The Ferrier Lecture 1998

The molecular biology of consciousness investigated with genetically modified mice

Jean-Pierre Changeux*

Institut Pasteur & Collège de France, 25 rue du Dr Roux, 75724 Cedex 15, Paris, France

The question is raised of the relevance of experimental work with the mouse and some of its genetically modified individuals in the study of consciousness. Even if this species does not go far beyond the level of ‘minimal consciousness’, it may be a useful animal model to examine the elementary building blocks of consciousness using the methods of molecular biology jointly with investigations at the physiological and behavioural levels. These building blocks which are anticipated to be universally shared by higher organisms (from birds to humans) may include: (i) the access to multiple states of vigilance, like wakefulness, sleep, general anaesthesia, etc.; (ii) the capacity for global integration of several sensory and cognitive functions, together with behavioural flexibility resulting in what is referred to as exploratory behaviour, and possibly a minimal form of intentionality. In addition, the contribution of defined neuronal nicotinic receptors species to some of these processes is demonstrated and the data discussed within the framework of recent neurocomputational models for access to consciousness.

Keywords: consciousness; exploratory behaviour; nicotinic receptors; genetically modified mice

David Ferrier (1878) in his book ‘The localization of cerebral diseases’ gives a brief report of the experiments he carried out on the ablation of the prefrontal lobes in the monkey. He mentions notable alterations in the ‘character and manners of the animals’ and notes that ‘instead of examining with interest and curiosity all what was happening in their field of observation, they remained apathic and weak, they were sleepy, responding only to sensations and impressions at the time’... ‘They had lost the capacity of intelligent observation and attention’. In the same book, Ferrier also reports the first observation of a frontal lobe lesion in humans, the famous case of Phineas Gage (see Damasio 1995). He notes that following his accident, changes in his emotional reactions and intellectual faculties took place: ‘he became nervous, irreverent... impatient... with an excessive obstination... capricious and undecided’. Ferrier further states that ‘what is true for the monkey is true for man’, thus introducing the important issue of the possible validity of animal models of human brain diseases. In the following years, the Italian neurologist, Luigi Bianchi (1921), went a step further. In his book ‘The mechanics of the brain: the function of the frontal lobes’, he states that ‘among the phenomenal factors of the activities of living organisms’ arises a ‘bond of coherence’, which ‘progresses with the development and complexity of living organisms and of their nervous system’. He refers to it as consciousness and further specifies that ‘the dawn of higher consciousness coincides with the apparition of the frontal lobes in the evolution of the brain ’, also noting their ‘inhibitory power’ and their capacity for ‘intellectual syntheses’.

Since these early times, the notion that non-humans may possess consciousness gave rise to contrasting attitudes. Following ‘Thomas Huxley’s (1874) provocative statement of his Belfast lecture that ‘we are conscious automata’, that ‘brutes’ share consciousness with humans and that ‘all states of consciousness in us, as in them, are immediately caused by molecular changes of brain substance’, non-human species have served as convenient experimental models for the scientific investigation of consciousness (Thorndike 1898; Yerkes 1916; Clark & Squire 1998; Jasper 1998; Hauser 2000; Dehaene & Changeux 2004; Koch 2004; Denton 2005). In particular, mice are easily accessible to pharmacological or genetic manipulations and carry out tasks that in humans involve awareness of the stimuli (e.g. Picciotto et al 1995, 1998; Maskos et al. 2005). Yet, it may be argued that mice lack the most characteristic features of consciousness and ‘what it is like to be a mouse’ may wildly differ from ‘what it is like to be a human being’ (see Nagel 1974). Accordingly, only humans would feel and introspect. Subjective—first person—experience would be the ‘hard problem’ to explain, which requires the invention of far-fetched bridging principles (Chalmers 1998).

The issue then becomes to what extent studying mice may help to understand consciousness in humans? Can the laboratory mouse become a relevant animal model to investigate the molecular biology of consciousness or at least some of its aspects?

A first point to make is that looking at the broad spectrum of living species, including humans and the developing infants, ‘we must question the monitoring...
of this mental activity’. As already stated by Preyer (1894) in his book ‘Mental development in the child’, there is not the least reason for assuming in advance that every human being comes into the world endowed with complete consciousness of self and ‘that at the moment of birth an immortal soul is in waiting, as it were, for the new and as yet unconscious citizen of the world in order to take possession of him forever’. Preyer also noted that in the newborn the different senses gradually develop, in a non-concomitant manner, the feeling of selfhood that he refers to as ‘cortical ego’ developing later. He further states that, as the last step, ‘out of knowledge (science) (he) has developed his conscience’. Unambiguously, ‘there are several grades of consciousness’. The development of these grades, or levels of consciousness, has been reinvestigated recently, in humans from the newborn to the adult (Lagercrantz et al. 2002; Johnson 2005; Lagercrantz 2005; Lee et al. 2005; Bartocci et al. 2006), as in the course of biological evolution (see Boakes 1984; Changeux 2002; Zeki 2003; Zhou et al. 2005).

Regrouping the evolutionary analysis of Barresi & Moore (1996), primarily based on social interactions (see also Trivers 1985), and the developmental data of Zelazo (1996, 2004), Lagercrantz (Lagercrantz et al. 2002; Bartocci et al. 2006) and others (Johnson 2005; Lee et al. 2005) on the human newborn, one may propose a breakdown of ‘consciousness’ into multiple-nested hierarchical levels. Beyond Edelman’s (1989) cleavage between what he refers to as ‘primary’ and ‘higher-order’ consciousness primarily on the basis of language use, one may suggest the following distinctions:

(i) A lowest level of ‘minimal consciousness’ for simple organisms, like rats or mice, is characterized by the capacity to display spontaneous motor activity and to create representations, for instance, from visual and auditory experience, to store them in long-term memory and use them, for instance, for approach and avoidance behaviour and for what is referred to as exploratory behaviour. The comparability between the action of self and others is reduced. The 25–30-week preterm fetus, according to Lagercrantz (2005), would have reached a stage of brain maturation analogous (though not identical) to a newborn rat/mouse; he/she may process tactile and painful stimuli in the sensory cortex (Bartocci et al. 2006), would perceive pain and thus would show signs of minimal consciousness;

(ii) ‘recursive consciousness’, which is present, for instance, in vervet monkeys (possibly also in some birds), manifests itself by functional use of objects and by proto-declarative pointing; organisms at this level may display elaborate social interactions, imitation, social referencing and joint attention; they possess the capacity to hold several mental representations in memory simultaneously, and are able to evaluate relations of self; they display elementary forms of recursivity in the handling of representations, yet without mutual understanding; along these lines, the newborn infant exhibits sensory awareness, ability to express emotions and processes mental representations (i.e. of a pacifier); he/she would already differentiate between self- and non-self touch (Rochat 2003) and imitate the tongue protrusion of an adult (Meltzoff 1990); and thus would have reached the stage of what Zelazo (2004) refers to as recursive consciousness;

(iii) explicit ‘self-consciousness’ develops in infants at the end of the second year, together with working and episodic memory and some basic aspects of language; it is characterized by self-recognition in mirror tests and by the use of single arbitrary rules with self–other distinction; to some extent chimpanzees might reach this level (see Boakes 1984);

(iv) ‘reflective consciousness’, theory of mind and full conscious experience, with first person ontology and reportability, reaches full development in humans and develops following 3–5 years in children.

The case of mice and rats is simple and unambiguous. These species do not go far beyond the level of minimal consciousness. The issue of self- or reflective consciousness is irrelevant. Nevertheless, the mouse may be useful in providing some elementary ‘building blocks’ (Searle 2000) of consciousness. As mentioned, it offers also the possibility of joint investigations at the behavioural, physiological and molecular levels, together with an easy access to a large spectrum of techniques (from molecular genetics to brain imaging), making possible the investigation of what I may refer to as the fundamentals of minimal consciousness. These fundamentals which are anticipated to be universally shared by higher organisms (from birds to humans, but also possibly some invertebrates; see Baars 2001) may include: (i) the access to multiple states of vigilance, like wakefulness, sleep, general anaesthesia, etc.; (ii) the capacity for global integration of several sensory and cognitive functions in goal-directed behaviour (see Hobson 1999), Bianchi’s ‘bond of coherence’, together with behavioural flexibility, and thus what may be referred to as a minimal form of ‘intentionality’ (see §§88). In his recent book, the ‘Quest for consciousness’, Christof Koch (2004) has reduced the issue of consciousness in animals to what he refers to as a ‘Turing test’ for consciousness. ‘Take your sensor–motor routine in some species and enforce a waiting period of a few seconds between the sensory input and the execution of an action. If the organism’s performance is only marginally affected by the delay, then the input must have been stored in some sort of intermediate buffer, implying some measures of consciousness’ (Koch 2004, p. 227). This definition, by its simplicity, is of significant interest; yet, as we shall see, it might not be sufficient, even in the case of the mouse, to specify the fundamentals of consciousness.
1. MODELLING CONSCIOUSNESS: THE NEURONAL WORKSPACE HYPOTHESIS

In an attempt to implement in neural terms the ‘intermediate buffer’ (Koch 2004) or ‘global integration’ (Changeux 1983; Hobson 1999) aspects of consciousness, neurocomputational models referred to under the generic name of neuronal workspace hypothesis (Dehaene et al. 1998, 2003; Dehaene & Changeux 2004, 2005) were proposed (figure 1). These models which extend earlier formal models of delayed-response and card-sorting tasks and of the Tower of London test (Dehaene & Changeux 1989, 1991, 1997) offer, in particular, neuronal implementations of Baars (1998) psychological global workspace. The proposed architectures separate, in a first minimal description (see also Norman & Shallice 1980), two distinct computational spaces, each characterized by particular patterns of connections. Subcortical networks and most of the cortex would be viewed as a mosaic of specialized processors, each attuned to a particular type of information processing, such as sensory (present), motor (future), long-term memory (past), evaluation (value, emotions), or attention focusing (selective amplification or inhibition). In spite of their diversity, processors share specialization, automaticity and fast feedforward and feedback processing (Dehaene et al. 2003) via a limited number of local and medium range connections. On top of this automatic non-conscious level, a distinct set of cortical ‘workspace’ neurons was postulated and characterized by their ability to send and receive projections to many distant areas through long-range excitatory axons. These neurons, mainly pyramidal neurons from cortical layers II and III, particularly dense in prefrontal, cingulate and parietal regions, would break the modularity of the cortex through the recruitment of sets of neurons in the underlying processors; they would allow the broadcasting of information, in a spontaneous and sudden manner—or ‘ignition’—to multiple neural targets, thus creating a global availability (or integration) that would be experienced in consciousness and would give rise to reportability. Other computational theories on consciousness, several of them consistent with the present proposal, have been recently reviewed (Dehaene & Changeux 2005; Maia & Cleeremans 2005).

An important feature of these neuronal models of cognitive functions (Dehaene & Changeux 1989, 1991; Dehaene et al. 1998) has been the deliberate incorporation into their architecture of systems of evaluation through reward neurons (Dehaene & Changeux 1989) and even of auto-evaluation (Dehaene & Changeux 1997; see also the algorithms of Barto et al. (1983) and Friston et al. (1994)) implemented in terms of limbic/mesencephalic aminergic neurons, with which the prefrontal cortex is densely interconnected, and to which dopaminergic (DA) and cholinergic, nicotinic systems may directly, or indirectly, contribute. More elaborate reward learning algorithms based on predictive Hebbian learning were also proposed (Montague et al. 1995, 1996), and their implementation studied both in silico and in vivo in the macaque (Schultz et al. 1997). In contrast with other views (Shallice 1988; Crick & Koch 2003; Koch 2004), a close link was thus proposed between the neural basis of ‘making a conscious mental effort’ and the modelling of reward neurons (Dehaene et al. 1998; Dehaene & Changeux 2000).

Simulations performed with the proposed neural architectures account for cognitive tasks which tax the prefrontal cortex, such as delayed-response tasks, the Stroop task, the attentional blink (Dehaene et al. 2003) or inattentional blindness, taking into account the ongoing spontaneous activity of the network (Dehaene & Changeux 2005). Yet, at this stage, many features which characterize the higher levels of consciousness, such as the recursive and reflective aspects of consciousness, have not been dealt with. On the other hand, a simple connectionist circuit initially proposed for self-evaluation (Dehaene & Changeux 1991, 1995) may serve as a basic ingredient to begin modelling features of explicit self-consciousness.

The mouse lies at a low level of the evolutionary scale with a rather reduced prefrontal cortex. Yet, in this species, long-range frontal and callosal connections are abundant, though with the distinctive feature (at variance with primate and cat) of large-scale maintenance of simultaneous projection by frontal cortical and callosal projection neurons in the adult (Mitchell & Macklis 2005). Also, in mice and rats, strong evidence exists for a relationship between working memory and prefrontal function and interaction (Thinus-Blanc 1996; Jones 2002). Moreover, in mice and rats, all the neuronal modulatory systems (cholinergic, DA, serotonergic, adrenergic, etc.) are present and functional in the regulation of states of wakefulness, attention but also in reward processes. The mouse may thus possess a sufficient number of building blocks (Thinus-Blanc 1996) to make the scientific investigation of the fundamentals of minimal consciousness of relevant interest.

2. THE NEURONAL NICOTINIC RECEPTORS: ALLOSTERIC MEMBRANE PROTEINS THAT MODULATE HIGHER BRAIN FUNCTIONS

Physiological, psychopharmacological and pathological evidence supports the concept of an implication of cholinergic systems in cognitive functions and consciousness awareness (Woolf 1997; Perry et al. 1999; Perry & Perry 2004) to the extent that Koch (2004) states that ‘if a single neurotransmitter is critical for consciousness, then it must be acetylcholine’ (p. 90). Of particular interest in the context of consciousness, cholinergic neurons form discrete nuclei that provide widely divergent projections throughout the brain (rev. Semba 2004):

(i) the basal forebrain neurons (including the nucleus basalis and medial septal nucleus) constitute the main afferent system to the cerebral cortex and thalamus, and its functions include arousal, selective attention and rapid eye movements (REM) sleep;

(ii) the pedunculopontine neurons and lateral dorsal tegmental nucleus widely project to the forebrain regions, such as the thalamus but also to the brainstem reticular formation; this system also
Figure 1. The neuronal workspace hypothesis: schematic of brain networks comprising multiple specialized processors and a central network of high-level areas interconnecting them by neurons with long-range axons (from Dehaene et al. 1998, 2003).
contributes to the control of sleep and wakefulness in mammals (Semba 1999); moreover, both nuclei provide the cholinergic innervation of the DA cells of the ventral tegmental area (VTA) and substantia nigra and play a role in the acquisition of drug-rewarded behaviour and drug addiction, in particular nicotine self-administration (DiChiara 2000; Robbins 2000; Balfour 2002; rev. Gutkin et al. 2006), together with rewarded cognitive tasks performance (Kobayashi & Isa 2002).

Last, intrinsic cholinergic neurons are also distributed in the cerebral cortex and striatum, together with a variety of cholinergic cells dispersed in several brain regions (Semba 2004).

Acetylcholine in the brain interacts with two distinct categories of pharmacological receptors: muscarinic receptors that belong to the broad family of G-protein-coupled receptors with seven transmembrane helices, and nicotinic acetylcholine receptors (nAChR), which are pentameric, transmembrane ligand-gated ion channels (figure 2). Muscarinic receptors control working memory, declarative memories, sustained visual attention and psychomotor speed (see Ellis et al. 2005), and drugs that antagonize muscarinic receptors cause hallucinations and reduce the level of consciousness (Perry et al. 1999; Perry & Perry 2004). On the other hand, nicotine acting on brain nAChRs enhances arousal, attention and working memory, causes EEG desynchronization and higher amounts of REM sleep (rev. Levin 2002). Moreover, nAChRs have been implicated in the mechanisms of general anaesthesia. Volatile anaesthetics and ketamine have been reported to inhibit nicotinic receptors (α4β2 and α3β4) at clinically relevant doses (Franks & Lieb 1994; Tassonyi et al. 2002; see §8a).

The nAChRs are among the best-characterized pharmacological receptors and their contribution to brain functions actively investigated (rev. Changeux & Edelstein 2005). Their potential role in the control of consciousness will thus exclusively be considered in the subsequent sections of this review.

Muscle nAChR was initially identified from fish electric organ (rev. Changeux 1981, 1990; Changeux & Edelstein 2005) as a protein of about 300 000 MW made up of five subunits of several different types. Extensive biochemical studies, including reconstitution experiments and expression studies from cloned subunit genes, have shown that...
this transmembrane oligomer carries several categories of binding sites: two distinct acetylcholine (nicotine and snake venom z-toxins) sites located at the interface between subunits, a unique ion channel located in the axis of pseudosymmetry perpendicular to the membrane plane and all the structural elements that mediate the signal transduction process elicited by the neurotransmitter acetylcholine: the activation and desensitization of the ion channel (Changeux & Edelstein 1998, 2005). Biochemical equilibrium studies, in vitro rapid kinetics and in vivo electrophysiological recordings in situ or from functional receptors reconstructed in frog oocytes from genetically engineered subunits (rev. Changeux & Edelstein 2005) support the extension to the nAChR of the scheme proposed by Monod, Wyman & Changeux (1965) (MWC) for regulatory enzymes (rev. Changeux 1981; Changeux & Edelstein 1998, 2005). In short, the signal transduction processes of activation and desensitization are viewed as resulting from: (i) the spontaneous occurrence of the receptor protein under discrete conformations in reversible equilibrium; (ii) the differential stabilization of one of these few conformations by the ligand for which it exhibits the highest affinity; (iii) the conservation of the molecular symmetry of the receptor oligomer in the course of these transitions; (iv) the differences in the states of activity of the discrete conformations involved: the active state is the only conformation where the channel is open, the desensitized state exhibits the highest affinity for agonists (and some antagonists); (v) the presence of multiple allosteric sites for physiological and/or pharmacological ligands (e.g. for local and general anaesthetics) through which ligand binding may differentially stabilize the conformations of the protein.

Determination of the X-ray crystal structure of the snail acetylcholine-binding protein (a homologue of the extracellular domain of the nAChR; Brejc et al. 2001; Celie et al. 2004), together with the cryo-electron microscopy data of the membrane domain collected on Torpedo marmorata muscle type nAChR (Unwin 2000, 2005), has recently led to refined three-dimensional models of the receptor molecule (Paas et al. 2005; Taly et al. 2005; Unwin 2005). These studies demonstrate that: (i) the ACh-binding sites are far distant from each other (ca 30 Å) and (ii) the signal transduction mechanism that links the neurotransmitter sites and the ion channel involves a global structural change of the quaternary and tertiary structure of the protein that unites the two distinct domains of each subunit. In the absence of X-ray structural data, the conformational transitions were explored in silico by a computer method referred to as normal mode analysis. The first mode (over the first 10) was found to produce a structural reorganization of the model molecule compatible with channel gating. A wide opening of the channel pore resulted from the concerted symmetrical quaternary twist motion of the protein (Paas et al. 2005; Taly et al. 2005; Cheng et al. 2006). The data were found consistent, in first approximation, with the MWC two-state model of allosteric transition.

Among the 17 genes encoding nAChR subunits that have been identified and cloned in mammals, nine (a2–a7 and b2–b4) are expressed in the brain. These subunits may co-assemble to form a wide variety of functional pentamers that are distributed throughout the brain with different pharmacological and electrophysiological properties. The most abundant combinations of subunits are: a4b2 containing nAChRs (a4b2*), which exhibit high affinity for acetylcholine and nicotine, do not bind z-bungarotoxin and are abundant in cerebral cortex and thalamus; a3b4, which is primarily present in the peripheral nervous system; and a7 homopentamers, which display a low affinity for acetylcholine, tightly bind z-bungarotoxin, desensitize very quickly and are abundant in the limbic systems and cerebral cortex (rev. Le Novère et al. 2002).

The existence of overlapping patterns of expression of these subunits has thus hampered the elucidation of the role of a particular subunit (or subunits) in brain functions until mice deleted for specific nAChR genes became available (Picciotto et al. 1995, rev. Champtiaux & Changeux 2004). In parallel, gene targeting and homologous recombination experiments were carried out to create mice carrying mutations in nAChR genes associated with different diseases in humans. Several of them are relevant to this review, since they cause idiopathic epilepsies (Steinlein et al. 1995) and, as we shall see, may affect states of consciousness (Champtiaux & Changeux 2004; Steinlein 2004).

3. The integration of neuronal nicotinic receptors to the neuronal workspace architecture

The prefrontal areas of the neocortex are known to play a key role in higher cognitive functions (see Levin 1992, 2002; Sacco et al. 2004), and were postulated to be essential components of the conscious neuronal workspace architecture (Dehaene et al. 1998; Changeux 2004; Dehaene & Changeux 2004; see Bianchi 1921). An eventual contribution of nAChRs to this network was investigated at the cellular and synaptic levels (figure 3).

Early work on synaptosomal preparations identified presynaptic nAChRs in vitro (rev. Wonnacott et al. 1995). Yet their cellular distribution and function in neuronal networks had not been established. In a first electrophysiological approach (Vidal & Changeux 1989), extracellular recordings of field potentials and unit discharges were taken from the superficial layers of slice preparations of the prelimbic area of the prefrontal cortex. They revealed excitatory effects of nicotine that were blocked by neuronal bungarotoxin, and thus possibly involved a3b2 or most likely a4b2 nAChRs. In subsequent experiments, still on in vitro slices (Vidal & Changeux 1993), intracellular recordings were made from pyramidal neurons located predominantly in layers II–III (where nicotine-binding sites were found concentrated; Clarke 1993), which belong to the long-axon neurons, subsequently postulated by the neuronal workspace hypothesis (Dehaene et al. 1998; Dehaene & Changeux 2004). Iontophoretic application of nicotine near the recording site increased the amplitude of

*Phil. Trans. R. Soc. B (2006)
channel blocker tetrodotoxin (TTX) was found to block these effects. The specificity of TTX is strictly limited to voltage sensitive Na$^+$ channels. These channels are thus mobilized in the coupling between nAChR and the terminal GABAsites release. The nAChRs involved, thus, are not located on the nerve ending. The term ‘preterminal’ was coined to qualify the occurrence of these nAChRs along the axon and the observation confirmed by other groups (McMahon et al. 1994; McGehee & Role 1996). The only plausible localizations of these axonal nAChRs are the ‘neurol de Ranvier’. Their precise role at this level is still not understood. As previously noted, the expansion of the neocortex and, in particular, of the prefrontal cortex in high mammals is accompanied by a massive increase of the white matter. A possible—but still highly speculative—function would be that they are the target of as yet unidentified cholinergic terminals that would exert global control of the state of activity of the long axons that compose the white matter and regulate, in a ‘paracrine’ or ‘volume’ manner, the states of consciousness mediated by the global workspace!

### 4. NICOTINIC RECEPTORS AND STATES OF VIGILANCE IN THE MOUSE

In mammals, including mice and men, two cholinergic systems exert activating effects on thalamocortical and cortical neurons: first, the pedunculopontine and laterodorsal tegmental nuclei that project to all thalamic nuclei, but have no direct projection to the cortex and second, the nucleus basalis that projects to the cortex and thalamic reticular nucleus. Acetylcholine release increases in the cortex and thalamic nuclei. Moreover, nicotine intake has been known to cause EEG desynchronization (Yamamoto & Domino 1965) and to increase arousal and attention (Lawrence et al. 2002; Levin 2002). Also, nicotine antagonists mccamylamine causes drowsiness and decreases the performance in attention-demanding tasks (Pickworth et al. 1997). Last, acetylcholine release increases in the rat prefrontal cortex during the performance of a visual attentional task (Passetti et al. 2000).

In order to evaluate the specific contribution of nAChRs to the regulation of sleep and wakefulness, mice lacking the β2 subunit (β2−/−; Picciotto et al. 1995, 1998) were investigated. The β2−/− mice are devoided of high-affinity nicotinic-binding sites in the brain (Picciotto et al. 1995; Zoli et al. 1998), and exhibit a loss or reduction of nicotine-elicited currents in neurons from various brain regions (rev. Klink et al. 2001). Yet, they survive and reproduce without difficulties. Their respiration was monitored by plethysmography (Cohen et al. 2002, 2005) and their cerebral electrical activity recorded as polygraphic EEG or EMG (Lena et al. 2004).
In a first series of experiments, the respiration of adult β2−/− individuals was followed by plethysmography (Cohen et al. 2002). During sleep, the episodes of quiet sleep manifest themselves by regular large-amplitude respiratory movements, which contrast with the irregular transient movements of wakefulness. Hypoxic challenge in β2−/− mice (compared to wild-type (WT)) an increased ventilatory response with a reduction of transient movements, revealing a contribution of β2−/− nAChR to awakening under these conditions (Cohen et al. 2002).

The plethysmography method was adapted to newborn mice and their breathing and arousal responses challenged, following exposure to controlled doses of nicotine during pregnancy. Newborn WT pups exposed to nicotine exhibited unstable breathing and impaired arousal, deficits remarkably similar to those detected in pups lacking β2−/− nAChR. Chronic nicotine exposure thus, like the lack of β2−/− nAChR, impairs the arousal response to distress of newborn mice (Cohen et al. 2005). The newborn mice maternally exposed to nicotine thus offer a plausible model of sudden infant death syndrome (SIDS), which prevalence is increased several folds with smoking mothers.

The organization of sleep was monitored by polygraphic recordings (EEG, EMG) in adult β2−/− versus WT (Lena et al. 2004) (figure 4). First nicotine (1–2 mg kg$^{-1}$, i.p.) produced an increase in the time spent in the awake state in WT mice and correspondingly caused a rebound increase in sleep during the second hour after the injection. These two effects of nicotine were not observed in β2−/− mice. In contrast, under normal conditions, the β2−/− mice displayed the same amounts of waking, non-rapid eye movement (NREM) sleep and REM sleep as their WT counterparts. The endogenous physiological activation of β2−/− nAChR thus does not contribute critically to overall sleep regulation, but plays a discrete role in shaping sleep episodes. In particular, the β2−/− mice exhibited longer REM sleep episodes and a reduced fragmentation of NREM sleep by events referred to as ‘microarousals’ characterized notably by a transient increase in EEG power and an EMG activation. The lower frequency of microarousals during NREM sleep observed in β2−/− mice resembles the deficit in putative arousals reported above after a hypoxic challenge (Cohen et al. 2002). These results indicate that nAChRs are involved in the phasic expression of arousal promoting mechanisms (Halasz 1998; Terzano & Parrino 2000) observed in NREM sleep. They thus control transitions of brain circuits from the sleeping conscious state.

5. NICOTINIC RECEPTORS AND NOCTURNAL FRONTAL LOBE EPILEPSIES

Autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE) was the first congenital epilepsy found in humans to segregate as a single gene disorder (Schaefèr et al. 1994) and to be caused by point mutations within the genes coding for the α4 and β2 subunits of the neuronal nAChR (Steinlein et al. 1995; De Fusco et al. 2000; Phillips et al. 2001). ADNFLE is characterized by nocturnal motor seizures, which tend to cluster, and often occur several times at night. Seizures are mainly occurring during non-REM sleep, either shortly after falling asleep, or in the early morning hours. The main age of onset is within the first or second decade of age and the frequency of nocturnal attacks tends to decrease with advanced age and even to disappear. The first ADNFLE mutation identified α4 Ser248Phe (Steinlein et al. 1995) hits precisely the particular amino acid which was first identified by chlorpromazine in Torpedo ion channel transmembrane segment MII (Giraudat et al. 1986; Hucho et al. 1986). Several other ADNFLE mutations have been identified in α4 and β2 genes (Steinlein 2004). Even though their physiological phenotype may vary from one mutation to the other, there seems to be a general agreement on their ‘gain of function’ character (Revah et al. 1991; Bertrand et al. 1992; Devillers-Thiéry et al. 1992; Itier & Bertrand 2002), a phenotype that is simply, but not exclusively, accounted for by the MWC model (Changeux & Edelstein 2005, see §8a).

Interestingly, the seizures tend to occur in a stage of light sleep, where intense cyber microarousals activity takes place (Terzano & Parrino 2000), and in some ADNFLE cases, a direct temporal correlation was found between microarousals and epileptic manifestations (Zucconi & Ferini-Strambi 2000). Mouse models of several of the ADNFLE mutations have been created for α4 (but also for another α7) subunits (Labarca et al. 2001; Broide et al. 2002; Gil et al. 2002) by knockin technology. Mice homozygous for the α7-L270 T mutation die shortly after birth, while heterozygous animals are viable, but more sensitive to the convulsant effects of nicotine (Broide et al. 2002; Gil et al. 2002). The α4 homologues of known human ADNFLE mutations have been produced and displayed enhanced susceptibility to epileptic seizures (Tapper et al. 2004; Rodrigues-Pinguet et al. 2005).

In conclusion, alteration of the allosteric transitions of the nAChR may cause the production of epileptic seizures. Discrete molecular properties of a receptor molecule thus ‘gate’, in a bottom–up manner, and the transitions between distinct brain states are known to affect the subject’s states of consciousness.

6. NICOTINIC RECEPTORS AND THE CONTENT OF CONSCIOUSNESS

Already in the early 1990s, a rich body of pharmacological and behavioural observations underlined the importance of cholinergic pathways and their nicotinic component in working memory and attentional processes that rely on the contribution of the prefrontal cortex (figure 5; Nordberg & Winblad 1986; Levin 1992). To evaluate the role of nAChRs in these cognitive—we might even say today ‘conscious’—tasks, the effect of nAChR blockers was tested in the rat on delayed-response tasks (Granon et al. 1994, 1995). The performance on two distinct tasks was compared. First, a delayed-non-matching-to-sample task (NMTS) was used as reference. The tasks depend on the tendency of...
the animal to spontaneously alternate, on a second run, to the branch it did not visit on the first time. This behaviour makes sense on an evolutionary basis, since the rat will not visit a place where it already consumed the food reward. As a consequence, the delayed-matching-to-sample task (MTS), which consists in visiting the arm where the food reward was, is an effortful task. As anticipated, lesions of the prefrontal cortex selectively impaired the performance of the MTS task, caused errors and perseverations, while microarousals increased, EEG pw decreased.

Figure 4. Nicotinic receptors and states of vigilance: consequences of the deletion of the β2 subunit on microarousals (from Lena et al. 2004) and on the respiratory arousal of newborn mice chronically exposed in utero to nicotine (nicotine exposure destabilizes breathing; from Cohen et al. 2005).
preserving the performance of the NMTS tasks (Granon et al. 1994). Interestingly, injection of neuronal bungarotoxin into the prelimbic area of the prefrontal cortex produced a significant performance deficit in the MTS, but not in the NMTS. In contrast, an inhibitor of the muscarinic receptor, scopolamine, impaired performance in both tasks. nAChRs in the prefrontal cortex are thus necessary to the successful performance of the delayed-response tasks, requiring effortful processing for response selection (Granon et al. 1995). Is this a sufficient demonstration of the Turing test for consciousness?

Mice do not pass MTS versus NMTS tasks as easily as rats, possibly because, as we shall see, of their compulsive tendency to explore. The specific contribution of particular species of nAChR to such complex cognitive functions was nevertheless examined in mutant versus WT mice (Picciotto et al. 1995, 1998; Granon et al. 2003). The β2−/− mice were selected because both high-affinity nicotine-binding sites and electrophysiological responses to nicotine were almost completely absent from the brain (Picciotto et al. 1995). The Morris watermaze evaluates spatial orientation learning. The performance of β2−/− mice in the test did not differ from that of WT mice when tested both on the visible platform task and on the hidden one, indicating an intact spatial memory (Picciotto et al. 1995).

Retention of an inhibitory avoidance response was then assessed because of its known sensitivity to nicotine administration. The mouse was placed in a well-lit chamber of a shuttle box, and the latency to enter the adjacent dark chamber measured. Upon entry to the dark chamber, a mild inescapable foot shock was delivered and nicotine (versus saline) injected into the mouse. After 24 hours, retention was assessed by measuring the latency to enter the dark chamber. Interestingly, treatment with nicotine after foot shock consistently aided retention and nicotine administration was completely ineffective in β2−/− mice. Control experiments confirmed that the β2−/− mice did not differ from their non-mutant siblings by sensory sensitivity, pain threshold or increased emotionality (see also Granon et al. 2003; S. Granon, unpublished work 2006). The effect of nicotine on memory retention may only indirectly concern the issue of consciousness in the mouse, yet an unexpected though relevant observation was made. Retention latency was significantly longer for mutant mice than for their non-mutant-injected siblings (Picciotto et al. 1995).

The learning abilities of the β2−/− versus WT mice were further investigated in a maze equipped with four dissociable arms. Animals learned to reach food from any location in a given configuration of the arms. The configuration of the arms was then changed and the starting point as well, without changing the position of the goal arm in the absolute space. Yet, here as well, quite unexpectedly, β2−/− mice learned the task more rapidly than WT mice (Granon et al. 2003). Explanation of this paradoxical behaviour, noted in both the multiple configuration maze and the avoidance learning task (Picciotto et al. 1995; which, in addition, becomes amplified with age; Zoli et al. 1999; Caldarone et al. 2000), came out from a closer analysis of the organization behaviour in β2−/− versus WT mice.

Exploratory activity is a spontaneous behaviour that does not involve explicit reinforcement (Renner 1990; Thinus-Blanc 1996; Poucet & Hermann 2001; even though it mobilizes, as we shall see, endogenous reward processes). It serves to gather and store spatial information, which allows allocentric coding of space, itself, necessary for flexible navigational processes. In one experiment, mice were placed in the centre of an empty arena and their spontaneous trajectories recorded with a camera fixed to the ceiling and connected to a video track system. The system was set up to break down each trajectory into navigation (large movements at fast speeds 14.4 cm s−1), fast exploration (small movements at a speed between 6.8 and 14.4 cm s−1) and slow exploration (small movements at a speed up to 6.8 cm s−1). A complementary ‘automatic’ analysis by ‘symbolic’ quantification of trajectories (Faure et al. 2003) was carried out in parallel. It included the definition of the position of the animal, its two-dimensional path and its instantaneous velocity. Statistical analysis revealed a modification in the organization of displacements in the β2−/− mice. The balance between navigation and exploration was shifted in the β2−/− mice in favour of navigation to the detriment of more precise exploration of the
environment. The symbolic analysis further revealed a reduction of the trajectories in the centre of the arena for the β2−/− mice, due to high-speed crossing without slowing down in the centre. In other words, the β2−/− mice were more rigid and exhibited less behavioural flexibility than the WT (Granon et al. 2003; Maskos et al. 2005), and the apparent faster learning in the above-mentioned situations simply results from the absence of time-consuming exploration of the apparatus used.

These aspects were further evaluated by placing the mice in more sophisticated environments. First, a single object was presented in the arena. No difference of exploratory behaviour was noted, indicating no gross sensory or memory impairment in the recognition process. But, when two objects were placed in the environment, the β2−/− mice exhibited an exploratory behaviour different from WT mice. Symbolic analysis revealed a striking reorganization of the transitions between sequences of action in the WT, which was absent in the β2−/− mice. WT mice increased the number of trajectories between the two objects, while β2−/− mice did not; in other words, they display a spontaneous flexibility of behaviour, an increased curiosity (Denton 2005), that requires the presence of β2*-nAChR (Granon et al. 2003). Last, the cognitive behaviour of the β2−/− mice was analysed further in conflict-behaviour paradigms that require strategic choice. Two situations were considered. In a first series of experiments, the multiple arms maze was used and the same learning protocol, except that objects were inserted in the maze as attentional distractors. WT mice reacted to novelty by increasing exploratory activity, whereas, again, β2−/− mice did not show this adaptative flexibility to a change in the environment.

In a last series of experiments, the sequences of so-called ‘social’ interactions were examined between a test resident mouse and a social intruder. In short, among the diverse types of interactions tested, the β2−/− mice showed significantly higher rates of approach behaviours, and lower rates of escape behaviours. In other words, the β2−/− mice display more ‘automatic’ interactions: faster rates of approach and impaired capacity for interrupting ongoing behaviour (Granon et al. 2003).

Lesion experiments (S. Granon unpublished work) reveal that mice with prefrontal lesions do not show significant alteration in spatial learning, but display evident deficit in conflict-resolution situations in the maze paradigm with an unexpected object. The prefrontal lesion creates a behavioural phenotype that displays several features in common with the loss of β2*-nAChR, which affects behavioural flexibility and ability to hierarchize competing motivations.

Along these lines, different related experimental paradigms, referred to as delay versus trace conditioning (Clark et al. 2002), were applied to the mouse by another group, yet without the above pharmacological and genetic analysis (Han et al. 2003a). They will be mentioned here for methodological reasons. In trace conditioning, a time gap was introduced between the end of the conditioned stimulus (such as a tone) and the unconditioned stimulus (like a foot shock), while in delay fear conditioning, the conditioned stimulus was immediately followed by the unconditioned stimulus. A visual distraction of flash of light was introduced at randomly selected intervals. Interestingly, the distraction interfered selectively with trace but not delay conditioning. Trace conditioning, which is thought as requiring conscious awareness, was expectedly found associated with increased neuronal activity in the anterior cingulate cortex and was selectively impaired by lesion of this structure in the medial prefrontal territories at variance with delay conditioning (Han et al. 2003a,b). Interestingly, similar experiments have been carried out in humans, in whom an explicit requirement for attentive awareness was demonstrated for the trace, but not delay conditioning task (Carter et al. 2003).

Moreover, the phenotypes observed with the β2−/− mice and characterized by alterations of behavioural flexibility and adaptive behaviours coupled with unimpaired memory and anxiety may reproduce the cognitive impairments observed in human diseases (Granon et al. 2003); for instance, in attention deficit hyperactivity disorders (ADHD; Granon & Changeux in press) and even in autism (Frith 2001; Andres 2002; Booth et al. 2003). Both groups of patients show a reduced activation of prefrontal and cingulate cortices (Frith 2001; Giedd et al. 2001), and autistic children further show a substantial decrease of α4β2-containing nAChR in their prefrontal cortex (Perry et al. 2001). The β2−/− mice thus might represent a useful animal model for the study of ADHD and autism (Granon et al. 2003; Granon & Changeux in press; but see §8).

In conclusion, mice under these experimental conditions do far more than to simply react to sensory information. They engage in complex extended behaviours geared towards far removed goals. They use processes that override or differentially select routines (Thinus-Blanc 1996) and, according to Denton (2005), orchestrate locomotor behaviours according to ‘conscious intentions’, a behaviour consonant with, though far more elaborate, than Koch’s Turing test for consciousness. Further studies are of course needed to further complement, I would even say justify, these conclusions (see §8).

7. THE JOINT RECOVERY OF EXPLORATORY BEHAVIOUR AND REWARD FUNCTION BY TARGETED RE-EXPRESSION OF nAChR SUBUNIT

To further understand the specific role of nAChR subunits in mediating the effects of nicotine and endogenous acetylcholine on higher cognitive functions in the mouse, a method of selective re-expression of a nAChR subunit by stereotoxic injection of a lentiviral vector in a defined territory of the brain of a knockout (KO) mouse was developed (Maskos et al. 2005). Lentivirus-based expression systems, initially developed for gene therapy purposes (Naldini et al. 1996), provide several advantages over other virus-based in vitro transgene expression strategies. In particular, they are capable of stable integration into the genome of non-dividing cells, such as neurons, facilitating the potential for long-term, stable transgene expression. The lentiviral expression vector was derived

*Phil. Trans. R. Soc. B (2006)*
from the pHR expression vectors of Naldini et al. (1996), yet with several modifications: it contained a bi-cistronic cassette simultaneously expressing the β2 subunit and the enhanced green fluorescent protein facilitating efficient detection of transduced cells (Maskos et al. 2002). The vector was stereotactically injected into the VTA of mice carrying β2-subunit deletions. The ligand-binding nAChR protein was found quantitatively re-expressed in DA-containing neurons of the VTA, using epibatidine as a ligand. Its functionality was assessed by recording in vivo the effect of nicotine on the electrophysiological activity of DA neurons in the VTA. The well-established rapid 1.5-fold increase in firing frequency upon intravenous injection of nicotine, which disappeared in the β2−/− mice, was recovered. However, instead of lasting on average nearly 10 min as in WT, this effect did not persist for more than 2 min. These data show that the re-expression of the β2 subunit exclusively in the VTA suffices to recover an effect of nicotine on DA neurons. Furthermore, using in vivo intracerebral microdialysis in awake, freely moving mice, nicotine-elicited DA release from the nucleus accumens (NuAcc) was found completely restored by the re-expressed β2 subunit in the β2−/− mice.

To assess whether the nicotine-elicited responses observed in the VTA and NuAcc following re-expression of the β2 subunit were sufficient to support nicotine reinforcement, an intra-VTA nicotine self-administration paradigm in the mouse, as described for morphine self-administration (David & Cazala 1994), was used. WT mice exhibited a clear nicotine-seeking and self-administration behaviour, which disappeared in the β2−/− mice. Interestingly, the lentivirus-injected mice recovered intra-VTA nicotine self-administration behaviour and improved their performance over learning sessions, yet they displayed a delay in the acquisition of self-administration and their discrimination performance was slightly lower than that observed for WT mice. Therefore, re-expression of β2−nAChR in the VTA is not only necessary, but also sufficient to re-establish sensitivity to nicotine reward in drug-naive mice.

Having established that VTA re-expressed β2−nAChR respond to nicotine in vivo, the spontaneous exploratory behaviour of the injected mice was examined. As mentioned, β2−/− mice exhibit modified spatio-temporal organization of displacements, with increased navigatory and decreased exploratory behaviour (Granon et al. 2003). Interestingly, the lentivirus-injected mice showed a selective restoration of exploration, bringing this measure up to the level of WT without a significant modification of navigation. In addition, they exhibited transitions between fast and slow movements in the central portion of the arena that were similar to WT. The targeted expression of β2−nAChR in the VTA thus generated, at least, a partial dissociation between exploratory and navigatory behaviour.

These results demonstrate that nAChRs in neurons originating in the VTA and/or their axonal projections suffice for the differential restoration of the cognitive executive function, as studied in our paradigm. This behaviour therefore appears to be mediated by endogenous acetylcholine acting on β2−nAChR expressed on VTA neurons and their axonal projections, in particular to the NuAcc. It suggests the implication of DA as the endogenous neuromodulatory substance for this executive function.

In conclusion, this study shows that re-expression of functional β2−nAChRs, specifically in the VTA and its axonal projections, suffices to simultaneously restore the self-administration of nicotine and the ability to perform an executive behaviour under endogenous cholinergic modulation, thus bringing a first experimental demonstration of the proposed theoretical mechanisms, implying a causal link between cognitive learning, motivation and reward processes (Dehaene & Changeux 1989, 2000; see also Damasio 1995). Moreover, it underlines a rather unclassical, yet predicted (Dehaene et al. 1998), but still hypothetical connection between reward neurons and consciousness.

8. IS THE MOUSE A USEFUL ANIMAL MODEL TO INVESTIGATE CONSCIOUSNESS?

The issue of mouse consciousness is not simply philosophical (Block 1990). The obvious limitation of the mouse model in the scientific study of consciousness is its reduction to a minimal level and thus of the impossibility to investigate some of its most characteristic human features like self-recognition and reflective consciousness. On the other hand, as we saw, it gives access to a large spectrum of methods, in particular to recombinant DNA technologies (gene deletion, gene exchange and targeted gene expression), which provide the means to explore, at the molecular, cellular, neuronal networks and behavioural levels what I referred to as the fundamentals of minimal consciousness. In this paper, data on the contribution of neuronal nAChRs to these fundamentals have been reviewed. The main conclusions may be summarized as follows.

(a) States of vigilance: sleep & arousal, epilepsy, general anaesthesia

Multiple states of vigilance (such as wakefulness, sleep, coma, general anaesthesia, epileptic seizures, etc.) and the regulation of their reversible transitions occur in the mouse as in all mammals, including humans; the circadian sleep–waking cycle is controlled in these species by brainstem reticular formation and intralaminar nuclei of the thalamus (Bogen 1995; Jones, E. 1998; Jones, B. 1998), with complex patterned releases of neuromodulatory substances. The use of mice deleted for the β2-subunit gene gave us the opportunity to demonstrate the positive contribution of β2−nAChR to the phasic expression of arousal promoting mechanisms by endogenous acetylcholine (Cohen et al. 2002; Lena et al. 2004). Similar phenotypes as those noticed in β2−/− mice are observed after chronic in utero exposure of the fetus to nicotine. These mice thus offer, in addition, a plausible animal model of SIDS, the prevalence of which is known to increase in smoking pregnant women (Cohen et al. 2005).

Furthermore, as discussed, knockin mice where point mutations were introduced in the α4-subunit gene by homologous recombination may constitute relevant

Phil. Trans. R. Soc. B (2006)
though still primitive animal models for human familial frontal lobe epilepsies (Labarca et al. 2001; Steinleijn 2004). Interestingly, these mutations cause rather paradoxical ‘gain of function’, pleiotropic, phenotypes characterized by higher affinity of acetylcholine and nicotine, loss of desensitization and conversion of the response to some nicotinic antagonists into an activatory response (Revah et al. 1991; Bertrand et al. 1992; Devillers-Thiéry et al. 1992). These complex phenotypes are simply accounted for by the MWC mechanism for signal transduction. Homologous gain of function mutations were also identified as responsible for congenital myasthenic syndromes (Engel & Sine 2005), and as well viewed as either editorially changing the intrinsic properties of the ion channel in the different conformation of the nAChR or altering the kinetic parameters of the transition from the resting to the active open-channel conformation (Edelstein et al. 1997; Changeux & Edelstein 2005). These observations unambiguously demonstrate that the allosteric properties of the nicotinic receptor protein may gate, when altered, the transition from a state of slow-wave sleep to an epileptic state of vigilance. In a more general manner, bottom-up causal determination may exist between the molecular dynamics of a critical receptor protein and the transitions between states of vigilance.

Last, nAChRs are inhibited by some volatile anaesthetics and ketamine (Tassonyi et al. 2002). On muscle nAChR, the latter acts mostly as open-channel blockers, but on neuronal nAChRs, it may also bind to allosteric sites located outside the ion channel where they inhibit activation and/or enhance desensitization. The direct role of brain nAChR (α4β2) as targets of general anaesthetics to cause losses of consciousness has been challenged (see Tassonyi et al. 2002). On the other hand, their potential contribution to the analgesic component of general anaesthesia has received support from both pharmacological and in vivo animal model experiments (Cordero-Erausquin et al. 2000; Cordero-Erausquin & Changeux 2001), even if GABAA receptors are thought to be a primary target for general anaesthetics. The mouse and its mutants may thus become convenient animal models to investigate the pharmacology of general anaesthesia. In many instance, both the chemistry and molecular biology of the various states of vigilance are currently under investigations.

(b) Content of consciousness: behavioural flexibility and the rudiments of ‘intentionality’ in the mouse?

Classically, in the waking conscious state, a ‘graded global integration’ of multiple cognitive functions is assumed to take place that yields an unified representation of the world, of the body and of the self’ (modified from Hobson 1999, see also Changeux 1983; Edelman 1989, 2002, 2004). This notion, which deals now with the content of consciousness, has been approached, still in a rather rudimentary form, by the qualitative and quantitative analysis of a spontaneous behaviour displayed by mice and rats and referred to as exploratory behaviour (Granon et al. 2003; Maskos et al. 2005). These behaviours were up to now investigated in the context of animal spatial cognition (Renner 1990; Thinus-Blanc 1996; Poucet & Hermann 2001), but without explicit reference to consciousness. I will take the risk to discuss here the notion that they might be plausibly relevant to the study of what I referred to as minimal or ‘level 1’ (Barresi & Moore 1996) consciousness. Exploratory behaviour expresses, in one way or another, a reaction to novelty. It may also, to some extent, implement the spontaneous (Dehaene & Changeux 2005) or basal activity of the brain (Gusnard & Raichle 2001). A mouse or a rat placed in a new environment displays intense investigatory activity: contact with objects, sniffing, stopping, rearing up on hind paws and so on. In humans, young children, for example, display a need to seize and touch any unfamiliar object and extensively visit their environment as soon as they can crawl and walk. Exploratory activity, in general, decreases over-time—it habituates—but it also dishabituates when some spatial relationships between the elements of the situation are modified. It is currently assumed that the organism detects spatial novelty by comparing the actual perceived arrangement and a representation or stored ‘internal model’ of the initial situation (Thinus-Blanc 1996). Intensive studies of the neurobiological bases of exploratory behaviour are consistent with a multistage processing of spatial inputs: from the visual cortices for the perception of local views, the associative parietal cortex for the extraction of spatial invariants, the prefrontal cortex for holding the information to plan and execute a trajectory and the hippocampus as a comparator (or detector) of mismatches between stored representations and currently perceived situations (Poucet 1993). In a more general manner, exploratory activity, as its habituation and dishabituation in reaction to spatial changes, is viewed as relying on multisensory perception of the environment matched with movement-related information or, in other words, as an ‘active information-seeking structure’ (Neisser 1976), in constant updating of its representations, a manifestation of behavioural flexibility (Thinus-Blanc 1996; Granon et al. 2003). In its very nature, exploratory behaviour would then be organized along an egocentric reference frame (e.g. go ahead, turn left) and sequentially ordered (go ahead, then turn left) bearing of the goal determined on the basis of ‘abstract’ exocentric representation of the physical world (see also Berthoz 2003). In a broader philosophical context, exploratory behaviour might be tentatively considered as a first sign of ‘intentionality’ in a simple animal species (for a discussion of classical literature, see Boakes 1984).

Indeed, according to the Viennese philosopher Brentano, intentionality would be the signature of consciousness as being the tractable feature of recognizing a difference between one’s own thoughts and the sensory information coming from the outside. In other words, intentional representations would differ from other purposive thoughts where the aim is clearly in view. Koch (2004), as mentioned earlier, has suggested as a Turing test for conscious behaviour in animals the ability to maintain information through a delay period. In light of what was mentioned about exploratory behaviour, one may further say that, with the course of exploration, as well as in the case of trace conditioning.
(Clark et al. 2002; Han et al. 2003a), the organism has to keep the information ‘on line’ through delays, but also to exploit delays to elaborate, select and evaluate plans, organize a programme of actions, which would ultimately let it respond in a natural environment to inner motivations as, for instance, Denton (2005) primordial emotions. Accordingly, the so-called minimal Turing test should be re-formulated as ‘the maintenance of information through delay plus selection of a relevant plan’ (Changeux 2005).

Denton (2005) has underlined that, in the wild, simple species like mice, and sometimes, under special circumstances, human beings themselves might be under the ‘totalitarian occupancy’ of primordial emotions which signals that their very existence is threatened. These primordial emotions basically organize the primitive functions necessary for survival: hunger, thirst, compulsion to sleep or to find a sex partner, which, in a natural environment, require constant efforts of attentive awareness. At some stage, the drive caused by some physical impairment or imperious sensation—for instance thirst—may create some organized effort to achieve a ‘particular purpose’, an intentional behaviour—for instance satisfy thirst. But, taking into account stored past memories, such as the map of the environment where, for example, a source of water would be located, or confronted with the unpredicted sight of a predator, the organism may have to display behavioural flexibility, define action priorities which match internal motivations and actual environment conditions to reach some kind of ‘mental synthesis’ and take the decision, for instance, to flee (or to keep going). For the evolutionary biologist and for the philosopher, the behaviour of the ‘simple-minded’ mouse may thus become an outstanding topic to approach the neurobiological bases as well as the origins of minimal consciousness. At least the question is raised!

More specifically, the observations that, in the mouse, the exploratory behaviour may be dissociated from the more automatic navigatory behaviour by the deletion of the β2-subunit gene of the nAChR (Granon et al. 2003; Maskos et al. 2005) illustrates: (i) the specialization of neural circuits engaged in such executive ‘conscious’ functions; (ii) the ‘gating’ of these functions (which mobilize, in particular, the prefrontal cortex) by nAChRs activated through endogenously released acetylcholine; last, (iii) the joint recovery of exploratory behaviour and reward functions by targeted re-expression of the β2 subunit in the VTA DA nucleus. This important result points to an issue, insufficiently recognized in altogether behavioural, psychological and philosophical studies on cognition in general and consciousness in particular, that a fundamental connection may exist between the activity of reward processes and the selection of plans, the maintenance of a goal, or in a philosophical and even more speculative manner, the content of consciousness in an intentional behaviour.

On the clinical side, these conclusions may also lead to a plausible interpretation of drug (nicotine) addiction. It may be viewed as an escape from the ‘voluntary’ control of drug-taking behaviour, for instance, as a consequence of the disconnection of a reciprocal-loop linking prefrontal cortex, DA neurons and striatum, thus uncovering the compulsive non-conscious aspect of drug addiction (for discussion, see Gutkin et al. 2006). Moreover, as noted earlier, the β2—/— mice which are compulsively navigating without pausing for exploration may offer an animal model for human ADHD behaviour, in which hyperactivity symptoms are known to improve with nicotine treatment (Shytle et al. 2002; rev. Granon & Changeux in press).

(c) Social relationships in the mouse?

Behavioural scientists together with philosophers have primarily considered intentionality in the framework of social relationships. Social organisms, including humans, represent intentional relations of themselves and other agents, yet at different levels. They unambiguously distinguish their own intentional relations (or first person information) from the qualitatively different informations available about another agent’s intentional relations (or third person information).

In this respect, one should remember that the analysis of the exploratory behaviour leads to the distinction between allocentric and exocentric motor behaviour (Thinus-Blanc 1996) pointing to the still highly speculative occurrence of a ‘self’ in the mouse, yet oriented towards the outer physical world. Furthermore, the so-called ‘social relationship’ we examined in WT and mutant mice were reduced to rather crude sequences of interactions between a resident mouse and a social intruder of the same sex. Higher rates of approach behaviour and lower rates of escape behaviour were noticed in β2—/— mice compared to WT, as if again the β2—/— mice were impaired in their behavioural flexibility for interrupting ongoing behaviours (Granon et al. 2003). No evidence was found, at this stage, with the mouse, for comparability between the actions of self and others, no sign of imitating goal-directed activity nor of ‘understanding’ the view point of others (Barresi & Moore 1996). In other words, the presently available evidence does not support the occurrence of authentic social relationships in the mouse. The human infant at birth seems already at a stage more advanced than the adult mouse in this respect. Indeed, as mentioned earlier, he or she may already distinguish between his own and others’ movements, in particular, by touch and he/she displays rudiments of imitations (Meltzoff & Gopnik 1993, but see Barresi & Moore 1996; Lagercrantz 2005). Thus, the mouse cannot be a good animal model to investigate intentional relations and social understanding, where highest level is reached exclusively in humans with the theory of mind. On the other hand, it may serve as baseline to define the neural circuits mobilized by these intentional relationships in higher mammals and humans. It may thus hardly become a reliable model of autism (at variance with what was suggested before). If autistic children show a strong need for stereotyped routines in daily life that evokes the dominant navigatory behaviour of the β2—/— mice, their most characteristic deficits concern the understanding of other minds (Leslie & Frith 1990; Baron-Cohen 1991). They do not show social pretend play, which in normal children develop during the first

Phil. Trans. R. Soc. B (2006)

Downloaded from http://rstb.royalsocietypublishing.org/ on October 10, 2016
2 years of life (Sigman & Ungerer 1984), neither do they express joint attention and adequate imitation. They fail standard theory of mind tasks. According to Barresi & Moore (1996), they would never attain an understanding of intentional relations at the level 2, ‘recursive consciousness’, level. In other words, following this reasoning, the WT mouse would already constitute some kind of animal model analogue of the autistic child.

(d) The neuronal workspace hypothesis, blindsight and the laboratory mouse? As discussed by Baars (2001) and Dehaene & Changeux (2003), in humans the standard observational index of consciousness is an accurate, verifiable report by pressing a button or any other voluntary response mediated or not by language. Reporting responses have also been used with some animal species. It is, for instance, the ‘commentary key’ report method developed by Weiskrantz (1991) and Cowey & Stoerig (1995) to demonstrate blindsight in macaque by forced-choice responses. Yet a still unanswered question is whether or not reportability would be demonstrated in simpler species like the mouse. Since it does not require any kind of social relationship, it looks plausible, but a reliable assay still needs to be invented. On the other hand, if the neuronal workspace hypothesis was initially designed to account for access to consciousness and reportability, it obviously deals with important features of minimal consciousness.

It accounts, for instance, for the active maintenance of ‘abstract rules’ through top-down amplification and the flexible control of tasks that require a novel interconnection of existing processors as it typically occurs in the Stroop task (rev. Dehaene & Changeux 2004). It deals, a fortiori, with the active maintenance of information during a delay period and, thus, with the ‘Turing test’ for conscious behaviour of Koch (2004) (see Han et al. 2003a). Even though the relevant simulations have not been carried out, the neuronal workspace architecture would adequately fit exploratory behaviour and offers an appropriate mechanism for the ultimate stage of spatial processing as introduced by Poucet’s model (1993), where unified location-independent representation with one unique reference direction is being built.

The model also offers simple explanation for the observed ‘disconnection’ between exploratory and navigatory behaviour. It may, for instance, result from a deficit of the VTA–NuAcc pathway activity, within the more elaborate circuit linking this system to the ‘prefrontal workspace’. Such an interpretation may hold as well for autism as long as it, primarily, affects primary consciousness and secondarily social understanding. Alternatively, as suggested (Changeux & Edelstein 2005), it may affect the development of the prefrontal workspace itself. The formation of the long-range axonal projection of the supragranular pyramidal neurons might be highly vulnerable to the process of epigenesis by selective stabilization of synapses (Changeux & Danchin 1976), as found with the visual pathways (Rossi et al. 2001; Grubb et al. 2003; Mrsic-Flogel et al. 2005). This interpretation may fit for the brain phenotype of the adult β2−/− mouse, but also, still in a speculative manner, for that of mentally retarded children, including autistic children (but see §8c), where postnatal cortical development would be altered by mutations affecting synapse formation and/or stabilization (X-fragile (Mandel & Biancalana 2004), autism (Bourgeron), Retts syndrome (Neul & Zoghbi 2004)).

Functional brain imaging from the mouse and its mutants are still at the stage of being developed (S. Suarez & S. Granon, unpublished work). In humans and monkeys, the available MRI and PET data are consistent with the observed link between cortical areas which are found active in conscious effortful tasks and areas particularly rich in workspace neurons with long-distance connections (i.e. prefrontal, parietal and superotemporal, cingulate entries; Goldman-Rakic 1988; Pardo et al. 1990; Cohen et al. 1997; Paus et al. 1998). These areas show the most intense and consistent spontaneous activity in the awake state (Gusnard & Raichle 2001; Mazoyer et al. 2001) and the greatest drop in metabolism during various types of transitions away from the awake state, during anaesthesia, sleep, coma or vegetative state (Maquet & Phillips 1998; Fiset et al. 1999; Laureys et al. 2000; Paus 2000; Balkin et al. 2002; Shulman et al. 2003). These areas also show the greatest contrast between conscious versus non-conscious processing of visual stimuli (Dehaene et al. 2001; Dehaene & Naccache 2001; Marois et al. 2004).

These studies should thus offer the opportunity to further dissect, already with the mouse, the above-mentioned multiplicity of states and of levels in consciousness. This might be of considerable help for the cognitive neuroscientist and the philosopher to clarify situations or experimental paradigms, in which the word consciousness is used with rather different meanings (for instance, Block 2005; Fiset et al. 1999; Paus 2000; Lamme 2003; Zeki 2003; Dehaene & Changeux 2004). The use of neurocomputational models might, in this respect, be of some help (Dehaene & Changeux 2004, 2005). Indeed, one may view the various graded states of consciousness (from deep anaesthesia and coma to full awareness) as directly related to the spontaneous activity of recurrent thalamo-cortical loops and reticular thalamic nuclei (see Llinas & Pare 1991; Steriade et al. 1993) under brainstem control. Simulations help to characterize several states of activity (Dehaene & Changeux 2005): (i) one of quiet vigilance with moderate spontaneous activity; I may tentatively suggest—at my own risk—that it suffices for the access of Zeki’s levels of ‘micro-’ and ‘macro-consciousness’ for colour perception or visual motion; (ii) another state that one may refer to as full conscious access occurs at higher levels of the neuromodulation parameter, where spontaneous ‘ignited’ states appear originating from higher association cortices and where sensory inputs become self-amplified above a threshold; again, it may account for access to Zeki’s ‘unified consciousness’. On the other hand, the model predicts that the content of consciousness builds up from the processing of both ‘exogenous’ sensory signal and ‘endogenous’ spontaneous activity as long as they give rise to above threshold ignition; in other words, considerable processing may remain...
below access to consciousness. This is expected to be the case for Zeki’s ‘micro- and macro-levels’ for processing of visual perception as well as for all processes of exploratory behaviour which in Poucet’s model do not reach the prefrontal cortex. These different states or levels in the WT mouse would therefore display electrophysiological signals of ignition in its neuronal workspace under conditions of exploration, while these ignition signals would disappear in the β2−/− mice. In other words, the β2−/− mice would be in a state of ‘quiet vigilance’ below threshold. Could it be called using Koch metaphor ‘a mouse zombie’?

In the nineteenth century (see Clarke & O’Malley 1968) it has been confirmed since the famous scientific studies on human consciousness, in particular to a re-evaluation of some of the key paradigms of the studies may have, as a principal consequence, given rise to consciousness. It may become a convenient model to consciousness. Would the mouse ever become the Drosophila of minimal consciousness?

I thank M. Besson, S. Dehaene, P. Faure, S. Granon, H. Lagercrantz and B. Molles for helpful comments. Support was provided by Centre National de la Recherche Scientifique, Collège de France, Ministère de la Recherche, Association pour la Recherche contre le Cancer, European Communities.

REFERENCES


Grubb, M. S., Rossi, F. M., Changeux, J. P. & Thompson, I. D. 2003 Abnormal functional organization in the dorsal lateral geniculate nucleus of mice lacking the beta2 subunit of the nicotinic acetylcholine receptor. Neuron 18, 8008–8018.


Huchu, F., Oberthur, W. & Lottspeich, F. 1986 The ion channel of the nicotinic acetylcholine receptor is formed by the homologous helices MII of the receptor subunits. FEBS Lett. 205, 137–142. (doi:10.1016/0014-5793(86)80881-X)


Pouget, B. 1993 Spatial cognitive maps in animals: new hypotheses on their structure and neural mechanisms. Psychol. Rev. 100, 163–182. (doi:10.1037/0033-295X.100.2.163)


