Major depression as a disorder of serotonin resistance: inference from diabetes mellitus type II

Bella Smolin¹, Ehud Klein², Yishai Levy¹ and Dorit Ben-Shachar²

¹ Department of Internal Medicine and ² Research Laboratory of Psychobiology, Department of Psychiatry – Rambam Medical Center, B. Rappaport Faculty of Medicine, Technion, Haifa, Israel

Abstract

The multifactorial nature of depression resembles that of other complex disorders such as diabetes mellitus or coronary artery disease. However, while for the latter disorders predisposing and risk factors have been identified, such knowledge is still scarce in depression. In this review we propose to use diabetes mellitus, for which characteristic milestones have been condensed to obesity–hyperinsulinaemia–insulin resistance–diabetes mellitus, as a conceptual analogical model. Based on this model we hypothesize that depression develops according to a similar pattern: prolonged psychological stress–hyperserotonism–serotonin resistance–major depression. We review extensive supporting evidence from human studies and animal models of depression, including stress involvement in the aetiology of depression, evidence for increased synaptic serotonin and decreased 5-HT₁A receptor activity. Conceptualizing the pathogenesis of depression as a multi-step process may inspire new concepts, which will eventually lead to delineation of additional preventive and therapeutic interventions similar to those currently practised in diabetes.

Received 17 July 2006; Reviewed 15 October 2006; Revised 22 November 2006; Accepted 23 December 2006; First published online 25 January 2007

Key words: Depression, diabetes type II, serotonin resistance.

Introduction

Despite its significant impact on morbidity burden, depression remains a disease with uncertain aetiology and pathophysiology. Estimates of its prevalence in Western countries vary from 15% to 25% for lifetime prevalence with point prevalence of at least 5% (Badamgarav et al., 2003; Druss et al., 2000). More than 35 yr after its formulation by Schildkraut, the biogenic amine hypothesis is still the core of our understanding of the biology of depression. This hypothesis claims that depression is caused by a central deficiency of biogenic amines, and that the clinical efficacy of antidepressants is exerted by correcting this deficiency (Schildkraut, 1969). Norepinephrine (NE) and serotonin (5-HT) have both been implicated in the pathophysiology of depression. A more comprehensive current understanding of the monoaminergic hypothesis implicates both presynaptic and postsynaptic components of 5-HT and NE neurotransmission as well as postsynaptic intracellular signalling cascades regulated by these monoamines. Since the introduction of the selective serotonin reuptake inhibitors (SSRIs) for the treatment of depression, 5-HT has become the focus of both basic and clinical research in this disorder (Mann, 1999; Owen and Nemerooff, 1994). An additional well-established finding in depression is a dysregulation of the HPA axis in a significant number of patients (Holsboer, 2000; Pariante and Miller, 2001; Stokes, 1995). The role of HPA axis disturbances in major depression can be understood alongside the alterations in monoamine transmission in light of the mutual effect the two systems exert on each other, but is beyond the scope of this review.

The great progress made in applying modern biological and imaging techniques in the research of psychiatric disorders, led to the perception that depression is a complex multi-factorial disorder affected by genetic, epigenetic and environmental factors similar to other complex medical disorders that
became ‘epidemic’ in the last century, such as diabetes mellitus, hypertension and coronary artery disease. However, while in the latter disorders predisposing and risk factors have been identified, allowing the assessment of milestones in their pathogenesis, such knowledge is still scarce in depression.

The last two decades have witnessed remarkable progress in our understanding of key events in the development of one of the afore-mentioned complex disorders namely non-insulin-dependent diabetes mellitus (NIDDM) also known as diabetes type II. The unravelling of the pathogenesis of diabetes, its predisposing conditions and risk factors led to the development of modern treatment and prevention strategies. In the present review we propose to use diabetes type II as a conceptual analogical model to further our understanding of key events in the pathogenesis of depression, which, as in the case of diabetes, may lead to improvement in the treatment and prevention of the disorder. Choosing a disease prototype as a model for drawing inferences regarding another disease is far from being a straightforward choice and is indeed more of a smart guess than a logically based proposal. However, we propose that, these very different diseases actually share similar features such as multi-step pathogenesis and a combination of environmental risk factors and genetic predisposition. This complex scenario can finally result in a dysfunction of a primary molecule such as insulin in diabetes and 5-HT and/or NE in depression. The characteristic steps leading to diabetes can be simplified to the following chain of events: obesity–hyperinsulinaemia and insulin resistance–diabetes mellitus. Based on this model we propose the hypothesis that depression develops according to a similar pattern: prolonged psychological stress–hyperserotonism and serotonin resistance–major depression. In this paper we will focus on the 5-HT network, a hallmark in the pathogenesis and treatment of depression, while not dismissing the important role of additional factors, such as the noradrenergic systems and the HPA axis.

Milestones in the development of the research of diabetes type II pathogenesis

Insulin deficiency was once thought to be the sole cause of diabetes. The earliest indication for coexistence of defects in insulin secretion as well as in its action was presented by Himsworth (1936) more than 70 yr ago. One of the first pieces of evidence for abnormal insulin activity came from a study in Pima Indians. This study demonstrated that insulin response to oral glucose and thereby plasma glucose concentration was dependent on body weight (Martin et al., 1980). Many other studies confirmed that insulin is increased rather than decreased in the presence of pathological high plasma glucose concentrations (Boyko et al., 1991; Haffner et al., 1986). Consequently the concept of insulin resistance emerged. Insulin resistance is a state of a reduced ability of insulin to lower plasma glucose levels. Extensive research demonstrated that hyperinsulinaemia precedes the development of type II diabetes (Haffner et al., 1990; Lillioja et al., 1988; Warram et al., 1990). These studies provide direct quantitative evidence that the progression from normal to impaired glucose tolerance is associated with the development of insulin resistance. It is currently believed that the prolonged insulin resistance with a compensatory hyperinsulinaemia is finally followed by a progressive loss of β-cell function leading through impaired glucose tolerance to overt diabetes mellitus. Evidence has accumulated that genetic factors are involved in the regulation of insulin action (LeRoith, 2002). However, as much as 50% of the variability in insulin-mediated glucose disposal has been attributed to differences in lifestyle (Bergman et al., 2003). Of pivotal importance is the finding that lifestyle modifications, such as low fat diet, physical activity and weight reduction, prevent the development and reduces the prevalence of diabetes by decreasing the resistance to insulin (Cheng et al., 2002; Harris and White, 2004; Ross et al., 2000, 2004).

In an attempt to extend the concept of insulin resistance in the development of diabetes to the development of major depression, psychological stress can serve as a conceivable analogue to obesity, likewise initiating the chain of events resulting in serotonin resistance and hyperserotonism, which finally lead to the disease.

Psychological stress, NE, 5-HT, and mood

Psychological stress is an activating stimulus, which triggers alterations both in the noradrenergic and serotonergic transmission leading to modulation of relevant neural circuits, such as the amygdala–hippocampus–prefrontal cortex, thereby mediating affective states such as depression (Flugge et al., 2004). An additional important target of psychological stress is the HPA axis, which modulates noradrenergic and serotonergic transmission and vice versa.

It was suggested that acute stress could activate the locus coeruleus (LC)–NE system causing an increase in NE release (Glavin et al., 1983; Tanaka et al., 1982). The activation of the LC in itself can lead to the inhibition of the raphe nucleus (RN)–5-HT system through
adrenoceptors in the medial and dorsal RN. The RN–5-HT system can further be inhibited by the prefrontal cortex efferents. Repeated stimuli can cause rapid desensitization of the LC with decreased NE phasic responses. The latter can induce disinhibition of 5-HT release, as LC and RN actions tend to be mutually inhibitory (Ressler and Nemeroff, 2000).

Multiple evidence suggests that, being an important homeostatic factor, 5-HT acts to trigger behavioural inhibition (i.e. freezing instead of flight or aggression) increasing the tolerance/resilience toward aversive experience (Dinan, 1996; Gold and Chousos, 2002). In light of the interactions between the noradrenergic and serotonergic systems during stress it is probable that more 5-HT is required to balance stress-induced over-activation. Thus, in the case of chronic psychological stress a state of prolonged ‘hyperserotonism’ can develop. Following an increase in 5-HT, a down-regulation of its receptors is anticipated in order to maintain homeostasis, which may lead to a serotonin resistance state. As a consequence a further increase in 5-HT release may occur to compensate for the decreased activity of 5-HT receptors. This vicious circle can lead to an exhaustion of the system and the development of a 5-HT deficiency state characteristic of depression (Figure 1). Such a scenario resembles the cascade of events characteristic of diabetes pathogenesis; initiating with obesity, followed by the development of insulin resistance and hyperinsulinaemia finally leading to an exhaustion of β-cell function. Depression can evolve along the axis of events of initial over-activation of 5-HT transmission followed by the exhaustion of the system leading to low 5-HT levels.

The model suggests prolonged stress as an essential factor in the pathogenesis of depression. In addition, epidemiological studies demonstrate an increased risk for developing depression in adulthood following early life stress (Famularo et al., 1992). Early adverse experiences may activate epigenetic processes that can unveil a pre-existing genetic vulnerability to stress and disease. However, the translation of such an early vulnerability to a disease phenotype may require further exposure to stressful life events (Heim and Nemeroff, 2001). Likewise, we suggest that serotonin resistance develops over years, eventually leading to pathological consequences in adulthood.

Our hypothesis that ‘hyperserotonism’ and ‘serotonin resistance’ are associated with the pathogenesis of depression is supported by a substantial number of experimental findings.

### Evidence for ‘hyperserotonism’ and ‘serotonin resistance’

#### Human studies

Direct evidence for abnormal 5-HT transmission in humans is sparse. In-vivo measurement of 5-HT levels in the human brain is not yet feasible and post-mortem measurement of neurotransmitters including 5-HT is fraught with variability related to post-mortem degradation. However, over the past two decades, a number of post-mortem (Arango et al., 1995; Austin et al., 2002; Mann et al., 2000) and neuroimaging (Malison et al., 1998; Meyer et al., 2001; Parsey et al., 2006; Willeit et al., 2000) studies found diminished availability of the 5-HT transporter in depression, thus providing an indirect evidence of a possible increase in 5-HT concentrations in the synapses.

Additional and more comprehensive evidence comes from genetic studies that focused on the association of a function relevant polymorphism in the 5-HT transporter and depression. A polymorphism in the promoter region of the human serotonin transporter (5-HTT) gene (SLC6A4) was detected and designated the 5-HTT gene-linked polymorphic region (5-HTTLPR). This polymorphism leads to modified promotor activity, when the short (‘s’) allele of the 5-HTTLPR is associated with lower transcriptional efficiency of the promotor compared with the long (‘l’) allele. Individuals with one or two copies of the short
allele of the 5-HTT promoter exhibited more depressive symptoms, diagnosable depression, and suicidality in relation to stressful life events than individuals homozygous for the long allele (Caspí et al., 2003; Kendler et al., 2005; Lesch et al., 1996). The short 5-HTTLPR allele is associated with lower transcriptional efficiency of the transporter, which inevitably has to result in less effective reuptake of 5-HT from brain synapses. Although the latter has not been directly proven, such a scenario is reasonable based on the core understanding of the SSRIs mechanism of action. This apparent contradiction, that a polymorphism causing a reduction in 5-HTT similar to SSRI action, is a negative trait predisposing to stress-related depression may be easily explained by our hypothesis on the essential role of a state of ‘hyperserotonism’ as a predisposing risk factor for depression. In line with the latter are the recent findings on the poorer treatment outcome of major depression patients (with the 5-HTTLPR short allele, compared to patients homozygous for the long allele), to various antidepressants including SSRIs, a combination of SSRIs and noradrenaline re-uptake inhibitors, tricyclic antidepressants and monoamine oxidase inhibitors (Arias et al., 2003; Lee et al., 2004; Murphy et al., 2004; Putzhammer et al., 2005; Smits et al., 2004).

According to our hypothesis ‘hyperserotonism’ is followed by ‘serotonin resistance’, i.e. decreased sensitivity of 5-HT receptors. Indeed, alteration in the 5-HT receptor population has been observed in brains of depressed patients (Blier and Ward, 2003). Among the various 5-HT receptors, the 5-HTT1A receptors are particularly relevant to the antidepressant and anxiolytic responses in human. They are located primarily presynaptically in the raphe nuclei, where they act as cell body autoreceptors to inhibit the firing rate of 5-HT neurons, as well as post-synaptically in limbic and cortical regions, where they transmit the 5-HT-associated information (Blier and Ward, 2003). In depressed patients 5-HTT1A receptors have been evaluated using positron emission tomography (PET) imaging or brain post-mortem specimens. A PET study by Sargent et al. (2000) found decreased antagonist binding to 5-HTT1A receptors in 15 non-medicated patients with major depressive disorder (MDD) across several cortical regions including medial temporal cortex, the temporal pole, orbitofrontal cortex, anterior cingulate cortex, insula and dorsolateral prefrontal cortex. Additional PET studies reported a persistent lower binding potential of brain 5-HTT1A and 5-HTT1A receptors in various cortical areas and/or the dorsal RN in patients with major depression (Bhagwagar et al., 2004; Meltzer et al., 2004; Messa et al., 2003; Sheline et al., 2004) and other mood disorders (Cleare et al., 2005; Neumeister et al., 2004; Sullivan et al., 2005).

Studies using post-mortem brain specimens from depressed patients are less consistent regarding the alteration in 5-HT receptor binding characteristics, reporting increases (Arango et al., 1995; Matsubara et al., 1991; Stockmeier et al., 1997, 1998) as well as no change (Arranz et al., 1994; Dillon et al., 1991; Lowther et al., 1997; Sher et al., 2003) in radioligand binding to 5-HTT1A receptors in the prefrontal cortex, hippocampus and the dorsal RN. The apparent discrepancy between the post-mortem studies and the PET studies has been partially attributed to a number of methodological confounders, which vary between both experimental approaches and are routinely better controlled in imaging than in post-mortem studies (see Stockmeier, 2003 for a detailed review).

**Animal studies**

A major obstacle in the research of psychiatric disorders is the lack of valid animal models. Many of the core symptoms of these disorders involve higher brain functions unique for humans and thus cannot be simply replicated in animals. In addition, the precise aetiiological factors and the underlying genetic abnormalities are still enigmatic, which adds to the difficulty to model the psychiatric diseases in a laboratory animal. Despite their limitation, animal behavioural and genetic models provided pivotal information on brain circuits, neurotransmission, intracellular mechanisms and mechanism of action of drugs. A broad range of evidence supporting the ‘hyperserotonism’ and ‘serotonin resistance’ hypothesis has been obtained from animal models of depression. Numerous studies have shown elevation of the concentration of 5-HT and its metabolites in several brain regions in models of stress and depression. Extracellular 5-HT levels have been monitored in more than 15 regions of the brain during a variety of spontaneous behaviours, and in response to several physiological, environmental, and behavioural stress manipulations. Although there are reports on a reduction or no change in extracellular 5-HT levels, the vast majority of these studies found increases (30–100%) in 5-HT release (Rueter et al., 1997) in almost all brain regions studied (Adell et al., 1997; Sher et al., 2003) in radioligand binding to 5-HTT1A receptors in the prefrontal cortex, hippocampus and the dorsal RN. The apparent discrepancy between the post-mortem studies and the PET studies has been partially attributed to a number of methodological confounders, which vary between both experimental approaches and are routinely better controlled in imaging than in post-mortem studies (see Stockmeier, 2003 for a detailed review).
knockout mice reported elevated extracellular concentrations of 5-HT, which was associated with increased susceptibility of mice to stress, and abnormal depressive and anxiety-like behavioural phenotypes (Collin et al., 2000; Holmes et al., 2002, 2003a,b; Lira et al., 2003; Tjumina et al., 2002). In Flinders Sensitive Line (FSL) rats, 5-HT and 5-hydroxyindoleacetic acid (5-HIAA) were elevated in limbic regions and normalized after chronic treatment with antidepressants (Zangen et al., 1997).

Concomitant with increased release of 5-HT, down-regulation of 5-HT receptors in general, and in the 5-HT$_{1A}$ receptor in particular, was observed in stress and genetic animal models of depression. Decreased 5-HT$_{1A}$ binding potential was observed in the raphe, amygdala, hippocampus, and cortical areas of depressed monkeys relative to non-depressed controls as well as in rats and tree shrews exposed to chronic unpredictable stress (Flugge et al., 1995, 1997; Lopez et al., 1998; Shively et al., 2006). In line with these findings 5-HT$_{1A}$ knockout mice demonstrated increased anxiety-like behaviour and a high tendency to avoid a novel fear-inducing environment (Heisler et al., 1998; Parks et al., 1998; Ramboz et al., 1998). A particularly important and more direct support for disturbed 5-HT receptor response to 5-HT, is the observation that 5-HT$_{1A}$ receptor-mediated responses to 5-HT in hippocampal cornu ammonis 1 (CA1) neurons is suppressed in stressed rats (Van Riel et al., 2003).

Moreover, in a genetic model of depression of FSL rats, an attenuated release of dopamine (DA) in the nucleus accumbens (NAc) in response to 5-HT application was observed (Zangen et al., 2001). The latter, was attributed to decreased sensitivity of the 5-HT$_{1A}$ and 5-HT$_2$ receptors known to be present on dopaminergic neurons in the NAc (De Deurwaerdere et al., 2005).

The studies described hitherto point to an increased 5-HT release and subsensitivity of 5-HT receptors particularly that of 5-HT$_{1A}$ in animal models of stress and depression. However, there are reports opposing our hypothesis of hyperserotonism and serotonin resistance both in human and animal models. For example, mixed findings including decreases of 5-HIAA in cerebrospinal fluid in depressed individuals were reported (Asberg et al., 1984). In FLS rats, along with the high tissue levels of 5-HT in the limbic regions of the brain (Shalit et al., 2003) and disturbed 5-HT receptor activity in the NAc (Zangen et al., 2001), which agree with our hypothesis, several studies showed increased rather than decreased serotonergic sensitivity assessed as the hypothermic response to 5-HT$_{1A}$ agonists (Overstreet et al., 1994; Wallis et al., 1988).

### Antidepressants in light of the serotonin resistance model

Our model suggesting a serotonin resistance state in the development of depression raises a question regarding the mode of action of SSRI antidepressants. We propose that SSRIs augment the already increased levels of synaptic 5-HT, leading to alleviation of symptoms by overcoming serotonin resistance. Alternatively, increased synaptic 5-HT levels may potentiate the resistance to 5-HT, thereby leading to lack of SSRIs’ efficacy or to an escape from their clinical effect, as observed in a considerable number of patients. A similar mode of action is currently accepted for sulfanylurea drugs in diabetes type II.

It is common knowledge that antidepressant drugs that act through the noradrenergic system are also effective in alleviating depressive symptoms. Obviously, our hypothesis cannot encompass the pathophysiology of all subsets of depressed patients. However, accumulating evidence suggest that the interaction between the serotonergic and adrenergic systems is critical for both the development of depression (as has been briefly described above) and the therapeutic mechanism of antidepressant treatment (Helms et al., 1986; Plaznik et al., 1994). For example, a stimulatory input of noradrenaline via $\alpha_2$-adrenoceptors on serotonergic neurons in the dorsal raphe, leads to increased 5-HT release from 5-HT neurons projecting to the hippocampus (Baraban and Aghajanian, 1980; Potter, 1996). Moreover, through $\alpha_2$-autoreceptors, low concentrations of noradrenergic agonists increase 5-HT neurotransmission by attenuating the release of endogenous noradrenaline (Maj et al., 1985). This modulation of serotonergic transmission can combine the clinical effectiveness of noradrenergic antidepressants and our model.

### Conclusions

The complexity and multifactorial nature of depression, as well as the interplay between heterogeneous molecular/genetic/epigenetic and environmental factors leading to a common clinical phenotype resemble that of other complex medical disorders such as diabetes mellitus and coronary artery diseases. However, unlike the case of depression, predisposing and risk factors have been identified for the latter disorders, leading to the delineation of milestones in their pathogenesis, and consequently to the development of new treatment and prevention strategies. By drawing an analogy from the currently accepted pathological cascade of events, which precede the onset of diabetes...
type II, i.e. obesity (life style)–hyperinsulinaemia–insulin resistance, we proposed a heuristic parallel model for the development of depression, prolonged stress (life style)–hyperserotonism–serotonin resistance. Indeed, extensive experimental evidence from human studies and animal models of depression support this chain of events in the pathogenesis of depression (Table 1). Thus, stressful life events are regarded as an important factor in the aetiology of depression. Most studies, although not all, from animal models of depression
demonstrate increased synaptic 5-HT levels and decreased 5-HT$_{1A}$ receptor activity. Human studies are obviously less conclusive, but the majority of the imaging studies agree on a down-regulation of 5-HT$_{1A}$ receptors in depression-relevant brain areas. We believe that an imaging study using $^{18}$F]MPPF, before and after fenfluramine challenge, comparing subjects who experienced stressful life events, subjects with major depression and normal controls could validate our model. $^{18}$F]MPPF is a selective and reversible antagonist to the 5-HT$_{1A}$ receptor, which is sensitive to increases in 5-HT level and therefore suitable for measuring receptor number as well as the extent of 5-HT release in response to fenfluramine. Moreover, the overall 5-HT activity can be inferred from the assessment of peripheral prolactin levels.

Our hypothesis suggests that a crucial step in the pathogenesis of depression is ‘serotonin resistance’. So far we suggested ‘stress’ and ‘hyperserotonism’ as the main cause of the latter. However, some patients may develop depression unrelated to psychological stress. Therefore, one has to consider other endogenous and exogenous factors, which have been reported to be abnormal in depression, as contributing to serotonin resistance (Figure 2). For example, there is evidence demonstrating that oestrogen alters the serotonergic system. Decreased testosterone levels were shown to be associated with a reduced number of 5-HT$_{1A}$ receptors, which was restored by testosterone substitution (Flugge et al., 1998; Ostlund et al., 2003). Moreover, activation of cytokines, which have been suggested to play a role in the aetiology of depression, is associated with an increase in brain 5-HT turnover (Pucak and Kaplin, 2005). Likewise, physical activity, sleep deprivation, nutritional manipulations and malnutrition are examples of multiple exogenous factors reported to modulate the activity of serotonergic neurons (Bailer et al., 2005; Blundell et al., 1995; Flugge et al., 1998; Greenwood et al., 2005; Senthivelan et al., 2006). The involvement of various factors in the

![Figure 2. Schematic description of the potential factors that can contribute to serotonin resistance: Acute stress or activating stimuli lead to initial activation of the LC–NE system and inhibition of RN–5-HT system. Prolonged stress or repeated stimuli lead to desensitization with decreased NE phasic responses and thereby to disinhibition of 5-HT release. Increased release of serotonin leads to down-regulation of serotonin receptors leading to serotonin resistance. ●, Enhanced response; ○, decreased response.](image-url)
development of 5-HT resistance is depicted in Figure 2 and is also true for the development of insulin resistance in diabetes type II, our prototype disease.

It is, however, noteworthy that our model is a simplification of the interaction between environmental genetic/epigenetic and biological factors involved in the development of the disease. Obviously, the model we suggest cannot encompass all aspects of the pathogenesis, pathophysiology and treatment of major depression. Nevertheless, the comprehension of the pathogenesis of depression as a multi-step process may inspire new ideas and concepts concerning the pathological processes involved in the development, remission and recurrence of depression.

This may eventually enable the development of novel prevention strategies as well as new treatment approaches aimed at various stages of the process. Indeed, in diabetes such a strategy led to the concept of the superior importance of lifestyle modification in an attempt to overcome risk factors and predisposing conditions. In addition, this brought to the development of new drugs such as rosiglitazone aimed at overcoming insulin resistance, along with the traditional therapies such as sulfonylurea which increases insulin production, or insulin per se. Likewise, considering stressful life events as a risk factor for depression may lead to lifestyle modification and emphasize the importance of psychotherapy or other behavioural interventions to curb life stressors. In addition, along with the current approach of elevation of 5-HT synaptic concentrations, referring to depression as a 5-HT resistant state can lead to the development of new drugs aimed directly at improving 5-HT activity, similar to novel drug treatments practised in diabetes.

Acknowledgements
None.

Statement of Interest
None.

References


Wallis E, Overstreet DH, Crocker AD (1988). Selective breeding for increased cholinergic function: increased...


