associated with cognition and the stress response (Kanarik et al., 2011; Shumake et al., 2000, 2001, 2002, 2003; Stone et al., 2007). Cytochrome oxidase c and c-Fos expression are both indices of oxidative phosphorylation, which is reduced as glucose is metabolized through aerobic glycolysis.

It seems likely that the increase in aerobic glycolysis in melancholia supports changes in synaptic connectivity. Aerobic glycolysis is associated with increased expression of genes involved in promoting synaptic plasticity (Goyal et al., 2014). This finding is consistent with evidence that dendritic spines, which regulate synaptic strength, contain glycolytic enzymes and the transporters for lactate (Pierre et al., 2009; Wu et al., 1997), yet rarely contain mitochondria (Li et al., 2004; Sheng and Hoogenraad, 2007).

In depressed patients, there is abundant evidence of altered connectivity in cortico-limbic structures, including the hippocampus and lateral PFC (Schilbach et al., 2014; Sheline et al., 2010; Steffens et al., 2011; Zeng et al., 2012; Zhou et al., 2010). Rodent models of melancholia confirm changes in hippocampal plasticity (Cai et al., 2013; Kim et al., 1996; Shors et al., 1989; Xu et al., 1997).

In summary, the melancholic brain appears to be reconfiguring to learn solutions to complex problems. The processes involved in this learning appear to be so energetically expensive that growth and reproduction are downregulated. We argue that the processes involved in making these trade-offs are coordinated by serotonin. We first discuss how serotonin is triggered in melancholia.

**Fig. 6.** Indirect evidence of aerobic glycolysis in major depression. The sample consists of unmedicated patients meeting diagnostic criteria for major depression with the melancholic subtype (black squares) and two samples of non-depressed controls (open circle and open square). The y-axis is the [15O]-water PET signal in the left VLPCF, and the x-axis is the score on the Hamilton Depression Rating Scale. Relative to the control groups, the [15O]-water PET signal is elevated in the melancholic patients, but the signal declines with the severity of depressive symptoms. Since aerobic glycolysis does not use oxygen, the pattern suggests that, among the melancholic patients, aerobic glycolysis increases with symptoms. The error bars on the control groups represent the standard error of the means. Reprinted with permission from Drevets et al. (1992).

**Fig. 7.** Further indirect evidence that aerobic glycolysis increases with depressive symptoms. The sample consists of eight unmedicated patients meeting DSM-IV criteria for major depression. The y-axis represents the degree to which cerebral blood flow (assessed by [15O]-water PET) is correlated with glucose consumption (assessed by [18F]-FDG PET). The x-axis represents severity of depressive symptoms as measured by the Hamilton Depression Rating Scale. Since a decline in the ratio of oxygen consumption to glucose consumption is considered evidence of aerobic glycolysis (Shulman et al., 2001a), the negative slope suggests that aerobic glycolysis increases with depressive symptoms. Reprinted with permission from Dunn et al. (2005).
5.2. The situational triggers of the melancholic state

While we have little understanding of the stressors that trigger melancholia in humans, there is greater understanding of how stressors affect the serotonergic system in rodent models of melancholia. The ventral region of the rat medial PFC (mPFCv) regulates depressive symptoms in response to inescapable shock and chronic social defeat (Amat et al., 2005; Covington et al., 2010). Inescapable shock triggers an increase in the homeostatic equilibrium of serotonin transmission (Section 2.2.1.3), which is mechanistically achieved by a glutamatergic projection from the mPFCv that synapses with GABA interneurons in the DRN (Varela et al., 2012). When the shock is escapable, this projection is activated, GABA interneurons limit the transmission of serotonin to the mPFCv to a brief pulse, and no depressive symptoms are triggered. When the shock is inescapable, the projection is not activated, the transmission of serotonin is sustained (i.e., the homeostatic equilibrium has been elevated), and depressive symptoms are triggered (Amat et al., 2005).

Inescapable shock may seem like a situation where sustained analysis is fruitless, because nothing can be done to avoid the shock. But this overlooks an obvious informational asymmetry: while the researcher knows that the rat cannot escape, the rat may not. In the inescapable shock paradigm, the rat learns that struggling or pressing a lever does not help it to escape the shock. The rat turns its attention to external cues, possibly evaluating the environment to better understand the situation (Lee and Maier, 1988; Minor et al., 1984). Similarly, in humans, depression is associated with a loss of control (Edwards and Weary, 1998; Jacobson et al., 1999; Lyubomirsky et al., 1999), and attention is directed toward regaining control (Weary et al., 1993).

5.3. Serotonin coordinates the mechanisms promoting rumination

Analytical rumination involves a number of sub-processes (Andrews and Durisko, in press; Andrews and Thomson, 2009; Barbic et al., 2014): (1) attention is oriented toward threats or problems; (2) interest in stimuli unrelated to the problem is reduced (anhedonia); (3) other energetically expensive activities are down-regulated; (4) working memory is allocated to the problem; and (5) working memory processes are less likely to be disrupted by problem-irrelevant stimuli (distraction-resistance). We discuss each of these processes, and how serotonin is involved in coordinating them in rodent models of melancholia.

5.3.1. The amygdala and orienting attention to the problem that triggered the episode

The orientation of attention to problems that triggered the episode likely involves heightened amygdala activity, which is consistently found in depressed patients (Whalen et al., 2002). The amygdala is associated with emotionally salient stimuli, but it is more generally involved in evaluating the environment with respect to important goals and orienting the brain toward situations for which additional information would be useful (Whalen et al., 2007). Thus, heightened serotonin transmission to the amygdala in rodent models of melancholia (Amat et al., 1998a; Blanchard et al., 1993) probably plays a role in directing attention to problems that the organism is facing.

5.3.2. The nucleus accumbens and anhedonia

Sustained analytical processing of complex problems may not offer immediate reward because they are not easy to solve. Attending to activities that do offer more of an immediate reward (e.g., eating, sex, companionship) would therefore tend to interfere with attempts to problem solve. Elsewhere, we have argued that anhedonia promotes uninterrupted analytical rumination by reducing the motivation to engage in hedonic activities (Andrews and Thomson, 2009). The nucleus accumbens regulates anhedonia and other motivation-related behaviors, and dopamine transmission is thought to play an important role in this region (Der-Avakian and Markou, 2012; Salamone et al., 2005). In the inescapable shock and the Flanders Sensitive Line models, indices of serotonin transmission are elevated in the nucleus accumbens, while dopamine transmission is reduced (Bland et al., 2003b; Zangen et al., 1997). Moreover, serotonin plays a causal role in anhedonia by regulating dopamine levels (Zangen et al., 2001).

5.3.3. The hypothalamus reallocates energy to rumination

The hypothalamus affects major metabolic processes, including growth, development, reproduction, immune function, and the stress response (Chrousos, 2009; Cyr and Eales, 1996; Sower et al., 2009; Tsang et al., 2014; Yang, 2010), and indices of serotonin transmission to the hypothalamus are elevated in rodent models of melancholia (Adell et al., 1988; Beita et al., 2005; Blanchard et al., 1993; Keeney et al., 2006). Chronic activation of the HPA axis suppresses both the growth and reproductive axes (Chrousos, 1998), so elevated serotonin transmission to this region probably plays a role in inhibiting appetite and sexual activity in melancholia. The hypothalamus also undoubtedly plays a role in the alterations in sleeping patterns, such as the increase in REM sleep.

5.3.4. The hippocampus and the allocation of working memory

The working memory system is an important function of the lateral PFC (Courtney et al., 1998). Amongst its many afferent projections, this region receives ipsilateral projections from the ventral CA1 region of the hippocampus (Cavada et al., 1983; Rose and Vanhoese, 1977). This synapse, henceforth referred to as the HC-PFC synapse, is susceptible to various forms of plasticity, including long-term potentiation (LTP) and long-term depression (LTD) (Hirsch and Crepel, 1990; Laroche et al., 1999).

Changes in the plasticity of the HC-PFC synapse subserve the working memory system (Burette et al., 2000; Laroche et al., 2000). For instance, several studies have found that LTD in this synapse is associated with improved performance on working memory tasks (Burette et al., 2000; Laroche et al., 2000). A broad reduction in the strength of this synapse, which may be accompanied by potentiation of a subset of hippocampal inputs to the PFC, may enhance the signal to noise ratio of transmitted information (Laroche et al., 2000), improving the transmission of task-relevant signals while reducing the transmission of distracting sensory-driven activity. Moreover, in tasks that do not require the long-term storage of newly acquired information, synaptic depression may also represent an ‘anti-consolidation’ signal. Since the putative function of analytical rumination is to learn how to manage stressors, long-term memory consolidation (mediated by hippocampal LTP) would seem most appropriate after processing is completed. Thus, LTP should be inhibited during ongoing analysis.

The synaptic changes in the hippocampus during depressed states are consistent with this working memory model of analytical rumination. First, LTD is generally upregulated in the CA1 region of the hippocampus in the inescapable shock model of depression, while LTP is inhibited (Kim et al., 1996; Shors et al., 1989; Xu et al., 1997). Moreover, a subset of synapses undergoes a more transient form of potentiation. In the chronic mild stress model of depression, the temporoammonic–CA1 pathway is selectively potentiated (Cai et al., 2013). Since potentiation of the temporoammonic–CA1 synapse promotes hippocampal output from the CA1, this mechanism may enable the transmission of pertinent signals from the hippocampus to working memory areas of the PFC. Thus, hippocampal LTD, along with potentiation of a subset of synapses, may enhance the signal-to-noise ratio of information transmitted to working memory areas of the PFC. This may promote...
analytical rumination by maintaining problem-relevant information in working memory.

Serotonin transmission to the hippocampus is elevated in many rodent models (Section 2.2). In these models, serotonin appears to coordinate the changes in synaptic plasticity. In chronic mild stress, serotonin induces a transient potentiation in the hippocampus and mPFC (Cai et al., 2013). Since this form of potentiation competes with the same pathways as LTP (Cai et al., 2013), LTP is reduced as well. Future research should test the generality of these findings in other rodent models of depression, particularly inescapable shock. Although the precise mechanism for triggering hippocampal LTD is unknown, we predict that it is triggered by a sustained elevation in serotonin transmission. For instance, elevated serotonin can trigger LTD in motor neurons (Garraway and Hochman, 2001b).

5.3.4.1. An energetic trade-off between synaptic plasticity and neurogenesis

An ongoing area of research attempts to understand hippocampal volume shrinkage during depression (Groves, 2007; Krishnan and Nestler, 2008). This seems to involve a reduction in brain-derived neurotrophic factor (BDNF), which promotes neurogenesis. For instance, humans and mice with a genetic variant that leads to lower expression of BDNF have smaller hippocampal volumes (Chen et al., 2006).

We suggest hippocampal neurogenesis is downregulated in melancholia to give energetic priority to the synaptic changes that underlie analytical rumination. A provocative study found that hippocampal neurogenesis in adult bonnet macaques is negatively associated with overall body weight (Perera et al., 2011b), suggesting that neurogenesis may be so metabolically expensive that it comes at the expense of the growth of other tissues. The most energetically expensive aspect of neurogenesis may be the hypoplastic phase where neurons undergo heightened LTP (Ge et al., 2007). Since the changes in synaptic plasticity that underlie analytical rumination are also energetically expensive (Section 5.1), it may be difficult for the brain to energetically support both neurogenesis and synaptic plasticity simultaneously. BDNF signaling tends to be inhibited in severe (but not mild) stressor models of depression (Bland et al., 2007; Larsen et al., 2010), which suggests that there is an increasing trade-off as stress becomes more severe. Since LTD involves a shrinkage of dendritic spines and elimination of synapses (Sheng and Ettürk, 2014; Zhou et al., 2004), neurogenesis may be mechanistically incompatible with the synaptic changes that underlie analytical rumination. Indeed, BDNF has an inhibitory effect on LTD in the hippocampus and visual cortex (Aicardi et al., 2004; Akaneya et al., 1996; Ikegaya et al., 2002; Rodrigues et al., 2014).

We further suggest that serotonin mediates this energetic trade-off. Rats lacking SERT express lower hippocampal BDNF levels, showing that BDNF signaling is under serotonergic control (Molteni et al., 2010). We predict that elevated serotonin transmission promotes the alterations in synaptic plasticity in melancholia and inhibits hippocampal neurogenesis.

5.3.5. The lateral PFC promotes distraction-resistance

To reduce interference with working memory processes, neurons in the left lateral PFC must fire continuously during the period of vulnerability to distraction (Courtney et al., 1998; D’Esposito et al., 2000; Dolcos et al., 2007; Funahashi et al., 1997; Jonides and Nee, 2006; Rao et al., 1997). However, sustained neuronal firing poses two problems, and aerobic glycolysis helps solve them.

First, the neurons need a source of energy that can support sustained firing. Neurons do not store glucose, and under many conditions they receive energy from the passive diffusion of glucose from nearby arterioles (Nehlig and Coles, 2007). When passive diffusion is the predominant means by which neurons obtain their energy, neuronal activity is positively correlated (or coupled) with blood flow. However, as neuronal activity becomes more intense and sustained, passive diffusion is not fast enough to support sustained neuronal firing (Pellerin et al., 2007; Shulman et al., 2001a,b). Instead, neurons rely increasingly on lactate produced by aerobic glycolysis in nearby astrocytes (Pellerin et al., 2007). The energy is produced at a much faster rate than is needed to support sustained neuronal firing (Shulman et al., 2001b), and the excess is dispersed through the bloodstream (Dienel, 2012).

The second problem caused by the sustained firing of left VLPFC neurons derives from the fact that the large majority (about 80%) of cortical neurons are glutamatergic (Somogyi et al., 1998). High levels of synaptic glutamate are toxic and can trigger apoptosis (Hara and Snyder, 2007). Thus, although sustained neuronal activity in the left VLPFC is crucial to prevent disruption of analysis, the accumulation of synaptic glutamate can trigger neuronal death. This problem is also solved by astrocytes, which clear glutamate from the synapse, convert it to a less toxic form (glutamine), and finally transport it back to the neuron where it is recycled by the neuron (Magistretti and Ransom, 2002). The energy needed to support glutamate-glutamine cycling probably comes from aerobic glycolysis (Magistretti and Ransom, 2002).

Aerobic glycolysis also has ramifications for blood flow patterns. Since aerobic glycolysis produces excess lactate and does not use oxygen (Shulman et al., 2001b), maintaining a tight coupling between neuronal activity and cerebral blood flow would be highly inefficient, since blood borne glucose and oxygen could be utilized elsewhere. Consequently, neuronal activity becomes decoupled from regional blood flow under conditions of prolonged, intense neuronal activity (Shulman et al., 2001a,b). Here again, astrocytes are mechanistically involved. They have processes that connect to nearby capillaries, and they are able to affect local blood flow (Magistretti and Ransom, 2002).

We examine how these processes are regulated by serotonin in rodents. However, the rat brain is packaged differently from the primate brain (Uylings et al., 2003). The rat mPFCv is the likely homologue to the human lateral PFC (Brown and Bowman, 2002; Kesner, 2000; Uylings et al., 2003). Indeed, the rat mPFCv is involved in regulating the resistance to distracting stimuli (Gisquet-Verrier and Delatour, 2006), and it regulates depressive symptoms in inescapable shock and chronic social defeat (Amat et al., 2005; Covington et al., 2010). Inescapable shock triggers elevated serotonin transmission to the mPFCv (Amat et al., 2005).

5.3.5.1. Serotonin and glutamatergic activity in the rodent mPFCv

Serotonin clearly regulates glutamatergic activity in the mPFCv (Puig and Gudelsky, 2011). For our purposes, the crucial issue is whether glutamatergic activity increases or decreases with a sustained elevation in serotonin transmission. In general, serotonin has a net excitatory effect on networks of cortical neurons, and this effect is mediated by activation of 5-HT1A receptors (Puig and Gudelsky, 2011). Serotonin produces a similar effect in motor neurons (Garraway and Hochman, 2001a; Harvey et al., 2006a,b; Liu et al., 2011). We predict that a sustained elevation in serotonin transmission triggers a tonic increase in mPFCv glutamatergic activity.

5.3.5.2. Serotonin and aerobic glycolysis

Serotonin stimulates aerobic glycolysis in various regions of the rodent brain (Darvesh and Gudelsky, 2003). Inescapable shock triggers aerobic glycolysis in the mPFCv, and this effect is mediated through the 5-HT1A heteroreceptor (Uehara et al., 2006).

5.3.5.3. Serotonin and glutamate clearance

We know of no direct tests of how serotonin affects glutamate clearance from the synapse, so the role of serotonin in glutamate-glutamine cycling is unclear.
5.3.5.4. Serotonin in vasoconstriction. Serotonin promotes vasoconstriction and vasodilation throughout the brain and the periphery, often mediated by the 5-HT$_{1A}$ heteroreceptor (Cohen et al., 1996).

5.4. The effects of ADMs on the melancholic energy allocation pattern

As we have articulated the problem, antidepressants must reverse the melancholic energy allocation pattern to reduce symptoms. To summarize this pattern, the symptoms of melancholia promote sustained cognition (analytical rumination) while growth and reproductive activity are suppressed. This cognitive activity requires altered synaptic plasticity in the HC–PFC synapse (increased LTD and serotonin-induced potentiation, decreased LTP) and sustained glutamatergic activity in the lateral PFC. These processes are so energetically expensive they require an upregulation in glycolysis, and cannot simultaneously support the growth of new neurons. Moreover, neurogenesis may functionally interfere with plasticity because many of the changes in synaptic connectivity require LTD and dendritic pruning.

The direct pharmacological properties of SSRIs do not appear to effect a reversal in this pattern. Rather, acute SSRI treatment tends to exacerbate it—LTP and BDNF signaling are inhibited, while serotonin-induced potentiation is increased (Cai et al., 2013; De Foubert et al., 2004; Shakesby et al., 2002). Reversal (increased LTP and BDNF signaling, and decreased serotonin potentiation) only occurs over chronic SSRI treatment (Bhagya et al., 2011; Cai et al., 2013; De Foubert et al., 2004). Altogether, the pattern suggests that the reversal is due to the brain’s compensatory responses to SSRI treatment.

Furthermore, we suggest that eventual reversal of the energy allocation pattern explains why promoting neurogenesis is crucial to symptom reduction in some models (Perera et al., 2011a; Santarelli et al., 2003). This finding appeared to support the hypothesis that reduced neurogenesis was a mechanistic cause of depression (Duman et al., 1997). However, further research showed that the ablation of neurogenesis is insufficient to trigger depressive symptoms (Jayatissa et al., 2009; Surget et al., 2008). This, and other findings, have cast doubt on the neurogenic hypothesis for depression (Groves et al., 2007; Mahar et al., 2014). Our approach elegantly explains this pattern. ADMs only reduce symptoms to the degree they induce a sufficiently strong compensatory response by the brain to suppress the allocation of energy devoted to sustained cognition. However, the ablation of hippocampal neurogenesis does not trigger depressive symptoms because it does not, by itself, promote sustained cognition.

6. Conclusion and future directions

The reigning paradigm conceptualizes depression as a state of reduced serotonin transmission. In this paper we have reviewed a large body of evidence indicating that the opposite appears to be true. For the depressive phenotypes we have considered—sickness behavior, starvation depression, and melancholia—serotonin transmission to multiple brain regions appears to be elevated. Others have suggested serotonin transmission is elevated in depression (Andrews and Thomson, 2009; Petty et al., 1994; Zangen et al., 1997), but this is the first in-depth review of the high serotonin hypothesis.

There are, of course, several caveats to this claim. First, serotonin cannot be simply described as an ‘upper’ or a ‘downer’. Evidence of elevated serotonin transmission can be found in both positive and negative states (Table 3). Alterations in other biochemical systems are probably needed to differentiate negative and positive mood states. For instance, depressive states seem to involve elevated serotonin and reduced dopamine transmission in the nucleus accumbens (Bland et al., 2003b; Zangen et al., 1997). Conversely, acute cocaine administration produces a euphoric mood and increases both extracellular serotonin and dopamine in the nucleus accumbens (Li et al., 1996).

Second, serotonin is probably a distal factor in the causal pathway that regulates depression and other energetically expensive states. Consequently, alterations in serotonin transmission are probably neither necessary nor sufficient to regulate depressive symptoms; researchers can affect depressive symptoms by altering more proximate mechanisms, such as glutamate transmission. Nevertheless, we propose that altered serotonin transmission is part of the evolved process by which depression and other energetically expensive states are regulated.

Third, the findings reviewed herein may not be generalizable to other depressed states. For instance, we did not consider atypical depression because it is heterogeneous (Stewart et al., 2007), and we lack valid non-human animal models for it. Still, we predict that serotonin transmission is also elevated in atypical depression phenotypes, because we suspect that they also involve the prolonged reallocation of energy.

Fourth, the evidence most relevant for evaluating the high and low serotonin hypotheses comes from studies on non-human animals. While some debate persists about the validity of animal models of depression, they are invaluable tools in the study of the neurobiology of depression (Berton et al., 2012). Nevertheless, the study by Barton and colleagues (2008) provided converging evidence of elevated serotonin transmission (indexed by 5-HIAA levels in the jugular vein) in patients meeting current diagnostic criteria for major depression.

Fifth, by necessity our review relies on indices of serotonin transmission. As discussed above, we are currently unable to measure serotonin in a living human brain without invasive techniques. However, even with the development of safe, non-invasive in vivo techniques for measuring serotonin concentrations, we would still have to rely on indices, such as the 5-HIAA/5-HT ratio. The development of techniques that would allow the direct measurement of serotonin transmission in the human brain is a long way off.

With these caveats in mind, the high serotonin and energy regulation hypotheses conjointly explain why depressive symptoms commonly worsen in acute treatment when serotonin levels are at their highest. They also explain the therapeutic delay as the result of the compensatory responses that attempt to restore energy homeostasis.

Future research should map out how the serotonergic system and depressive symptoms change over acute, chronic, and more prolonged SSRI treatment, and after discontinuation of the treatment. Since serotonin has state-dependent effects, it is important to control for the baseline state. We suggest using a well-developed model of depression, such as inescapable shock or chronic social defeat, as the baseline state. Once the compensatory changes have been compellingly mapped out, we predict that disrupting the compensatory responses will prevent or delay the antidepressant effect.

Finally, the energy regulation hypothesis suggests many potential lines of research that could be important in understanding what serotonin is doing in depressed states. We highlight two such areas. First, in addition to serotonin, melancholia involves the heightened, sustained secretion of cortisol. Serotonin and cortisol both affect aerobic glycolysis and oxidative phosphorylation. Aerobic glycolysis occurs in the cytosol and the endproduct is lactate, which must be converted back to pyruvate before it can go through oxidative phosphorylation in the mitochondrion, so there is a trade-off between the two processes (Andersson et al., 2003; Pfeiffer et al., 2001). Serotonin and cortisol may interact to regulate and influence the balance between these processes, and this should be investigated.
Second, we have been unable to find any research on how serotonin enters the mitochondrion. That strikes us as a glaring gap in our knowledge. Since aerobic glycolysis and oxidative phosphorylation occurs in different intracellular compartments, it may be important to understand how the intracellular distributions of serotonin are regulated. At stake is the interpretation of neuroimaging studies of depression, which depend upon the balance between aerobic glycolysis and oxidative phosphorylation (Section 5.1).

In summary, we propose that depressed states are high serotonin phenomena, which challenges the prominent role the low serotonin hypothesis continues to have in depression research (Albert et al., 2012). We also propose that the direct serotonin-enhancing effects of antidepressants disturb energy homeostasis and worsen symptoms. We argue that symptom reduction, which only occurs over chronic treatment, is attributable to the compensatory responses of the brain attempting to restore energy homeostasis. Understanding the true relationship between serotonin and depressed states will be important in understanding the etiology of those states and developing effective treatments.

Acknowledgments

The authors gratefully acknowledge the following people for comments, conversations or correspondence that influenced the manuscript: Marie Banich, Dick Day, Denys DeCatanzaro, Zac Durisko, Steve Gangestad, Ed Hagen, Chris Lowry, Steve Maier, and Onkar Marway. Three anonymous reviewers provided thoughtful comments that greatly improved the manuscript.

References


