

Brain states and circuit mechanisms underlying sleep and general anesthesia

by

Laura D. Lewis  
B.Sc., McGill University (2008)

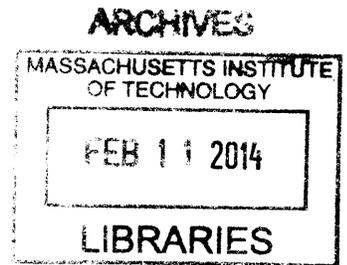
Submitted to the Department of Brain and Cognitive Sciences  
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2014



© 2014 Massachusetts Institute of Technology. All rights reserved.

Signature of Author .....  
Department of Brain and Cognitive Sciences  
Dec. 13, 2013

Certified by .....  
Emery N. Brown, M.D., Ph.D.  
Professor of Computational Neuroscience, Edward Hood Taplin Professor of Medical  
Engineering, and Warren M. Zapol Professor of Anaesthesia  
Thesis Supervisor

Accepted by .....  
Matthew A. Wilson, Ph.D.  
Sherman Fairchild Professor of Neuroscience  
Director of Graduate Education for Brain and Cognitive Sciences



# **Brain states and circuit mechanisms underlying sleep and general anesthesia**

by

Laura D. Lewis

Submitted to the Department of Brain and Cognitive Sciences on December 13, 2013  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Neuroscience

## **Abstract**

During sleep and general anesthesia, the brain enters a state of decreased arousal and consciousness is transiently suspended. How this transition occurs is a fundamental and unsolved question in neuroscience. The neural dynamics that disrupt consciousness have not been identified, and the circuit mechanisms that generate these dynamics remain unknown. Furthermore, understanding the neural basis of sleep and anesthesia is key to improving clinical monitoring of patients undergoing general anesthesia and to advancing treatments of sleep disorders and neurological conditions such as coma. In this thesis, I combine intracranial electrophysiology in human subjects with optogenetic manipulation of thalamocortical circuits in mice to identify the neural dynamics underlying sleep and anesthesia. I first show that loss of consciousness during propofol general anesthesia is associated with the abrupt onset of slow oscillations that disrupt cortical networks. I then demonstrate that activation of the thalamic reticular nucleus generates slow wave activity and decreases arousal state, identifying a causal mechanism that generates physiological and behavioral signs of sleep. Finally, I study patients undergoing deep general anesthesia at levels corresponding to medically induced coma, and show that this state is marked by local cortical dynamics consistent with impaired cerebral metabolism. Taken together, these results identify a set of neural dynamics associated with unconscious states, and demonstrate specific mechanisms for how they disrupt brain function. These findings provide new insight into the neuroscience of arousal states, and suggest clinical approaches that could improve patient care.

Thesis supervisor: Emery Brown, M.D., Ph.D.

Title: Professor of Computational Neuroscience, MIT

Edward Hood Taplin Professor of Medical Engineering, MIT

Warren M. Zapol Professor of Anaesthesia, Harvard Medical School

Anesthetist, Massachusetts General Hospital

## Acknowledgments

I am incredibly grateful to the many people who have given me so much help and support. First, Emery Brown, my thesis advisor, has been an incredible mentor over the past five years. His scientific insight, creativity, and generosity made my graduate school experience exciting and fulfilling. Second, Patrick Purdon, who advised me on all the clinical studies, has taught me so much, and I feel lucky to have worked with such a brilliant scientist and engineer. I'm also very grateful to my thesis committee members Chris Moore, for his conceptual insights and extremely useful advice, and Matt Wilson for many discussions that helped shape our experiments and analyses.

I would also like to thank the patients who participated in our studies and their families. These studies could not have taken place without their generosity.

A huge thanks goes to our many collaborators at MGH who helped execute this research. Most importantly, Syd Cash has devoted many hours of his time to these projects and has provided insightful perspective into each study we carried out – these studies would never have been possible without his help. Jake Donoghue, Mia Borzello, and Jason Naftulin were incredibly helpful with running experiments, and I can't imagine trying to do this without them. I am also very grateful to all of our clinical collaborators: Emad Eskandar, Bob Peterfreund, Kristy Nordstrom, Kara Houghton, Seun Akeju, and the many clinical staff who have helped throughout the past few years.

A highlight of grad school has been working in the Brown and Purdon labs, filled with so many amazing people. Aaron Sampson was integral to getting the clinical studies running, and his energy and positivity made our early morning recordings way more fun. Francisco Flores has been a great collaborator, helping me to transition into the world of optogenetics and finding solutions to every problem. Rob Haslinger was a great friend and teacher, who spent many hours helping me learn statistical techniques. Eran Mukamel and ShiNung Ching provided invaluable guidance on analyzing our intracranial data, and Veronica Weiner contributed extremely helpful datasets. The many other lab members provided a great environment to work and learn, including Pavitra Krishnaswamy, Demba Ba, Christa Van Dort, Behtash Babadi, Mike Prerau, Katie Hartnack, Kara Pavone, Gabriel Obregon, and Sage Chen. Finally, thanks to Sheri Leone for all of her help with the many issues that popped up!

Our animal studies would not have been possible without the assistance of Mike Halassa, whose work and insight brought a new dimension to our experiments. Ralf Wimmer, Phill Brunetti, Dan Zachs, and Steve Ramirez were all incredibly generous with their time in helping me learn experimental techniques.

A huge community of people have been incredibly supportive throughout this process. Lorna Gibson has been a wonderful friend and mentor throughout the past five years. Denise Heintze was an incredible source of support for students. I'm incredibly grateful to my dad, my mom and Andrew, and Xavier and Dix, for being an awesome and supportive family. And to all my friends, especially Masha Westerlund, Rodrigo Garcia, Meg Krench, Kean Jaime-Bustamante, John McCoy, Sangyu Xu, Chris Saenz, Elias Issa, Danielle Feldman, Retsina Meyer, and Ali Horowitz, for being so great and making grad school a happy place.

Finally, thanks to Jakob Voigts, for being an incredible scientist and person. He has been so wonderful over the past five years and made every day better.

# Table of Contents

## Chapter 1: Introduction

1.1 The neuroscience of unconscious states .....	6
1.2 General anesthesia in the operating room.....	7
1.3 Molecular mechanisms of general anesthesia.....	9
1.4 Brainstem and subcortical arousal circuits .....	10
1.5 Electroencephalogram correlates of sedation and unconsciousness.....	13
1.6 Theories of unconsciousness.....	15
1.7 Overview of the thesis .....	18
1.8 References.....	20

## Chapter 2: The neural correlates of propofol-induced loss of consciousness

2.1 Abstract.....	27
2.2 Introduction.....	28
2.3 Results.....	30
2.4 Discussion .....	53
2.5 Experimental Procedures .....	57
2.6 Author Contributions .....	63
2.7 References.....	63

## Chapter 3: The thalamocortical circuit mechanisms controlling arousal state

3.1 Abstract.....	68
3.2 Introduction.....	69
3.3 Results.....	71
3.4 Discussion .....	84
3.5 Experimental Procedures .....	89
3.6 Author Contributions .....	95
3.7 References.....	95

## Chapter 4: The neurophysiology of burst suppression in propofol-induced coma

4.1 Abstract.....	100
4.2 Introduction.....	102
4.3 Results.....	104
4.4 Discussion.....	114
4.5 Experimental Procedures .....	119
4.6 Author Contributions .....	125
4.7 References.....	125

## Chapter 5: Conclusions

5.1 Summary of the thesis.....	128
5.2 Implications for the neuroscience of arousal states .....	128
5.3 Implications for clinical care .....	133
5.4 Conclusion and future directions .....	134
5.5 References.....	137

# Chapter 1: Introduction

## 1.1 The neuroscience of unconscious states

Sleep and general anesthesia produce a profound but reversible change in nearly every aspect of brain function, extinguishing both our responses to the outside world and our internal state of consciousness. Remarkably, transitions between the awake and unconscious state can occur in just a few seconds, suggesting that the brain can switch between multiple states without major changes to its anatomical structure. How this phenomenon takes place is an unresolved question in neuroscience: what changes in the brain when a person loses consciousness, and what mechanism produces this transition into unconsciousness?

The most familiar state of unconsciousness is sleep, which is naturally occurring and has restorative effects for cognitive and physiological function. However, there are also many other ways that consciousness can be disrupted, such as during general anesthesia, absence seizures, and coma. Each of these states of decreased arousal is marked not only by diminished awareness, but also by a collection of physiological and neurological effects that underlie the change in behavior. While there are some neurophysiological phenomena that can be observed across sleep, anesthesia, and coma, overall these states are heterogeneous; even a slight change in the dosage of a single drug can produce a qualitatively different brain state. A major challenge for the neuroscience of unconscious states has been to identify which of these many neurophysiological changes is causally related to an animal's level of arousal or awareness. Most studies have compared the awake state with a deeply anesthetized state, blurring the many different gradations that can occur between sedation, loss of consciousness, and profound inactivation. This approach cannot identify the distinct dynamics that produce unconsciousness, versus a deeper, coma-like state.

Extensive work has characterized the brainstem and neuromodulatory pathways that can modulate arousal state and cause unconsciousness, reviewed briefly in Section 1.4. Furthermore, anesthetic-induced loss of consciousness is achieved through known molecular mechanisms, reviewed in Section 1.3. However, how these low-level effects

lead to states of decreased awareness remains unclear. Sleep, general anesthesia, and coma all disrupt the neural dynamics that produce consciousness, but the precise mechanism by which they do so is unknown. To understand the neural basis of states of decreased arousal, we must: 1) characterize the systems-level neural dynamics that underlie sleep, general anesthesia, and coma; 2) determine how these dynamics affect cognitive function; and 3) identify the circuit mechanisms that create these altered states. These questions require an integration of both human and animal studies, as the link between neurophysiology and awareness is best studied in humans, but the causal mechanisms can only be directly tested in animal models.

## **1.2 General anesthesia in the operating room**

### **1.2.1 Physiological effects of general anesthesia**

Although arousal state is naturally modulated in the sleep-wake cycle, a more rapid and potent control of arousal was provided by the advent of general anesthesia. The first public demonstration of general anesthesia, using ether, revolutionized the practice of surgery by enabling physicians to induce a profound state of unconsciousness. A broad array of drugs have now been developed that are used in various combinations to achieve the five essential components of general anesthesia: unconsciousness, amnesia, analgesia, paralysis, and stability of physiological systems (autonomic, cardiovascular, thermoregulatory, and respiratory (Brown et al., 2010)).

Perhaps the most important aspect of general anesthesia is that it is reversible. At the end of a surgery, drug administration is stopped and the patient will gradually return to the awake state. Because anesthesia is lifted passively, through cessation of drug, rather than actively (e.g. by administering an antidote), the half-life of a drug is an important consideration for clinical use. Propofol has become a commonly used anesthetic drug in part due to its rapid clearance times, enabling patients to wake up within minutes of when the infusion is turned off. Due to its widespread clinical use and relatively well-understood molecular effects, propofol is the primary anesthetic drug studied in this thesis.

Propofol is a generally safe and effective drug for producing unconsciousness, but is also accompanied by a range of other side effects that may be undesirable or potentially dangerous. Its physiological effects include decreased blood pressure and respiratory depression (Gold et al., 1987), and patients undergoing propofol anesthesia are generally intubated to preserve airway function. Propofol also causes nausea and vomiting, although to a lesser degree than many other anesthetic drugs (Hofer et al., 2003).

Anesthetic drugs are a highly diverse group, and other drugs cause a range of different physiological and cognitive effects. Inhaled agents, such as sevoflurane and desflurane, are commonly used for induction and maintenance of general anesthesia, and typically lower blood pressure and heart rate and carry a risk of tachycardia (Yildirim et al., 2004). Ketamine raises blood pressure, preserves respiratory function, and induces dissociative feelings and hallucinations during induction and emergence (Reich and Silvay, 1989). Dexmedetomidine is primarily used for sedation, and causes profound drowsiness. It can only be delivered as a gradual infusion, as a bolus dose can cause hypotension and bradycardia (Coursin et al., 2001).

The diversity of physiological effects across anesthetic agents highlights the fact that many of these drugs act through different molecular mechanisms and thus alter brain function through a variety of pathways. Furthermore, these differences illustrate that physiological states cannot be reliably linked to brain function, as the association between physiological variables and unconsciousness is not consistent across drugs or even across patients receiving the same drug.

### **1.2.2 Clinical monitoring during general anesthesia**

Although anesthetic drugs take effect primarily in the central nervous system, most anesthesiologists do not directly monitor brain function during general anesthesia. Instead, physiological variables such as blood pressure, oxygen saturation, and heart rate, and the presence of reflexes are used to infer the depth of anesthesia. These measures cannot provide precise information about whether a person is awake or unconscious, and intraoperative awareness remains a serious risk of general anesthesia, occurring in approximately 26,000 patients in the United States each year (Sebel et al., 2004).

However, simply increasing the dosage of anesthetic drugs is not a feasible solution, as general anesthesia is accompanied by many other risks. While it is typically a safe procedure, it is so frequently used (in millions of patients per year) that it nevertheless causes undesirable side effects in tens of thousands of patients. Common risks include nausea and vomiting, which can also cause respiratory distress and lead to more serious side effects. Death from general anesthesia is extremely rare, ranging around 8 deaths per million patients, but it is nevertheless a contributing cause in approximately 300 deaths annually in the United States (Li et al., 2009).

In an attempt to reduce both the risk of intraoperative awareness and excessive drug dosages, some anesthesiologists have begun using the bispectral index (BIS), a system that processes the electroencephalogram (EEG) and outputs a single number intended to index depth of anesthesia. In theory, this monitoring system could provide a method to titrate patients to a desired level of unconsciousness and signal when dosages are too high or too low. However, multiple recent studies have shown that the BIS cannot reliably detect unconsciousness, does not decrease anesthetic dosage, and is not effective at reducing intraoperative awareness (Avidan et al., 2011; 2008; Kaskinoro et al., 2011). An improved system for patient monitoring is therefore needed. To precisely assess consciousness, we must identify and characterize the brain activity patterns that cause unconsciousness, and develop measures that can detect the presence of these patterns in the operating room.

### **1.3 Molecular mechanisms of general anesthesia**

Anesthetic drugs act at the molecular level by binding cell receptors and thereby altering neuronal activity. The specific molecular mechanisms are highly variable across anesthetics, with different classes of drugs affecting different receptors to different degrees. However, a major target is receptors for gamma-aminobutyric acid (GABA), the primary inhibitory neurotransmitter in the brain. Many anesthetic drugs potentiate the effects of the GABA-A receptor (Franks, 2008), suggesting that enhancing GABAergic synaptic transmission is a potent mechanism for causing unconsciousness. However, major exceptions such as nitrous oxide, ketamine, and xenon, have little effect on

GABAergic transmission and instead act through other molecular pathways. In addition, most drugs and the inhaled agents in particular are nonselective, binding to many different types of receptors at clinically relevant concentrations (Rudolph and Antkowiak, 2004).

By comparison, propofol is a relatively selective drug at the molecular level, acting primarily as a GABA-A agonist. Propofol binds GABA-A receptors and lowers the rate of decay of GABA-mediated chloride conductances, leading to enhanced inhibition (Concas et al., 1991; Orser et al., 1994). At low doses, propofol only potentiates GABA-A currents, but at high doses may also directly activate those currents (Orser et al., 1994). This enhancement of GABA-A function is one of the key mechanisms by which propofol induces unconsciousness, as mice with a point mutation in the GABA-A receptor are highly resistant to propofol (Jurd et al., 2003). However, propofol also affects several other receptors to a lesser degree, including glycine receptors (Pistis et al., 1997), AMPA receptors (Haines et al., 2008), and hyperpolarization-activated cyclic-nucleotide-gated channels (Chen et al., 2005; Ying et al., 2006). GABA-A receptors are located throughout the central nervous system, providing multiple potential circuit mechanisms through which propofol could exert its hypnotic effects.

## **1.4 Brainstem and subcortical arousal circuits**

### **1.4.1 Ascending arousal systems**

The brainstem contains multiple structures that promote wakefulness through distinct neuromodulatory pathways. Cholinergic systems play a major role in generating wake states, and acetylcholine levels throughout cortex correlate with the sleep-wake cycle, with high levels of acetylcholine during the awake state and during rapid eye movement (REM) sleep (Saper et al., 2005). Cholinergic inputs arise from brainstem structures and the basal forebrain, innervating thalamus and cortex (Jones, 2004). Cholinergic inputs also disinhibit thalamic neurons by inhibiting the thalamic reticular nucleus (McCormick and Prince, 1986), a source of GABAergic inhibitory projections to central thalamus.

The noradrenaline system is a second major activating pathway, with the locus coeruleus (LC) sending noradrenergic inputs to nearly the entire central nervous system (Berridge, 2008). LC has long been known to play a key role in maintaining arousal, as its firing rates correlate strongly with arousal state, and drugs that impair noradrenergic transmission (such as the  $\alpha$ 2-adrenergic agonist dexmedetomidine) induce profound sedation. More recently, optogenetic activation of LC was shown to induce sleep-to-wake transitions, confirming a selective role for this structure in maintaining the awake state (Carter et al., 2010).

In addition to these two neuromodulatory systems, dopaminergic and orexinergic pathways also support arousal. Recent work has highlighted their effectiveness at inducing awake states, as optogenetic activation of orexin pathways induces sleep-to-wake transitions (Adamantidis et al., 2007), and stimulation of dopaminergic pathways can arouse even anesthetized animals (Taylor et al., 2013). Taken together, these pathways demonstrate that the brain has multiple redundant arousal systems (Jones, 2003), possibly reflecting evolutionary pressures given the importance of the waking state for survival.

Due to the many different arousal circuits that generate wake states, sleep must involve the suppression of several pathways. The hypothalamus in particular plays an important role in regulating activity throughout the ascending arousal system. The ventrolateral preoptic area of the hypothalamus sends GABAergic inhibitory projections to most of these arousal-promoting regions, including the cholinergic, dopaminergic, and noradrenergic pathways. Increasing GABAergic synaptic transmission could therefore impair multiple branches of the ascending arousal system, providing a possible mechanism for suppression of arousal during sleep and general anesthesia (Brown et al., 2011; Lee and Dan, 2012).

#### **1.4.2 Thalamocortical circuits**

The central thalamus plays a major role in arousal state, receiving extensive input from the brainstem arousal systems and projecting throughout cortex, where it modulates ongoing activity (Steriade et al., 1993). Most sensory information, excepting olfaction, is transmitted to cortex through synapses in the thalamus. Thalamocortical neurons are

therefore referred to as ‘relay’ cells, passing sensory information up to higher-level areas for further processing (Sherman and Guillery, 1996). Consequently, modulation of thalamic activity has substantial impact on cortical processing, as it controls which information is successfully transmitted to cortex (McCormick and Bal, 1994). However, this transmission of information is not unidirectional. Cortex sends broad and strong projections back to thalamus as both direct excitatory inputs and indirect inhibitory inputs. This network enables complex interactions between thalamus and cortex, allowing the two structures to act as a feedback loop (Nicolelis and Shuler, 2001).

The thalamocortical loop plays an important role in generating the EEG patterns of non-rapid eye movement (NREM) sleep: it is involved both in sleep spindles, a 12-15 Hz rhythm with a waxing-and-waning power envelope, and in slow wave activity (0.5-4 Hz). These rhythms have been suggested to result from the oscillatory properties of individual thalamic cells. Thalamocortical neurons inherently oscillate at delta (1-4 Hz) rhythms when inhibited due to the interaction of a hyperpolarization-activated cation current and a low-threshold calcium current (McCormick and Pape, 1990; McCormick and Huguenard, 1992). Thalamic reticular neurons, which provide inhibitory input to thalamus, can inherently oscillate at spindle frequencies and also interact with thalamic neurons at that frequency (Steriade et al., 1987). Extensive work has characterized the involvement of these circuits in generating the EEG signatures of NREM sleep (McCormick and Bal, 1997).

In addition to its role in sleep oscillatory dynamics, thalamic activity is strongly correlated with loss of consciousness. Thalamic activity declines prior to cortical activity during sleep onset (Magnin et al., 2010), and is suppressed during general anesthesia and in disorders of consciousness (Alkire et al., 2000; Schiff, 2008), leading researchers to suggest that decreased thalamic activity could play a causal role in cortical and behavioural decreases in arousal.

However, direct experimental manipulation of thalamic neurons has not established a conclusive relationship between thalamic activity and arousal. Optogenetic activation of thalamic neurons induces an aroused, desynchronized state in cortex (Poulet et al., 2012), suggesting a potential role for thalamus in regulating arousal state. However, lesioning thalamus does not disrupt the awake cortical state (Constantinople

and Bruno, 2011), nor does optogenetically silencing it (Zagha et al., 2013). The role of thalamus and thalamic inhibition in modulating arousal state thus remains unclear (Mashour and Alkire, 2013).

An important regulator of the thalamocortical circuit is the thalamic reticular nucleus (TRN), a shell of GABAergic neurons that envelop the thalamus (Houser et al., 1980). The TRN is the primary source of inhibitory input to thalamus, and thus has potential to modulate excitatory input from thalamus to cortex. However, electrical stimulation cannot be used to selectively manipulate TRN due to its proximity to thalamic nuclei, and its role in generating sleep states therefore remained unexplored until recently. TRN is now known to play a causal role in generating sleep spindles (Halassa et al., 2011), but its effects on behavior and arousal states remain unknown.

In addition to its potential role in sleep, thalamic inhibition may be one mechanism underlying propofol general anesthesia, as well as other GABAergic anesthetic drugs. Propofol enhances the effects of inhibitory input to cortical neurons (Kitamura et al., 2003), and of the GABAergic projections from TRN to thalamus (Ying and Goldstein, 2005), providing two direct sources of inhibition. In addition, propofol's effects on brainstem structures would be expected to further inhibit thalamus, by inhibiting the ascending arousal systems and thus reducing excitatory input to thalamus. The role of thalamus in both sleep and anesthesia-induced unconsciousness therefore merits further investigation.

## **1.5 Electroencephalogram correlates of sedation and unconsciousness**

### **1.5.1 EEG effects of propofol sedation and anesthesia**

Propofol causes striking changes in the EEG, inducing specific patterns that can vary depending on the dose and the susceptibility of individual patients. Prior to loss of consciousness, patients may experience paradoxical excitation, in which they exhibit euphoria, purposeless movements, incoherent speech, and a beta (13-25 Hz) oscillation in the EEG (Brown et al., 2011). During maintenance of propofol general anesthesia, a collection of oscillatory dynamics are commonly observed: a large-amplitude slow (0.1-1 Hz) and/or delta (1-4 Hz) oscillation (Jäntti and Sloan, 2008), an alpha (~10 Hz)

oscillation that is primarily observed across frontal regions (Supp et al., 2011), and an increase in low gamma (25-40 Hz) power (Murphy et al., 2011).

At deep levels of general anesthesia, many different drugs all converge on a common EEG signature: burst suppression (Akrawi et al., 1996). Burst suppression is marked by an alternation between isoelectric suppressions lasting seconds, and high-amplitude bursts. In a routine surgical procedure, burst suppression usually signals unnecessarily high levels of anesthetic, and this EEG pattern can therefore signify that dosage should be decreased. However, burst suppression is induced clinically to treat certain neurological conditions, such as status epilepticus (Claassen et al., 2002) and traumatic brain injury (Kelly et al., 1999). While the mechanism underlying burst suppression is not known, it appears to have neuroprotective effects in these cases, and propofol infusions are typically maintained for 24 hours or more to ensure a prolonged state of burst suppression.

### **1.5.2 EEG dynamics during drowsiness and sleep**

Similarly to the diverse brain states described above, sleep induces a range of different EEG dynamics. People transition through multiple sleep stages over the course of the night, visible as distinct patterns in the EEG. These patterns have been categorized as discrete stages, although in reality they may reflect a continuous gradation. Multiple sleep staging systems have been developed (Roth, 1961), but all begin with loss of the alpha rhythm seen during awake, eyes-closed recordings, marking the sleep onset period (Ogilvie, 2001). Next, sleep spindles appear, a 12-15 Hz oscillation with a waxing-and-waning power envelope, lasting under a second. Low-frequency power increases initially with the appearance of K-complexes, large amplitude deflections corresponding to an isolated down-state in which neuronal activity is silenced (Cash et al., 2009). Later, slow-wave activity becomes continuous and marks a periodic alteration between up and down states, reflecting rhythmic suppression and recovery of neuronal activity. This marks the deepest stage of non-rapid-eye-movement (NREM) sleep, and these stages of NREM alternate with rapid-eye-movement (REM) sleep repeatedly over the sleep cycle. REM sleep is marked by an active, desynchronized, low-amplitude EEG, and is correlated with

dreaming (Hobson, 2009); it is therefore not included in this thesis as a state of unconsciousness.

### **1.5.3 EEG features of coma**

As with general anesthesia, coma is a broad term encompassing many different brain states. Many types of neurological conditions can result in coma, including anoxia, hypoxia-ischemia, encephalitis, and traumatic brain injury. The resulting EEG can take on many forms as well, and EEG patterns are now used as a prognostic method to assess whether patients will recover from coma (Synek, 1988). Delta activity is commonly observed, but does not translate readily into a specific prognosis. In contrast, coma with theta and alpha oscillations usually signals a relative good prognosis. An exception is ‘alpha coma’, marked by an anteriorized alpha rhythm superimposed on a delta rhythm (somewhat similar to propofol), which is nearly always fatal (Tomassen and Kamphuisen, 1986). Periodic patterns with epileptiform activity also carry a poor prognosis (Young, 2000). Burst suppression, which is caused by deep general anesthesia, also occurs in certain types of coma. Although burst suppression is induced clinically as a treatment for certain brain injuries, when it occurs in the absence of drugs it signals an extremely poor prognosis (Brenner, 2005; Young, 2000).

## **1.6 Theories of unconsciousness**

### **1.6.1 Decreased neuronal activity**

Several theories have been proposed for how general anesthetic drugs produce unconsciousness. An initial hypothesis was simply that increasing anesthetic levels causes an increase in inhibition, and therefore a decrease in neuronal firing rates. A depression of cortical neuronal activity would then be expected to cause unconsciousness, by impairing all cortical function. This theory was based on consistent observations of decreased firing rates in cortical neurons caused by many different anesthetic drugs, including propofol (Andrada et al., 2012), isoflurane (Hentschke et al., 2005), and ketamine (Antkowiak, 1999). However, no experiment has shown that the decrease in firing rates occurs simultaneously with loss of consciousness, or with any specific

behavioural effect. In the absence of such studies, it is possible that decreased firing rates correlate with sedation rather than unconsciousness, or alternatively that firing rates decrease after loss of consciousness and reflect a deepening of anesthetic levels rather than a correlate of behavioural state. This theory also lacks explanatory power, as it is not clear how a simple decrease in firing rates would lead to the progressive breakdown of higher cognitive function and memory observed during gradual inductions and sedation.

### **1.6.2 A thalamic switch**

A second hypothesis is that general anesthetics activate a thalamocortical switch that controls consciousness. Multiple different drugs have been shown to reduce thalamic activity, leading some to suggest that this could be a potential unifying mechanism for unconsciousness (Alkire et al., 2000). However, this hypothesis has yet to be verified experimentally: there are no neurophysiology studies characterizing how thalamic activity changes immediately at the transition into unconsciousness, nor any manipulations demonstrating a causal role for this mechanism. A debate remains over whether the thalamus is a switch that can control consciousness, or whether its decrease in activity simply reflects a decrease of excitatory feedback from cortex, as cortical activity declines during unconsciousness (Alkire et al., 2008).

The thalamic mechanism has also been suggested to play a major role in coma, as neuronal damage to the central thalamus is correlated with disorders of consciousness (Schiff, 2008). Somewhat counterintuitively, the GABA-A agonist zolpidem has produced arousal and transient recovery in coma patients (Brefel-Courbon et al., 2007; Cohen and Duong, 2008). However, this drug is also proposed to act through increasing thalamic activity, as the GABA-A agonism could inhibit globus pallidus, and thereby disinhibit thalamus (Schiff, 2010).

### **1.6.3 A core consciousness-generating set of cortical areas**

A recent theory, which is also compatible with the two previous hypotheses, is that unconsciousness is due to inactivation of specific cortical regions that are required for consciousness. This idea was proposed after neuroimaging studies found that cingulate, posterior parietal, and precuneal areas are consistently inactivated during

general anesthesia (Fiset et al., 1999; Kaisti et al., 2002), leading to the hypothesis that these core regions are required for consciousness (Alkire et al., 2008). A related theory suggests that frontoparietal interactions in particular are a key element of consciousness, and that anesthetics act by disrupting these interactions (Ku et al., 2011; Lee et al., 2009). However, there is no causal evidence demonstrating that these areas are functionally important for awareness, as opposed to simply being particularly sensitive to anesthetic drugs. It is also theoretically challenging to explain how a small subregion of cortex could generate consciousness.

### **1.6.3 Breakdown of binding through gamma oscillations**

A more general theory for how cortical disruption could produce unconsciousness focuses on gamma (~40 Hz) oscillations. Gamma has been proposed to implement ‘binding’, enabling information to enter consciousness by synchronizing neuronal assemblies in distributed cortical regions (Engel and Singer, 2001). Conversely, general anesthesia has been suggested to operate through disruption of gamma synchrony, resulting in ‘cognitive unbinding’ (Mashour, 2004). Some studies have indeed found that gamma coherence is altered during general anesthesia (Imas et al., 2005; John et al., 2001). If this hypothesis is correct, directly manipulating gamma oscillations should alter arousal state, providing an empirically testable prediction that could support this theory. However, a systems-level explanation of anesthetic-induced unconsciousness should also include a circuit mechanism that disrupts gamma synchrony and causes this loss of binding.

### **1.6.4 Loss of information integration**

In an effort to develop a unified theory for consciousness, Tononi proposed the idea of consciousness as information integration (Tononi, 2004; 2005). In this framework, consciousness is related to the number of possible states in an integrated system. Unconsciousness would then be produced by either a disruption of the system (impairing the integration) or by a decrease in the number of possible states (impairing the information.) This theory is agnostic as to the neural mechanism that would impair information integration during general anesthesia, although several possibilities have

been suggested. These include disruption of integration due to loss of gamma synchrony, and disruption of information due to the stereotyped, bistable states in burst suppression (Alkire et al., 2008). However, it is unclear whether the theory of information integration inherently provides any falsifiable predictions that could be experimentally tested.

### **1.6.5 Empirical approaches to validating theories**

These hypotheses range from empirically testable to purely philosophical. For example, the suggestions that a specific set of cortical regions are needed in order to be conscious could be tested through lesion studies and optogenetic inactivation of those areas. In contrast, the concept of information integration may be a useful philosophical construct, but does not suggest a specific mechanism that could be experimentally validated or falsified.

Two sets of experiments are needed to conclusively examine the question of how anesthetic drugs produce unconsciousness. First, neurophysiological studies should be performed during the transition into unconsciousness, to determine which effects occur simultaneously with loss of consciousness rather than during deep general anesthesia. Most studies have only compared the awake state with a fully anesthetized state, or a lightly anesthetized state with a deeply anesthetized state, and these comparisons cannot assess the neural dynamics specifically associated with loss of consciousness. Second, causal manipulations are needed in order to directly test whether these proposed mechanisms can indeed induce changes in arousal. In addition, it is possible that different drugs may act through different mechanisms, as the variety in molecular effects suggests that there could also be a range of circuit- and systems-level mechanisms that produce unconsciousness. Several theories could therefore be correct, but might only explain a subset of the drugs or pathologies that affect arousal state.

## **1.7 Overview of the thesis**

Significant work has identified the molecular-level mechanisms of general anesthetics and their resulting effects on subcortical structures. However, unconsciousness occurs well before brainstem responses are extinguished, and it is likely

that anesthetic effects on cortical dynamics are responsible for disrupting higher-level cognitive function. Furthermore, the diversity of ways one can become unconscious (e.g. sleep, many anesthetic drugs, many types of coma) and the range of EEG phenomena observed across these states suggests that there are many different mechanisms that can produce a state of decreased arousal. Although several theories exist as to the neural basis of unconsciousness (Section 1.6), empirical evidence has not yet confirmed any of these hypotheses. Furthermore, most studies have not differentiated between the brain states associated with loss of consciousness, and those that occur during deep, coma-like states of general anesthesia. A more fine-grained approach is needed to identify how brain states and function are altered across different states of decreased arousal.

Two major questions are unresolved: first, what are the cortical states induced by sleep and general anesthesia? And second, what are the circuit mechanisms that produce these states? This thesis seeks to answer these questions, which should provide substantial scientific insight into the cortical basis of awareness and unconsciousness. Furthermore, it will also form an important step towards developing clinical systems to monitor patients during general anesthesia, improving patient outcomes and reducing the rate of intraoperative awareness.

The first part of the thesis focuses on the brain states associated with loss of consciousness. In Chapter 2, I find that loss of consciousness occurs simultaneously with the onset of a slow (0.1-1 Hz) oscillation in cortex. This oscillation is asynchronous across the brain, and is associated with hundreds of milliseconds of silence in cortical neurons. This spatial pattern means that neurons will be silenced at different times in different brain regions, effectively fragmenting cortical networks and disrupting information transfer across the brain. These results demonstrate that the cortical state associated with propofol-induced unconsciousness is local slow waves.

Given this robust correlation between slow waves and unconsciousness, I next searched for the circuit mechanism that generates slow waves, and tested whether manipulating this circuit alone is sufficient to control arousal state. In Chapter 3, I show that tonic optogenetic activation of the thalamic reticular nucleus generates cortical slow wave activity and decreases behavioural signs of arousal. These effects were fast-onset (<100 ms), suggesting a direct engagement of the circuit responsible for these

oscillations. Furthermore, slow wave activity was induced in local cortical areas, providing an explanation for how this brain state could be controlled separately across different cortical regions. Taken together, Chapters 2 and 3 identify a brain state associated with unconsciousness and describe the circuit mechanism that generates this state.

In Chapter 4, I focus on the much deeper state of anesthesia that occurs during medically-induced coma. I study the EEG pattern of burst suppression, and find that it can occur in isolated parts of cortex, in contrast to previous assumptions in the literature. Furthermore, I show that the bursts recover the spectral dynamics of the underlying brain state, suggesting that normal brain activity resumes intermittently but is then interrupted by a prolonged suppression. These characteristics are consistent with the hypothesis that burst suppression is caused by local decreases in cerebral metabolism. This chapter therefore provides a detailed characterization of the brain state that occurs during medically-induced coma, and suggests decreased cerebral metabolism as an underlying neurophysiological mechanism.

Taken together, Chapters 2-4 identify the brain states associated with unconsciousness and coma and the circuit mechanisms that generate these states. In Chapter 5, I discuss the implications of these results for the neuroscience of sleep and general anesthesia, and highlight potential clinical applications to improve patient monitoring and reduce anesthetic risks.

## 1.8 References

- Adamantidis, A.R., Zhang, F., Aravanis, A.M., Deisseroth, K., and de Lecea, L. (2007). Neural substrates of awakening probed with optogenetic control of hypocretin neurons. *Nature* 450, 420–424.
- Akrawi, W.P., Drummond, J.C., Kalkman, C.J., and Patel, P.M. (1996). A comparison of the electrophysiologic characteristics of EEG burst-suppression as produced by isoflurane, thiopental, etomidate, and propofol. *Journal of Neurosurgical Anesthesiology* 8, 40–46.
- Alkire, M., Haier, R., and Fallon, J. (2000). Toward a unified theory of narcosis: brain imaging evidence for a thalamocortical switch as the neurophysiologic basis of

- anesthetic-induced unconsciousness. *Consciousness and Cognition* 9, 370–386.
- Alkire, M., Hudetz, A., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876.
- Andrada, J., Livingston, P., Lee, B.J., and Antognini, J. (2012). Propofol and etomidate depress cortical, thalamic, and reticular formation neurons during anesthetic-induced unconsciousness. *Anesthesia & Analgesia* 114, 661–669.
- Antkowiak, B. (1999). Different actions of general anesthetics on the firing patterns of neocortical neurons mediated by the GABA(A) receptor. *Anesthesiology* 91, 500–511.
- Avidan, M.S., Jacobsohn, E., Glick, D., Burnside, B.A., Zhang, L., Villafranca, A., Karl, L., Kamal, S., Torres, B., O'Connor, M., et al. (2011). Prevention of intraoperative awareness in a high-risk surgical population. *New England Journal of Medicine* 365, 591–600.
- Avidan, M.S., Zhang, L., Burnside, B.A., Finkel, K.J., Searleman, A.C., Selvidge, J.A., Saager, L., Turner, M.S., Rao, S., Bottros, M., et al. (2008). Anesthesia awareness and the bispectral index. *New England Journal of Medicine* 358, 1097–1108.
- Berridge, C.W. (2008). Noradrenergic modulation of arousal. *Brain Research Reviews* 58, 1–17.
- Brefel-Courbon, C., Payoux, P., Ory, F., Sommet, A., Slaoui, T., Raboyeau, G., Lemesle, B., Puel, M., Montastruc, J.-L., Demonet, J.-F., et al. (2007). Clinical and imaging evidence of zolpidem effect in hypoxic encephalopathy. *Annals of Neurology* 62, 102–105.
- Brenner, R.P. (2005). The interpretation of the EEG in stupor and coma. *Neurologist* 11, 271–284.
- Brown, E.N., Lydic, R., and Schiff, N.D. (2010). General anesthesia, sleep, and coma. *New England Journal of Medicine* 363, 2638–2650.
- Brown, E.N., Purdon, P.L., and Van Dort, C.J. (2011). General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual Review of Neuroscience* 34, 601–628.
- Carter, M.E., Yizhar, O., Chikahisa, S., Nguyen, H., Adamantidis, A., Nishino, S., Deisseroth, K., and de Lecea, L. (2010). Tuning arousal with optogenetic modulation of locus coeruleus neurons. *Nature Neuroscience* 13, 1526–1533.
- Cash, S.S., Halgren, E., Dehghani, N., Rossetti, A.O., Thesen, T., Wang, C., Devinsky, O., Kuzniecky, R., Doyle, W., Madsen, J.R., et al. (2009). The human K-complex represents an isolated cortical down-state. *Science* 324, 1084–1087.

- Chen, X., Shu, S., and Bayliss, D.A. (2005). Suppression of Ih contributes to propofol-induced inhibition of mouse cortical pyramidal neurons. *Journal of Neurophysiology* 94, 3872–3883.
- Claassen, J., Hirsch, L.J., Emerson, R.G., and Mayer, S.A. (2002). Treatment of refractory status epilepticus with pentobarbital, propofol, or midazolam: a systematic review. *Epilepsia* 43, 146–153.
- Cohen, S.I., and Duong, T.T. (2008). Increased Arousal in a Patient with Anoxic Brain Injury After Administration of Zolpidem. *American Journal of Physical Medicine & Rehabilitation* 87, 229–231.
- Concas, A., Santoro, G., Serra, M., Sanna, E., and Biggio, G. (1991). Neurochemical action of the general anaesthetic propofol on the chloride ion channel coupled with GABAA receptors. *Brain Research* 542, 225–232.
- Constantinople, C.M., and Bruno, R.M. (2011). Effects and mechanisms of wakefulness on local cortical networks. *Neuron* 69, 1061–1068.
- Coursin, D.B., Coursin, D.B., and Maccioli, G.A. (2001). Dexmedetomidine. *Current Opinion in Critical Care* 7, 221–226.
- Engel, A.K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences* 5, 16–25.
- Fiset, P., Paus, T., Daloze, T., Plourde, G., Meuret, P., Bonhomme, V., Hajj-Ali, N., Backman, S., and Evans, A. (1999). Brain mechanisms of propofol-induced loss of consciousness in humans: a positron emission tomographic study. *Journal of Neuroscience* 19, 5506.
- Franks, N.P. (2008). General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. *Nature Reviews Neuroscience* 9, 370–386.
- Gold, M.I., Abraham, E.C., and Herrington, C. (1987). A controlled investigation of propofol, thiopentone and methohexitone. *Canadian Journal of Anaesthesia* 34, 478–483.
- Haines, M., Mao, L.M., Yang, L., Arora, A., Fibuch, E.E., and Wang, J.Q. (2008). Modulation of AMPA receptor GluR1 subunit phosphorylation in neurons by the intravenous anaesthetic propofol. *British Journal of Anaesthesia* 100, 676–682.
- Halassa, M.M., Siegle, J.H., Ritt, J.T., Ting, J.T., Feng, G., and Moore, C.I. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. *Nature Neuroscience* 14, 1118–1120.
- Hentschke, H., Schwarz, C., and Antkowiak, B. (2005). Neocortex is the major target of sedative concentrations of volatile anaesthetics: strong depression of firing rates and increase of GABAA receptor-mediated inhibition. *European Journal of*

Neuroscience 21, 93–102.

- Hobson, J.A. (2009). REM sleep and dreaming: towards a theory of protoconsciousness. *Nature Reviews Neuroscience* 10, 803–813.
- Hofer, C.K., Zollinger, A., Büchi, S., Klaghofer, R., Serafino, D., Bühlmann, S., Buddeberg, C., Pasch, T., and Spahn, D.R. (2003). Patient well-being after general anaesthesia: a prospective, randomized, controlled multi-centre trial comparing intravenous and inhalation anaesthesia. *British Journal of Anaesthesia* 91, 631–637.
- Houser, C.R., Vaughn, J.E., Barber, R.P., and Roberts, E. (1980). GABA neurons are the major cell type of the nucleus reticularis thalami. *Brain Research* 200, 341–354.
- Imas, O.A., Ropella, K.M., Ward, B.D., Wood, J.D., and Hudetz, A.G. (2005). Volatile anesthetics disrupt frontal-posterior recurrent information transfer at gamma frequencies in rat. *Neuroscience Letters* 387, 145–150.
- Jääntti, V., and Sloan, T. (2008). EEG and anesthetic effects. *Handbook of Clinical Neurophysiology* 8, 77–93.
- John, E., Prichep, L., Kox, W., Valdes-Sosa, P., Bosch-Bayard, J., Aubert, E., Tom, M., DiMichele, F., and Gugino, L. (2001). Invariant reversible QEEG effects of anesthetics. *Consciousness and Cognition* 10, 165–183.
- Jones, B.E. (2003). Arousal systems. *Frontiers in Biosciences* 8, s438–s451.
- Jones, B.E. (2004). Activity, modulation and role of basal forebrain cholinergic neurons innervating the cerebral cortex. *Progress in Brain Research*, 157–169.
- Jurd, R., Arras, M., Lambert, S., Drexler, B., Siegwart, R., Crestani, F., Zaugg, M., Vogt, K.E., Ledermann, B., Antkowiak, B., et al. (2003). General anesthetic actions in vivo strongly attenuated by a point mutation in the GABA(A) receptor beta3 subunit. *Faseb Journal* 17, 250–252.
- Kaisti, K.K., Metsähonkala, L., Teräs, M., Oikonen, V., Aalto, S., Jääskeläinen, S., Hinkka, S., and Scheinin, H. (2002). Effects of surgical levels of propofol and sevoflurane anesthesia on cerebral blood flow in healthy subjects studied with positron emission tomography. *Anesthesiology* 96, 1358–1370.
- Kaskinoro, K., Maksimow, A., Långsjö, J., Aantaa, R., Jääskeläinen, S., Kaisti, K., Särkelä, M., and Scheinin, H. (2011). Wide inter-individual variability of bispectral index and spectral entropy at loss of consciousness during increasing concentrations of dexmedetomidine, propofol, and sevoflurane. *British Journal of Anaesthesia* 107, 573–580.
- Kelly, D.F., Goodale, D.B., Williams, J., Herr, D.L., Chappell, E.T., Rosner, M.J., Jacobson, J., Levy, M.L., Croce, M.A., Maniker, A.H., et al. (1999). Propofol in

- the treatment of moderate and severe head injury: a randomized, prospective double-blinded pilot trial. *Journal of Neurosurgery* *90*, 1042–1052.
- Kitamura, A., Marszalec, W., Yeh, J.Z., and Narahashi, T. (2003). Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. *Journal of Pharmacology and Experimental Therapeutics* *304*, 162–171.
- Ku, S.-W., Lee, U., Noh, G.-J., Jun, I.-G., and Mashour, G.A. (2011). Preferential Inhibition of Frontal-to-Parietal Feedback Connectivity Is a Neurophysiologic Correlate of General Anesthesia in Surgical Patients. *PLoS ONE* *6*, e25155.
- Lee, S.-H., and Dan, Y. (2012). Neuromodulation of brain states. *Neuron* *76*, 209–222.
- Lee, U., Kim, S., Noh, G., Choi, B., Hwang, E., and Mashour, G. (2009). The directionality and functional organization of frontoparietal connectivity during consciousness and anesthesia in humans. *Consciousness and Cognition* *18*, 1069–1078.
- Li, G., Warner, M., Lang, B.H., Huang, L., and Sun, L.S. (2009). Epidemiology of anesthesia-related mortality in the United States, 1999-2005. *Anesthesiology* *110*, 759–765.
- Magnin, M., Rey, M., Bastuji, H., Guillemant, P., Mauguière, F., and Garcia-Larrea, L. (2010). Thalamic deactivation at sleep onset precedes that of the cerebral cortex in humans. *Proceedings of the National Academy of Sciences* *107*, 3829.
- Mashour, G.A. (2004). Consciousness unbound: toward a paradigm of general anesthesia. *Anesthesiology* *100*, 428–433.
- Mashour, G.A., and Alkire, M.T. (2013). Consciousness, anesthesia, and the thalamocortical system. *Anesthesiology* *118*, 13–15.
- McCormick, D.A., and Bal, T. (1994). Sensory gating mechanisms of the thalamus. *Current Opinion in Neurobiology* *4*, 550–556.
- McCormick, D.A., and Bal, T. (1997). Sleep and arousal: thalamocortical mechanisms. *Annual Review of Neuroscience* *20*, 185–215.
- McCormick, D.A., and Pape, H.C. (1990). Properties of a hyperpolarization-activated cation current and its role in rhythmic oscillation in thalamic relay neurones. *The Journal of Physiology* *431*, 291–318.
- McCormick, D.A., and Prince, D.A. (1986). Acetylcholine induces burst firing in thalamic reticular neurones by activating a potassium conductance. *Nature* *319*, 402–405.
- McCormick, D.A., and Huguenard, J.R. (1992). A model of the electrophysiological properties of thalamocortical relay neurons. *Journal of Neurophysiology* *68*,

1384–1400.

- Murphy, M., Bruno, M., Riedner, B., Boveroux, P., Noirhomme, Q., Landsness, E., Brichant, J., Phillips, C., Massimini, M., and Laureys, S. (2011). Propofol anesthesia and sleep: a high-density EEG study. *Sleep* 34, 283.
- Nicolelis, M.A., and Shuler, M. (2001). Thalamocortical and corticocortical interactions in the somatosensory system. *Progress in Brain Research* 130, 90–110.
- Ogilvie, R. (2001). The process of falling asleep. *Sleep Medicine Reviews* 5, 247–270.
- Orser, B.A., Wang, L.Y., Pennefather, P.S., and MacDonald, J.F. (1994). Propofol modulates activation and desensitization of GABAA receptors in cultured murine hippocampal neurons. *The Journal of Neuroscience* 14, 7747–7760.
- Pistis, M., Belelli, D., Peters, J.A., and Lambert, J.J. (1997). The interaction of general anaesthetics with recombinant GABAA and glycine receptors expressed in *Xenopus laevis* oocytes: a comparative study. *British Journal of Pharmacology* 122, 1707–1719.
- Poulet, J.F.A., Fernandez, L.M.J., Crochet, S., and Petersen, C.C.H. (2012). Thalamic control of cortical states. *Nature Neuroscience* 15, 370–372.
- Reich, D.L., and Silvey, G. (1989). Ketamine: an update on the first twenty-five years of clinical experience. *Canadian Journal of Anaesthesia* 36, 186–197.
- Roth, B. (1961). The clinical and theoretical importance of EEG rhythms corresponding to states of lowered vigilance. *Electroencephalography and Clinical Neurophysiology* 13, 395–399.
- Rudolph, U., and Antkowiak, B. (2004). Molecular and neuronal substrates for general anaesthetics. *Nature Reviews Neuroscience* 5, 709–720.
- Saper, C.B., Scammell, T.E., and Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature* 437, 1257–1263.
- Schiff, N.D. (2008). Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Annals of the New York Academy of Sciences* 1129, 105–118.
- Schiff, N.D. (2010). Recovery of consciousness after brain injury: a mesocircuit hypothesis. *Trends in Neurosciences* 33, 1–9.
- Sebel, P.S., Bowdle, T.A., Ghoneim, M.M., Rampil, I.J., Padilla, R.E., Gan, T.J., and Domino, K.B. (2004). The Incidence of Awareness During Anesthesia: A Multicenter United States Study. *Anesthesia & Analgesia* 99, 833–839.
- Sherman, S.M., and Guillery, R.W. (1996). Functional organization of thalamocortical

- relays. *Journal of Neurophysiology* 76, 1367–1395.
- Steriade, M., Domich, L., Oakson, G., and Deschênes, M. (1987). The deafferented reticular thalamic nucleus generates spindle rhythmicity. *Journal of Neurophysiology* 57, 260–273.
- Steriade, M., McCormick, D.A., and Sejnowski, T.J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685.
- Supp, G.G., Siegel, M., Hipp, J.F., and Engel, A.K. (2011). Cortical hypersynchrony predicts breakdown of sensory processing during loss of consciousness. *Current Biology* 21, 1988–1993.
- Synek, V.M. (1988). Prognostically important EEG coma patterns in diffuse anoxic and traumatic encephalopathies in adults. *Journal of Clinical Neurophysiology* 5, 161–174.
- Taylor, N.E., Chemali, J.J., Brown, E.N., and Solt, K. (2013). Activation of D1 dopamine receptors induces emergence from isoflurane general anesthesia. *Anesthesiology* 118, 30–39.
- Tomassen, W., and Kamphuisen, H.A. (1986). Alpha coma. *Journal of the Neurological Sciences* 76, 1–11.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience* 5, 42.
- Tononi, G. (2005). Consciousness, information integration, and the brain. *Progress in Brain Research*, 109–126.
- Yildirim, H., Adanir, T., Atay, A., Katircioğlu, K., and Savaci, S. (2004). The effects of sevoflurane, isoflurane and desflurane on QT interval of the ECG. *European Journal of Anaesthesiology* 21, 566–570.
- Ying, S.-W., and Goldstein, P.A. (2005). Propofol-block of SK channels in reticular thalamic neurons enhances GABAergic inhibition in relay neurons. *Journal of Neurophysiology* 93, 1935–1948.
- Ying, S.-W., Abbas, S.Y., Harrison, N.L., and Goldstein, P.A. (2006). Propofol block of I(h) contributes to the suppression of neuronal excitability and rhythmic burst firing in thalamocortical neurons. *European Journal of Neuroscience* 23, 465–480.
- Young, G. (2000). The EEG in coma. *Journal of Clinical Neurophysiology*.
- Zagha, E., Casale, A.E., Sachdev, R.N.S., McGinley, M.J., and McCormick, D.A. (2013). Motor cortex feedback influences sensory processing by modulating network state. *Neuron* 79, 567–578.

## **Chapter 2: The neural correlates of propofol-induced loss of consciousness<sup>1</sup>**

### **2.1 Abstract**

The neurophysiological mechanisms by which anesthetic drugs cause loss of consciousness are poorly understood. Anesthetic actions at the molecular, cellular, and systems levels have been studied in detail at steady-states of deep general anesthesia. However, little is known about how anesthetics alter neural activity during the transition into unconsciousness. We recorded simultaneous multi-scale neural activity from human cortex, including ensembles of single neurons, local field potentials, and intracranial electrocorticograms, during induction of general anesthesia. We analyzed local and global neuronal network changes that occurred simultaneously with loss of consciousness. We show that propofol-induced unconsciousness occurs within seconds of the abrupt onset of a slow (<1 Hz) oscillation in the local field potential. This oscillation marks a state in which cortical neurons maintain local patterns of network activity, but this activity is fragmented across both time and space. Local (<4 mm) neuronal populations maintain the millisecond-scale connectivity patterns observed in the awake state, and spike rates fluctuate and can reach baseline levels. However, neuronal spiking occurs only within a limited slow oscillation phase window, and is silent otherwise, fragmenting the time course of neural activity. Unexpectedly, we found that these slow oscillations occur asynchronously across cortex, disrupting functional connectivity between cortical areas. We conclude that the onset of slow oscillations is a neural correlate of propofol-induced loss of consciousness, marking a shift to cortical dynamics in which local neuronal networks remain intact but become functionally isolated in time and space.

---

<sup>1</sup> The findings in this chapter were previously published in Lewis, Weiner, Mukamel, Donoghue, Eskandar, Madsen, Anderson, Hochberg, Cash, Brown, and Purdon (2012). *PNAS* 109(49):3377-3386.

## 2.2 Introduction

General anesthesia is a drug-induced reversible coma commonly initiated by administering a large dose of a fast-acting drug to induce unconsciousness within seconds (Brown et al., 2010). This state can then be maintained as long as needed to execute surgical and many non-surgical procedures. One of the most widely used anesthetics is propofol, an intravenous drug that enhances GABAergic inhibitory input to neurons (Bai et al., 1999; Brown et al., 2011; Rudolph and Antkowiak, 2004), with effects in cortex, thalamus, brainstem, and spinal cord (Alkire et al., 1995; Fiset et al., 1999; Kungys et al., 2009). Despite the understanding of propofol's molecular actions, it is not clear how these effects at molecular targets affect single neurons and larger-scale neural circuits to produce unconsciousness.

The effects on macroscopic dynamics are noticeable in the electroencephalogram (EEG), which contains several stereotyped patterns during maintenance of propofol general anesthesia. These patterns include increased delta (0.5-4 Hz) power (Murphy et al., 2011; Steriade et al., 1993b); increased gamma (25-40 Hz) power (Murphy et al., 2011); an alpha (~10 Hz) rhythm (Cimenser et al., 2011; Feshchenko et al., 2004; Supp et al., 2011) that is coherent across frontal cortex; and burst suppression, an alternation between bursts of high-voltage activity and periods of flat EEG lasting for several seconds (Akrawi et al., 1996; Ching et al., 2012). In addition, slow oscillations (<1 Hz) have been well characterized in deeply anesthetized animals, and they are associated with an alternation of the neuronal membrane potential between UP (depolarized) and DOWN (hyperpolarized) states (Contreras and Steriade, 1995; Steriade et al., 1993b).

Although these patterns are consistently observed, it is unclear how they are functionally related to unconsciousness under general anesthesia. Most studies have focused on a deep steady-state of general anesthesia, and have not used a systematic behavioral measure to track the transition into unconsciousness. This steady-state approach cannot distinguish between patterns that are characteristic of a deeply anesthetized brain and those that arise at the onset of unconsciousness. Unconsciousness can occur in tens of seconds (Brown et al., 2011) whereas many neurophysiological features continue to fluctuate for minutes after induction and are highly variable between different levels of general anesthesia (Bennett et al., 2009; Brown et al., 2010).

Therefore, identifying the specific dynamics associated with loss of consciousness requires an examination of the transition into unconsciousness, linking neurophysiology with behavioral measures.

In addition, the dynamic interactions between cortical areas that underlie these EEG oscillations are not well understood, as few studies have simultaneously recorded ensembles of single neurons and oscillatory dynamics from sites distributed across the brain. Consequently, how propofol acts on neural circuits to produce unconsciousness remains unclear. A leading hypothesis suggests that anesthetics disrupt cortical integration (Alkire et al., 2008; Mashour, 2004). Identifying the mechanism by which this disruption might occur requires a better understanding of how the spatial and temporal organization of neural dynamics evolves during induction of unconsciousness.

To address this question, we investigated both neuronal and circuit-level dynamics in the human brain during induction of unconsciousness with propofol. We obtained simultaneous recordings of single units, local field potentials, and intracranial electrocorticograms over up to 8 cm of cortex, enabling us to examine neural dynamics at multiple spatial scales with millisecond-scale temporal resolution. We used a behavioral task to identify within seconds the time at which patients became unresponsive to auditory stimuli, which we defined as loss of consciousness (LOC).

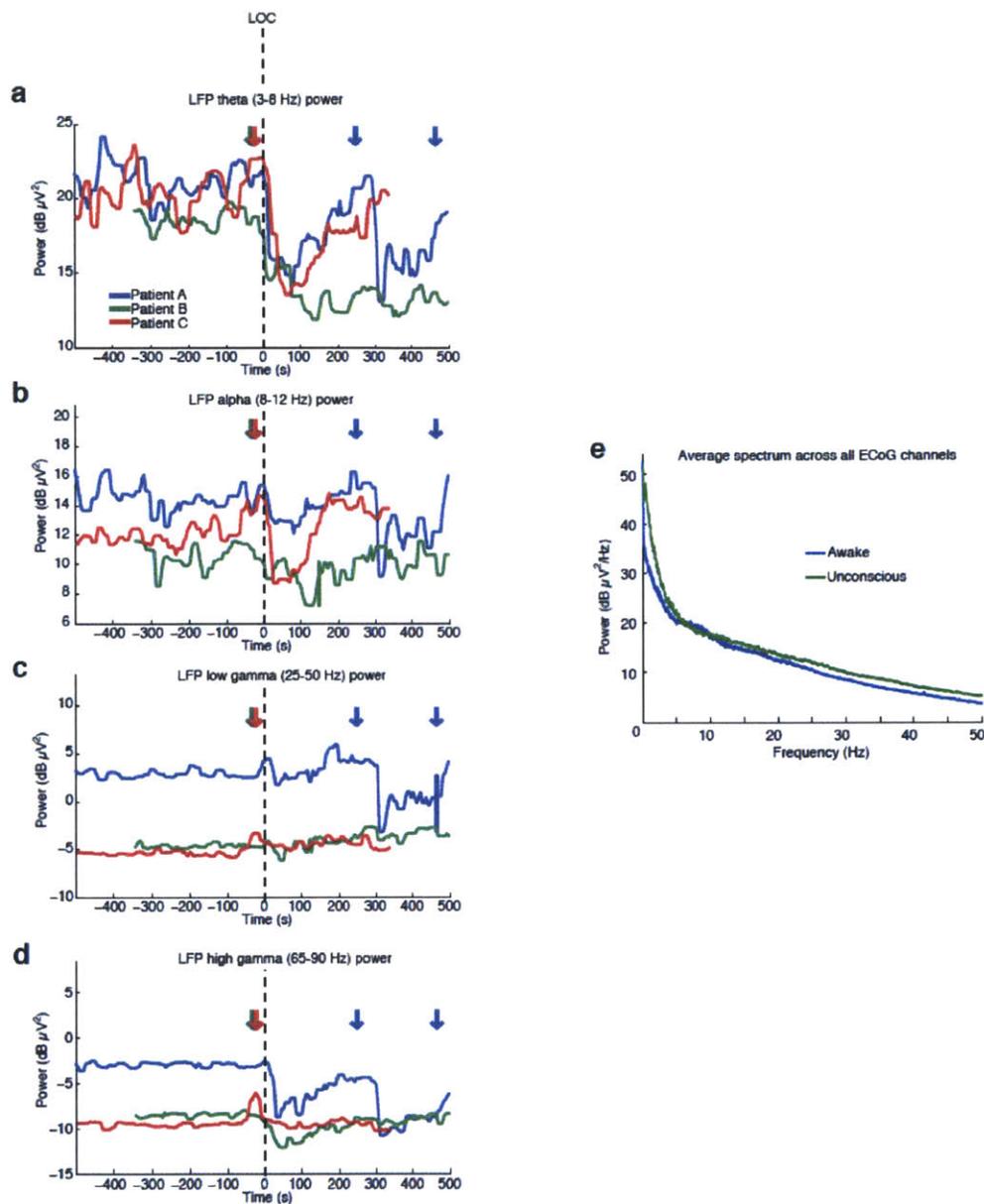
Our results reveal a set of neurophysiological features that accompany loss of consciousness that, together with previously reported effects (Contreras and Steriade, 1995; Murphy et al., 2011; Steriade et al., 1993b), enable a multi-scale account of this profound shift in brain state. We find that LOC is marked by the abrupt onset of slow oscillations (0.1-1 Hz) in the local field potential. Power in the slow oscillation band rises sharply at LOC and maintains this increase throughout the post-LOC period. Neuronal spiking becomes coupled to the local slow oscillation within seconds of LOC: spiking occurs only in short intervals of activity that are interspersed with suppression lasting hundreds of milliseconds, periodically interrupting information processing. These periods in which activity may occur are not simultaneous across the brain, implying that information transfer between distant ( $>2$  cm) cortical networks is impaired. Cortical networks are therefore fragmented both temporally and spatially, disrupting both local and long-range communication. However, small-scale ( $<4$  mm) functional connectivity

measures remain similar to the conscious state and neuronal spike rates can recover to baseline levels after LOC despite continued unresponsiveness. This demonstrates that short periods of normal spike dynamics can still occur during unconsciousness. We conclude that the slow oscillation is a fundamental component of propofol-induced unconsciousness, marking a functional isolation of cortical regions while significant connectivity is preserved within local networks.

## **2.3 Results**

### **2.3.1 Rapid loss of consciousness after propofol bolus**

We recorded single units (n=198), local field potentials (LFP), and intracranial electrocorticograms (ECoG) in three patients undergoing intracranial monitoring for surgical treatment of epilepsy. Single units and LFPs were recorded from a 96-channel microelectrode array (Truccolo et al., 2011) implanted in temporal cortex for research purposes. We recorded throughout induction of general anesthesia by bolus administration of propofol, before planned neurosurgery to remove the electrodes. Patients performed an auditory task requiring a button press in response to stimuli. All patients completely ceased responding to the task within 40 seconds of propofol administration and remained unresponsive for the remainder of the recording period, lasting 5-10 minutes post-LOC. LOC was defined as the onset of this period of unresponsiveness to auditory stimuli. To acknowledge the fact that LOC could have occurred at any point between the last response and the failure to make the next response, LOC was defined as the interval beginning one second before the first missed stimulus, up until the second missed stimulus (5 seconds total). We then compared spectra across all ECoG channels in the pre- and post-LOC periods. In agreement with previous scalp EEG studies of healthy subjects (Murphy et al., 2011), we found that average spectra in the post-LOC period differed significantly from that in the pre-LOC period: slow (0.1-1 Hz) and gamma (25-40 Hz) power increased in the unconscious state (Fig. 2.1e). These results suggested that propofol acted as expected in these patients, and did not reveal any gross disruption of GABA networks.



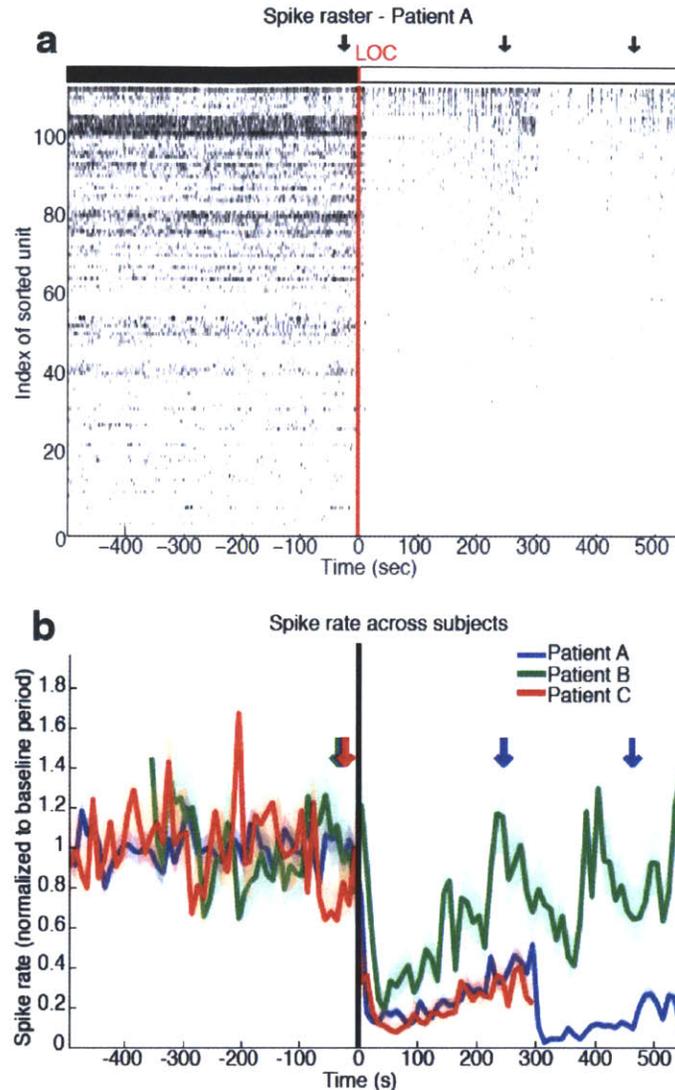
**Figure 2.1 Bandpower changes relative to LOC.** Power in different frequency bands from a representative microelectrode LFP, for each patient. None of the bands show a strong change at LOC that is then maintained throughout the post-LOC period. Dashed line indicates LOC and arrows are times of propofol delivery ( $\pm 20$  seconds). A) Theta power transiently decreases after LOC. B) Alpha power transiently decreases after LOC. C) Low gamma power is variable across patients, and tends to increase minutes after LOC. D) High gamma power is variable across patients. E) Overall spectrum averaged across the grid of ECoG electrodes from all patients. Blue line shows spectrum in the 90 seconds before LOC; green line shows spectrum in the 90 seconds after LOC. The unconscious state shows an increase in both low-frequency (0.1-4 Hz) and gamma (25-50 Hz) power.

### **2.3.2 Spike rates are highly variable after loss of consciousness**

To determine the relationship between spike rate changes and LOC, we first examined the overall spike rate in a local network of cortical neurons. Consistent with propofol's enhancement of GABAergic signaling, widespread suppression of spiking was observed after LOC. In each patient, the spike rate across the population of units decreased significantly 0-30 seconds post-LOC (Fig. 2.2, Table 2.1, range of values is the range across patients). Mean spike rates across all units reached a minimum 35-85 seconds after LOC, having decreased 81-92% from the baseline awake state. However, spike rates subsequently recovered over several minutes (Fig. 2.2b). At 4 minutes post-LOC, the rate across the entire population of units varied widely, ranging from 33% of baseline in Patient A, up to 117% of baseline in Patient B. At this 4-minute post-LOC period, individual units also displayed a wide range of spike rates, with some as high or higher than baseline: only 35.2% of units still had spike rates significantly below baseline, whereas 55.1% of units were not significantly different and 9.7% of units had significantly increased spike rates. We conclude that propofol rapidly causes a near-complete suppression of cortical spiking, but this suppression is transient, and after several minutes many individual neurons recover to baseline spike rates. The fluctuation in spike rates across time, which could have come about from changing propofol blood levels, demonstrates that brain state is dynamic after LOC. However, subjects remained unconscious throughout this period despite widely varying spike rates, suggesting that unconsciousness is not strictly associated with gross spike rate changes.

	Spike rate decrease	Low gamma power (25-50 Hz) increase	Slow power (0.1-1 Hz) increase	Modulation index increase
PatA	[10 15] sec	[10 15] sec	[5 10] sec	[-5 0] sec
PatB	[25 30] sec	[170 175] sec	[0 5] sec	[5 10] sec
PatC	[0 5] sec	[200 205] sec	[-5 0] sec	[0 5] sec

**Table 2.1: Time at which each feature becomes different from baseline.** Time is relative to LOC and calculated in 5 second bins. Only the slow oscillation reliably increases within ~5 seconds of LOC, occurring within a 5 second bin that either overlaps with LOC or is adjacent to a bin containing LOC. The modulation index, measuring coupling between spikes and the slow oscillation phase, also increases within a 5 second bin of LOC. Spike rates and gamma power are affected later, and the timing of those changes is more variable across patients.



**Figure 2.2: Propofol induction of unconsciousness causes a sharp drop and slow recovery of spike rates.** Arrows indicate approximate times of propofol administration ( $\pm 20$  seconds). A) Example spike raster from patient A. Units are sorted by post-LOC spike rate. Red line indicates time of LOC. A second drop in spike rate is visible after a second propofol bolus. B) Bayesian state-space estimate of population spike rate by patient, locked to LOC (black line). Population spike rate is normalized to pre-LOC period. Shaded region shows 95% confidence intervals. All spike rates drop within 0-30 seconds of LOC and then begin rising  $\sim 1$  minute after LOC (second drop in patient A occurs after second propofol bolus).

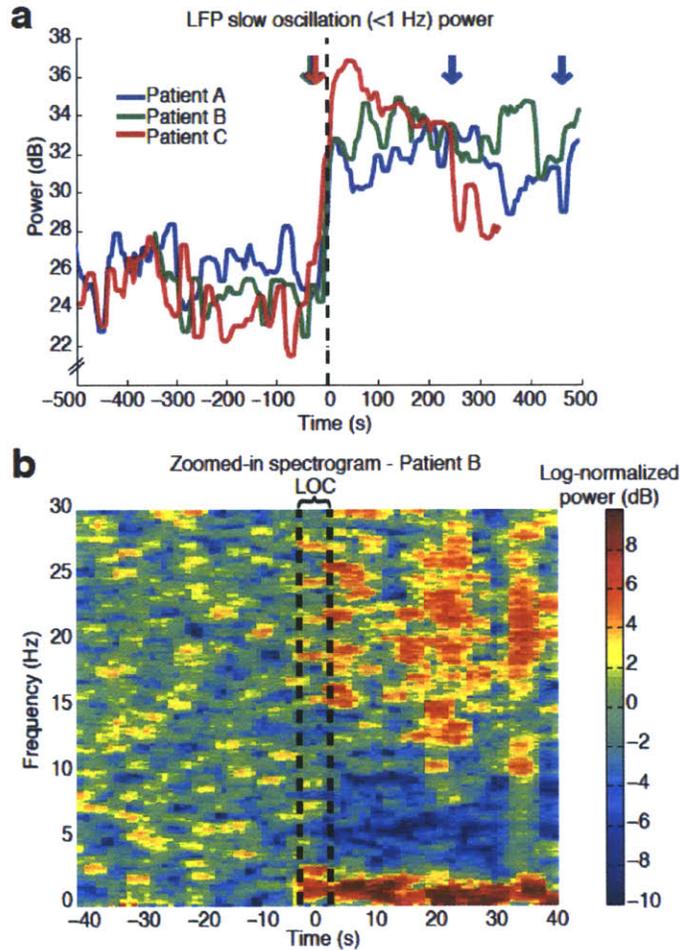
### 2.3.3 Spiking activity is organized into periods of activity and quiescence after loss of consciousness

Given that mean spike rates did not exhibit a fixed relationship with state of consciousness, we examined whether unconsciousness was instead associated with a

change in the temporal structure of spiking. We observed that spiking activity across the population of units occurred in short periods of activity that were interrupted by periods of silence. To conservatively estimate the amount of time with no spike activity, we binned spikes from all units into 400 millisecond bins. We found that 63% of bins contained no spikes, significantly more than simulated neurons with a constant rate (33%,  $p < 0.001$  for each patient, Pearson's chi-square test). We therefore concluded that cortical networks can be highly active during unconsciousness, but this activity is concentrated in short periods that are followed by profound suppression.

#### **2.3.4 The unconscious state is marked by a rapid increase and stable maintenance of power in the slow oscillation band**

The slow oscillation is known to modulate neuronal spiking (Contreras and Steriade, 1995; Steriade et al., 1993b), and we therefore examined the timecourse of its onset relative to LOC. Before LOC, power in the slow oscillation band (0.1-1 Hz) was stable (s.d.  $< 7\%$  in each patient pre-LOC). At LOC, power in the slow oscillation band increased abruptly by 35-70% (Fig. 2.3a,b), and this power increase occurred within one 5-second window of LOC in all patients (Table 2.1). The slow oscillation power then persisted at this high level for the remainder of the recording, with 99.0% of the post-LOC time bins having higher slow oscillation power than any time bin during baseline (Fig. 2.3a). We therefore concluded that power in the slow oscillation band is modulated simultaneously with LOC, and is preserved thereafter despite large fluctuations in spike rate.



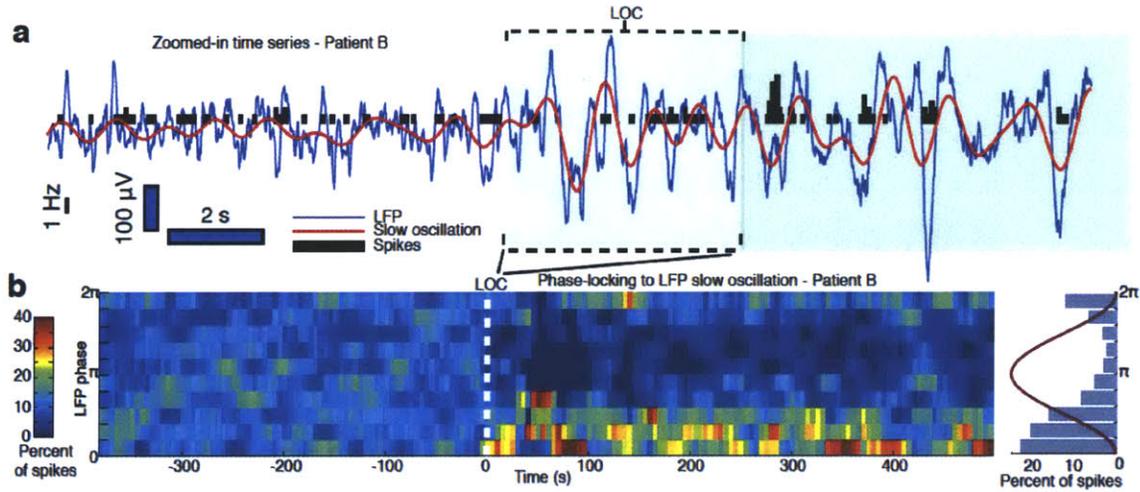
**Figure 2.3: The slow oscillation develops abruptly at loss of consciousness (LOC) and is maintained thereafter.** Dashed black line indicates LOC in all panels. A) Slow oscillation (0.1-1 Hz) power in the LFP from a representative microelectrode across time, where time is computed relative to LOC. Line color indicates which patient the data is from. Slow oscillation power increases sharply at LOC in all patients and remains higher than baseline throughout the post-LOC period. B) Zoomed-in spectrogram from a representative microelectrode in Patient B, where power is normalized within each frequency band to the pre-LOC period. The abrupt and stable power increase after LOC is specific to the slow oscillation band.

We next examined other frequency bands to investigate whether the power change at LOC was specific to the slow oscillation band, or whether other frequency bands showed a similar relationship. Although power in the >10 Hz range increased slowly after LOC, theta (3-8 Hz) power showed the opposite trend, decreasing 20-30% after LOC (Fig. 2.2b, Fig. 2.1). In addition, power in all these bands continued to undergo modulations for several minutes, rather than maintaining a consistent change after LOC

(Fig. 2.1), perhaps due to differences in propofol dosage during the maintenance phase. The stable increase in power at LOC was therefore specific to the slow oscillation band. These results demonstrated that both spike rates and many oscillatory features (gamma, alpha, theta) were highly variable after LOC. In contrast, slow oscillation power increased abruptly at LOC and remained elevated throughout the rest of the recording (Fig. 2.3a). We therefore concluded that onset of power in the slow oscillation band is associated with the transition into unconsciousness, whereas other oscillatory features do not reach a steady state until minutes later, and may reflect dynamic neural shifts at varying concentrations of propofol.

### **2.3.5 Neuronal spiking becomes phase-coupled to the slow oscillation at loss of consciousness**

Studies of deeply anesthetized animals have shown that neuronal spike activity is coupled to the phase of the slow oscillation (Contreras and Steriade, 1995; Steriade et al., 1993b; 1993a). We examined whether this spike-phase relationship developed immediately at LOC and whether it was consistently maintained thereafter. In each patient, population spike activity after LOC was significantly phase-coupled to the LFP slow oscillation (0.1-1 Hz), with 46.6% of spikes from all units occurring near the trough of the slow oscillation, during a phase of 0 to  $\pi/2$  (max. spiking at a phase of  $\pi/20$  to  $4\pi/20$ ). Phase-coupling developed within seconds of LOC (between -2.5 and 7.5 seconds; Table 2.1) and persisted throughout the ensuing changes in spike rate (Fig. 2.4). Spikes were also phase-coupled to the slow oscillation in the nearest ECoG channel but at a significantly different phase (max. phase=0 to  $\pi/10$ ;  $p<0.001$ , Kolmogorov-Smirnov test), suggesting that the LFP slow oscillation has a different relationship to spiking than the nearby, larger-scale ECoG recording. These results support the hypothesis that spikes become phase-coupled to the slow oscillation at LOC.



**Figure 2.4: Spikes become phase-coupled to the slow oscillation at loss of consciousness (LOC).** Dashed line indicates LOC in all panels. A) Example LFP from a representative microelectrode in patient B. Filtered slow oscillation is overlaid in red, and the mean spike rate across all units is in black, showing onset of slow oscillation at LOC. The LOC period is shaded in light gray and the post-LOC period in dark gray. B) Left panel: Phase-coupling of all single units in Patient B to their local LFP slow oscillation, where color indicates the percent of spikes in a given phase bin. Plot demonstrates that phase-coupling begins at LOC. Right panel: Red line shows a sinusoid to indicate slow oscillation phase. Histogram shows the phase distribution of all post-LOC spikes, which are coupled to the rising phase of the slow oscillation.

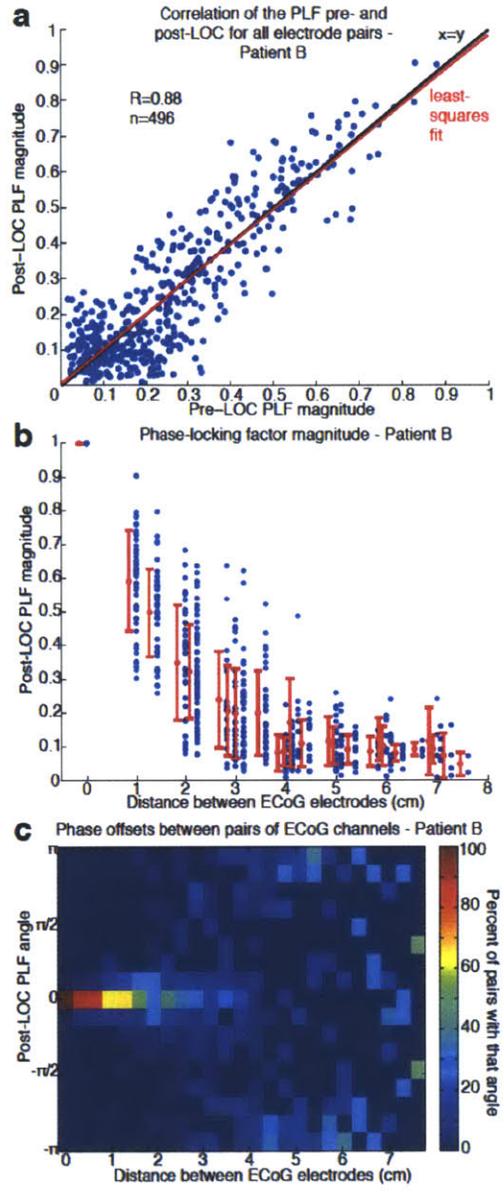
When examining individual units, most (67.2% of the 183 units with post-LOC spiking) were significantly phase-coupled to the LFP slow oscillation ( $p < 0.05$ , Pearson's chi-square test). When this analysis was restricted to units with post-LOC spike rates over 0.1 Hz, 94.0% of units had significant phase coupling ( $p < 0.05$ ,  $n = 50$ , Pearson's chi-square test). Of the units without significant phase-coupling, 65.0% also showed peak spiking activity within a phase of  $0$  to  $\pi/2$ , demonstrating that most units had the same phase-coupling trend. These results demonstrated that after LOC nearly all spiking activity is tightly coupled to the slow oscillation phase, and is suppressed for a large portion of the slow oscillation cycle. We refer to these periods of high spiking as 'ON' states and the silent periods as 'OFF' states to remain consistent with previous work using only extracellular recordings (Nir et al., 2011; Vyazovskiy et al., 2009). Due to the alternation of ON and OFF states, spike activity was limited to periods of a few hundred milliseconds, interrupted by periods of silence that can also last hundreds of milliseconds.

We therefore concluded that the slow oscillation marks a temporal fragmentation of cortical spiking that occurs at LOC.

### **2.3.6 The slow oscillation impairs information transfer between distant cortical regions**

Given that post-LOC spiking is periodically interrupted within a cortical region, we investigated whether communication across distant areas was also affected. We examined slow dynamics across the grid of ECoG electrodes in the two patients (A and B) for whom we had at least three minutes of post-LOC ECoG data. Since spiking was strongly coupled to slow oscillation phase, we examined how this phase varied across the brain to infer the relative timing of neuronal activity in different cortical regions.

We quantified the phase relationships between different cortical regions using the phase-locking factor (PLF), which characterizes the phase offset between two oscillations over a period of time (Lachaux et al., 1999). The PLF magnitude ranges between 0 and 1, and quantifies the stability of the phase offset (1 reflects constant phase offset, 0 represents variable phase offset). The PLF angle indicates the average phase offset. We calculated the PLF between every pair of ECoG channels on the grid (8x4 or 8x8 cm, n=96 total electrodes) to determine the relationship between local and distant slow oscillations. We found that the PLF magnitude was conserved between the pre- and post-LOC states (Fig. 2.5a, correlation coefficient  $R=0.66$ , pat. A;  $R=0.88$ , pat. B;  $p<10^{-50}$  for each, t-test), with a small but significant increase in PLF magnitude post-LOC (mean increase=0.02-0.07,  $p<0.01$ , Wilcoxon signed rank test). This result was consistent with previous findings that low-frequency correlations in neural activity are maintained after loss of consciousness (Breshears et al., 2010; Vincent et al., 2007), and suggests that LOC is associated with only a slight shift in the strength of phase relationships between slow oscillations in different areas.



**Figure 2.5: Slow oscillations in distant ECoG channels have variable phase offsets.**

The phase-locking factor (PLF) characterizes the stability of the phase offset between two oscillations over a period of time, selected as either pre- or post- loss of consciousness (LOC). The PLF magnitude ranges between 0 and 1, where 1 reflects constant phase offset, and 0 represents variable phase offset. The PLF angle indicates the average phase offset. A) PLF magnitude between each pair of ECoG electrodes, with pre-LOC PLF on x-axis, and post-LOC PLF on the y-axis. Pre- and post-LOC PLF magnitude are highly correlated. Red line marks the line of best fit to the data. B) PLF magnitude during the post-LOC recording, plotted according to the distance between the electrodes in each pair. The PLF magnitude decreases significantly with distance, reflecting higher variability in phase offsets between distant ECoG electrodes. Red lines show mean (+/- s.d.) of the PLF magnitude at all electrode pairs with that distance. C) 2-D histogram of the PLF angle between all ECoG pairs post-LOC, showing that the mean phase offset

between distant channels is more variable than between nearby channels, with values as large as  $\pi$ .

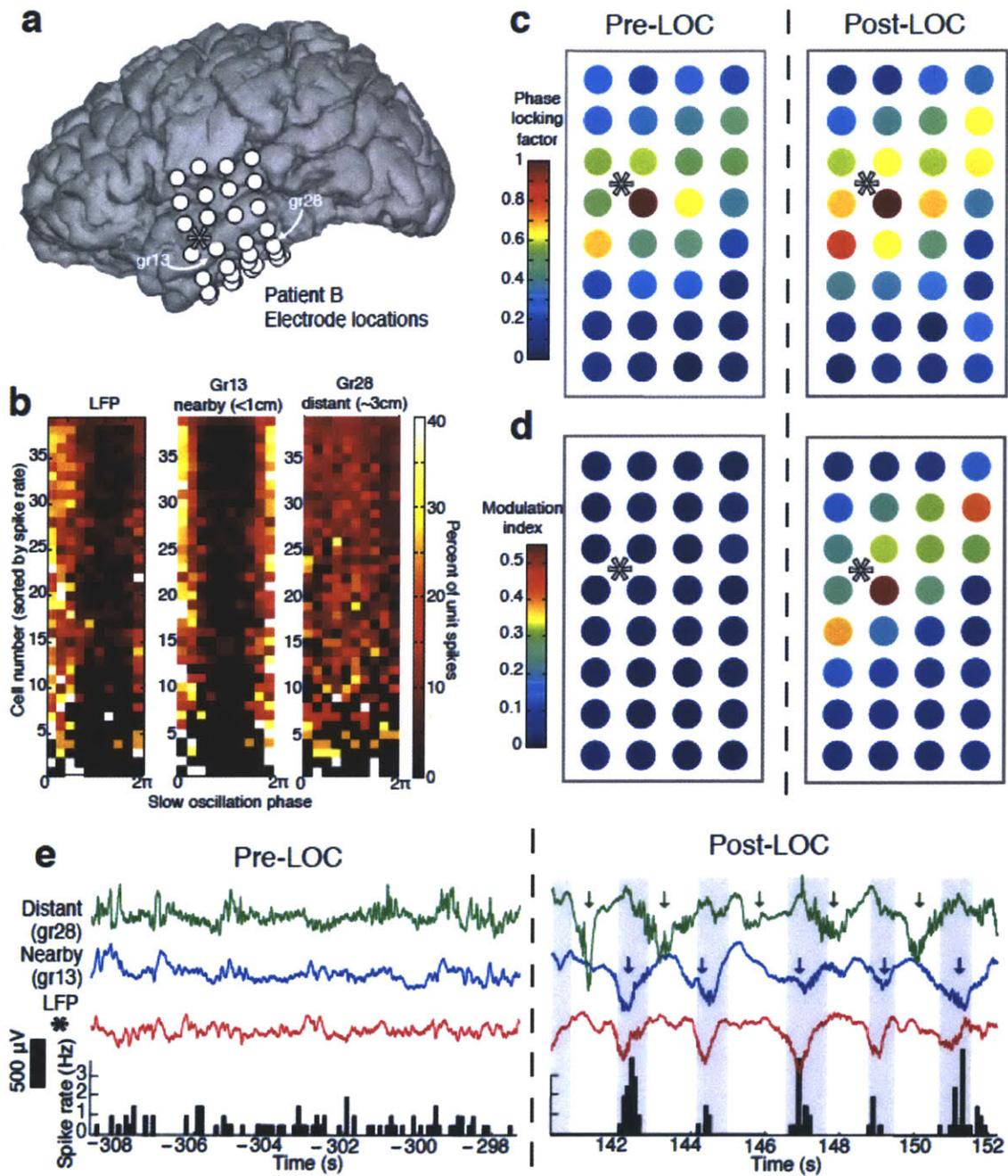
We next examined how the PLF varied with distance, to determine whether slow oscillations in different cortical regions were at different phases. The PLF magnitude dropped significantly with distance (Fig. 2.5b, 2.6c,  $R=-0.61$ , pat. A;  $R=-0.82$ , pat. B;  $p<10^{-6}$  for each), demonstrating that the phase offsets between distant slow oscillations were variable. We also examined the mean phase offsets (PLF angle): mean phase offsets between distant channels varied across a wide range, spanning 0 to  $\pi$  (Fig. 2.5, 2.7). Because a phase offset of just  $\pi/4$  corresponds to a lag of approximately 250 ms, slow oscillations in distant ECoG channels therefore had substantial timing differences. These results demonstrated that distant slow oscillations were often at different phases than the local oscillation, and these phase differences were not stable across time.

To examine how these phase offsets would affect neuronal activity, we examined the phase relationship between local spiking and slow oscillations measured across the ECoG grid. We measured spike phase-coupling as a modulation index (MI) quantifying the Kullback-Liebler divergence, in bits, between the observed phase distribution and a uniform distribution (see Experimental Procedures). A large MI indicates a strong relationship between local spiking and ECoG phase, whereas an MI of zero indicates no relationship. In the pre-LOC period, MI values were consistently small across all ECoG channels (Fig. 2.6d, MI range: 0.001-0.04 bits), demonstrating that slow oscillation phase was not associated with strong suppression of spiking in the pre-LOC period. After LOC, the MI was significantly more variable across channels (range: 0.006-0.62 bits,  $p<0.01$  in each patient, Levene's test). Spikes were strongly phase-coupled to the slow oscillation in the nearest ECoG channel, and this relationship declined significantly with distance (Fig. 2.6b,d, Fig. 2.7;  $R=-0.40$ , pat. A;  $R=-0.68$ , pat. B;  $p<0.001$  in each patient).

Taken together, our analysis of phase-phase and spike-phase coupling show that the post-LOC state is characterized by periodic and profound suppression of spiking coupled to the local slow oscillation phase, and that this phase is not consistent across cortex. Given the strong relationship between phase and ON/OFF periods, this result suggests that after LOC, ON periods in distant ( $>2$  cm) cortical regions occur at different times (Fig. 2.6e, right panel). In contrast, low-frequency oscillations in the pre-LOC state

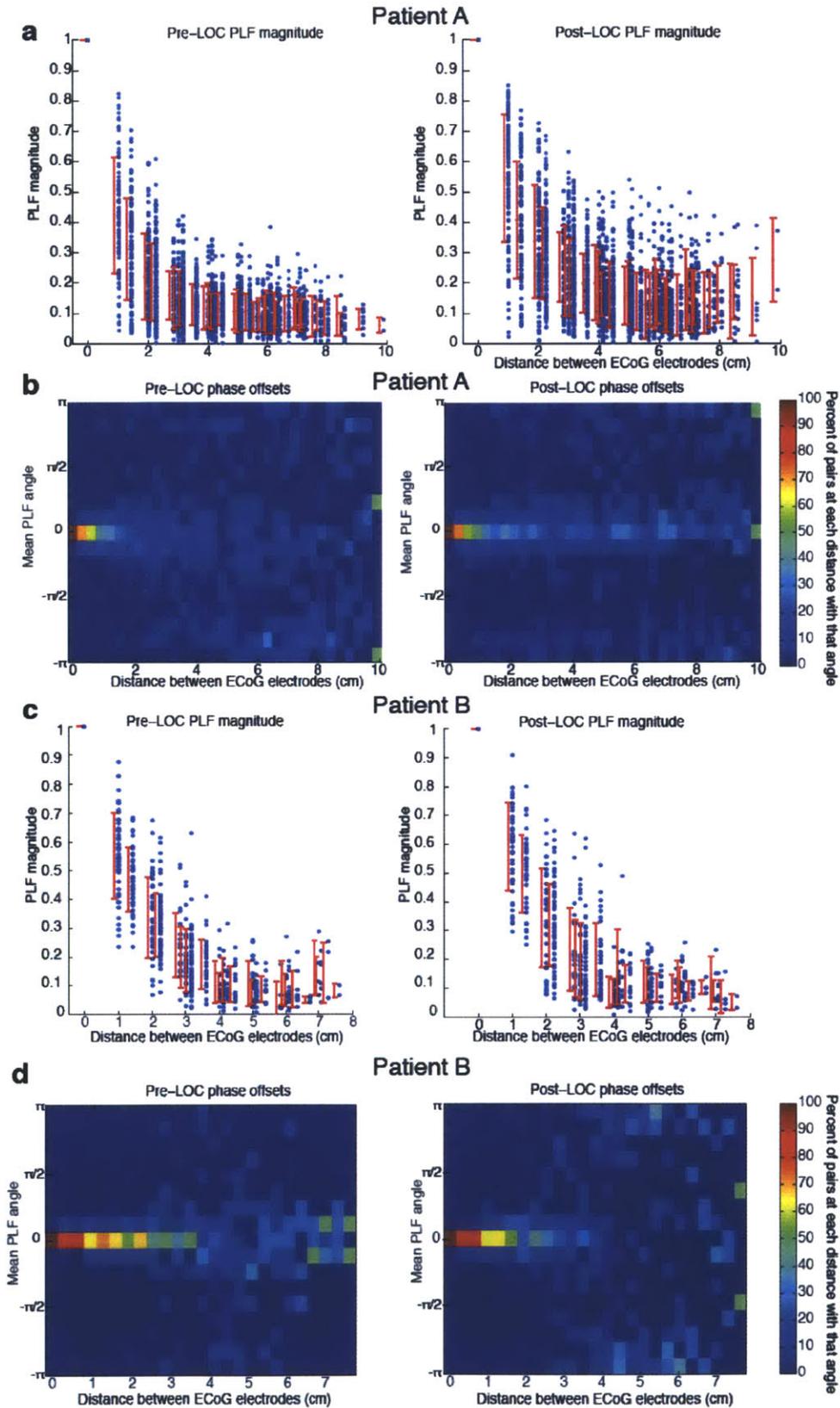
are not associated with strong suppression of spiking, so neurons are able to fire at any phase of local or distant slow oscillations despite the presence of phase offsets (Fig. 2.6e, left panel). The combination of phase offsets and strong phase-coupling of spikes that occurs at LOC is therefore expected to produce a disruption of communication between distant cortical areas, as one cortical area will frequently be profoundly suppressed while another area is active.

Although spikes were not strongly phase-coupled to distant slow oscillations during the post-LOC period, there were several electrodes located more than 3 cm from the spike recording site that showed a statistically significant relationship (Fig. 2.8). In these cases, phase-coupling was weak and the phase of maximal spiking was shifted, consistent with our conclusion that distant cortical regions are unlikely to have simultaneous ON periods. However, this finding raises the possibility that despite the asynchrony of slow oscillations across the brain, there might still be a link between slow oscillations in distant cortical regions. Given the observed phase offsets (which ranged up to  $\pi$ ), such coupling would frequently occur over hundreds of milliseconds, and would not reflect precisely timed inputs and interactions. Overall, these analyses support the conclusion that distant cortical regions were frequently at a suppressed phase of the slow oscillation while the local network was active. Activity within a cortical area was therefore isolated, impairing communication between distant regions.



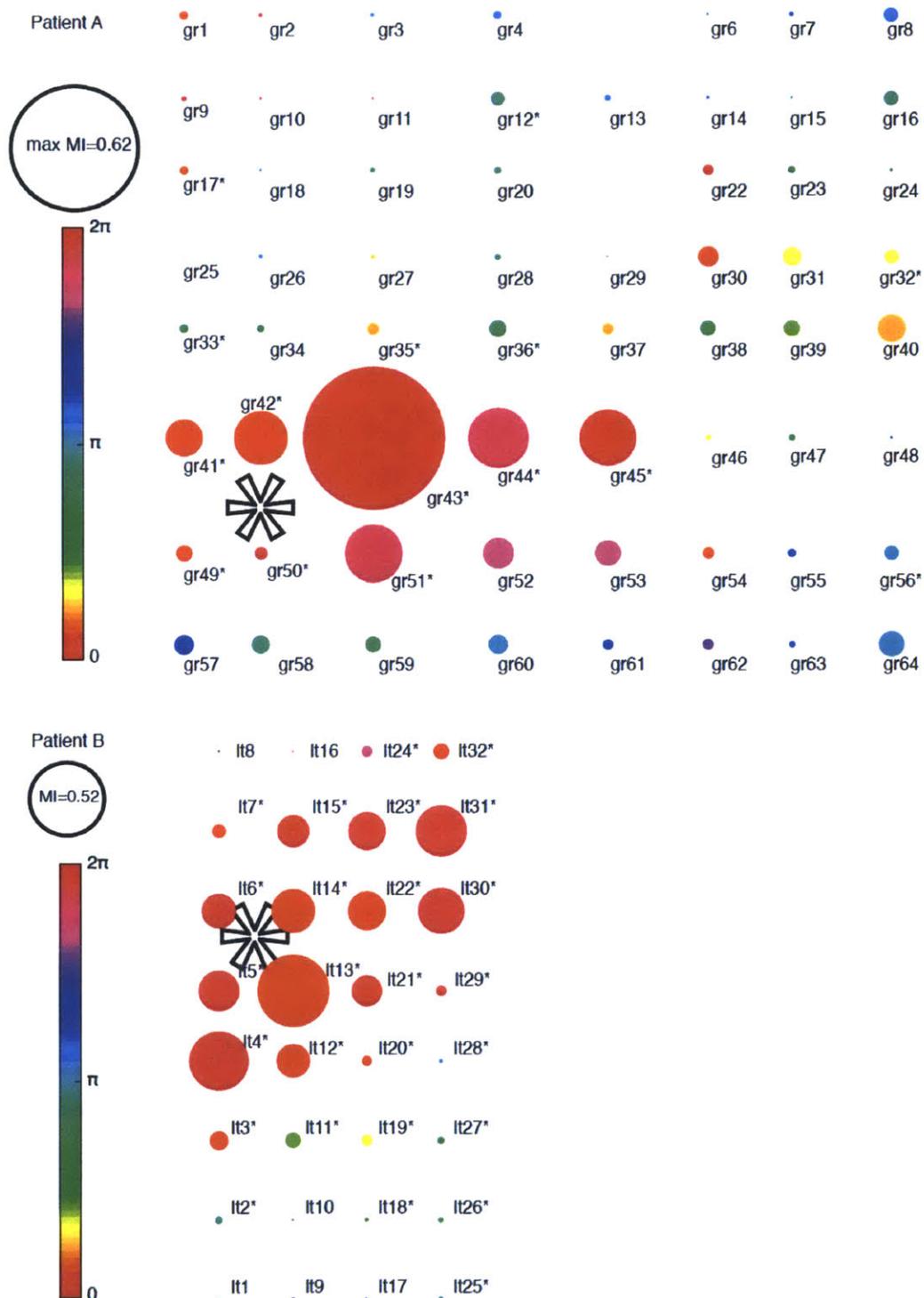
**Figure 2.6: After loss of consciousness (LOC), slow oscillations are asynchronous across cortex and are associated with ON/OFF states; therefore, distant cortical areas are frequently at a suppressed phase during local ON periods.** A) Position of ECoG and microelectrode recordings in Patient B. Each white circle marks the location of an ECoG electrode, and the microelectrode where spikes were recorded is marked with a star. B) Phase histograms for every single unit in Patient B. Each panel displays the spike coupling to phase in a different recording site. Units are arranged by post-LOC spike rate (highest rate at top of plot), and the same phase-coupling trend is visible across all units to the phase of the LFP and nearby ECoG (gr13). In contrast, the slow oscillation in the distant ECoG channel (gr28) does not have the same phase relationship to local

spiking. C) Magnitude of the phase-locking factor (PLF) between every ECoG electrode relative to the ECoG closest to the spike recordings (gr13). The PLF drops with distance in both the pre- and post-LOC states, showing that distant areas have variable phase offsets relative to the local recording. D) Modulation index (MI) quantifying the strength of the spike-phase relationship. The MI is consistently low in the pre-LOC state, demonstrating the absence of a strong spike-phase relationship. Post-LOC, the MI is high only in local ECoG recordings, demonstrating that spikes are strongly phase-coupled to local slow oscillations and this relationship weakens with distance. E) Traces from a nearby ECoG (blue), distant ECoG (green), representative LFP channel (red), and mean spike rate across all units (black). Arrows mark the times that the ECoG slow oscillation is at phase zero (when local spike rates are expected to be high; see Fig. 3). Gray shading marks the times at which the LFP slow oscillation phase is between  $-\pi/4$  and  $3\pi/4$ , when most spiking occurs. Plots show that post-LOC, local spikes occur in ON periods that typically overlap with the zero phase in the nearby ECoG channel, but frequently do not overlap with the zero phase in the distant ECoG channel.



**Fig. 2.7: Slow oscillations in distant ECoG channels have variable phase offsets.** The phase-locking factor (PLF) characterizes the stability of the phase offset between two

oscillations over a period of time. The PLF magnitude ranges between 0 and 1, where 1 reflects constant phase offset, and 0 represents variable phase offset. The PLF angle indicates the average phase offset. A and C) PLF magnitude between every electrode pair during the pre-LOC and post-LOC recording periods for patients A and B, plotted according to the distance between the electrodes in each pair. The PLF magnitude decreases significantly with distance, reflecting higher variability in phase offsets between distant ECoG electrodes. Red lines show mean and standard deviation of the PLF magnitude at all electrode pairs with that distance. B and D) 2-D histogram of the PLF angle between all ECoG pairs, in the pre- and post-LOC periods, showing that the mean phase offset between distant channels is more variable than between nearby channels, with values as large as  $\pi$ .



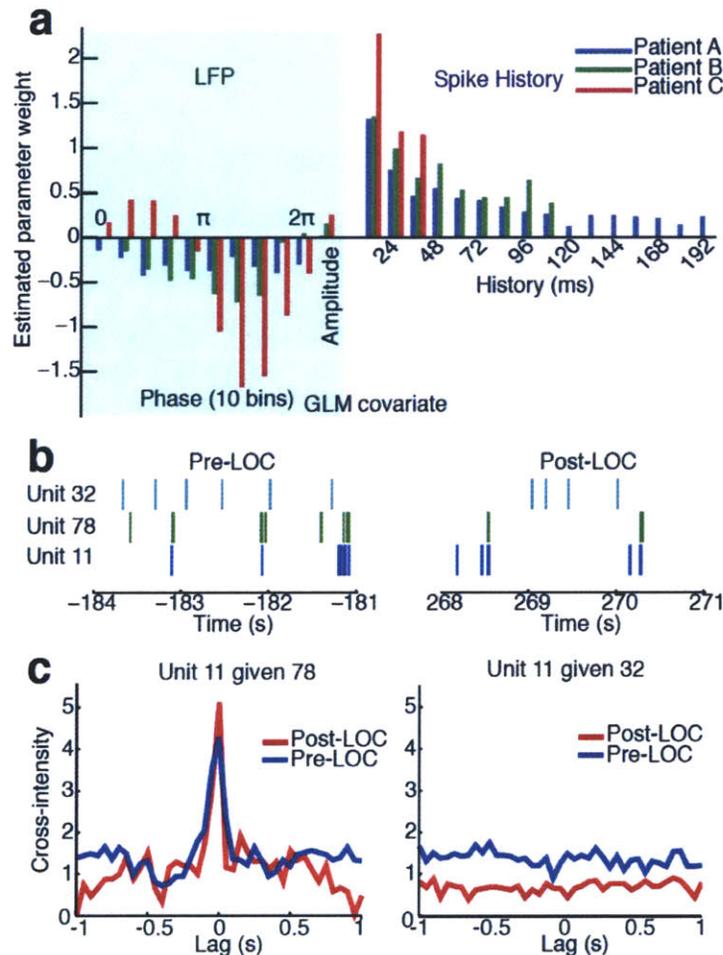
**Fig. 2.8: Phase-locking across the ECoG grid.** Star indicates location of spike recordings. Color of circles indicates the phase of the slow oscillation at each site that is associated with maximal local spiking, and the size of the circle is the MI at that electrode. Stars next to channel labels indicate that spiking at the microelectrode array is significantly modulated relative to the slow oscillation phase from that ECoG recording site.

### **2.3.7 Local network structure is preserved after loss of consciousness**

Having observed interruptions in local activity and disruption of long-range communication, we examined whether connectivity within the local cortical network was also impaired. We fit a generalized linear model (GLM) to spike activity from the ensemble of units, to test whether spiking could be predicted by the slow oscillation phase alone or whether the history of local network activity also contributed (Truccolo et al., 2005). We used the Bayesian Information Criterion (Schwarz, 1978) to select the number of covariates to include in the model. In each patient, we found that this model included >30 milliseconds of population spike history (Fig. 2.9a). Ensemble spike history therefore predicted future spiking, demonstrating that although cortical activity was limited to brief ON periods, inter-unit structure existed within these periods. This pattern resembled the pre-LOC state, in which recent spike history (0-48 ms) was predictive of future spikes while more distant spike history contributed less. This result suggests that after LOC, cortical activity is not reduced to disordered spiking during ON periods. Instead, significant structure is maintained between nearby neurons during their brief periods of activity.

Structure between single units was further reflected in a peak in the cross-intensity function between several pairs of units, demonstrating millisecond-scale synchronization of spike activity (Fig. 2.9b,c). To examine whether pairwise synchronization persisted after LOC, we analyzed the cross correlation between the 15 units with the highest spike rates in patient A. Of the 103 pairs (excluding pairs recorded on the same electrode), 21 were significantly correlated pre-LOC ( $p < 0.05$ , exact Poisson test relative to baseline from shuffled spikes, Bonferroni correction for multiple comparisons). After LOC, 71.4% of these pairs remained significantly correlated. In contrast, only 18.3% of pairs that were not correlated pre-LOC became correlated post-LOC, a significantly smaller number ( $p < 10^{-5}$ , Fisher's exact test). This result demonstrated that after LOC, pairs of units tended to retain the same correlation structure that they had before LOC, whether it was the presence or absence of a correlation. Taken together, both the GLM and paired correlation results show that significant inter-unit

connectivity is maintained within post-LOC ON periods. This suggests that the dominant change after LOC is the isolation of cortical networks, while aspects of local network structure may remain unaltered.



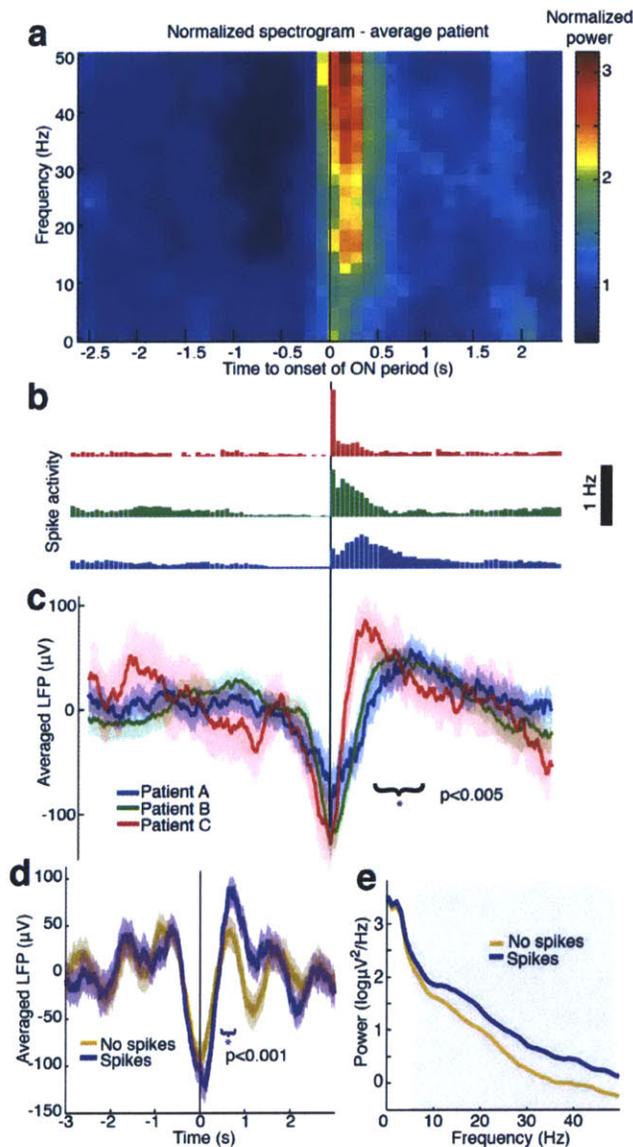
**Figure 2.9: Spikes occur in brief ON periods that maintain inter-unit structure.** A) Parameter estimates from the best GLM model for population spiking after loss of consciousness (LOC). For each patient, the best model includes information from both LFP phase and recent population spike history. B) Time-series example: units 11 and 78 spike together; unit 32 has a similar spike rate but does not. These units preserve the same correlation structure after LOC that they had before LOC. C) Example cross-intensity functions (square-root estimate) pre- and post-LOC, showing that the pre-LOC inter-unit structure persists after LOC.

### **2.3.8 Spiking activity is associated with modulations in slow oscillation shape and higher frequency power**

The mechanisms underlying the slow oscillation are debated (Blethyn et al., 2006; Crunelli and Hughes, 2009; Sanchez-Vives and McCormick, 2000; Steriade et al., 1993a; Timofeev et al., 2000). Therefore, we examined the relationship between spike activity and slow oscillation shape in greater detail. We calculated an average LFP triggered at the beginning of ON periods. The triggered average demonstrated that ON periods begin at the minimum of the LFP slow oscillation (Fig. 2.10b,c). In addition, the LFP slow oscillation was asymmetric (Fig. 2.10c), with a higher peak after spiking than before spiking (mean difference = 40.7  $\mu\text{V}$ ,  $p < 10^{-5}$ ,  $p < 0.005$  for each patient, t-test). We tested whether this asymmetric shape occurred on all cycles of the slow oscillation, or was specific to cycles with high spike activity. We compared cycles of the LFP slow oscillation that contained spikes to cycles that did not, matching the amplitudes of the slow oscillation minimum. Cycles that were not associated with spikes were symmetric (mean difference = 0.3  $\mu\text{V}$ ,  $p > 0.9$ , t-test), whereas those that were associated with many spikes produced a higher peak after spiking (Fig. 2.10d; mean difference = 32.2  $\mu\text{V}$ ,  $p < 0.001$ ,  $p < 0.05$  for each patient, t-test). This asymmetry did not extend to the nearby ECoG recording (Fig. 2.11), suggesting that the relationship between spike activity and slow oscillation shape is a highly local effect limited to less than a centimeter (i.e., the spacing in the ECoG grid). These results demonstrated that high spike rates are associated with an increased slow oscillation peak in the LFP, potentially reflecting enhanced suppression after spike activity.

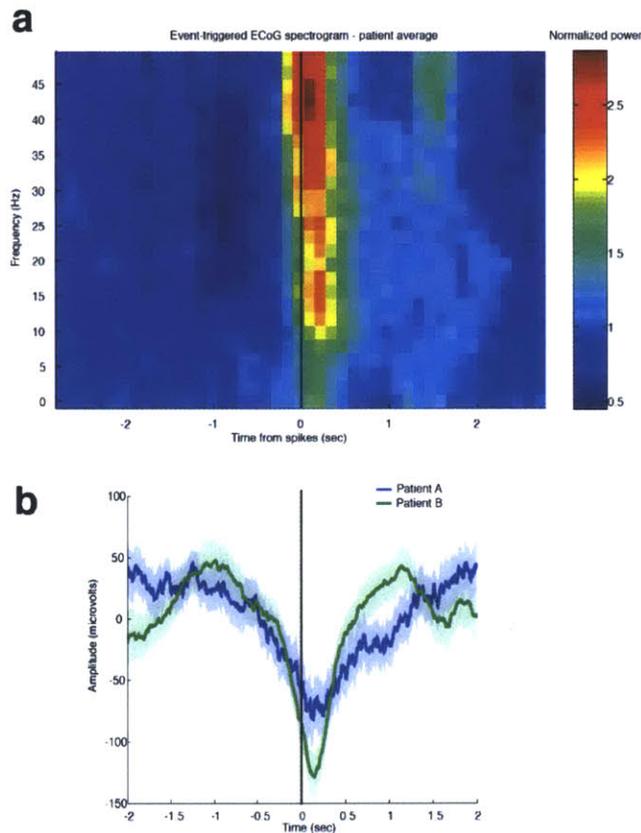
Because low gamma (25-50 Hz) power also increased after LOC (Fig. 2.1, Table 2.1), we examined its relationship to spike activity as well. ON periods were associated with significantly increased broadband (0-50 Hz) power in the LFP and ECoG (Fig. 2.10a, 2.11,  $p < 0.05$ , F-test, Bonferroni correction for multiple comparisons across frequencies). LFP power in alpha, beta, and gamma bands was significantly higher in slow oscillation cycles with high spike activity than in cycles with low spike activity (Fig. 2.10e;  $p < 0.05$ , F-test, Bonferroni correction for multiple comparisons across frequencies). These results showed that in addition to the slow changes in gamma power occurring over minutes (Fig. 2.1), gamma power also fluctuated at the timescale of the

slow oscillation (0.1-1 Hz) and was higher during ON periods. We therefore concluded that after loss of consciousness, power in the low gamma range is associated with high local spike rates. This result suggested that the gradual increase in gamma power after LOC may be related to the post-LOC fluctuations in spike rate rather than reflecting dynamics induced specifically at LOC.



**Figure 2.10: Spike activity is associated with modulations in slow oscillation morphology and gamma power.** A) LFP spectrogram (top panel) triggered at onset of ON periods (black line). Power is normalized within each band. B) Population spike histograms for each patient, demonstrating spike activity locked to ON period detection times. C) LFP average time-series triggered at ON period onset times, showing an increased LFP peak after spike activity compared to before spike activity. Shaded region is approximate 95% confidence intervals. The slow oscillation peak after spiking is

significantly higher than before spiking. D) LFP triggered at troughs of the slow oscillation. Purple trials are cycles of the slow oscillation associated with many spikes, and brown trials are cycles that were not associated with any detected spikes (n=153 cycles in each condition). Only trials associated with high spike rates show an asymmetric LFP peak. Shaded region is approximate 95% confidence intervals. F) Power spectra of LFP time-series triggered at the trough of the slow oscillation. Gray region indicates significant difference in >10 Hz power during trials with high spike rates.



**Fig. 2.11: ON period-locked effects in the ECoG channel near the spike recordings.** A) Triggered spectrogram locked to the onset of ON periods, averaged across patients A and B, with power normalized within each frequency band. There is increased high-frequency power during spiking activity. B) ON period-triggered ECoG recording: the ON periods begin just before the slow oscillation trough, and are not associated with an asymmetrical peak after spiking.

## **2.4 Discussion**

### **2.4.1 Isolation of local cortical networks at loss of consciousness**

In summary, we found that rapid induction of unconsciousness using propofol causes the human brain to undergo an abrupt change in network dynamics, tipping into a new state in which neuronal activity is coupled to slow oscillations in the local field potentials. Neural dynamics can be highly variable during the unconscious period, as spike rates and most oscillatory patterns continue to fluctuate for minutes after loss of consciousness (Figs. 2.1, 2.2). The slow oscillation has a markedly different pattern: it develops simultaneously with LOC and maintains this increase thereafter (Figs. 2.3, 2.4). Spiking activity is constrained to brief time periods coupled to the phase of the slow oscillation (Figs. 2.4, 2.6), interrupting information processing within a cortical area. Moreover, these brief activity periods are phase-shifted across cortex (Figs. 2.5, 2.6, 2.7), limiting activity spatially, as different cortical areas are likely to be active at different times. However, functional connectivity within the local network is preserved (Fig. 2.9). These results suggest that the slow oscillation prevents both sustained information processing within an area and communication between distant cortical areas. General anesthetics have been proposed to cause unconsciousness by disrupting cortical integration (Alkire et al., 2008; Mashour, 2004). Our results are consistent with this hypothesis and demonstrate that in the case of propofol, spatiotemporal slow oscillation dynamics may mediate the breakdown of communication by isolating local cortical networks.

### **2.4.2 Slow oscillations and functional connectivity**

The slow dynamics reported here may also underlie the observation that gamma coherence decreases after propofol-induced loss of consciousness, particularly across distant cortical regions (Alkire et al., 2008; John et al., 2001). We found that spiking activity is strongly associated with gamma power (Fig. 2.10), and spiking is unlikely to occur simultaneously in distant cortical regions due to the asynchronicity of slow oscillations across the brain (Fig. 2.5). Slow oscillations may therefore impair coupling of gamma oscillations between cortical areas, and this effect could produce gamma oscillations that are not coherent over long distances.

Low-frequency spatial correlations in fMRI (Vincent et al., 2007) and ECoG (Breshears et al., 2010), sometimes used to assess functional connectivity, have been found to remain invariant after loss of consciousness under propofol. Our analysis of the PLF magnitude, which has a similar spatial distribution before and after LOC (Fig. 2.5a), corroborates this observation. Our studies show that while the low-frequency spatial relationships remain similar before and after LOC, the functional properties of low frequency oscillations change at LOC, grouping spiking into brief ON states that are disjoint across space.

#### **2.4.3 Potential circuit mechanisms underlying the slow oscillation**

The mechanisms underlying the slow oscillation are unclear. One hypothesis is that the slow oscillation is cortically generated (Sanchez-Vives and McCormick, 2000; Steriade et al., 1993a; Timofeev et al., 2000), while others suggest that it results from an interaction between cortex, thalamus, and thalamic reticular nucleus (Blethyn et al., 2006; Crunelli and Hughes, 2009). The relationship identified here between spike activity and slow oscillation shape suggests that cortical spiking may have a causal role in the slow oscillation. Spikes predict a high amplitude peak in the LFP slow oscillation (Fig. 2.10d), but this effect does not extend to the ECoG recordings, which integrate activity from a larger population of neurons. The highly local nature of this effect suggests that cortical spiking may directly affect the local slow oscillation. One possible mechanism is that pyramidal neuron spiking during ON periods excites GABAergic interneurons, whose inhibitory actions are enhanced by propofol, driving the local network into a more hyperpolarized state. Another possibility is that spike activity may drive disfacilitation of cortical neurons, a mechanism which has been demonstrated in slow-wave sleep (Timofeev et al., 2001). These effects could be consistent with either the cortical or corticothalamic hypothesis.

#### **2.4.4 Slow oscillations in general anesthesia and in sleep**

The slow oscillation during propofol-induced unconsciousness shares several features with slow waves during sleep: in both states, spike activity is coupled to a local slow oscillation that is not synchronous across the brain (Nir et al., 2011). The

asynchronicity observed here (Figs. 2.5, 2.6e) contrasts with previous observations in anesthetized animals (Amzica and Steriade, 1995), and is most likely due to the increased spatial sampling provided by the 8-cm grid of intracranial electrodes. In addition, the preservation of pre-LOC neuronal network properties after LOC (Fig. 2.9) is consistent with the hypothesis that cortical UP states during sleep have similar dynamics to the waking state (Destexhe et al., 2007).

However, the observed patterns under propofol also show striking differences to patterns during sleep. The onset of the slow oscillation during induction of general anesthesia was abrupt (Figs. 2.3, 2.4), accompanying rapid loss of consciousness due to the bolus administration of propofol. Since general anesthesia is typically induced with a bolus of propofol, this abrupt transition into the slow oscillation is likely to occur in the majority of clinical patients when they lose consciousness during general anesthesia. During sleep, the slow oscillation develops over minutes, consistent with the gradual nature of the transition into sleep (Achermann and Borbely, 1997; Ogilvie, 2001). In both cases, slow oscillation dynamics temporally track loss of consciousness, further supporting the proposal that the slow oscillation represents a breakdown of cortical communication.

In addition, we found that periods of spike activity were brief (Fig. 2.6e), whereas sleep is characterized by persistent spiking with brief periods of suppression during slow wave events (Bersagliere and Achermann, 2010; Cash et al., 2009; Cserecsa et al., 2010; Nir et al., 2011). This observation was recently corroborated in a study of sleeping and anesthetized cats (Chauvette et al., 2011). A difference in the ratio of UP and DOWN states could provide one explanation for why propofol creates a more profound disruption of consciousness than sleep: there is less temporal overlap in neuronal spiking between different cortical regions, more reliably preventing the organization of large-scale population activity. Furthermore, recent findings that isolated OFF states in sleep-deprived rodents are associated with behavioural impairment (Vyazovskiy et al., 2011) are consistent with the hypothesis that the spatial and temporal properties of OFF states affect cortical function.

#### **2.4.5 Potential role of slow oscillations in unconsciousness**

Our results show that the slow oscillation appears abruptly at the onset of propofol-induced loss of consciousness. In addition, we demonstrate that the slow oscillation marks a state in which neuronal networks are fragmented, impairing both local and long-range communication. However, the scope of this study, performed in human subjects, does not allow us to explicitly test the causal role of the slow oscillation in unconsciousness. Furthermore, this experiment defines loss of consciousness as loss of voluntary response, but cannot disambiguate whether the ability to respond may be suppressed prior to loss of consciousness. Given the slow oscillation's association with prolonged and asynchronous periods of near-complete suppression of neuronal activity, it seems unlikely that it is compatible with conscious processing. However, other possible mechanisms for propofol-induced unconsciousness include coherent frontal alpha rhythms that limit thalamocortical function (Ching et al., 2010; Cimenser et al., 2011; Supp et al., 2011). Future work in animal models could test whether the slow oscillation is sufficient to produce unconsciousness.

#### **2.4.6 Generalizability of propofol-induced slow oscillations**

A limitation of this study is that we enrolled patients with epilepsy, and it is possible that their cortical networks differed due to seizure foci or medication history. However, several factors support the hypothesis that these results generalize to the healthy brain. First, the microelectrodes were located at least 2 cm from the seizure focus in each patient, and histology did not reveal any disruption of the local network, suggesting that the LFPs and single units were recorded from healthy cortex. Second, the overall effects of propofol were highly consistent with those observed in healthy subjects: unconsciousness was associated with increased slow oscillation power and increased gamma power, in strong agreement with previous studies (Murphy et al., 2011). These results suggest that propofol acted typically in these patients' brains. Finally, we report statistics for each individual patient, and show that the timing of the slow oscillation onset and its relationship to spiking were replicated across patients despite their individual clinical profiles. As epilepsy is a heterogeneous disease with different cortical origins, the high consistency of these results suggests that the effects reported here are

not due to the presence of epilepsy. These three observations suggest that our results are not a product of an epileptic brain, but rather reflect a true neural correlate of loss of consciousness that is likely to generalize to the healthy brain. Future studies in patients with pathologies other than epilepsy could further address this issue.

#### **2.4.7 Role of slow oscillations in other brain states**

Other anesthetic drugs, such as ketamine and dexmedetomidine, operate through different molecular and neural circuit mechanisms than propofol (Brown et al., 2011). The study presented here provides a framework in which to further study these other drugs, and identify how they influence neural dynamics to produce altered states of arousal. In addition, further work will be needed to explore how these findings may relate to other conditions, as slow wave activity is a common feature in conditions such as coma (Young, 2000) and complex-partial seizures (Blumenfeld et al., 2004), and those slow waves share some characteristics with the slow oscillations studied here. Although a unitary mechanism for unconsciousness under general anesthesia is technically possible, it is more likely that a variety of mechanisms exist that can produce unconsciousness (Rudolph and Antkowiak, 2004). We have shown here that the slow oscillation is a fundamental component of propofol-induced unconsciousness, marking an impairment of cortical integration at both the local and global scale.

## **2.5 Experimental Procedures**

### **2.5.1 Data collection**

Three patients with epilepsy intractable to medication were implanted with intracranial subdural electrocorticography electrodes for standard clinical monitoring (AdTech). Informed consent was obtained from all patients in accordance with the local institutional review board. ECoG electrode placement was determined solely by clinical criteria, and the electrodes were located in temporal, frontal, and parietal cortices. Individual ECoG electrodes within a grid were spaced 1 cm apart. In addition, a 96-channel NeuroPort microelectrode array with 1.0 mm long electrodes (BlackRock Microsystems) was implanted into the superior (Patient B) or middle (Patients A, C)

temporal gyrus, to record LFPs and ensembles of single units for research purposes. In each patient, the Neuroport array was located at least 2 cm from the seizure focus. There was no evidence of disruption in local network structure based on the firing properties of the neurons or the post-resection histological examination of the area around the array. All recordings were collected at the beginning of surgery to explant the electrodes. Anesthesia was administered as a bolus dose of propofol according to standard clinical protocol. All propofol doses were based on the anesthesiologist's clinical judgment rather than the research study considerations. Patient A received 3 boluses (130mg, 50mg, 20mg), Patient B received 1 (200mg), and Patient C received 1 (150mg). After induction, patients were transferred to a continuous intravenous infusion of propofol to maintain anesthetic levels. Throughout the induction period, patients responded to auditory stimuli (pre-recorded words and the patient's name) with a button press, and stimuli were presented every 4 seconds in order to obtain precision for LOC time on the order of seconds. LOC time was defined as the [-1 4] second period surrounding the first stimulus after the patient completely ceased responding. Spike sorting was carried out according to standard procedures (Lewicki, 1998) with Offline Sorter and produced 198 single units for further analysis. LFPs were referenced to a wire distant from the microelectrode array, and collected with hardware filters bandpassing between 0.3-7500 Hz with a sampling rate of 30 kHz. LFPs were then low-pass filtered at 100 Hz and resampled to 250 Hz. For display, raw time-series were lowpass filtered with a finite-impulse response filter of length 4464 of unit gain between 0 and 40 Hz, and attenuation of more than -300 dB above 42 Hz. ECoG was collected with a sampling rate of either 250 Hz (Patients B and C) or 2000 Hz (Patient A) in which case it was low-pass filtered at 100 Hz and resampled to 250 Hz. ECoG recordings were referenced to an intracranial reference strip channel when available (Patient A) and otherwise to an average reference. In patients A and B, ECoG recordings were collected throughout. In patient C, the microelectrode recordings were collected throughout but the ECoG recording ended ~100 seconds after LOC and therefore slow oscillation phase-coupling significance could not be assessed in ECoG channels, as the spike rate was near zero during this time. Two ECoG grid channels were rejected in patient A due to large artifact. All data were exported to Matlab (Mathworks) for analysis with custom software.

### **2.5.2 Spike rate analysis**

Spike rates and confidence intervals were computed with Bayesian state-space estimation (Smith et al., 2010). To minimize any error due to unstable recordings, the spike rate analysis excluded units that were not confidently detected throughout the entire baseline period (8.1%). The computed spike rate effects were similar when these units were included. Periods of silence were compared to a simulated Poisson distribution of equal rate over each 10 second period, and significance was assessed for each patient with a chi-square test relative to that distribution.

### **2.5.3 Spectral analysis**

Spectrograms were calculated with multitaper methods using the Chronux toolbox. Spectrograms were calculated with sliding windows of 5 seconds and 3 tapers computed every 1 second, for a time-bandwidth product of 2 and spectral resolution of 0.4 Hz. Bandpower measures were computed with smoother taper settings to collapse across frequencies (10 seconds every 5 seconds, 39 tapers, time-bandwidth product of 20 and spectral resolution of 2 Hz), and then median-filtered across 5 time bins to remove brief transient effects due to low-frequency artifacts. Triggered spectrograms were calculated with a window of 0.5 seconds every 0.125 seconds, using 3 tapers for a time-bandwidth product of 2 and spectral resolution of 4 Hz. Pre- and post-LOC spectra were computed across 90 seconds with 9 tapers for a resolution of 0.056 Hz, and error bars (marked by shaded region) were taken from the Chronux theoretical computation method at  $p=0.05$ .

Power changes post-LOC were computed as the percent change in the 30-60 second post-LOC period relative to the 30-60 second pre-LOC period. Ranges in the text reflect ranges across patients. The slow oscillation was extracted by applying a symmetric finite impulse response bandpass filter of length 4464 with unit gain from 0.1 to 1 Hz and attenuation of more than -50 dB from 0-0.85 and 1.15-125 Hz. Due to hardware filter settings with a highpass at 0.3 Hz, the power contribution below 0.3 Hz was minimized. Phase was extracted with a Hilbert transform. Statistical testing of triggered spectrograms was done by taking a ratio of each chi-square distribution, and

significance was calculated as an F-test with a Bonferroni correction for multiple comparisons across frequencies. For comparing spectra during and before an ON period, power spectra from 250 ms after ON period onset were compared to spectra from 250 ms before ON period onset. Averaged LFP waveforms were compared by pre-selecting a time period and performing a t-test on the mean amplitude values within that interval. When comparing the waveform height before and after spiking, a t-test was performed comparing the mean amplitude in the [-750 -500] to the [500 750] millisecond time windows locked to ON period onset or slow oscillation minimum.

#### 2.5.4 Phase modulation

Significance for single unit phase-coupling was computed with a chi-square test on the binned phase distribution. The analysis was performed a second time on cells with spike rates only above 0.1 Hz, ensuring that there were at least 5 expected spikes per phase bin. Strength of phase modulation was computed as a modulation index (Tort et al., 2010) (MI), adapted to quantify the Kullback-Liebler divergence of the phase histogram from the uniform distribution, measured in bits. Spike phase was split into a phase histogram ( $p$ ) of 10 bins, and MI was computed as  $\sum_{i=1}^{10} p_i \log_2 p_i + \log_2 10$ . We also computed the chi-square statistic as an alternative measure, yielding similar results. MI significance for each ECoG channel was calculated by shuffling the entire spike train randomly between 2 and 10 seconds, and calculating a shuffled MI over 2000 random shifts. The empirical MI was then compared to the shuffled MI with a significance level of 0.05 and a Bonferroni correction for multiple comparisons across channels. For LFP phase analysis, each single unit was compared to its local LFP channel. The time-varying phase modulation was computed with a window of 20 seconds sliding every 5 seconds. To assess the phase of maximal spiking relative to the ECoG slow oscillations, the phase of spiking was divided into 20 bins and then the mode of the phase histogram was reported.

### 2.5.5 Timing of spike rate and spectral power changes relative to LOC

We tested spike rates and spectral power to determine the first time bin in which these features differed significantly from the baseline period prior to propofol administration. We compared every time point starting 30 seconds before LOC to a baseline of spike rates or spectral features from the 3-minute baseline period immediately preceding it. To assess spike rate significance, we used a Bayesian hierarchical analysis, in which each post-baseline time point was compared to samples drawn from the Gaussian distribution of the baseline period and tested for a significant difference. This baseline sampling distribution was computed with the same state-space algorithm used to calculate spike rates (Smith et al., 2010). For determining the time at which power at a given frequency differed significantly, we used an analogous method but replaced the Gaussian sampling distribution with a chi-square distribution, which is the appropriate distribution for power measures. The time bin in Table 2.1 lists the earliest point at which the value had a 95% probability of being higher than the baseline period. We could not construct a similar sampling test for the MI because its distribution is not known, so we did not resample the baseline and instead reported the time at which MI became higher than the mean of the baseline period plus two standard deviations. For all of these measures, 5-second non-overlapping bins were used to identify the time at which changes occurred relative to LOC, which is the [-2.5 2.5] second period.

### 2.5.6 Phase-locking factor

The phase-locking factor (PLF) was computed in order to obtain a time-varying measure of phase offsets between slow oscillations. The phase of the slow oscillation was extracted as described in the ‘Spectral analysis’ section. For each time point, we then computed a  $z(t) = \exp(-i * (\varphi_A(t) - \varphi_B(t)))$  where  $\varphi_A(t)$  is the phase of one ECoG slow oscillation each time point and  $\varphi_B(t)$  is the phase of another ECoG slow oscillation. The PLF was then calculated as the mean of  $z(t)$  across the pre-LOC periods and across the post-LOC period. To assess the variability of phase offsets, we calculated the magnitude of the PLF. The distribution of PLF magnitude was assessed by plotting the mean and standard deviation of the PLF magnitude across each pair of ECoG channels separated by a given distance (distance between channels computed geometrically across the grid). To

determine the mean value of the phase offset across time we calculated the angle of the PLF. We then plotted the distribution of mean phase offsets across all pairs of channels separated by a given distance, by taking a 2-D histogram of PLF angle values for all electrode pairs.

### **2.5.7 Generalized linear model fitting**

A generalized linear model (GLM) was fit to ensemble spiking using custom software that performed regression with Truncated Regularized Iteratively Re-weighted Least Squares (Komarek, 2005; Truccolo et al., 2005) (TR-IRLS) and using the Bayesian Information Criterion to select the best model. Using the Akaike's Information Criterion also yielded a significant contribution of spike history. The GLM was constructed to predict ensemble spiking, which was defined as a series of 12 ms bins that contained a 1 if any spikes from any units occurred in that period, and a 0 otherwise. 10 covariates were used to represent the range of possible LFP phase values. Amplitude was normalized to range between zero and one. Since individual unit spike rates are low, the history dependence terms in this model predominately reflect interactions between units. The presented version is with 12 ms bins of spike history; similar results were obtained when using 4 or 8 ms bins. We excluded the minute surrounding LOC to ensure that any correlation between pre-LOC and post-LOC analyses was not due to bias from adjacent recordings during the LOC transition.

### **2.5.8 Single unit correlations**

Single units with high post-LOC spike rates were selected for correlation analysis in order to ensure sufficient spikes to assess the significance of their correlations. The minute surrounding LOC was excluded to reduce bias that could result from comparing adjacent recordings. Correlations between single units were computed relative to a shuffled baseline, in order to examine fine time-scale synchronization beyond the changes in population spike rate induced by the slow oscillation. Spike times were randomly shuffled 200 times, between 50 and 500 milliseconds, to obtain a baseline of correlated spike rate without millisecond-level timing information. Paired correlations were then tested for significance between -100 to 100 milliseconds, with  $p < 0.05$  using a

Bonferroni correction for multiple comparisons across lags. Correlations were judged significant if they had a  $p < 0.05$  departure from the Poisson distribution of spike occurrence predicted by the shuffled baseline. The relationship between pairs of single units was visualized with the square-root of the estimate of the cross-intensity function (Brillinger, 1976). Fisher's exact test was performed in R statistical software.

### **2.5.9 Detecting initiation of ON periods**

ON periods were detected by binning spikes from all units in 50 ms time bins and then setting a threshold to detect local peaks in the spike rate. The threshold was determined manually for each patient after visually checking to ensure adequate detection, as the number of units and thus expected population spike rates differed in each patient. After detection, the first spike within 300 ms of ON period detection was taken as the initiation time, and spike histograms verified that these times represented initiation of spiking. These ON period initiation times were then used for subsequent analysis of slow oscillation spectra and waveform morphology.

## **2.6 Author Contributions**

Patrick Purdon, Emery Brown, and Sydney Cash designed the experiment; Leigh Hochberg developed the recording system; Veronica Weiner collected the data; Emad Eskandar, William Anderson, and Joseph Madsen performed the surgeries; Eran Mukamel and Jacob Donoghue contributed to the analysis; Laura Lewis designed and performed the data analysis and wrote the manuscript with Patrick Purdon. All authors commented on the manuscript.

## **2.7 References**

- Achermann, P., and Borbely, A. (1997). Low-frequency (<1 Hz) oscillations in the human sleep encephalogram. *Neuroscience* 81, 213–222.
- Akrawi, W.P., Drummond, J.C., Kalkman, C.J., and Patel, P.M. (1996). A comparison of the electrophysiologic characteristics of EEG burst-suppression as produced by isoflurane, thiopental, etomidate, and propofol. *Journal of Neurosurgical Anesthesiology* 8, 40–46.

- Alkire, M.T., Haier, R.J., Barker, S.J., Shah, N.K., Wu, J.C., and Kao, Y.J. (1995). Cerebral metabolism during propofol anesthesia in humans studied with positron emission tomography. *Anesthesiology* 82, 393–403.
- Alkire, M., Hudetz, A., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876.
- Amzica, F., and Steriade, M. (1995). Short-and long-range neuronal synchronization of the slow (<1 Hz) cortical oscillation. *Journal of Neurophysiology* 73, 20.
- Bai, D., Pennefather, P.S., MacDonald, J.F., and Orser, B.A. (1999). The general anesthetic propofol slows deactivation and desensitization of GABA(A) receptors. *Journal of Neuroscience* 19, 10635–10646.
- Bennett, C., Voss, L., Barnard, J., and Sleight, J. (2009). Practical use of the raw electroencephalogram waveform during general anesthesia: the art and science. *Anesthesia & Analgesia* 109, 539.
- Bersagliere, A., and Achermann, P. (2010). Slow oscillations in human non-rapid eye movement sleep electroencephalogram: effects of increased sleep pressure. *Journal of Sleep Research* 19, 228–237.
- Blethyn, K.L., Hughes, S.W., Tóth, T.I., Cope, D.W., and Crunelli, V. (2006). Neuronal basis of the slow (<1 Hz) oscillation in neurons of the nucleus reticularis thalami in vitro. *Journal of Neuroscience* 26, 2474–2486.
- Blumenfeld, H., Rivera, M., McNally, K., and Davis, K. (2004). Ictal neocortical slowing in temporal lobe epilepsy. *Neurology*.
- Breshears, J.D., Roland, J.L., Sharma, M., Gaona, C.M., Freudenburg, Z.V., Tempelhoff, R., Avidan, M.S., and Leuthardt, E.C. (2010). Stable and dynamic cortical electrophysiology of induction and emergence with propofol anesthesia. *Proceedings of the National Academy of Sciences* 107, 21170–21175.
- Brillinger, D. (1976). Estimation of the second-order intensities of a bivariate stationary point process. *Journal of the Royal Statistical Society. Series B (Methodological)* 38, 60–66.
- Brown, E.N., Lydic, R., and Schiff, N.D. (2010). General anesthesia, sleep, and coma. *New England Journal of Medicine* 363, 2638–2650.
- Brown, E.N., Purdon, P.L., and Van Dort, C.J. (2011). General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual Review of Neuroscience* 34, 601–628.
- Cash, S.S., Halgren, E., Dehghani, N., Rossetti, A.O., Thesen, T., Wang, C., Devinsky, O., Kuzniecky, R., Doyle, W., Madsen, J.R., et al. (2009). The human K-complex represents an isolated cortical down-state. *Science* 324, 1084–1087.

- Chauvette, S., Crochet, S., Volgushev, M., and Timofeev, I. (2011). Properties of slow oscillation during slow-wave sleep and anesthesia in cats. *Journal of Neuroscience* *31*, 14998–15008.
- Ching, S., Cimenser, A., Purdon, P.L., Brown, E.N., and Kopell, N.J. (2010). Thalamocortical model for a propofol-induced alpha-rhythm associated with loss of consciousness. *Proceedings of the National Academy of Sciences* *107*, 22665–22670.
- Ching, S., Purdon, P.L., Vijayan, S., Kopell, N.J., and Brown, E.N. (2012). A neurophysiological-metabolic model for burst suppression. *Proceedings of the National Academy of Sciences* *109*, 3095–3100.
- Cimenser, A., Purdon, P.L., Pierce, E.T., Walsh, J.L., Salazar-Gomez, A.F., Harrell, P.G., Tavares-Stoeckel, C., Habeeb, K., and Brown, E.N. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proceedings of the National Academy of Sciences* *108*, 8832–8837.
- Contreras, D., and Steriade, M. (1995). Cellular basis of EEG slow rhythms: a study of dynamic corticothalamic relationships. *Journal of Neuroscience* *15*, 604–622.
- Crunelli, V., and Hughes, S.W. (2009). The slow (<1 Hz) rhythm of non-REM sleep: a dialogue between three cardinal oscillators. *Nature Neuroscience* *13*, 9–17.
- Csercsa, R., Dombovari, B., Fabo, D., Wittner, L., Eross, L., Entz, L., Solyom, A., Rasonyi, G., Szucs, A., Kelemen, A., et al. (2010). Laminar analysis of slow wave activity in humans. *Brain* *133*, 2814–2829.
- Destexhe, A., Hughes, S.W., Rudolph, M., and Crunelli, V. (2007). Are corticothalamic “up” states fragments of wakefulness? *Trends in Neurosciences* *30*, 334–342.
- Feshchenko, V.A., Veselis, R.A., and Reinsel, R.A. (2004). Propofol-induced alpha rhythm. *Neuropsychobiology* *50*, 257–266.
- Fiset, P., Paus, T., Daloze, T., Plourde, G., Meuret, P., Bonhomme, V., Hajj-Ali, N., Backman, S., and Evans, A. (1999). Brain mechanisms of propofol-induced loss of consciousness in humans: a positron emission tomographic study. *Journal of Neuroscience* *19*, 5506.
- John, E., Prichep, L., Kox, W., Valdes-Sosa, P., Bosch-Bayard, J., Aubert, E., Tom, M., DiMichele, F., and Gugino, L. (2001). Invariant reversible QEEG effects of anesthetics. *Consciousness and Cognition* *10*, 165–183.
- Komarek, P. (2005). Making logistic regression a core data mining tool. Robotics Institute.
- Kungys, G., Kim, J., Jinks, S.L., Atherley, R.J., and Antognini, J.F. (2009). Propofol produces immobility via action in the ventral horn of the spinal cord by a

- GABAergic mechanism. *Anesthesia & Analgesia* 108, 1531–1537.
- Lachaux, J.P., Rodriguez, E., Martinerie, J., and Varela, F.J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping* 8, 194–208.
- Lewicki, M.S. (1998). A review of methods for spike sorting: the detection and classification of neural action potentials. *Network* 9, R53–R78.
- Mashour, G.A. (2004). Consciousness unbound: toward a paradigm of general anesthesia. *Anesthesiology* 100, 428–433.
- Murphy, M., Bruno, M., Riedner, B., Boveroux, P., Noirhomme, Q., Landsness, E., Brichant, J., Phillips, C., Massimini, M., and Laureys, S. (2011). Propofol anesthesia and sleep: a high-density EEG study. *Sleep* 34, 283.
- Nir, Y., Staba, R.J., Andrillon, T., Vyazovskiy, V.V., Cirelli, C., Fried, I., and Tononi, G. (2011). Regional slow waves and spindles in human sleep. *Neuron* 70, 153–169.
- Ogilvie, R. (2001). The process of falling asleep. *Sleep Medicine Reviews* 5, 247–270.
- Rudolph, U., and Antkowiak, B. (2004). Molecular and neuronal substrates for general anaesthetics. *Nature Reviews Neuroscience* 5, 709–720.
- Sanchez-Vives, M.V., and McCormick, D.A. (2000). Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nature Neuroscience* 3, 1027–1034.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* 6, 461–464.
- Smith, A.C., Scalou, J.D., Wirth, S., Yanike, M., Suzuki, W.A., and Brown, E.N. (2010). State-space algorithms for estimating spike rate functions. *Computational Intelligence and Neuroscience* 426539.
- Steriade, M., Contreras, D., Curro Dossi, R., and Nunez, A. (1993a). The slow (<1 Hz) oscillation in reticular thalamic and thalamocortical neurons: scenario of sleep rhythm generation in interacting thalamic and neocortical networks. *Journal of Neuroscience* 13, 3284.
- Steriade, M., Nunez, A., and Amzica, F. (1993b). A novel slow (<1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *Journal of Neuroscience* 13, 3252–3265.
- Supp, G.G., Siegel, M., Hipp, J.F., and Engel, A.K. (2011). Cortical hypersynchrony predicts breakdown of sensory processing during loss of consciousness. *Current Biology* 21, 1988–1993.
- Timofeev, I., Grenier, F., and Steriade, M. (2001). Disfacilitation and active inhibition in the neocortex during the natural sleep-wake cycle: an intracellular study.

Proceedings of the National Academy of Sciences 98, 1924–1929.

- Timofeev, I., Grenier, F., Bazhenov, M., Sejnowski, T.J., and Steriade, M. (2000). Origin of slow cortical oscillations in deafferented cortical slabs. *Cerebral Cortex* 10, 1185–1199.
- Tort, A.B.L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of Neurophysiology* 104, 1195–1210.
- Truccolo, W., Donoghue, J.A., Hochberg, L.R., Eskandar, E.N., Madsen, J.R., Anderson, W.S., Brown, E.N., Halgren, E., and Cash, S.S. (2011). Single-neuron dynamics in human focal epilepsy. *Nature Neuroscience* 14, 635–641.
- Truccolo, W., Eden, U.T., Fellows, M.R., Donoghue, J.P., and Brown, E.N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology* 93, 1074–1089.
- Vincent, J.L., Patel, G.H., Fox, M.D., Snyder, A.Z., Baker, J.T., Van Essen, D.C., Zempel, J.M., Snyder, L.H., Corbetta, M., and Raichle, M.E. (2007). Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447, 83–86.
- Vyazovskiy, V.V., Olcese, U., Hanlon, E.C., Nir, Y., Cirelli, C., and Tononi, G. (2011). Local sleep in awake rats. *Nature* 472, 443–447.
- Vyazovskiy, V.V., Olcese, U., Lazimy, Y.M., Faraguna, U., Esser, S.K., Williams, J.C., Cirelli, C., and Tononi, G. (2009). Cortical firing and sleep homeostasis. *Neuron* 63, 865–878.
- Young, G. (2000). The EEG in coma. *Journal of Clinical Neurophysiology* 17, 473–485.

## Chapter 3: The thalamocortical circuit mechanisms controlling arousal state<sup>2</sup>

### 3.1 Abstract

During sleep and states of decreased arousal, cortical neurons exhibit rhythmic activity in the delta (0.5-4 Hz) range. The circuit mechanism that underlies this dynamic and its causal link to decreased arousal is not known. Here, using a combination of optogenetics and *in vivo* electrophysiology, we show that tonic activation of thalamic reticular nucleus (TRN) rapidly produces delta waves in the associated cortex. The induced oscillations resemble those seen in sleep, as cortical units undergo periods of silence that are phase-locked to the delta wave. Importantly, the induced delta waves are associated with a significant reduction in arousal state, decreasing the animals' motor activity and increasing non-REM sleep. The induced oscillations were limited to local cortical areas, demonstrating that delta generation is supported by corticotopic circuitry. These results show an essential role for TRN and thalamic inhibition in regulating local cortical state and arousal.

---

<sup>2</sup> The findings in this chapter were under review at the time of thesis submission. (Lewis, Voigts, Flores, Wilson, Halassa, and Brown, *submitted*)

## 3.2 Introduction

Modulation of arousal is one of the central aspects of behavior. Arousal varies both at long timescales (minutes to hours) such as in circadian sleep regulation, and at short timescales (seconds) during shifts in alertness and transitions in and out of sleep. Mechanisms for slow arousal modulation have been greatly explored, including hypothalamic regulation of sleep homeostasis and the orexin system (Saper et al., 2005; Tsunematsu et al., 2011). However, the mechanism underlying fast transitions to decreased arousal states is not known.

A key marker of decreased arousal is the cortical delta (1-4 Hz) oscillation, which occurs both during non-REM sleep (Huber et al., 2000) and in awake animals during low vigilance states and sleep deprivation. Brief delta waves in local cortical regions correlate with behavioral deficits on sub-second timescales (Vyazovskiy et al., 2011). The negative deflection of the delta wave marks periods of suppression in cortical neurons (OFF periods) lasting hundreds of milliseconds (Ji and Wilson, 2006; Vyazovskiy et al., 2009), and these brief offline periods are a candidate mechanism for decreased arousal. Delta waves are thus correlated with both behavioral decreases in arousal and disruption of cortical activity, but it is not known whether there is a central circuit mechanism that both generates delta and regulates arousal. Furthermore, no mechanism has been identified that generates the spatially isolated delta waves that occur during drowsiness, known as ‘local sleep’ (Vyazovskiy et al., 2011).

A central modulator of corticothalamic feedback that could initiate these dynamics is the thalamic reticular nucleus (TRN), a subcortical structure that provides powerful inhibition to dorsal thalamic nuclei. The TRN is a thin sheath of GABAergic neurons that surrounds the thalamus and inhibits thalamic relay cells (Guillery and Harting, 2003; Pinault, 2004). TRN has been implicated in sensory processing (Deleuze and Huguenard, 2006; Hartings et al., 2003), attentional gating (Crick, 1984; McAlonan et al., 2008), and state modulation (McCormick and Bal, 1997; Steriade, 2000) - and is uniquely positioned to selectively and rapidly modulate cortical state. TRN has a causal role in initiating sleep spindles (Halassa et al., 2011), and molecular genetic manipulation of TRN conductances reduces EEG sleep rhythms (Cueni et al., 2008; Espinosa et al.,

2008), indicating a role for thalamocortical feedback in cortical sleep oscillations. However, direct manipulations of thalamic activity have yielded conflicting results: nonspecific activation of multiple thalamic nuclei (including TRN) increases time spent in sleep (Kim et al., 2012); selectively stimulating thalamus induces an aroused, desynchronized cortical state (Poulet et al., 2012); but lesioning thalamus does not disrupt the desynchronized state (Constantinople and Bruno, 2011). These findings suggest a complex involvement of thalamus in regulating behavioral arousal, which could be mediated through the TRN. In addition, many anesthetic drugs such as propofol act to enhance GABAergic synaptic transmission and thus potentiate the effects of TRN activity (Brown et al., 2011), suggesting that it could be a component of the mechanism by which they induce an unconscious state (Franks, 2008). However, the role of TRN in generating the low-frequency oscillatory dynamics characteristic of low arousal states is not known, and the behavioral significance of such cortical dynamics has not been causally tested.

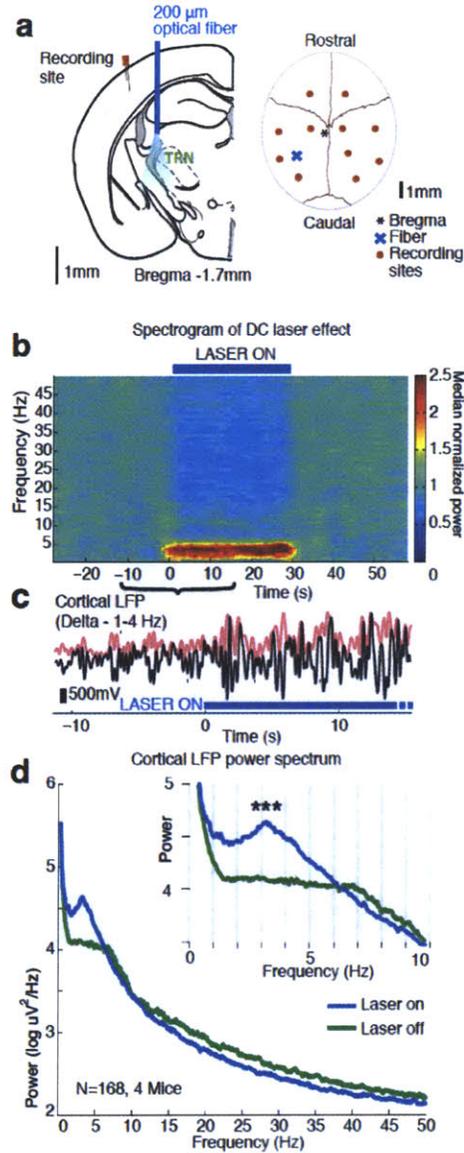
Here, we optogenetically activated TRN, and found that this manipulation rapidly induces local sleep-like thalamocortical delta waves and a state of decreased arousal. Tonic activation of TRN in awake animals produced delta waves in the associated cortical region, together with phase-locked periods of silence in cortical neurons (OFF periods). This manipulation also produced a state-dependent decrease in arousal state: awake animals exhibited less motor activity and spent more time in non-REM sleep, whereas anesthetized animals exhibited a suppression of cortical activity and a shift in dynamics favoring OFF periods. We find that the net effect of TRN stimulation is to decrease thalamic firing, suggesting that TRN may induce a lower arousal state by decreasing tonic thalamic input to cortex, facilitating the onset of delta waves. Furthermore, TRN and other thalamic neurons are phase-locked to the induced oscillations, suggesting that TRN, thalamus, and cortex are all engaged in the delta rhythm. We conclude that tonic depolarization of TRN rapidly modulates cortical state and controls the animals' arousal, by inducing a rhythmic suppression of thalamic activity. The spatial characteristics and rapid timescale (<100 ms) of these effects show that local oscillatory dynamics between thalamus and cortex are a central mechanism for modulation of arousal.

### 3.3 Results

To investigate the contribution of thalamic input to ongoing cortical dynamics, we used optogenetic targeting of TRN neurons. We used previously characterized transgenic mice (Zhao et al., 2011) in which channelrhodopsin2 (ChR2) expression was under the control of the vesicular GABA transporter (VGAT-ChR2). In these mice, the TRN exhibits preferential expression of ChR2 compared to surrounding subcortical regions (Halassa et al., 2011), allowing us to manipulate TRN activity using implanted optical fibers. We implanted a total of eight mice with fibers targeting TRN (Fig. 3.1a).

#### 3.3.1 Tonic activation of TRN produces cortical delta waves.

We first tested how tonic activation of TRN affected neural dynamics in the cortex of awake head-fixed mice. We implanted four mice with stereotrodes distributed across cortex and an optical fiber targeting the somatosensory sector of TRN (Fig. 3.1a). To examine the timecourse of how thalamic manipulations affect cortical dynamics, we activated TRN tonically for 30 second epochs. Tonic TRN activation produced a substantial increase in delta power in the local field potential (LFP) of ipsilateral somatosensory cortex (Fig. 3.1b-d). This power increase was specific to the delta (1-4 Hz) band, which increased by 2.56 dB (95% confidence interval (CI)=[2.13 2.97]) during laser stimulation. In contrast, beta and gamma (15-50 Hz) power decreased slightly (Fig. 3.1d, median=-1.03 dB, CI=[-1.24 -0.84]). The increase in delta power was rapid and robust: the appearance of delta waves was already evident in the first second of TRN activation (change=1.12 dB, CI=[0.48 1.76]) and the increase persisted throughout the stimulation period (Fig. 3.1b).



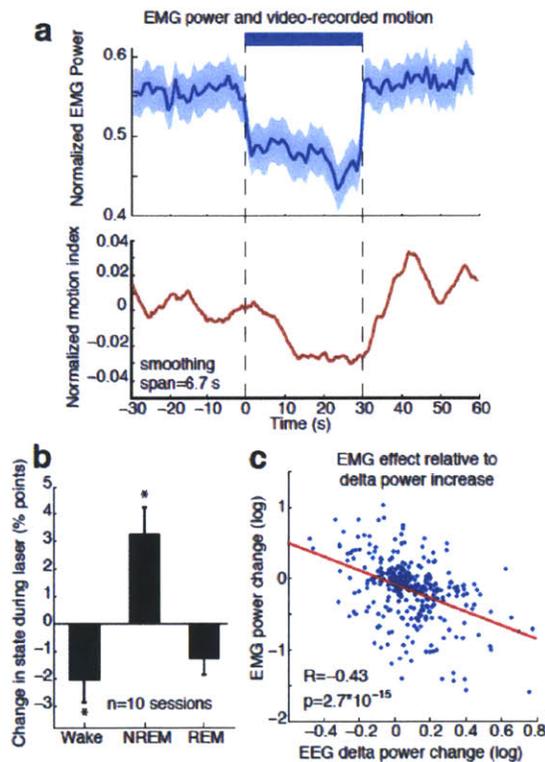
**Figure 3.1: Tonic optogenetic stimulation of thalamic reticular neurons produces delta waves in ipsilateral somatosensory cortex.** A) Diagram of surgery: fiber is implanted into left TRN, and stereotrodes are implanted in multiple sites across cortex. B) Spectrogram showing average effect in ipsilateral somatosensory cortex across 168 trials (4 mice, 8 sessions): TRN stimulation causes a rapid increase in delta power that persists throughout the stimulation period. Power is normalized to the 30 second pre-stimulus period. C) Example trace from a single trial, showing the LFP filtered between 1-4 Hz (black line), and the instantaneous delta amplitude (red line). D) Average spectrum of LFP in somatosensory cortex: during tonic optogenetic activation of TRN, this cortical site demonstrates an increase in delta (1-4 Hz) power and a decrease in beta and gamma (12-50 Hz) power. Gray region shows zoomed-in plot of delta power increase, asterisks indicate significance at  $\alpha=0.001$ .

### 3.3.2 Tonic TRN activation produces a behavioral state of decreased arousal

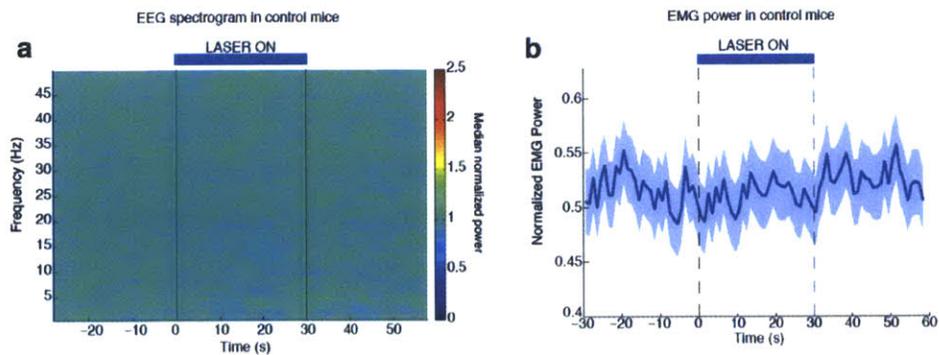
Delta waves are a major feature of non-REM sleep (Pace-Schott and Hobson, 2002), so we next investigated whether TRN activation produced behavioral signs of sleep. A key marker of sleep is the substantial decrease in motor activity, which can be detected through electromyographic (EMG) recordings. We recorded from two mice with neck-muscle EMG and frontal electroencephalograph (EEG) electrodes to test whether TRN activation caused sleep-like behavior. We used the same tonic stimulation in freely behaving mice, enabling the mouse to sleep or explore depending on its arousal state. We found that EMG power decreased significantly during TRN stimulation (Fig. 3.2a, mean=-1.08 dB, CI=[-1.36 -0.81]), indicating that stimulation caused the animals to become less active. The decrease in EMG power was significant within 1 second of laser onset, demonstrating rapid modulation of behavioral state. To further confirm this finding, we recorded videos of seven mice and used an automatic video scoring algorithm to quantify the amount of motion exhibited by the mice (see Experimental Procedures). In agreement with the EMG findings, we found that motion decreased significantly during TRN activation (Fig. 3.2a, bottom panel, decrease in 58.0% of trials, CI=[53.1 62.7]). To ensure that the decrease in motor activity was truly due to Chr2 activation and not merely the presence of light, we performed control experiments in littermates that did not carry the Chr2 gene and found no delta or EMG effect (Fig. 3.3). These results demonstrated that TRN activation causes a rapid decrease in arousal state, evident by a decline in motor activity.

We next investigated whether the behavioral effect was simply a decrease in motion during the awake state, or whether the mice were also sleeping more during TRN activation. We used a semi-automated sleep scoring algorithm to calculate the total time spent in sleep, non-REM, and REM states (see Experimental Procedures). We applied this method to the two mice with EEG and EMG recordings and found that tonic TRN activation caused mice to spend significantly less time awake (median=-2.0%, CI=[-3.6 -0.01]) and significantly more time in non-REM sleep (median=3.3%, CI=[1.0 5.7]) (Fig. 3.2b). TRN activation thus produced a shift in sleep dynamics such that mice were biased towards a state of non-REM sleep.

To examine the relationship between the induced delta waves and the decrease in motor activity, we computed the correlation between the decrease in EMG power and the increase in EEG delta power. We found that the EEG and EMG effects were significantly negatively correlated on the single trial level (Fig. 3.2c), with a correlation coefficient of -0.43 (CI= [-0.53 -0.33]). This correlation was significantly higher than the correlation at randomly shuffled times, ( $p < 0.05$ , bootstrap) demonstrating that the decrease in arousal was specifically associated with the optogenetically induced delta waves. Taken together, these results establish a causal link between cortical delta waves and decreased arousal state through TRN-induced thalamic inhibition.



**Figure 3.2: Tonic TRN activation produces a decrease in arousal state.** A) Top panel: Mean EMG power locked to laser onset shows that EMG power decreases significantly during TRN activation ( $n=315$  trials, 8 sessions, 2 mice). Bottom panel: Mean smoothed motion (6.67 second moving average) detected in video: animals' motion decreases significantly during optogenetic stimulation ( $n=421$  trials, 9 sessions, 7 mice). B) Mean arousal state during TRN activation: mice spend significantly more time in non-REM sleep and significantly less time in the awake state. Stars indicate significant effects at  $\alpha=0.05$ . C) Individual trial correlation shows that the decrease in EMG power is correlated with the TRN-induced increase in EEG delta power ( $n=315$  trials, 8 sessions, 2 mice).

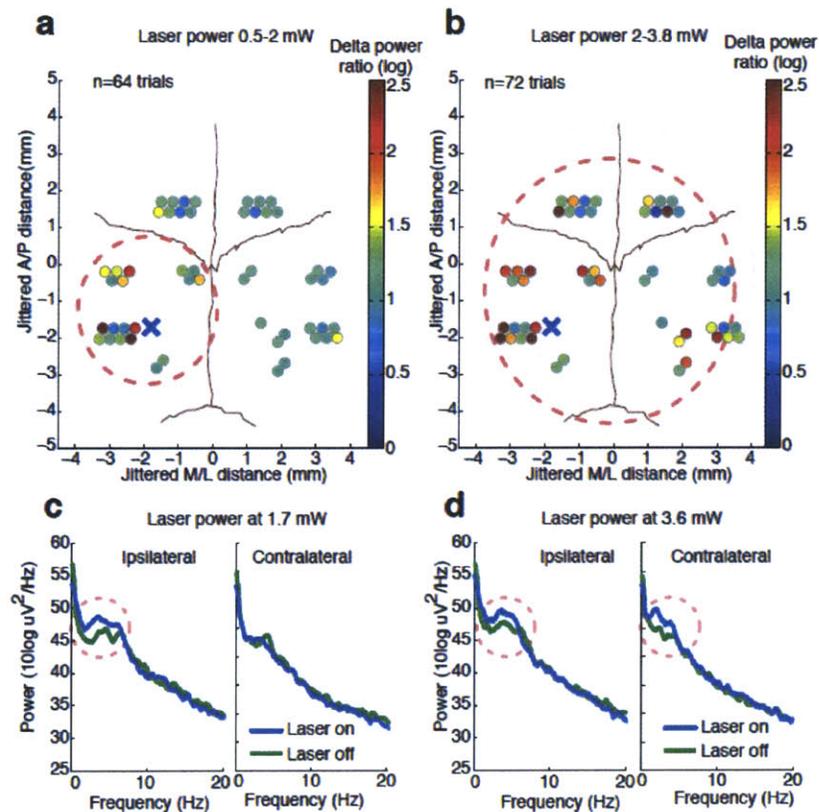


**Figure 3.3: Delta waves and decreased arousal depend on Chr2 expression.** A) Control mice that are negative for Chr2 do not exhibit delta waves during TRN stimulation (n=494 trials, 10 sessions, 10 mice). B) EMG power does not decrease during laser stimulation in control mice, confirming that the behavioral effect is not due to a nonspecific effect of light (n=494 trials, 10 sessions, 3 mice).

### 3.3.4 TRN activation selectively controls a local ipsilateral cortical region.

We demonstrated that tonic unilateral activation of TRN causes a decrease in arousal state, marked by both neurophysiological (cortical delta waves) and behavioral (decreased motor activity) effects. Recent findings have shown that cortical delta waves are observed locally in awake sleep-deprived animals, and that this ‘local sleep’ correlates with decreased performance on cognitive tasks (Vyazovskiy et al., 2011). Given that our experimental preparation allowed us to manipulate the degree of TRN activation and subsequent thalamic inhibition, we used it to ask whether this mechanism could support local delta generation. We leveraged our ability to record local field potentials across cortex in individual mice (Fig. 3.1a) to investigate the spatial spread of induced delta waves. We recorded in four awake head-fixed mice with fibers targeting the somatosensory sector of TRN, and found that low laser power consistently enhanced local delta power in ipsilateral S1 (Fig. 3.4a). Across all electrodes in the ipsilateral posterior quadrant (Fig. 3.4a, red circle), 9 out of 20 (45%) recording sites showed a significant increase in delta power during tonic activation ( $p < 0.05$ , signed-rank test with Bonferroni correction). In contrast, only 2 out of the 33 (6%) recording sites in other cortical regions (e.g. contralateral or frontal) showed a significant increase in delta power, demonstrating that delta waves were selectively induced in a local ipsilateral cortical region (Fig. 3.4a,c).

We next tested whether activating a larger region of TRN could produce delta waves across larger regions of cortex. We found that on trials using more than 2 mW of laser power (i.e. with light spreading to recruit a larger region of TRN (Bernstein et al., 2008)), delta waves were induced across a large cortical area. 10 out of 20 (50%) of electrodes in the associated cortex and 11 out of 33 (33%) electrodes in other regions showed a significant increase in delta power, including electrodes in frontal and contralateral regions (Fig. 3.4b,d). We therefore concluded that tonic activation of a small population of TRN neurons produces delta waves in a local ipsilateral cortical region, and that the extent of TRN activation controls the spatial spread of cortical delta waves. Local activation of TRN thus controls an aligned region of cortex, showing that local thalamocortical loops support the spatially restricted delta waves that occur in local sleep.



**Figure 3.4: Degree of TRN stimulation can induce either local or global cortical delta waves.** A and B) Circles represent electrodes, and their color indicates the size of the delta power increase when laser is on (total n=136 trials, 4 mice). At low powers (<2 mW), delta power increases only in electrodes near ipsilateral somatosensory cortex (red dashed circle). At high powers (>2 mW) that activate larger regions of TRN, delta power increases across multiple cortical areas, including frontal cortex and contralateral cortex

(red dashed circle). Distances are jittered so that all electrodes can be displayed. Blue 'X' indicates placement of laser fiber. C) Example spectra from one mouse at low laser power in electrodes ipsilateral and contralateral to the laser fiber (n=10 trials): delta effect is seen in ipsilateral cortex but not in contralateral cortex. D) Example spectra at high laser power (n=9 trials): delta waves are generated in both ipsilateral and contralateral cortex.

### 3.3.5 Cortical units phase-lock to induced delta waves and undergo OFF periods

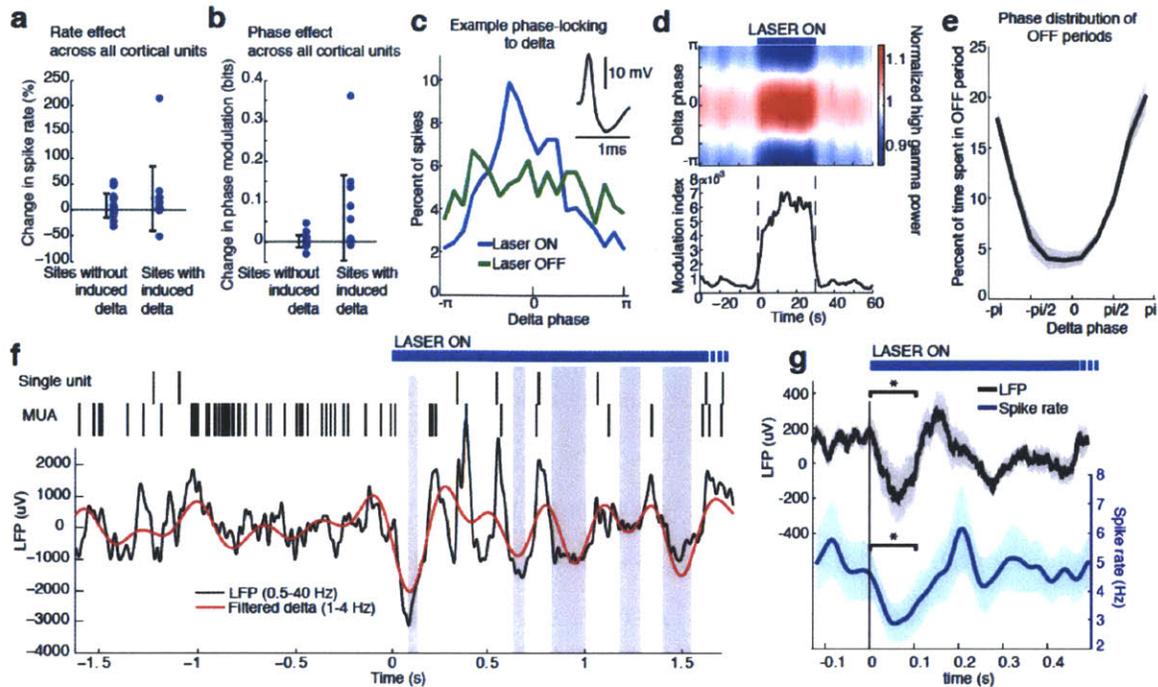
Delta waves are a common feature of sleep, where they mark an alternation of cortical spiking between activated (ON) and inactivated (OFF) states (Vyazovskiy et al., 2009). We therefore investigated whether the TRN-induced delta waves shared the sleep-like property of modulating cortical spiking. We identified 31 cortical units across the distributed recording sites and tested whether they were modulated by local delta wave activity. We found that while TRN stimulation did not significantly change firing rates in cortical units (Fig. 3.5a, median=-0.05 Hz, CI=[-0.17 0.09]), cortical units became strongly phase-locked to induced delta oscillations, similar to cortical activity during natural sleep (Vyazovskiy et al., 2009). We quantified delta phase-locking using the modulation index (MI), which measures the divergence between a unit's phase distribution and the uniform distribution. In electrodes with induced delta waves, cortical units became significantly more phase-locked to delta (Fig. 3.5b,c, MI change=0.044 bits, CI=[0.0012 0.112], n=13 units). In contrast, increased phase-locking was not observed in units from electrodes with no induced delta rhythm (Fig. 3.5b, MI change=0.0002 bits, CI=[-0.004 0.007], n=18 units). Across all units, the increase in spike phase-locking was strongly correlated with the increase in LFP delta power (R=0.77, CI=[0.57 0.88].)

To examine a larger-scale sample of neural activity, we also examined the phase modulation of high gamma (70-100 Hz) power, which is strongly correlated with multi-unit spiking activity (MUA) (Ray and Maunsell, 2011). In agreement with the single unit findings, we found that high gamma power rapidly became strongly phase-modulated by delta waves during TRN activation (Fig. 3.5d, MI change=0.0015, CI=[0.0009 0.0022]), indicating that neuronal activity was broadly locked to the induced delta waves.

The fact that cortical neurons become phase-locked to the induced delta waves suggested that they might be undergoing periods of silence (OFF periods). To explicitly test this, we selected electrodes that contained both MUA activity and a TRN-induced

delta wave and computed the instantaneous firing rate to automatically detect OFF periods (Fig. 3.5f). We found that during TRN activation, cortical neurons spent 13.1% of the time in OFF periods, significantly more than in the awake state (7.04%,  $p < 0.001$  in each mouse, signed-rank test) and significantly more than would be expected to occur randomly in simulated neurons with the same rate (3.09%, CI=[0.04 6.15]). In addition, we found that these OFF periods occurred predominantly during the negative deflection of the delta wave (Fig. 3.5e,  $p < 0.001$ ). The average duration of OFF periods was 122 ms (quartiles=[69 146] ms) and they occurred at a mean frequency of 0.998/second, consistent with natural sleep (Vyazovskiy et al., 2009). We therefore concluded that the induced delta waves are sleeplike, marking an oscillatory pattern in which cortical neurons undergo periods of silence lasting tens or hundreds of milliseconds.

To determine the timescale of the shift into sleeplike dynamics, we computed the mean LFP and spike rate locked to laser onset, across all cortical units with LFPs that displayed an induced delta wave. The LFP underwent a negative-going deflection in the first 100 ms of laser stimulation (Fig. 3.5g, top), and this wave was accompanied by a significant decrease in spike rate (Fig. 3.5g, bottom). These findings demonstrated that the effect of TRN activation is rapid, driving a delta wave and cortical suppression in less than a hundred milliseconds, and thereby inducing an abrupt transition into a new cortical state in which neurons undergo rhythmic OFF periods.



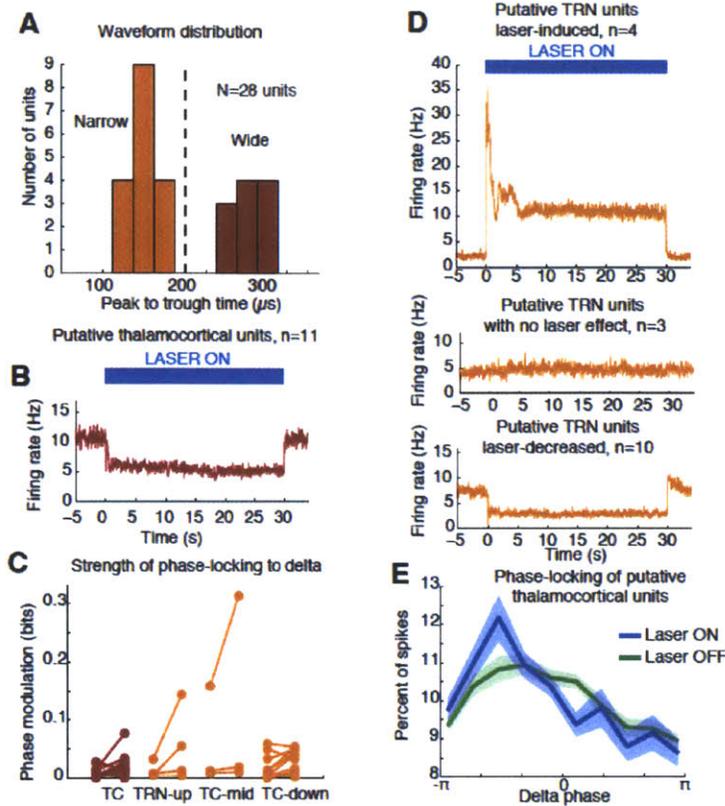
**Figure 3.5: Cortical units undergo OFF periods that are phase-locked to the delta waves during TRN activation.** A) Rate effect across all cortical units, categorized by strength of delta power increase in that channel. On average, there is no significant change in spike rate for either group. Error bars show st. dev. B) Phase-locking effects across all cortical units show that units on channels with induced delta waves become phase-locked to delta during TRN stimulation. Error bars show st. dev. C) Phase distribution of spikes from an example cortical unit recorded on a channel with a 3.4 dB delta power increase during TRN activation: unit becomes phase-locked to the delta wave. D) Phase distribution of high gamma (70-100 Hz) power relative to delta (1-4 Hz) oscillation shows that high gamma power becomes rapidly phase-locked to delta waves during TRN stimulation. E) Phase distribution of all OFF periods shows that they occur during the trough of the delta waves. F) Example trace from somatosensory cortex: optogenetic TRN stimulation rapidly induces delta waves that are associated with OFF periods in cortical activity (gray shaded regions mark automatically detected OFF periods). G) Mean spike rate and LFP locked to laser onset: the induced delta trough and phase-locked cortical inhibition are observed within 100 ms of laser onset. Stars indicate significant ( $p < 0.05$ ) decrease in LFP voltage and mean spike rate in the first 100 ms. Triggered LFP and units are averaged across cortical electrodes with a delta power increase ( $n = 14$  channels), shaded region is std. err.

### **3.3.6 Heterogeneous ensemble effects within thalamic reticular nucleus reduce thalamic firing rates**

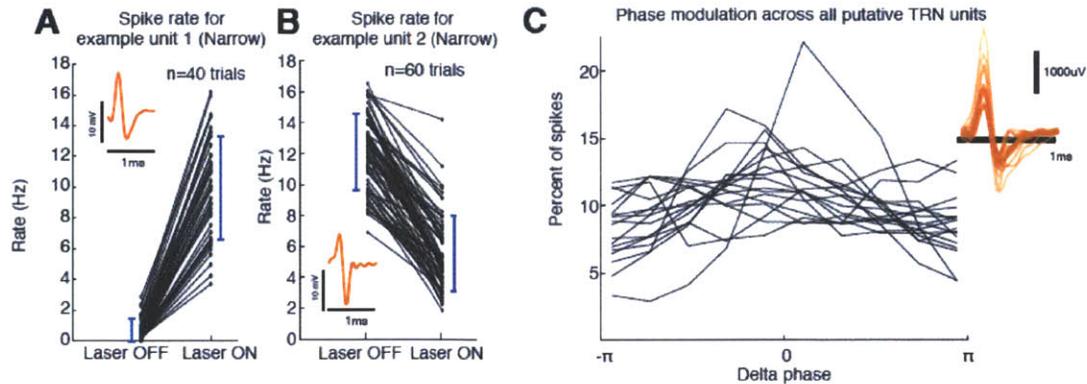
Tonic TRN stimulation produced striking cortical and behavioral effects that were locally defined, suggesting that the key circuit mechanism was through thalamus, which is the main target of TRN outputs and has corticotopic projections that could support local control of cortex. We therefore analyzed the firing rates of putative TRN and thalamocortical (TC) cells during tonic laser stimulation. We isolated 28 units from four mice and used spike waveforms to distinguish between TRN and TC neurons. The waveform distribution was bimodal (Fig. 3.6a), with ‘Narrow’ units (peak-to-trough time under 200 ms), and ‘Wide’ units (peak-to-trough time above 200 ms). Narrow waveforms are typically characteristic of TRN GABA-ergic fast-spiking inhibitory neurons (Houser et al., 1980; Lam and Sherman, 2011), so we used these units (n=17) to infer the activity of TRN, and the ‘Wide’ units to infer TC neuron activity (Wang et al., 2010). We found that nearly all Wide units (9/11, 81.8%) showed a significant decrease in spike rate during laser stimulation, whereas no units significantly increased their spike rate (Fig. 3.6b). Thalamic activity was therefore consistently suppressed by TRN activation.

In contrast to the TC neurons, optogenetic stimulation had heterogeneous effects on the firing rates of Narrow (putative TRN) units. Although some units increased their spike rates as predicted (Fig. 3.6d, 3.7a, 4/17 units, 23.5%), we also found 3 units with no significant change (17.7%), and an even larger number of units that decreased their firing rate significantly (10/17 units, 58.8%, Fig. 3.6d, 3.7b). These findings are consistent with previous reports of a heterogeneous population of neurons within the TRN (Contreras et al., 1992; Lee et al., 2007), and show that optogenetic activation shifts the internal circuit dynamics to preferentially enhance firing in only a subpopulation of the neurons. These results suggest that optogenetic stimulation of TRN shifts the balance of activity within the local population of TRN neurons to a regime that is heterogeneous within TRN, but nevertheless strongly activates a subset of neurons and causes consistent inhibition of thalamic activity. In this scenario, inhibition of thalamus would depend upon a nonlinearity that enables increased firing rates in only a fraction of TRN neurons to decrease thalamic firing.

Whether thalamic inhibition alone can generate sleep states has been debated: although thalamic activation induces wake states (Poulet et al., 2012), lesioning thalamus does not produce sleep states (Constantinople and Bruno, 2011). We therefore investigated whether TRN and TC neurons were engaged in the induced delta rhythms. We indeed observed that subcortical units increased their phase-locking to the LFP delta (Fig. 3.6c). Phase-locking in putative TRN units was diverse: laser stimulation increased overall phase-locking to the delta oscillation (Fig. 3.6c, increase=0.026 bits, CI=[0.005 0.041]), but with substantial variation across units (Fig. 3.6c, Fig. 3.7c), again possibly due to a heterogeneous population of neurons within TRN with varied functional roles, or to differential spread of light across neurons. In contrast, putative TC neurons consistently increased their phase-locking during stimulation (Fig. 3.6c, increase=0.011 bits, CI=[0.001 0.021]) with most spikes occurring around  $\pi/2$  (Fig. 3.6e). Due to the geometry of TRN, it was not possible to determine whether the delta oscillation in subcortical LFPs originated from TRN or other thalamic nuclei. We therefore concluded that both TRN and TC neurons participate in the delta oscillation, but future studies, likely with intracellular recordings, will be needed to determine the precise dynamics within TRN relative to thalamus. Nonetheless, our results suggest that thalamic neurons are actively engaged in delta rather than undergoing a simple decrease in activity.



**Figure 3.6: Heterogeneous effects in TRN neurons produce suppression of thalamic neurons.** A) Histogram of waveform parameters show a bimodal distribution of peak-to-trough time across subcortical units. Units with peak-to-trough times under  $200 \mu\text{s}$  were categorized as Narrow (putative TRN), and units over  $200 \mu\text{s}$  were categorized as Wide (putative thalamocortical). B) Putative thalamocortical (Wide) units consistently decrease their firing rates during laser stimulation. C) Phase-locking effects across all subcortical units show that most become phase-locked to the delta oscillation during TRN stimulation. Different unit types do not show consistent differences in phase-locking properties. D) Heterogeneous firing rates in TRN during stimulation: 4 units increase their firing rates whereas 10 units decrease their firing rates. The modulation in firing rate is strongly time-locked to laser onset and offset. E) Putative thalamocortical neurons consistently increase their phase-locking during TRN stimulation, peaking around  $-\pi/2$ .

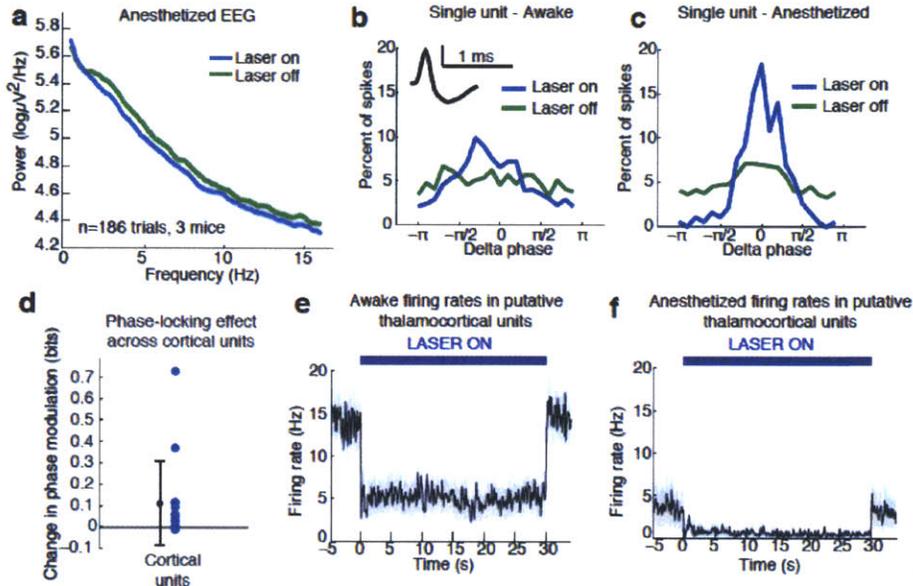


**Figure 3.7: Heterogeneous effects of optical stimulation across putative TRN units, but consistent effects within units.** A) Example Narrow unit with a consistent large firing rate increase during optical stimulation. B) Example Narrow unit with a consistent large firing rate decrease during optical stimulation. C) Phase-locking across all Narrow units to the delta rhythm during optical stimulation. Both the magnitude and preferred phase is variable across units.

### 3.3.3 TRN does not induce delta in anesthetized mice.

Our results showed that TRN activation decreases arousal state in awake mice. We next examined whether TRN can exert similar effects in anesthetized mice; in other words, does tonic TRN activation induce delta when the animal is already in a state of decreased arousal? To assess this, we analyzed EEG recordings from isoflurane anesthetized mice. We found that the baseline delta power in the EEG was high and there was no further increase in delta power during TRN activation (Fig. 3.8a), suggesting that the ability of TRN to generate delta waves was saturated. Instead the EEG showed a broadband (0.5-50 Hz) decrease in power (-0.53 dB, CI=[-0.69 -0.37]), demonstrating a generalized quieting of cortical activity. If the TRN activation were indeed causing mice to enter a more deeply anesthetized state, it would be expected to cause a decrease in ON periods and an increase in OFF periods. In support of this hypothesis, we found that TRN activation caused cortical units to increase their phase-locking to delta (Fig. 3.8b-d, median change=0.06 bits, CI=[0.018 0.189]), while their firing rates decreased (median change = -0.09 Hz (-5.5% of baseline), CI=[-0.22 -0.02]). In addition to this increased rhythmicity in cortical neurons, the fraction of time spent in OFF periods increased by 4.02% (CI=[1.9 6.2%]). We next analyzed putative thalamocortical neurons, and found that firing rates were suppressed to even lower levels by TRN stimulation during

anesthesia (Fig. 3.8e,f). We concluded that the anesthetized cortex is shifted into an even deeper state by TRN activation: not by inducing delta, but rather by modulating the dynamics of a delta wave that is already present and thereby prolonging the duration of the periodic suppressions.



**Figure 3.8: TRN stimulation deepens neuronal suppression during isoflurane anesthesia.** A) During isoflurane anesthesia, the delta waves appear to be saturated and are not increased by TRN stimulation. Instead, broadband power decreases, suggesting a shift in dynamics that favours the inactivated state. B) Example effect of TRN activation in an awake mouse: cortical unit becomes phase-locked to induced delta waves. C) Example effect in anesthetized mouse: cortical unit is already phase-locked to delta waves during isoflurane anesthesia, and TRN activation causes the phase-locking to become sharper. D) Phase-locking effects across all cortical units during isoflurane anesthesia: units become significantly more modulated by delta waves when TRN is activated. Error bars are st. dev. E) Putative thalamic units in the awake mouse are inhibited during TRN stimulation. F) Putative thalamic units in the anesthetized have a baseline firing rate equivalent to the awake, TRN-stimulated mouse. TRN stimulation induces an even larger suppression of thalamic activity.

### 3.4 Discussion

In rodents, states of decreased arousal such as sleep are marked by a cortical delta (1-4 Hz) oscillation, but both the circuit mechanisms that induce these states and their causal link to arousal are unknown. In this study, we identified a local thalamocortical

circuit that generates both delta rhythms and reduced arousal. Specifically, we found that tonic TRN activation mediates an increase in thalamic inhibition and produces sleep-like cortical delta waves whose spatial spread depends on the extent of TRN activation. This electrophysiological effect is strongly correlated with an optogenetically-induced reduction in behavioral arousal. This is the first study to show rapid reduction of arousal state (less than a second) by a forebrain circuit mechanism.

### **3.4.1 Delta waves are generated by local corticotopic circuits**

We find that TRN can selectively induce delta oscillations in local cortical regions. This result reinforces recent findings suggesting that sleep contains dynamics that are differentiated across cortex rather than a globally homogeneous cortical state (Krueger et al., 2008). Awake sleep-deprived rats also exhibit delta waves and OFF periods in local cortical areas (Vyazovskiy et al., 2011). Our results show that localized depolarization in TRN can produce such local oscillations, and could therefore underlie the fragmented cortical delta waves observed during sleep deprivation. In addition, local cortical OFF states have been observed during sleep (Nir et al., 2011) and general anesthesia (Lewis et al., 2012) in human subjects, demonstrating that OFF periods frequently occur locally even when slow-wave activity is present throughout cortex. The observed asynchronous slow waves in these unconscious states could be due to a global activation of TRN, producing delta waves throughout cortex, but different cortical regions are associated with specific thalamocortical circuits that enable them to undergo separate and asynchronous oscillations. Finally, the local control that TRN exerts over cortex provides evidence for how TRN could modulate attention across sensory modalities, by suppressing arousal in specific cortical regions. This finding supports the theory that TRN could function to modulate attention by gating thalamic transmission of sensory information to cortex (Crick, 1984). The finding that TRN can independently control limited corticothalamic circuits therefore suggests it could serve as a central circuit mechanism to regulate specific cortical regions, modulating both attention and arousal.

### **3.4.2 TRN-induced thalamocortical delta oscillation shows immediate onset**

In natural behavior, animals can rapidly transition between arousal states. We find that the delta waves induced by depolarization of TRN are initiated abruptly, suggesting we directly engaged circuit dynamics rather than slower neuromodulatory effects. Cortical activity is suppressed within 100 ms of laser onset, and the delta power increase in the LFP can be detected within a second. This pattern suggests that tens of milliseconds of TRN activation are sufficient to inhibit thalamic input to cortex and produce an OFF state. The dynamics at laser offset are similarly abrupt, with delta waves vanishing with a second. TRN can therefore serve as a rapid modulator of arousal state. This finding is compatible with established neuromodulatory sleep circuits that act at longer timescales (Pace-Schott and Hobson, 2002), such monoaminergic arousal pathways (Saper et al., 2005), as these neuromodulators affect TRN activity as one component of arousal regulation. TRN thus engages a fast-acting circuit for arousal control, demonstrating that thalamocortical loops can rapidly control cortical arousal state.

### **3.4.3 TRN supports both delta and spindle oscillations**

Here we used tonic activation of TRN, possibly mimicking naturally occurring physiological states in which neuromodulators such as acetylcholine modulate TRN activity. This tonic paradigm induced delta waves and modulated cortical state in awake mice without affecting power in the spindle band (7-15 Hz). Interestingly, strong phasic activation of TRN induces spindles during non-REM sleep but not in the awake state (Halassa et al., 2011). This contrast may be due to differences in thalamic firing mode. Thalamic neurons can alternate between burst firing and tonic firing depending on their level of hyperpolarization (Sherman, 2001). In the pulsed case (Halassa et al., 2011), high amplitude TRN depolarization hyperpolarizes thalamic relay neurons sufficiently that they fire a burst and initiate spindles upon release of the stimulation. The fact that spindles were limited to non-REM sleep suggests that they require thalamic neurons to already be hyperpolarized. In this study, tonic TRN activation at lower light levels produces a decrease in thalamic firing over the 30 second stimulation period. Phasic and tonic modulation of TRN activity therefore produce qualitatively different sleep

oscillations, suggesting that changes in the dynamics of inputs to TRN could underlie shifts between different stages of sleep.

#### **3.4.4 TRN cells respond heterogeneously to tonic optical stimulation**

Interestingly, we observed a heterogeneous response of TRN neurons to activation via ChR2 (Fig. 3.6): only a fraction of units (24%) exhibit significant tonic increase in rate, while 18% are unaffected and 59% are suppressed. However, the units that were activated underwent a strong increase in firing rate and thalamic neurons were consistently suppressed, suggesting that activation of a subset of TRN neurons is sufficient to inhibit thalamus. This finding is consistent with reports that TRN contains heterogeneous subclasses of neurons that display different firing patterns across sleep and wake states (Barrionuevo et al., 1981). These dynamics may be explained by the extensive chemical and electrical synapses present within TRN (Deleuze and Huguenard, 2006; Landisman et al., 2002; Long et al., 2004) that enable TRN neurons to reciprocally inhibit each other, and could serve to limit overall firing rates. In addition, multiple other factors could contribute to this heterogeneity: differences in ChR2 expression levels across cells, distance from the optical fiber, or activation of GABAergic projections to the TRN from other regions. Nevertheless, activity in thalamocortical neurons was consistently inhibited, suggesting that the circuit implements a nonlinearity such that strongly activating a subset of TRN neurons is sufficient to decrease thalamic firing. A possibility is that the local TRN interconnectivity implements a fast-acting homeostatic circuit which internally limits activity during tonic depolarization, so that even ChR2 activation strong enough to affect broad regions of TRN does not lead to excessive inhibitory output to thalamus, allowing it to oscillate at delta frequencies rather than being consistently suppressed.

#### **3.4.5 Circuit mechanisms underlying the induced thalamocortical delta waves**

The circuit mechanism that generates slow wave activity during sleep and anesthesia remains a topic of debate (Crunelli and Hughes, 2009; Destexhe and Contreras, 2011; McCormick and Bal, 1997). Our results show that delta waves can be produced by depolarizing TRN, and suggest they may be generated through overall inhibition of

thalamic input to cortex. TRN depolarization shifts activity to favor the TRN neurons that inhibit the associated region of thalamus. In the absence of this thalamic input to cortex, which drives the desynchronized cortical state (Poulet et al., 2012), cortex and thalamus jointly enter a delta oscillation in which activity is periodically suppressed. Previous studies have demonstrated that cortex can maintain an awake state even when thalamus is lesioned or inactivated (Constantinople and Bruno, 2011; Zagha et al., 2013); our results therefore suggest that delta waves in the intact brain require the involvement of cortical, TRN, and TC neurons in a coordinated rhythm. Furthermore, TRN stimulation did not induce delta in the anesthetized animal, when thalamic T-type calcium channels are blocked (Todorovic and Lingle, 1998). Our results therefore suggest that inhibition of thalamus is key to generating delta waves, but requires both an intact thalamocortical loop and TRN-mediated inhibition rather than complete optogenetic suppression. We suggest that reduction of thalamic input to cortex is responsible for delta waves, although we cannot exclude alternate routes of state modulation through other intermediary areas. However, any intermediate area would need to be corticotopically organized (as is thalamus) due to the spatially defined nature of the induced delta waves.

#### **3.4.6 TRN activation as a component of general anesthesia**

The finding that TRN activation induces delta oscillations and decreases arousal could explain a subset of the effects of GABAergic drugs used for general anesthesia, such as propofol. In human subjects, propofol induces a large increase in low-frequency (0.1-4 Hz) power (Murphy et al., 2011), and this slow wave induction has been suggested as a potential mechanism for unconsciousness (Lewis et al., 2012; Massimini et al., 2009). Propofol is a GABA-A agonist (O'Shea et al., 2000), suggesting that it could increase low-frequency power by increasing inhibition of thalamus (Alkire et al., 2000). Decreased thalamic activity has also been implicated in disorders of consciousness, and may be a potent mechanism for inducing decreased arousal (Schiff, 2008). Modulation of thalamic activity may therefore be an important component of general anesthesia. These findings suggest that selectively activating both TRN as well as additional targets of general anesthesia could induce a similar coma-like state. Similarly, activating broad

regions of TRN in combination with other sleep-promoting centers could provide a way to optogenetically engineer sleep and other arousal states.

### **3.4.7 TRN controls local cortical arousal state.**

We conclude that tonic depolarization of TRN leads to heterogeneous effects within TRN itself, shifting dynamics to activate a subset of TRN neurons that in turn inhibit thalamic activity. This tonic inhibition of thalamic neurons rapidly modulates cortical state and controls the animals' arousal. The induction of delta waves is fast ( $\ll 1$  sec) and has local cortical topography, demonstrating that TRN controls local oscillatory dynamics between thalamus and cortex, and that this interaction forms a central mechanism for modulation of arousal.

## **3.5 Experimental Procedures**

### **3.5.1 Cortical Implants**

In order to deliver light to TRN, all mice were implanted with a 2.1 NA fiber of 200 micron diameter targeting left TRN (1.8 mm lateral, -0.8 to -1.7 mm posterior relative to bregma; 2.1 mm deep). Two types of electrode implant were performed: a cortical implant with stereotrodes distributed across different cortical sites (McNaughton et al., 1983); and a subcortical implant, with moveable stereotrodes targeted to TRN. For the cortical implants, stereotrodes were made from pairs of 12.5 micron nichrome wire gold plated to  $\sim 300$  kOhm (California Fine Wire, Grover Beach CA). Electrodes were attached to small sections of plastic tubing cut to defined depth offsets and inserted by hand in 11 recording sites distributed across the cortex (Fig. 3.1a) at depths of 400, 500, 600, or 1300 microns, as defined by the length of the electrode extending from the plastic tubing. To calculate laser power within the brain, the laser power was first measured outside of the brain, and then this value was scaled to account for diminished power after passing through the fiber, measured before the implant surgery. For three of the four cortical mice, power was estimated at 80% of that measured outside the brain, and the fourth was estimated at 20% due to a misalignment of the fiber that reduced light transmission. For surgery, mice were anesthetized with 1% isoflurane and individual

holes were drilled for electrode and optical fiber insertion. Electrodes were inserted by hand and the optical fiber was placed using a stereotaxic arm. A total of eight mice positive for ChR2 and three mice negative for ChR2 were implanted.

### **3.5.2 Subcortical implants**

Hyperdrive bodies were designed in 3D CAD software (SolidWorks, Concord, MA) and stereolithographically printed in Accura 55 plastic (American Precision Prototyping, Tulsa, OK). Each hyperdrive was loaded with 6-8 individual, independently movable microdrives made of a titanium screw cemented to a 21-gauge cannula. Each microdrive was loaded with 1-3, 12.5 micron nichrome stereotrodes (California Fine Wire Company, Grover Beach, CA), which were pinned to a custom-designed electrode interface board (EIB) (Sunstone Circuits, Mulino, OR). Two EMG wires, two EEG wires and one ground wire (A-M systems, Carlsborg, WA), were also affixed to the EIB. An optical fiber targeting TRN (Doric Lenses, Quebec, Canada) was glued to the EIB. TRN targeting was achieved by guiding stereotrodes and optical fiber through a linear array (dimensions  $\sim 1.1 \times 1.8$  mm) secured to the bottom of the hyperdrive by cyanoacrylate. For surgery, mice were anesthetized with 1% isoflurane and placed in a stereotaxic frame. For each animal, five stainless-steel screws were implanted in the skull to provide EEG contacts (a prefrontal site and a cerebellar reference), ground (cerebellar), and mechanical support for the hyperdrive. A craniotomy of size  $\sim 3.0 \times 2$  mm was drilled with a center coordinate of (M/L 2.5mm, A/P -0.5mm). The implant was attached to a custom-designed stereotaxic arm, rotated 15 degrees about the median and lowered to the craniotomy. Stereotrodes were lowered slightly at the time of implantation ( $<500$  microns) and implanted into the brain.

### **3.5.3 Data acquisition**

Electrophysiology data was acquired on a Neuralynx (Neuralynx, Bozeman MT) system with a 32 kHz sampling rate. Full sampling was used to record spikes, detected with a manually set voltage threshold. LFPs were collected with a lowpass filter between 0.1 and 0.3 Hz and a highpass between 2000 and 9000 Hz. EMG was collected with a highpass filter of 10 Hz to prevent data saturation. All electrophysiology data was

exported to MATLAB (Mathworks, Natick MA), and LFPs and EMGs were then lowpass filtered offline at 500 Hz and downsampled to 1000 Hz sampling rate. Spike sorting was performed with custom software (Simpleclust, <http://moorelab.github.com/simpleclust>), using standard waveform features to classify spikes. Spikes that could not be assigned to a well-defined cluster were labeled as multi-unit activity, and triphasic waveforms were excluded as fibers of passage. Awake recordings were carried out in either a head-fixed setup or in a clear plastic bowl. All asleep recordings were performed in the bowl to ensure that animals could sleep. Anesthetized recordings were performed with isoflurane in 100% oxygen, in which drug levels were increased if the animal showed any signs of motion, and decreased when the EEG showed burst suppression, for an average range of 0.6% to 1% isoflurane. Analysis of EEG power in anesthetized trials was performed after anesthesia was induced with at least 1.5% isoflurane. Experiments began after mice had lost the righting reflex and isoflurane was maintained with at least 0.5% isoflurane throughout the stimulation period. Anesthetic levels were varied manually to stay within a lightly anesthetized range, by decreasing levels if the EEG showed burst suppression and increasing levels if mice showed any sign of movement. In sessions with automated motion quantification, two video cameras were mounted at two orthogonal angles to enable automated motion capture.

#### **3.5.4 Laser stimulation**

Light was delivered as 30 second DC stimulation periods, followed by at least 30 seconds (typically 60-90 seconds) with no stimulation. Light was maintained at constant levels throughout a single 30 second period. For experiments comparing different laser strengths, the laser output was varied within a single session, but not within a single 30 second stimulation period. Recording sessions were limited to no more than 60 stimulation trials, totaling approximately 1-2 hours.

#### **3.5.5 Spectral analysis**

Spectra were computed with the Chronux toolbox using 19 tapers over 30 second windows for a spectral resolution of 0.33 Hz. Spectrograms were computed with 5 tapers in 5 second sliding windows every 1 second, for a spectral resolution of 0.6 Hz.

Normalized spectrograms were computed by first taking the median power across all trials, and then dividing the power in each frequency by the mean of the power at that frequency during the 30 second pre-stimulus window. Error bars across multiple sessions (e.g. Fig. 3.1e) show the standard error of the mean across all trials. Statistical testing was done by taking the sum of power within a band of interest on individual trials, and then comparing power during the 30 second stimulation against power during the 30 seconds immediately preceding TRN stimulation. Change in power is reported as the median in dB, and 95% confidence intervals for the median effect were computed by inverting the Wilcoxon signed-rank test using R statistical software (<http://www.r-project.org>). For analyses across electrodes, the change in power was computed for each electrode and the Bonferroni correction was applied for multiple comparisons across electrodes.

### **3.5.6 Behavioral analysis**

All trials used for behavioral analysis were collected while mice behaved freely in a clear plastic bowl. Recording sessions lasted 1-2 hours and were performed during the day. Two video cameras were mounted at two orthogonal angles to enable automated motion capture. EMG effects were calculated using the Chronux toolbox to determine power in the 10-200 Hz band in non-overlapping windows of 1 second width. Power was summed across all frequencies within the band to obtain a single measure of EMG power. Statistical testing was performed with the Wilcoxon signed-rank test, comparing EMG power within each laser trial with the EMG power in the associated pre-stimulus period. EMG power was normalized to range between 0, corresponding to the 20<sup>th</sup> percentile of the raw data, and 1, corresponding to the 80<sup>th</sup> percentile, to approximate the range of values seen between sleep and wake states. The rapid onset of the EMG effect was assessed by comparing EEG power in the second prior to laser onset with power in the second following laser onset. Overall changes in behavior were calculated by comparing the [-30 0] baseline period to the [0 30] laser-induced period.

The correlation between EEG and EMG power was calculated by computing the correlation coefficient between the change in delta (1-4 Hz) power in the EEG, calculated as the difference in the [-30 0] and [0 30] periods relative to laser onset, and correlating it with the change in EMG power across those same periods. Statistical significance was

tested by performing a bootstrap with 1000 iterations on the paired power data, taking the 97.5% percentile of these resampled values, and testing whether its absolute value was larger than the correlation computed on a randomly shuffled set of times. This provided a test at significance level  $\alpha=0.05$  of whether the correlation between EEG and EMG effects was significantly higher during laser trials than would be expected during baseline conditions. Automated motion scoring was computed using automated custom software written in Matlab that calculated the optical flow for each frame in the video via the Horn–Schunck method (Horn and Schunck, 1981). The points of maximal motion in each camera view were used to compute a motion vector in the horizontal plane. The motion vector was normalized by its mean to avoid artifact due to variations in lighting conditions and camera placement. The magnitude of the motion vector for each frame was then smoothed using a moving average filter ( $\sigma=200$  frames/6.67 s) and used as a proxy for the magnitude of animal's overall motion. In order to ensure objective assessment of sleep, semi-automated sleep scoring was performed using an algorithm that first computes the instantaneous amplitude of the EMG between 60-200 Hz, smooths with a Gaussian filter of 50 ms, and then uses a manually entered threshold to identify wake states as those with at least 5 seconds of data above the threshold. Second, it computes the ratio of <4 Hz and 4-16 Hz power in the EEG, smooths with a Gaussian filter, and uses a manually entered threshold to segment non-REM and REM states. Sessions where the automated algorithm could not achieve a separation of sleep and wake states were not included in the analysis. The user selecting the EEG and EMG thresholds was blind to the timing of laser stimulation during the sleep scoring procedure. As with spectral effects, changes were reported as the median power change in dB, and 95% confidence intervals were computed by inverting the Wilcoxon signed-rank test.

### **3.5.7 Single unit analysis**

Statistical comparisons of firing rates were performed using confidence intervals derived from the Wilcoxon signed-rank test to quantify the median difference between each neuron's pre-stimulus ([-30 0] s) and stimulus-induced ([0 30] s) spike rates. To compute phase modulation, the instantaneous phase was calculated from the local LFP channel by bandpass filtering the LFP between 1-4 Hz with a finite impulse response

filter, taking the Hilbert transform, and then extracting the angle. The phase distribution of individual units relative to the delta phase was quantified using the modulation index (MI), adapted from phase-amplitude measurements (Tort et al., 2010), which measures the Kullback-Liebler distance between the observed phase distribution and a uniform distribution. The MI for spikes was computed over 10 phase bins as

$$\sum_{i=1}^{10} p_i \log_2 p_i + \log_2 10,$$

where  $p_i$  is the proportion of spikes falling within a given phase

bin. The MI for gamma power was computed over 100 phase bins using a sinusoid fit to model the amplitude of the gamma oscillation. The effect of TRN activation on unit phase modulation was assessed as in the firing rate case, with 95% confidence intervals derived from a Wilcoxon signed-rank test comparing each neuron's MI values in the [-30 0] and [0 30] periods. When comparing cortical units on channels with and without a delta effect, normalized delta power changes were computed by dividing delta (1-4 Hz) power by total 0-50 Hz power in the [-30 0] and [0 30] periods. Electrodes with a normalized delta power increase of at least 2% during TRN activation were labeled as having a delta effect. Subcortical units were divided into two categories based on the time between the peak and trough of the waveform. Narrow (<200 ms peak-to-trough) units were further subdivided into three categories based on their spike rate response to the laser, computed by taking the difference of their spike rates in the [0 30] s period vs. the [-30 0] period. Phase modulation was computed relative to the local LFP for each unit. To account for sign reversals due to electrode placement or referencing and ensure consistent phase measurements across units, LFPs were flipped such that the laser-induced deflection was negative across all channels. Confidence intervals for the median change in phase-locking across units was computed with a bootstrap, using 1000 draws (resampling with replacement) for each group, computing the difference in the median, and taking confidence intervals from the corresponding percentiles of the resampled values.

### 3.5.8 OFF period analysis

To detect OFF periods, we combined all multi-unit and single-unit activity on a single channel into a point process representation, and then smoothed with a Gaussian

kernel with a standard deviation of 20 ms to approximate an instantaneous firing rate in the units surrounding that electrode. OFF periods were labeled as any period of at least 50 ms with an estimated firing rate of zero. To verify that OFF periods were occurring at a greater rate than would happen by random chance, we also computed OFF periods on simulated data with the same mean firing rate as the experimental data. The simulated data was generated by taking the interspike intervals throughout the recording period, fitting a gamma distribution to these intervals, and then generating a new spike train from that gamma distribution with the same number of spikes as the original dataset. The OFF periods were then calculated with the same method for the simulated data. Statistical testing for OFF periods was performed across trials within each session: the percent of time spent in an OFF period during laser stimulation was compared to the percent of time spent in an OFF period in the 30 seconds preceding laser stimulation with the Wilcoxon signed-rank test. The percent of time in OFF periods was compared to simulated data by running the simulation 1000 times and testing whether the experimental value was greater than the 97.5<sup>th</sup> percentile of the simulated value.

### **3.6 Author Contributions**

Laura Lewis, Jakob Voigts, and Emery Brown conceived the experiments; Michael Halassa, Francisco Flores, and Matthew Wilson contributed to experimental design; Jakob Voigts, Francisco Flores, and Michael Halassa contributed to performing experiments; Jakob Voigts contributed to data analysis; Laura Lewis collected data, analyzed data, and wrote the manuscript with Jakob Voigts. All authors commented on the manuscript.

### **3.7 References**

Alkire, M., Haier, R., and Fallon, J. (2000). Toward a unified theory of narcosis: brain imaging evidence for a thalamocortical switch as the neurophysiologic basis of anesthetic-induced unconsciousness. *Consciousness and Cognition* 9, 370–386.

- Barrionuevo, G., Benoit, O., and Tempier, P. (1981). Evidence for two types of firing pattern during the sleep-waking cycle in the reticular thalamic nucleus of the cat. *Experimental Neurology* 72, 486–501.
- Bernstein, J.G., Han, X., Henninger, M.A., Ko, E.Y., Qian, X., Franzesi, G.T., McConnell, J.P., Stern, P., Desimone, R., and Boyden, E.S. (2008). Prosthetic systems for therapeutic optical activation and silencing of genetically-targeted neurons. *Proceedings of the Society of Photo-Optical Instrumentation Engineers* 6854, 68540H.
- Brown, E.N., Purdon, P.L., and Van Dort, C.J. (2011). General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual Review of Neuroscience* 34, 601–628.
- Constantinople, C.M., and Bruno, R.M. (2011). Effects and mechanisms of wakefulness on local cortical networks. *Neuron* 69, 1061–1068.
- Contreras, D., Curro Dossi, R., and Steriade, M. (1992). Bursting and tonic discharges in two classes of reticular thalamic neurons. *Journal of Neurophysiology* 68, 973–977.
- Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences* 81, 4586–4590.
- Crunelli, V., and Hughes, S.W. (2009). The slow (<1 Hz) rhythm of non-REM sleep: a dialogue between three cardinal oscillators. *Nature Neuroscience* 13, 9–17.
- Cueni, L., Canepari, M., Luján, R., Emmenegger, Y., Watanabe, M., Bond, C.T., Franken, P., Adelman, J.P., and Lüthi, A. (2008). T-type Ca<sup>2+</sup> channels, SK2 channels and SERCAs gate sleep-related oscillations in thalamic dendrites. *Nature Neuroscience* 11, 683–692.
- Deleuze, C., and Huguenard, J.R. (2006). Distinct electrical and chemical connectivity maps in the thalamic reticular nucleus: potential roles in synchronization and sensation. *Journal of Neuroscience* 26, 8633–8645.
- Destexhe, A., and Contreras, D. (2011). The fine structure of slow-wave sleep oscillations: from single neurons to large networks. In *Sleep and Anesthesia*, A. Hutt, ed. (Springer Science+Business Media), p. 258.
- Espinosa, F., Torres-Vega, M.A., Marks, G.A., and Joho, R.H. (2008). Ablation of Kv3.1 and Kv3.3 potassium channels disrupts thalamocortical oscillations in vitro and in vivo. *Journal of Neuroscience* 28, 5570–5581.
- Franks, N.P. (2008). General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. *Nature Reviews Neuroscience* 9, 370–386.
- Guillery, R.W., and Harting, J.K. (2003). Structure and connections of the thalamic

- reticular nucleus: Advancing views over half a century. *Journal of Comparative Neurology* 463, 360–371.
- Halassa, M.M., Siegle, J.H., Ritt, J.T., Ting, J.T., Feng, G., and Moore, C.I. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. *Nature Neuroscience* 14, 1118–1120.
- Hartings, J.A., Temereanca, S., and Simons, D.J. (2003). State-dependent processing of sensory stimuli by thalamic reticular neurons. *Journal of Neuroscience* 23, 5264–5271.
- Horn, B.K.P., and Schunck, B.G. (1981). Determining optical flow. *Artificial Intelligence* 17, 185–203.
- Houser, C.R., Vaughn, J.E., Barber, R.P., and Roberts, E. (1980). GABA neurons are the major cell type of the nucleus reticularis thalami. *Brain Research* 200, 341–354.
- Huber, R., Deboer, T., and Tobler, I. (2000). Topography of EEG dynamics after sleep deprivation in mice. *Journal of Neurophysiology* 84, 1888–1893.
- Ji, D., and Wilson, M.A. (2006). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience* 10, 100–107.
- Kim, A., Latchoumane, C., Lee, S., Kim, G.B., Cheong, E., Augustine, G.J., and Shin, H.-S. (2012). Optogenetically induced sleep spindle rhythms alter sleep architectures in mice. *Proceedings of the National Academy of Sciences* 109, 20673–20678.
- Krueger, J.M., Rector, D.M., Roy, S., Van Dongen, H.P.A., Belenky, G., and Panksepp, J. (2008). Sleep as a fundamental property of neuronal assemblies. *Nature Reviews Neuroscience* 9, 910–919.
- Lam, Y.-W., and Sherman, S.M. (2011). Functional organization of the thalamic input to the thalamic reticular nucleus. *Journal of Neuroscience* 31, 6791–6799.
- Landisman, C.E., Long, M.A., Beierlein, M., Deans, M.R., Paul, D.L., and Connors, B.W. (2002). Electrical synapses in the thalamic reticular nucleus. *Journal of Neuroscience* 22, 1002–1009.
- Lee, S.H., Govindaiah, G., and Cox, C.L. (2007). Heterogeneity of firing properties among rat thalamic reticular nucleus neurons. *The Journal of Physiology* 582, 195–208.
- Lewis, L.D., Weiner, V.S., Mukamel, E.A., Donoghue, J.A., Eskandar, E.N., Madsen, J.R., Anderson, W.S., Hochberg, L.R., Cash, S.S., Brown, E.N., et al. (2012). Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. *Proceedings of the National Academy of Sciences* 109, E3377–E3386.

- Long, M.A., Landisman, C.E., and Connors, B.W. (2004). Small clusters of electrically coupled neurons generate synchronous rhythms in the thalamic reticular nucleus. *Journal of Neuroscience* 24, 341–349.
- Massimini, M., Tononi, G., and Huber, R. (2009). Slow waves, synaptic plasticity and information processing: insights from transcranial magnetic stimulation and high-density EEG experiments. *European Journal of Neuroscience* 29, 1761–1770.
- McAlonan, K., Cavanaugh, J., and Wurtz, R.H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature* 456, 391–394.
- McCormick, D.A., and Bal, T. (1997). Sleep and arousal: thalamocortical mechanisms. *Annual Review of Neuroscience* 20, 185–215.
- McNaughton, B.L., O'Keefe, J., and Barnes, C.A. (1983). The stereotrode: a new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. *Journal of Neuroscience Methods* 8, 391–397.
- Murphy, M., Bruno, M., Riedner, B., Boveroux, P., Noirhomme, Q., Landsness, E., Brichant, J., Phillips, C., Massimini, M., and Laureys, S. (2011). Propofol anesthesia and sleep: a high-density EEG study. *Sleep* 34, 283.
- Nir, Y., Staba, R.J., Andrillon, T., Vyazovskiy, V.V., Cirelli, C., Fried, I., and Tononi, G. (2011). Regional slow waves and spindles in human sleep. *Neuron* 70, 153–169.
- O'Shea, S.M., Wong, L.C., and Harrison, N.L. (2000). Propofol increases agonist efficacy at the GABA(A) receptor. *Brain Research* 852, 344–348.
- Pace-Schott, E.F., and Hobson, J.A. (2002). The neurobiology of sleep: genetics, cellular physiology and subcortical networks. *Nature Reviews Neuroscience* 3, 591–605.
- Pinault, D. (2004). The thalamic reticular nucleus: structure, function and concept. *Brain Research Reviews* 46, 1–31.
- Poulet, J.F.A., Fernandez, L.M.J., Crochet, S., and Petersen, C.C.H. (2012). Thalamic control of cortical states. *Nature Neuroscience* 15, 370–372.
- Ray, S., and Maunsell, J.H.R. (2011). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *Plos Biology* 9, e1000610.
- Saper, C.B., Scammell, T.E., and Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature* 437, 1257–1263.
- Schiff, N.D. (2008). Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Annals of the New York Academy of Sciences* 1129, 105–118.

- Sherman, S.M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. *Trends in Neurosciences* 24, 122–126.
- Steriade, M. (2000). Corticothalamic resonance, states of vigilance and mentation. *Neuroscience* 101, 243–276.
- Todorovic, S.M., and Lingle, C.J. (1998). Pharmacological properties of T-type Ca<sup>2+</sup> current in adult rat sensory neurons: effects of anticonvulsant and anesthetic agents. *Journal of Neurophysiology* 79, 240–252.
- Tort, A.B.L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of Neurophysiology* 104, 1195–1210.
- Tsunematsu, T., Kilduff, T.S., Boyden, E.S., Takahashi, S., Tominaga, M., and Yamanaka, A. (2011). Acute optogenetic silencing of orexin/hypocretin neurons induces slow-wave sleep in mice. *Journal of Neuroscience* 31, 10529–10539.
- Vyazovskiy, V.V., Olcese, U., Hanlon, E.C., Nir, Y., Cirelli, C., and Tononi, G. (2011). Local sleep in awake rats. *Nature* 472, 443–447.
- Vyazovskiy, V.V., Olcese, U., Lazimy, Y.M., Faraguna, U., Esser, S.K., Williams, J.C., Cirelli, C., and Tononi, G. (2009). Cortical firing and sleep homeostasis. *Neuron* 63, 865–878.
- Wang, Q., Webber, R.M., and Stanley, G.B. (2010). Thalamic synchrony and the adaptive gating of information flow to cortex. *Nature Neuroscience* 13, 1534–1541.
- Zagha, E., Casale, A.E., Sachdev, R.N.S., McGinley, M.J., and McCormick, D.A. (2013). Motor cortex feedback influences sensory processing by modulating network state. *Neuron* 79, 567–578.
- Zhao, S., Ting, J.T., Atallah, H.E., Qiu, L., Tan, J., Gloss, B., Augustine, G.J., Deisseroth, K., Luo, M., Graybiel, A.M., et al. (2011). Cell type-specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nature Methods* 8, 745–752.

## Chapter 4: The neurophysiology of burst suppression in propofol-induced coma<sup>3</sup>

### 4.1 Abstract

Burst suppression is an electroencephalogram pattern that consists of a quasi-periodic alternation between isoelectric 'suppressions' lasting seconds or minutes, and high-voltage 'bursts.' It is characteristic of a profoundly inactivated brain, occurring in conditions including hypothermia, deep general anesthesia, infant encephalopathy and coma. It is also used in neurology as an electrophysiological endpoint in pharmacologically induced coma for brain protection after traumatic injury and during status epilepticus.

Classically, burst suppression has been regarded as a 'global' state with synchronous activity throughout cortex. This assumption has influenced the clinical use of burst suppression as a way to broadly reduce neural activity. However, the extent of spatial homogeneity has not been fully explored due to the challenges in recording from multiple cortical sites simultaneously. The neurophysiological dynamics of large-scale cortical circuits during burst suppression are therefore not well understood.

To address this question, we recorded intracranial electrocorticograms from patients who entered burst suppression while receiving propofol general anesthesia. The electrodes were broadly distributed across cortex, enabling us to examine both the dynamics of burst suppression within local cortical regions and larger-scale network interactions. We found that in contrast to previous characterizations, bursts could be substantially asynchronous across the cortex. Furthermore, the state of burst suppression itself could occur in a limited cortical region while other areas exhibited ongoing continuous activity. In addition, we found complex temporal structure within bursts, which recapitulated the spectral dynamics of the state preceding burst suppression, and evolved throughout the course of a single burst.

---

<sup>3</sup> The findings in this chapter were previously published as Lewis, Ching, Weiner, Peterfreund, Eskandar, Cash, Brown, and Purdon (2013). *Brain* 136(9):2727-2737.

Our observations imply that local cortical dynamics are not homogenous, even during significant brain inactivation. Instead, cortical and, implicitly, subcortical circuits express seemingly different sensitivities to high doses of anesthetics that suggest a hierarchy governing how the brain enters burst suppression, and emphasize the role of local dynamics in what has previously been regarded as a global state.

These findings suggest a conceptual shift in how neurologists could assess the brain function of patients undergoing burst suppression. First, analyzing spatial variation in burst suppression could provide insight into the circuit dysfunction underlying a given pathology, and could improve monitoring of medically-induced coma. Second, analyzing the temporal dynamics within a burst could help assess the underlying brain state. This approach could be explored as a prognostic tool for recovery from coma, and for guiding treatment of status epilepticus. Overall, these results suggest new research directions and methods that could improve patient monitoring in clinical practice.

## 4.2 Introduction

The anesthetized brain, though profoundly inactivated, is nevertheless characterized by rich electrophysiological dynamics. At deep levels of general anesthesia, the brain reaches a state of burst suppression, in which periods of high voltage brain activity (bursts) alternate with periods of isoelectric quiescence (suppressions). This alternation is quasi-periodic, and the suppression periods become longer in duration as the brain becomes more inactivated, lasting for seconds or even minutes (Amzica, 2009; Brown et al., 2010). Burst suppression appears to be a fundamental characteristic of the deeply anesthetized brain, and can occur in a range of conditions including hypothermia (Stecker et al., 2001), deep levels of general anesthesia (Akrawi et al., 1996), certain infant encephalopathies (Ohtahara and Yamatogi, 2003), and coma (Young, 2000). Clinically, it is commonly used as an electrophysiological endpoint in neurological intensive care following traumatic injury (Hall and Murdoch, 1990) and in the treatment of status epilepticus (Claassen et al., 2002; Rossetti et al., 2004). However, despite the presence of burst suppression in this broad range of inactivated brain states, its biophysical mechanisms are poorly understood.

Burst suppression has classically been viewed as a homogenous brain state. This perspective has been derived from EEG studies in which bursts and suppressions have been shown to occur concurrently across the scalp (Clark and Rosner, 1973). However, because scalp EEG is spatially blurred, the underlying dynamics are not fully understood. In vivo studies in anesthetized animals have helped to identify the potential cellular correlates of burst suppression, showing that although nearly all cortical neurons are inhibited during suppression periods, a subset of thalamocortical neurons can continue firing at delta frequencies (Steriade et al., 1994). This work established the involvement of different cortical and subcortical cell-types within both burst and suppression. However, a broader mechanism for burst suppression has not been determined, and the notion of burst suppression as a global state remains.

In search of a more detailed and complete mechanistic understanding, recent studies have suggested that burst suppression is associated with enhanced excitability in cortical networks (Kroeger and Amzica, 2007; Ferron et al., 2009). These studies implicate extracellular calcium as a correlate for the switches between burst and

suppression. Ching et al. (2012) have proposed an alternative mechanism, using computational methods, where burst suppression manifests in a state of reduced neuronal activity and cerebral metabolism. In such a state, insufficient production of ATP in local cortical networks can gate neuronal potassium channels, leading to suppression of action potentials. Such a mechanism accounts for the general features of burst suppression previously observed, as well as its occurrence under multiple etiologies, and also predicts a specific frequency structure for the neuronal activity within each burst.

A crucial element that has yet to be studied is how large-scale cortical networks function during burst suppression. Little is known about these systems-level circuits in burst suppression, partly due to the technical challenges of sampling multiple cortical sites simultaneously. Furthermore, it is not clear if burst suppression in the human brain matches what is described in animal models. Animal studies have typically been done over millimeter-scale areas of cortex, and it is possible that larger-scale recordings in human cortex could reveal spatial differences in burst suppression across cortex. Such spatial differences could identify certain brain regions to be differentially sensitive during burst suppression and pharmacologically induced coma, impacting strategies for clinical monitoring and patient care.

Here, we examine the cortical dynamics underlying burst suppression, and test whether their spatiotemporal properties are consistent with predictions from previous animal and computational studies. Specifically, we test: 1) whether the spatial distribution of burst suppression is homogenous across cortex, and 2) whether the temporal structure of the state preceding burst suppression is replicated within bursts and decelerates throughout bursts, as suggested by a computational model (Ching et al., 2012). We tested these hypotheses by recording intracranial electrocorticograms (ECoG) from five patients with intractable epilepsy who entered a state of burst suppression while undergoing general anesthesia for clinical treatment. We recorded from subdural grid, strip, and depth electrodes distributed throughout temporal, parietal, and frontal cortex. We investigated spatial dynamics in the four patients with grid electrodes, and temporal dynamics in all five patients. This combination of broad spatial coverage and excellent temporal resolution enabled an in-depth examination of the spatiotemporal dynamics of burst suppression.

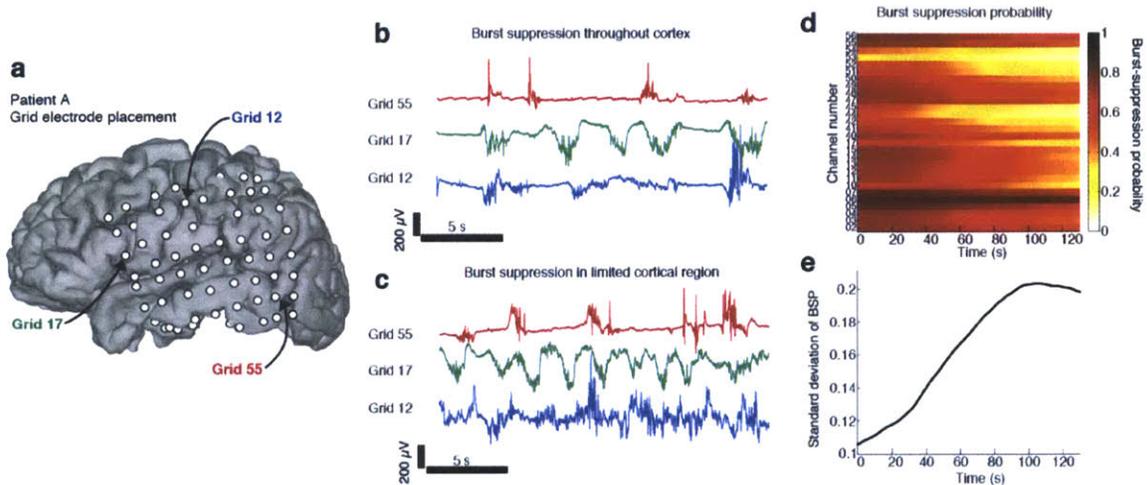
## 4.3 Results

### 4.3.1 Spatially isolated states of burst suppression.

We first examined the spatial dynamics of burst suppression in four subjects who each had a subdural grid of electrodes spaced 1 cm apart and spanning up to 11 cm of cortex (233 electrodes across 4 patients). This broad spatial sampling allowed us to examine how burst suppression dynamics varied across cortex. In contrast to the prevailing assumptions, we found clear cases in which the burst suppression pattern was localized. We computed the burst suppression probability (BSP) across all grid electrodes, which ranges from 0 to 1 and indicates the probability of a suppression period at a given point in time (Chemali et al., 2011). We found that the standard deviation of the BSP across the recording sites ranged from 0.04 to as high as 0.2. Across the entire post-induction period in these four patients, the mean range of the BSP across electrodes at a given point in time was 0.46, demonstrating a large average difference in burst dynamics across channels. These measures indicated that burst suppression dynamics could often diverge substantially across different cortical areas, with different cortical regions exhibiting different propensities for suppression.

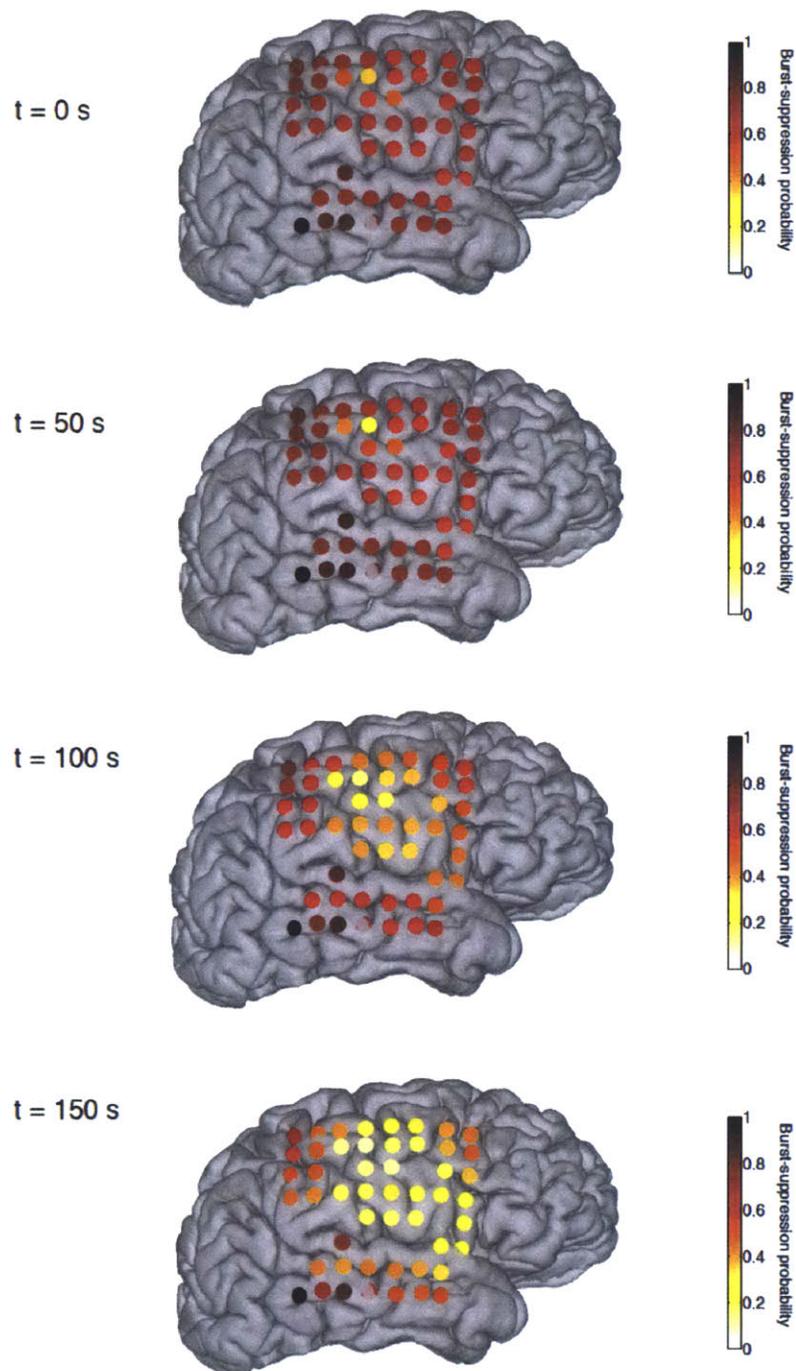
Given that the burst suppression probability could vary widely across cortex, we investigated whether the state of burst suppression itself could be limited to a restricted cortical region. We identified periods in which a subset of channels exited burst suppression, defined as an interval of at least 30 seconds in which the channel did not undergo a suppression. We found that in three of the four subdural grid patients, a subset of channels exited burst suppression while others remained in deep burst suppression, with BSPs over 0.5 (Fig. 4.1b, c, Fig. 4.2). To quantify the amount of time spent in spatially isolated burst suppression, we identified the total time over all patients in which some channels were deeply suppressed, while other channels remained nearly continuous. The epochs of deep suppression, defined as possessing any 3 channels with a BSP over 0.5, amounted to 15.9 minutes. Of this time, 4.6 minutes (28.7%) included at least 3 other channels that had a BSP less than 0.2, that is, at least 3 channels that remained in a lighter state of suppression.

These findings demonstrated that one region of cortex can be in a state of burst suppression, while neighboring cortical regions exhibit continuous activity characteristic of a lighter stage of anesthesia. Burst suppression can therefore occur in limited cortical regions, and does not necessarily reflect a cortex-wide phenomenon.



**Figure 4.1: The state of burst suppression can be limited to a local cortical region.** A) Reconstructed MRI for Patient A, showing grid electrode locations. Arrows mark the channels that are displayed in panels B and C. B) Example time-series in different cortical regions where all channels are in burst suppression, but bursts are asynchronous in different regions. C) Example from later in recording in same regions as panel A: channel 55 is in burst suppression, whereas channels 17 and 12 are not. The state of burst suppression is therefore not necessarily cortex-wide. D) Zoomed-in example from Patient B of the burst suppression probability (BSP) changing over time: initially most channels have a high BSP, but then a subset of channels exit burst suppression (e.g. channel 44) while other channels maintain high BSPs (e.g. channel 5 remains in burst suppression with a BSP above 0.5). E) Standard deviation of the BSP across all channels from panel B: the increasing standard deviation demonstrates that the BSPs in different cortical regions are becoming uncoupled as they diverge into different states.

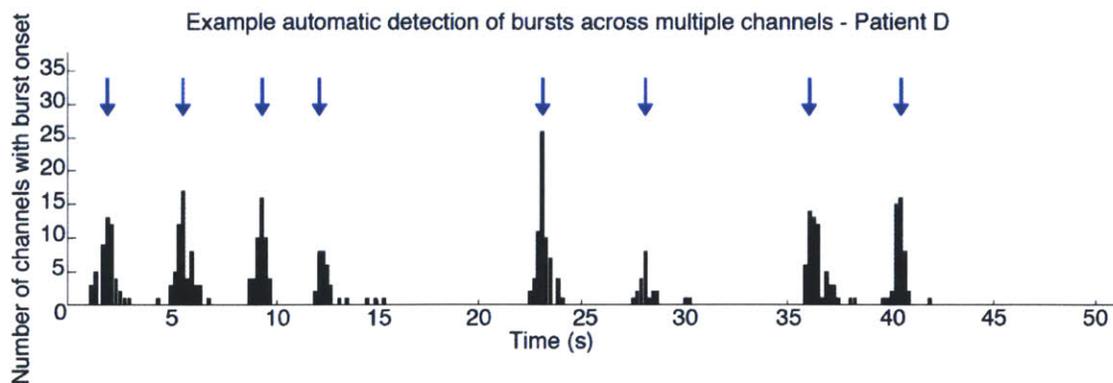
### Burst suppression probability across time



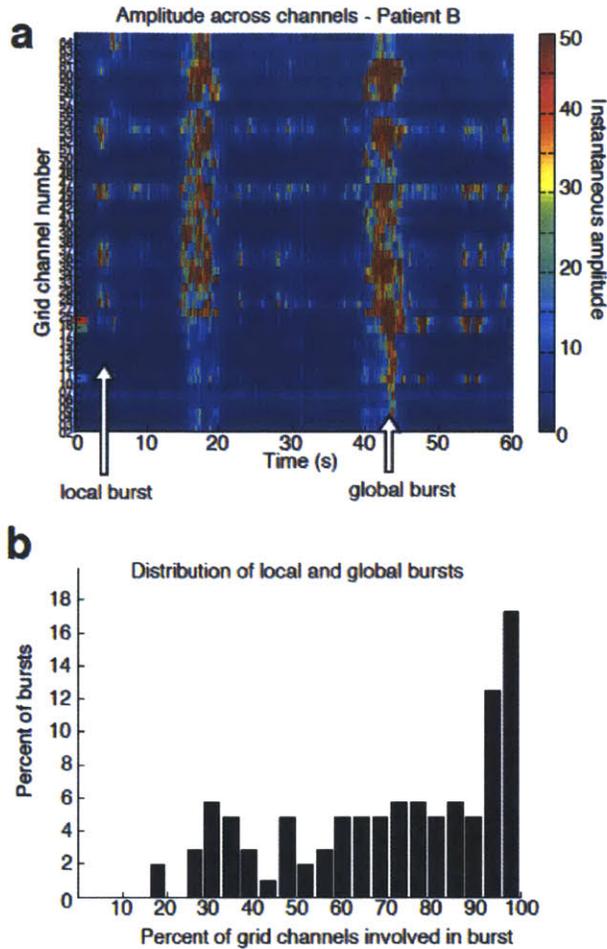
**Figure 4.2: Example of local burst suppression probability (BSP) in Patient B.** At time 0, the entire brain is in burst suppression; 150 seconds later, the middle and anterior grid contacts have exited burst suppression, whereas many posterior and temporal recordings remain in deep burst suppression.

### 4.3.2 Asynchronous bursts across cortex.

We next asked whether there could be spatially isolated burst dynamics even when all of cortex is in burst suppression. Specifically, we examined the spatial distribution of individual bursts in order to test whether bursts were sometimes constrained to a limited cortical region, as was suggested by observation of raw traces (Fig. 4.1b). We analyzed each burst individually and identified which grid channels participated in the burst by selecting those with nearby burst onset times (Fig. 4.3). We found that bursts frequently occurred locally: bursts were observed in only a subset of the grid channels while other regions remained suppressed (Fig. 4.4a). These local bursts were interspersed with global bursts that occurred across all channels, demonstrating that the state of burst suppression was present across cortex but that individual bursts could occur asynchronously (Fig. 4.4a). The median percentage of channels involved in a single burst was only 76% (quartiles: 52%-94%, Fig. 4.4b), demonstrating that local bursts made up a substantial portion of total bursting. We therefore concluded that bursts can be either global or local, and that local bursts reflect activation in a limited cortical area while other regions continue to be suppressed. The fact that suppressions can continue in one region despite high-amplitude bursts in a neighboring region suggests that a profoundly inactivated state can be confined to specific cortical regions.



**Figure 4.3: Method for quantifying number of channels in each burst.** Algorithm bins burst onsets across all channels into 200 ms bins, and then finds local maxima of at least 5 channels. Blue arrows mark the detected multi-channel bursts; all channels with a burst onset within 1.5 seconds of the arrow are counted as participating in a burst.



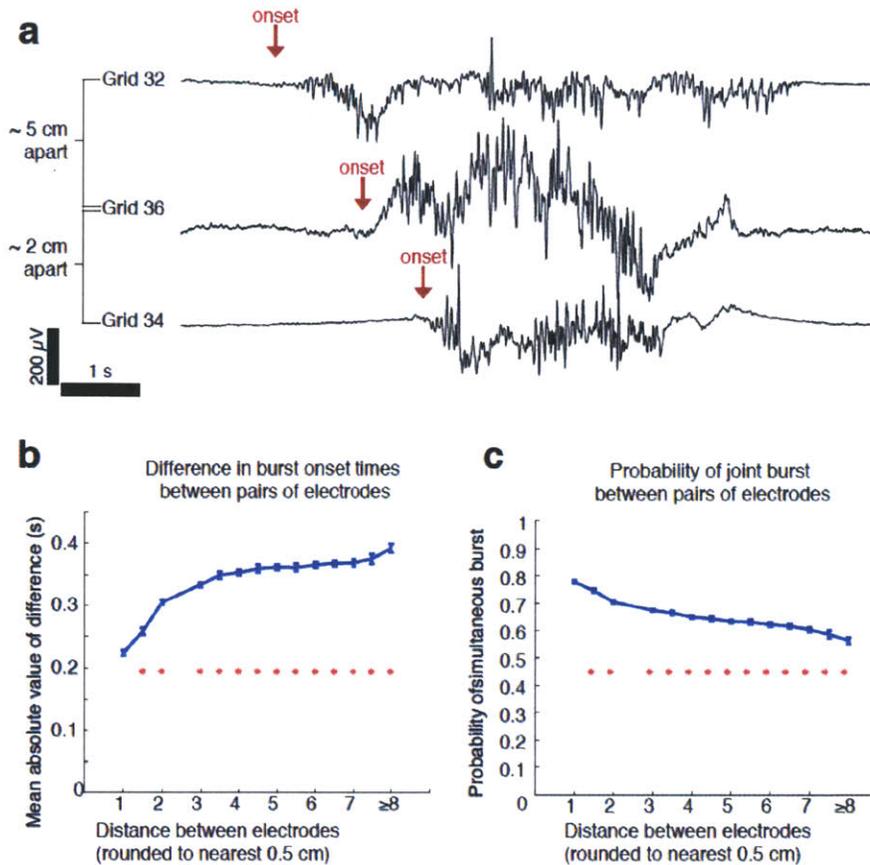
**Figure 4.4: Bursts can occur in limited areas of cortex.** A) Instantaneous amplitude across all grid channels. The burst at 42 sec. involves all channels, but the burst at 5 sec. occurs in only a small subset of electrodes, indicating that a limited cortical region is bursting. B) Histogram shows the number of grid channels participating in each burst across four patients: many bursts are global, but there is a long leftward tail to the distribution, demonstrating the frequent occurrence of local bursts.

### 4.3.3 Timing of burst onsets vary across cortex.

Despite the presence of spatially localized bursts, it is also clear that many bursts occur broadly across cortex, as nearly a third of bursts (31%) occurred in over 90% of channels (Fig. 4.4b). We therefore tested whether these ‘global’ bursts across multiple channels began simultaneously in each channel, or whether there were consistent time lags between distant channels. We compared burst onset times between every pair of electrodes, and found that mean differences in burst onset time were significantly correlated with the distance between pairs of channels. Onset time differences were larger

between more distant pairs of channels (Fig. 4.5b), increasing from  $225 \pm 83$  ms (mean  $\pm$  s.d.) in adjacent (1 cm) channels to  $368 \pm 107$  ms (mean  $\pm$  s.d.) in channels separated by more than 4 cm (difference=143 ms,  $p < 10^{-5}$ ). Because the difference in burst times was correlated with distance ( $R=0.30$ ,  $p < 10^{-5}$ ), these timing differences could not be attributed simply to noise in the automated segmentation algorithm; even if the entire difference between adjacent channels was due to segmentation noise, an additional 143 ms would remain as the mean difference in burst timing between distant channels. To further ensure that these results were not an artifact of our burst detection algorithm, we repeated this analysis using the variance of the raw signal to detect burst onsets (see Methods) (Löfhede et al., 2008), and replicated the correlation of timing differences with distance (difference = 153 ms,  $R=0.27$ ,  $p < 10^{-5}$ ). This result therefore demonstrated that even a single ‘global’ burst can be locally differentiated, and can begin hundreds of milliseconds apart in different cortical regions (Fig. 4.5a).

We further tested this spatial heterogeneity by computing the probability that two different electrodes would simultaneously be in a bursting state. We found that adjacent electrodes had a 78.1% probability of being in a burst state simultaneously, whereas more distant electrodes shared only a 62.0% chance of simultaneous bursting (Fig. 4.5c,  $p < 10^{-5}$ ). As with the differences in burst timing, this result demonstrated that burst onsets are asynchronous across cortex, with significant lags between distant cortical regions participating in a simultaneous burst.

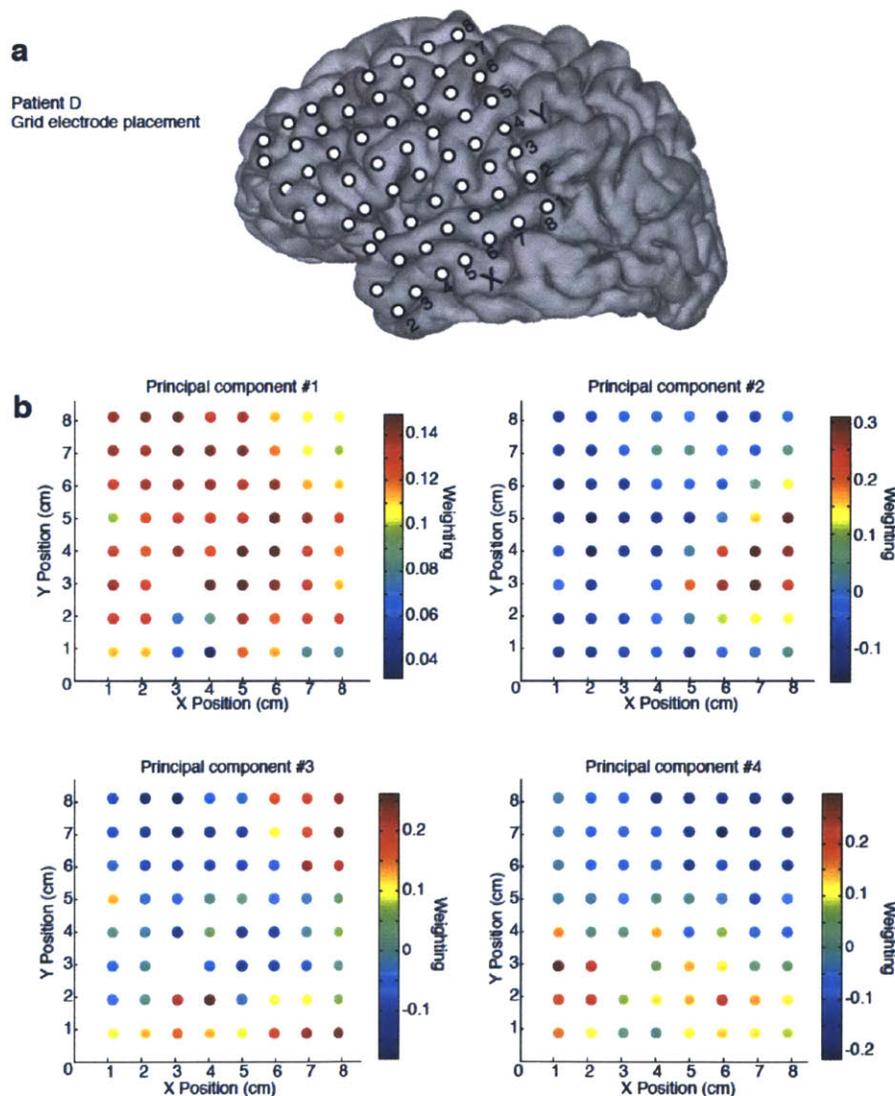


**Figure 4.5: Burst timing is heterogeneous across cortex.** A) Example trace from Patient C: the burst in channel 32 starts hundreds of milliseconds before the bursts in channels 36 and 34. B) Plot of mean difference in burst onset times between electrodes shows that there are substantial timing differences in burst onsets between distant electrodes, with distant electrodes showing larger gaps in burst timing. Blue lines are mean and standard error, red stars mark distances that are significantly different than pairs 1 cm apart ( $n=233$  electrodes in 4 patients). C) Probability that two electrodes are both bursting (total time where both electrodes burst, normalized by total time that either electrode has a burst.) Probability decreases with distance, demonstrating that distant electrodes are less likely to be simultaneously in a burst state.

#### 4.3.4 Anatomically clustered burst dynamics.

Taken together, these results demonstrate that there can be substantial heterogeneity in bursting dynamics across the cortex, and suggest that bursts are spatially clustered. To explicitly test for spatial clustering of bursts, we performed a principal components analysis on the burst state across all grid electrodes. We found that 78% of the variance could be explained by the first 4 components in each patient. To test whether spatial clustering was present in these first 4 principal components (across all 4 patients, a

total of 16 components), we compared the spatial derivative of the estimated components to a randomly shuffled grid. This shuffling analysis demonstrated that 15 out of the 16 components were significantly spatially clustered ( $p < 0.05$ , Fig. 4.6), supporting the hypothesis that clusters of anatomically close cortical areas tended to share burst properties. We therefore concluded that although burst suppression is sufficiently correlated across cortex to produce a seemingly synchronous pattern in scalp EEG recordings, the underlying dynamics exhibit substantial local heterogeneity.



**Figure 4.6: Principal components analysis demonstrates that bursts are spatially clustered.** A) Reconstruction of the grid electrode placement in patient D. B) Each panel shows one of the first four principal components from Patient D, and the colour variation

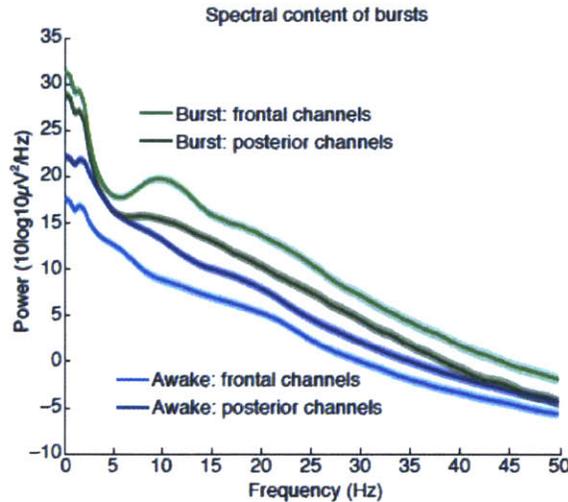
across the grid demonstrates how burst probability is locally differentiated. Each of the components is significantly spatially clustered ( $p < 0.05$ ), demonstrating that burst properties are anatomically clustered and differ in distant cortical regions.

#### **4.3.5 Recovery of local dynamics within each burst.**

This demonstration of spatially differentiated dynamics suggested that bursts and suppressions depend on local cortical state. This finding could be compatible with a previously described model for the generation of burst suppression, which proposes that a depressed cerebral metabolism could lead to burst suppression by producing a slow cycle in ATP levels (Ching et al., 2012). This model makes specific predictions about the spectral content within individual bursts: first, that they will recover the dynamics of the state immediately preceding burst suppression, and second, that the recovered oscillatory features will decelerate through the course of each burst.

We therefore tested whether the spectral content of bursts returned to the pre-burst suppression state – in this case, a lighter stage of propofol general anesthesia. Propofol general anesthesia produces two striking features in the EEG: a large increase in low-frequency (0.1-4 Hz) power (Murphy et al., 2011; Lewis et al., 2012; Purdon et al., 2013), and an alpha (~10 Hz) rhythm that occurs predominantly in frontal regions (Cimenser et al., 2011; Murphy et al., 2011; Supp et al., 2011; Purdon et al., 2013). We computed the within-burst spectrum across all ECoG channels to test whether these features were present. As this analysis did not require spatial information, we expanded our dataset to include an additional subject with implanted depth electrodes, as well as including all depth and strip electrodes from the first four patients (total of 374 electrodes in 5 patients). We classified each channel as ‘frontal’ if it was anterior to the central sulcus, as defined by visual inspection of reconstructed MRI images (Dykstra et al., 2012), and ‘posterior’ otherwise. Spectra were then calculated for the two seconds following burst onset, and averaged separately for frontal and posterior channels. The resulting spectra were highly consistent with previous reports of lighter stages of general anesthesia: posterior channels had a strong power component at slow frequencies, and frontal channels had both increased slow power as well as a pronounced alpha oscillation (Fig. 4.7). Comparing these spectra to a baseline recording within the same patients

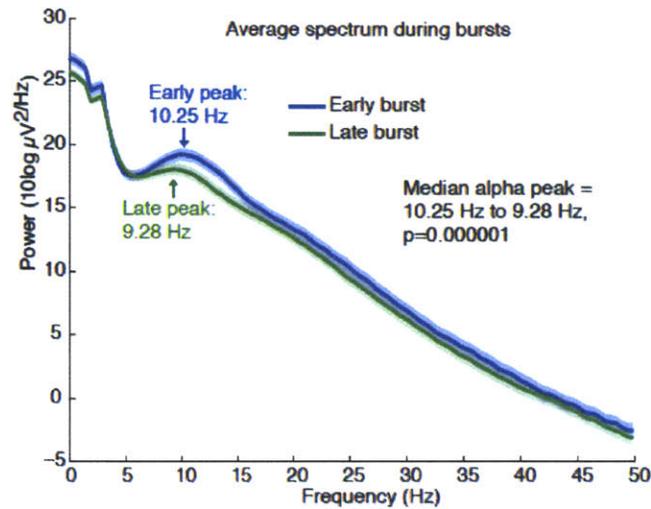
before any anesthesia was administered confirmed that these changes marked a departure from the awake state. Within-burst dynamics were therefore consistent with the dynamics observed during lighter stages of propofol general anesthesia, in agreement with the prediction of the metabolism-based model.



**Figure 4.7: Bursts recover the spectral dynamics of propofol general anesthesia.** Average spectra (+/- std. err.) within a burst, across all channels (n=374 electrodes in 5 patients), categorized by anatomical location. Plot shows that the bursts contain increased slow power relative to the awake state, and a frontal alpha oscillation.

#### 4.3.6 Deceleration of frequency structure during bursts.

Given that the propofol-induced alpha rhythm resumed during bursts, we next tested whether its frequency decelerated throughout a burst. We first selected all channels with a peak in power in the alpha band, defined as higher power in the alpha (8-14 Hz) band than in the theta (4-7 Hz) band. The timecourse of alpha dynamics was examined by comparing the early (0-1.5 s) and late (1.5-3 s) components of bursts, restricting the analysis to bursts lasting at least 3 seconds. The spectra showed that alpha rhythms did decelerate throughout a burst, with a peak frequency dropping from 10.25 Hz in the early period to 9.28 Hz in the late period of the burst (Fig. 4.8,  $p < 10^{-5}$ ). Burst dynamics were therefore not only variable across cortex, but also exhibited consistent patterns *within* a single burst, and these patterns aligned precisely with the predictions of the metabolism-based model.



**Figure 4.8: The alpha rhythm decelerates over the course of the burst.** Average spectra across all channels with an alpha oscillation (n=160 electrodes in 5 patients) show that there is a significant decrease in peak frequency between the early (0-1.5 s) and late (1.5-3 s) portions of a burst.

## 4.4 Discussion

### 4.4.1 Local dynamics in burst suppression

In this paper we have shown local cortical dynamics in the state of burst suppression induced by propofol general anesthesia. Specifically, our results establish that (1) bursts and suppressions can occur in a limited cortical region while continuous activity persists in other areas; (2) even when all of cortex undergoes a ‘global’ burst, there are significant differences in the timing of onset of bursts between disparate cortical regions, (3) that, within each burst, the frequency structure matches the brain state that was present prior to the onset of burst suppression; and (4) this frequency structure decelerates through the course of each burst. Taken together, these findings suggest that burst suppression is highly dependent on local cortical dynamics, as the state evolves both across time and across different cortical areas.

### 4.4.2 Implications for neurological disease

Given that burst suppression is both a symptom of neurological conditions – for instance, in post-anoxic coma – and is induced as a treatment for conditions such as status epilepticus and traumatic brain injury, these findings could have significant impact on

clinical practice. In particular, detection of the spectral content within each burst could reveal the neural dynamics that remain intact when not interrupted by the suppression epochs. For instance, bursts may contain activity synonymous with general anesthesia as observed here, or they could be morphologically similar to epileptiform patterns associated with seizure. Monitoring the transition between these two patterns could reveal the underlying brain state, and a shift in the spectral content of bursts could signal an opportunity to lift a pharmacologically induced coma. Similarly, tracking the complexity of burst dynamics could also aid in evaluation of coma recovery and of brain development in early neonates. Future clinical studies could examine in detail how the spectral content of the EEG during burst suppression may be useful for predicting patient outcomes.

In addition to monitoring spectral dynamics within a burst, the spatial heterogeneity of burst suppression has implications for our understanding of neurological disease, and could impact clinical treatment. First, these results indicate that patients exhibiting burst suppression may in fact have substantial local variation in brain function. Neurologists may therefore wish to examine spatial differences in burst suppression to ascertain whether specific cortical regions are more susceptible to circuit dysfunction, as inactivation in different brain structures may be a function of underlying pathology. Furthermore, these results suggest that medically-induced coma, as used for treatment of status epilepticus and traumatic brain injury, should be monitored across multiple cortical regions and the treatment adjusted accordingly, as dynamics in one brain region may not fully reflect the ongoing state. In addition, the ability to observe and characterize local expression of suppression epochs could allow for more precise tracking of anesthetic induction and emergence, and of hypothermia induced during surgery (Michenfelder and Milde, 1991). Specialized monitoring systems could be designed that exploit EEG spatial patterns to enable superior control of drug dosages when inducing burst suppression to control status epilepticus or for treatment of traumatic brain injury, ensuring that a desired burst suppression ratio is achieved throughout the brain rather than at a single cortical site.

#### **4.4.3 Balance between local and global dynamics**

Burst suppression has previously been viewed as a global phenomenon, with synchronous bursts occurring simultaneously across cortical areas. In our studies, we indeed observed high correlation of bursts across cortex, demonstrating that on average bursts are broadly synchronous. However, we also identified substantial local variation in burst dynamics. Burst timing differed consistently across cortex, with larger timing offsets between bursts in distant regions. In addition, both bursts and suppressions frequently occurred locally, limited to a small cluster of electrodes while other cortical regions were in a different state. One possible explanation for this could be local variation in cerebral metabolism: when metabolism is globally depressed, bursts can spread across cortex, producing a gradient of timing differences; whereas when metabolic rates are more varied in different regions, they may enter dissociated states with different burst suppression probabilities and different refractory periods, leading to spatially isolated bursts and suppressions. This interaction could resolve the contrast inherent in these results, as this mechanism would produce dynamics in which bursts are often correlated but can nevertheless demonstrate substantial local variation.

#### **4.4.4 Subcortical circuit mechanisms**

Our data and results have centered on the cortical dynamics of burst suppression. However, subcortical structures are undoubtedly an important determinant in the expression of burst suppression in the brain. Indeed, the local differences we have shown in cortical measurements are suggestive of nontrivial subcortical participation in each burst and suppression. The state of burst suppression can be viewed as a severe reduction in the ability of cortical neurons to sustain continued processing. Whether the reason is protective, for instance by metabolic mechanisms, or otherwise, the neurons in question simply cannot fire for prolonged periods of time. In contrast, previous research on the cellular correlates of burst suppression has shown that certain subcortical populations, namely thalamic reticular and relay cells, may exhibit ongoing activity even during cortical suppressions (Steriade et al., 1994). The generation of individual bursts is thought to be caused by input from these relay neurons once cortical post-suppression refractory periods subside (Kroeger and Amzica, 2007). The extent to which burst

suppression is expressed differentially in the cortex may thus be a reflection of the integrity of specific thalamocortical networks. In this scenario, the dynamic range in some subcortical loops – and the efficacy of ascending and descending excitation – can remain largely intact, despite existing in a significantly inactivated brain. These differences suggest that there are differential sensitivities of cortical regions and their associated functions to anesthetic drugs at high concentrations, hypothermia and diffuse brain injury.

#### **4.4.5 Relationship to neuronal and metabolic mechanisms**

Our results are consistent with the neuronal and metabolic mechanisms proposed in recent computational work (Ching et al., 2012). In that model, it was suggested that lowered cerebral metabolism leads to periods of suppression, but that the activity within each burst recovers the oscillatory dynamics of the state preceding burst suppression. An alternative hypothesis is that bursts are due to cortical hyperexcitability (Amzica 2009; Ferron et al., 2009). In the case of propofol general anesthesia, the EEG prior to burst suppression contains two distinct rhythms: a slow (0.1-1 Hz) oscillation that is asynchronous across cortex, and an alpha (~10 Hz) rhythm that is highly coherent across frontal electrodes (Supp et al, 2011; Lewis et al., 2012; Purdon et al., 2013). The slow oscillation contains EEG deflections that mark brief (<1 s) periods of local cortical neuron inactivation. These inactivated periods occur both during sleep (Cash et al., 2009, Nir et al., 2011) and general anesthesia (Lewis et al., 2012), and correlate with loss of consciousness. In this study, we found that bursts indeed replicated the EEG signatures of lighter stages of general anesthesia: they exhibited both a slow oscillation and a frontal alpha oscillation that decelerated throughout the burst, as predicted by the decreased cerebral metabolism model (Ching et al., 2012). The fact that slow oscillations were contained within bursts suggests that burst suppression may be due to prolonged epochs of suppression overriding the ongoing cortical state. Bursts would then reflect a transient recovery in which the oscillatory rhythms characteristic of the preceding state (i.e. the slow and alpha oscillations) resume. This theory is additionally consistent with the fact that patients remain anesthetized during bursts, as their EEG continues to reflect the signatures of propofol general anesthesia. Our analyses therefore suggest that the main

emergent feature of burst suppression may in fact be the suppression, which acts as an intermittent but prolonged interruption of ongoing cortical activity. The burst content could then serve as a readout of the previous cortical state, which could provide useful clinical information when monitoring patients during burst suppression.

The spatial heterogeneity we observe here is also consistent with the metabolic model. In particular, it would follow from the model that bursts and suppressions may be shorter or longer in different brain regions depending on regional variations in perfusion, local network activity, ATP concentration, and metabolic state. These spatial results are also compatible with a calcium-based mechanism for burst suppression. Namely, it has been suggested that transient increases and decreases in extracellular calcium, leading to synaptic disfacilitation, are a key determinant in suppression duration (Amzica, 2009). Again, such a mechanism would naturally lead to local variability due to calcium distribution and expression. Taken together, our results support a model in which burst suppression is driven by local variations in cortical dynamics, and are consistent with the hypothesis that suppressions are caused by decreased cerebral metabolism (Ching et al., 2012). Nevertheless, our results do not definitively verify a mechanism for burst suppression, and future experiments combining cortical and subcortical recordings with direct metabolic measurements will be needed to determine the precise molecular mechanisms involved.

#### **4.4.6 Future directions**

There remain limitations in the current study that could be addressed by future work. In particular, the recordings were collected in patients with epilepsy, rather than in healthy subjects. However, the consistency of our results across patients, despite their heterogeneous clinical backgrounds, suggests that these findings are not an artifact of their clinical histories. In addition, our data replicate the burst suppression characteristics reported in previous studies, such as quasiperiodic bursts and increasing BSP with increasing anesthetic dose, suggesting that the burst suppression dynamics in these patients do not substantially differ from those of the typical brain. Moreover, the majority of electrodes in this analysis were not immediately adjacent to the seizure focus, but rather in cortex that was presumed healthy and not resected during clinical treatment. A

second consideration is that we have only analyzed burst suppression induced by propofol general anesthesia. Given the similarity of burst suppression across different etiologies, we expect that these results would generalize to other neurological conditions that produce burst suppression, but future studies in patients with different etiologies will be needed to verify this experimentally. In addition, future studies will be needed to investigate whether other causes of burst suppression (i.e. other anesthetic drugs, coma, and hypothermia) produce similar spatiotemporal dynamics to those observed here. A final limitation is the lack of scalp EEG recordings for comparison with the intracranial recordings. Future studies will be needed to determine how the spatial structure detected in intracranial electrodes translates to recordings at the scalp, which will be important for clinical applications.

#### **4.4.7 Understanding brain function through burst suppression**

Our findings provide new insight into the neurophysiology of the profoundly inactivated brain. Despite trends towards synchronous activity, local cortical dynamics vary across time and space, and can lead to uncoupled burst suppression states across cortex. These results demonstrate previously unknown complexity in neural circuit dynamics during deep general anesthesia, and suggest new roles for cortical and subcortical structures in producing neurophysiological diversity during profound neural inactivation. These findings indicate that burst suppression in neurological conditions could benefit from an examination of how cortical activity varies within bursts and across electrodes, as these dynamics may be highly variable. In addition, they suggest future clinical studies to investigate how analysis of the spatiotemporal structure of burst suppression patterns could improve patient monitoring and the effectiveness of clinical treatments.

## **4.5 Experimental Procedures**

### **4.5.1 Data acquisition**

We enrolled five patients with epilepsy intractable to medication, who were implanted with intracranial electrocorticography (ECoG) electrodes for standard clinical

monitoring (AdTech Inc, Racine Wisconsin). Informed consent was obtained from all patients in accordance with the local institutional review board. Electrode placement was determined solely by clinical criteria. One patient was implanted only with depth electrodes and the other four had a combination of depth electrodes and subdural grid and strip electrodes, with 1 cm spacing between electrode contacts. Recordings were collected throughout induction of general anesthesia using propofol, at the beginning of a surgery to explant the electrodes. Relevant clinical information, including anesthetic protocol and recording sites, can be found in Tables 4.1 and 4.2. A portion of one recording from one patient was previously reported in a separate analysis of slow oscillations (Lewis et al., 2012). ECoG data was recorded with a sampling rate of 2000 Hz, lowpass filtered at 100 Hz and resampled to 250 Hz. For all analyses of spatial dynamics (Figs. 4.1-4.6), grid channels were referenced with a Laplacian montage in which the average of all available neighboring channels (up to four nearest neighbors) was subtracted, in order to minimize spatial spread of the signals. When analyzing temporal structure across all channels (grid, strip, and depth electrodes; Figs. 4.7-4.8), they were referenced in a bipolar scheme as the depth and strip electrodes were positioned in 1-dimensional arrays. Channels with large artifacts or with periods of signal saturation were excluded from the analysis. All data were exported to Matlab (Mathworks) for further analysis with custom software.

Patient	Age	Gender	Weight	Location of epileptogenic area	Etiology	Electrodes
A	52	F	65 kg	Temporal lobe	Cortical dysplasia	64-contact grid covering left temporal cortex, extending partially over frontal and parietal regions; 4-contact strip on left occipital cortex; 3 4-contact strips spanning left anterior, middle, and posterior subtemporal cortex; 2 8-contact depth electrodes targeting anterior and posterior left hippocampus.
B	32	M	67 kg	Frontal lobe	Unknown	64-contact grid over temporal and parietal cortex; 4-contact strip extending to frontal cortex; 8-contact strip along parietal cortex; 4-contact strip on subtemporal cortex; two 8-contact depths into medial temporal cortex; one occipital 8-contact depth.
C	23	F	55 kg	Frontal lobe	Unknown	64-contact grid over frontal, temporal, and parietal cortex; two 4-contact subfrontal strips; two 4-contact interhemispheric frontal strips; 8-contact medial temporal depth; 8-contact cingulate depth; 8-contact prefrontal depth
D	34	F	112 kg	Frontal lobe	Cortical dysplasia	64-contact grid over frontal, temporal, and parietal cortex; 4-contact frontopolar strip; 4-contact subfrontal strip; 4-contact subtemporal strip; two 8-contact medial temporal depths; two 8-contact frontal depths
E	29	M	75 kg	Temporal lobe	Post-traumatic Epilepsy	4 8-contact depth electrodes targeting left and right temporal & hippocampal cortex; 6 6-contact depth electrodes in subfrontal, cingulate, and posterior frontal cortex

**Table 4.1:** Clinical information for each study participant.

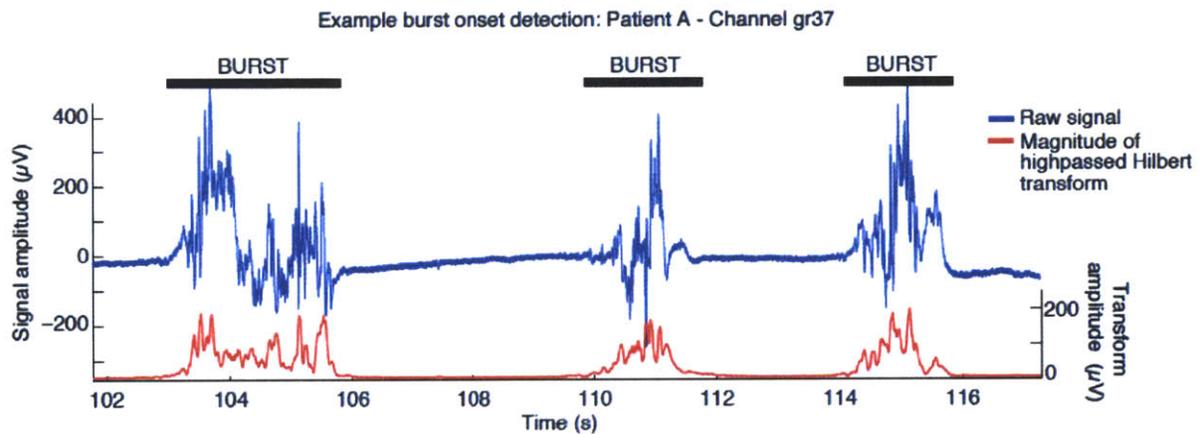
Patient	Propofol bolus	Propofol infusion	Comedication	Mean arterial pressure	CO2
A	130 mg (2 mg/kg)	None	150 mcg fentanyl ~1 min post-induction; 2 mg midazolam~10 mins pre-induction	88	35
B	Total of 350 mg (5.2 mg/kg) over four boluses	None	0.05 mcg/kg/min remifentanyl throughout induction	131	49
C	Total of 200 mg (3.64 mg/kg) over two boluses	None	250 mcg fentanyl simultaneously with propofol	83	21
D	None	Gradual infusion with varying rate, for a total of 3.3 mg/kg over 14 minutes	None	81	31
E	Total of 400 mg (5.3 mg/kg) over 4 boluses	None	None	70	29

**Table 4.2:** Anesthetic protocol for each study participant.

#### 4.5.2 Segmentation of burst suppression

In each patient, a period of burst suppression was manually identified and extracted for further analysis. We used an automated method to segment bursts and suppressions (Fig. 4.9). The method first required manual labeling of unambiguous suppression periods in the first 60 seconds of the recording. The data was then transformed in three steps: 1) signals were highpass filtered with a finite impulse response filter of length 2206, with a gain of 0 from 0-2.55 Hz and a gain of 1 from 3-125 Hz. 2) the Hilbert transform of the transformed signal was used to calculate the instantaneous amplitude, and 3) the instantaneous amplitude was smoothed with a moving average filter with a span of 50 samples (200 ms). These transformations yielded a continuous measure approximating high-frequency power. The value of this measure during the manually labeled suppression periods was used to set a threshold for burst detection (mean plus four standard deviations of the value during manually-labeled suppressions). Threshold crossings lasting over 500 ms were labeled as bursts, and burst terminations were labeled when the measure returned below threshold for 500 ms. We used 500 ms as a computational requirement for threshold crossings but manually confirmed that our method successfully detected the slow timescale shifts characteristic

of burst suppression. In particular, we noted that the median duration of suppressions was 4.76 s, with an inter-quartile range of 3.76-7.31 s. To ensure that our results on burst timing were not an artifact of the burst detection algorithm, we implemented an alternative variance-based method. In this method, the variance of the raw signal was computed in 100 ms sliding windows and this measure replaced the instantaneous amplitude as the segmentation threshold.



**Fig 4.9: Example of automated burst segmentation.** Algorithm transforms raw signal (blue) to compute the instantaneous amplitude above 3 Hz (red). It sets a threshold based on manually identified suppression periods, and then labels threshold crossings longer than 0.5 s as bursts.

#### 4.5.3 Comparisons of burst timing

The difference in burst onset times was taken between every pair of electrodes in the grid. For each burst onset in a given electrode, the burst occurring closest in time in every other electrode was selected if it occurred within 1 s of the first burst. The absolute value of this timing difference was then calculated, and averaged across all pairs of electrodes in the grid. Timing differences were statistically compared across different distances of electrode separation using the Wilcoxon ranksum test.

The joint probability of bursting in two electrodes was computed for each pair of electrodes by calculating the amount of time that both electrodes were simultaneously in a burst state, and then normalizing by the total amount of time that either electrode was in a burst state. As above, significant changes in joint bursting probability at different distances were calculated using the Wilcoxon ranksum test.

#### **4.5.4 Identification of local bursts**

We plotted burst onsets across all channels and found that burst onsets were visibly clustered across channels (Fig. 4.3), enabling an automated selection of multi-channel bursts using a simple threshold. Clusters of burst onsets were identified when at least 5 channels had a burst onset within a 200 ms bin. The number of channels involved in each burst was then computed by counting the number of channels that demonstrated a burst onset within 1.5 seconds of the main cluster, to ensure that all channels were counted even if burst onset was substantially delayed.

#### **4.5.6 Spectral analysis of bursts**

The spectral content of bursts was analyzed using multitaper spectral estimation, computed with the Chronux toolbox (Bokil et al., 2010). Within-burst dynamics were analyzed by selecting bursts lasting at least 3 seconds, and running a triggered spectral analysis at the onset of those bursts. Spectra were estimated with a  $T = 2$ -second window, a time-bandwidth product of  $TW = 3$ , and 5 tapers, yielding a spectral resolution of 1.5 Hz. An analogous calculation was performed on the baseline awake period by taking a triggered spectrum of an equal number of windows spaced 4 seconds apart. Error bars were computed as the standard error of the spectra across channels.

Comparisons of early and late portions of the burst were performed by selecting bursts lasting at least 3 seconds, and dividing them into two 1.5-second windows, marked 'early' and 'late'. The analysis was restricted to channels with an alpha peak, defined as channels where the maximum power in the 8-14 Hz range was higher than the maximum power in the 4-7 Hz range. For each portion, the spectrum was calculated in a  $T = 1.5$  second window, with time-bandwidth product  $TW = 4$ , with 7 tapers, yielding a spectral resolution of 2.67 Hz. The peak alpha power was then identified as the frequency with the highest power lying between 8 and 14 Hz. Statistical testing was performed by identifying the difference between the early and late peak in each channel, and then performing a Wilcoxon signed-rank test on the difference across all channels. Plots show the average power across channels and error bars show the standard error.

## 4.6 Author Contributions

Laura Lewis, Patrick Purdon, Emery Brown, Sydney Cash, and ShiNung Ching designed the research; Veronica Weiner contributed to data collection; Emad Eskandar performed the surgeries; Robert Peterfreund contributed to anesthetic protocols and delivery; Laura Lewis collected data, performed the data analysis, and wrote the manuscript with ShiNung Ching and Patrick Purdon. All authors commented on the manuscript.

## 4.7 References

- Akrawi, W.P., Drummond, J.C., Kalkman, C.J., and Patel, P.M. (1996). A comparison of the electrophysiologic characteristics of EEG burst-suppression as produced by isoflurane, thiopental, etomidate, and propofol. *Journal of Neurosurgical Anesthesiology* 8, 40–46.
- Amzica, F. (2009). Basic physiology of burst-suppression. *Epilepsia* 50, 38–39.
- Bokil, H., Andrews, P., Kulkarni, J. E., Mehta, S. & Mitra, P. P (2010). Chronux: A platform for analyzing neural signals. *Journal of Neuroscience Methods* 192, 146–151.
- Brown, E.N., Lydic, R., and Schiff, N.D. (2010). General anesthesia, sleep, and coma. *New England Journal of Medicine* 363, 2638–2650.
- Chemali, J.J., Wong, K.F.K., Solt, K., and Brown, E.N. (2011). A state-space model of the burst suppression ratio. *Conference proceedings of the IEEE Engineering in Medicine and Biology Society. 2011*, 1431–1434.
- Ching, S., Purdon, P.L., Vijayan, S., Kopell, N.J., and Brown, E.N. (2012). A neurophysiological-metabolic model for burst suppression. *Proceedings of the National Academy of Sciences* 109, 3095–3100.
- Cimenser, A., Purdon, P.L., Pierce, E.T., Walsh, J.L., Salazar-Gomez, A.F., Harrell, P.G., Tavares-Stoeckel, C., Habeeb, K., and Brown, E.N. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proceedings of the National Academy of Sciences* 108, 8832–8837.
- Claassen, J., Hirsch, L.J., Emerson, R.G., and Mayer, S.A. (2002). Treatment of refractory status epilepticus with pentobarbital, propofol, or midazolam: a systematic review. *Epilepsia* 43, 146–153.
- Clark, D.L., and Rosner, B.S. (1973). Neurophysiologic effects of general anesthetics. I. The electroencephalogram and sensory evoked responses in man. *Anesthesiology*

38, 564–582.

- Dykstra, A.R., Chan, A.M., Quinn, B.T., Zepeda, R., Keller, C.J., Cormier, J., Madsen, J.R., Eskandar, E.N., and Cash, S.S. (2012). Individualized localization and cortical surface-based registration of intracranial electrodes. *Neuroimage* 59, 3563–3570.
- Ferron, J.-F., Kroeger, D., Chever, O., and Amzica, F. (2009). Cortical inhibition during burst suppression induced with isoflurane anesthesia. *Journal of Neuroscience* 29, 9850–9860.
- Hall, R., and Murdoch, J. (1990). Brain protection: physiological and pharmacological considerations. Part II: The pharmacology of brain protection. *Canadian Journal of Anaesthesia* 37, 762–777.
- Kroeger, D., and Amzica, F. (2007). Hypersensitivity of the anesthesia-induced comatose brain. *Journal of Neuroscience* 27, 10597–10607.
- Lewis, L.D., Weiner, V.S., Mukamel, E.A., Donoghue, J.A., Eskandar, E.N., Madsen, J.R., Anderson, W.S., Hochberg, L.R., Cash, S.S., Brown, E.N., et al. (2012). Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. *Proceedings of the National Academy of Sciences* 109, E3377–E3386.
- Löfhede, J., Löfgren, N., Thordstein, M., Flisberg, A., Kjellmer, I., and Lindcrantz, K. (2008). Classification of burst and suppression in the neonatal electroencephalogram. *Journal of Neural Engineering* 5, 402–410.
- Michenfelder, J.D., and Milde, J.H. (1991). The relationship among canine brain temperature, metabolism, and function during hypothermia. *Anesthesiology* 75, 130–136.
- Murphy, M., Bruno, M., Riedner, B., Boveroux, P., Noirhomme, Q., Landsness, E., Brichant, J., Phillips, C., Massimini, M., and Laureys, S. (2011). Propofol anesthesia and sleep: a high-density EEG study. *Sleep* 34, 283.
- Ohtahara, S., and Yamatogi, Y. (2003). Epileptic encephalopathies in early infancy with suppression-burst. *Journal of Clinical Neurophysiology* 20, 398–407.
- Purdon, P.L., Pierce, E.T., Mukamel, E.A., Prerau, M.J., Walsh, J.L., Wong, K.F.K., Salazar-Gomez, A.F., Harrell, P.G., Sampson, A.L., Cimenser, A., et al. (2013). Electroencephalogram signatures of loss and recovery of consciousness from propofol. *Proceedings of the National Academy of Sciences* 110, E1142–E1151.
- Rossetti, A.O., Reichhart, M.D., Schaller, M.-D., Despland, P.-A., and Bogousslavsky, J. (2004). Propofol treatment of refractory status epilepticus: a study of 31 episodes. *Epilepsia* 45, 757–763.

- Stecker, M.M., Cheung, A.T., Pochettino, A., Kent, G.P., Patterson, T., Weiss, S.J., and Bavaria, J.E. (2001). Deep hypothermic circulatory arrest: II. Changes in electroencephalogram and evoked potentials during rewarming. *Annals of Thoracic Surgery* *71*, 22–28.
- Steriade, M., Amzica, F., and Contreras, D. (1994). Cortical and thalamic cellular correlates of electroencephalographic burst-suppression. *Electroencephalography and Clinical Neurophysiology* *90*, 1–16.
- Supp, G.G., Siegel, M., Hipp, J.F., and Engel, A.K. (2011). Cortical hypersynchrony predicts breakdown of sensory processing during loss of consciousness. *Current Biology* *21*, 1988–1993.
- Young, G.B. (2000). The EEG in coma. *Journal of Clinical Neurophysiology* *17*, 473–485.

## **Chapter 5: Conclusions**

### **5.1 Summary of the thesis**

This thesis integrates clinical studies in human subjects with optogenetic manipulations in mice to identify the neural basis of states of decreased arousal. I focus on two specific effects of general anesthesia: the dynamics that cause the transition into unconsciousness, and the neurophysiology of the deeply anesthetized brain. In Chapter 2, I show that propofol-induced loss of consciousness occurs at the onset of a slow oscillation that isolates local cortical networks. I next show that increasing activity in the thalamic reticular nucleus can create cortical slow-wave activity and decrease arousal through modulation of thalamocortical activity, identifying a potential mechanism for low-frequency EEG patterns during both general anesthesia and sleep (Chapter 3). Finally, in Chapter 4 I show that burst suppression during deep general anesthesia consists of local suppression superimposed on the ongoing EEG dynamics, and suggest local variation in cerebral metabolism as the neurophysiological mechanism (Chapter 4). This work identifies the brain states that underlie loss of consciousness and medically induced coma, and provides evidence for specific mechanisms that generate these states.

### **5.2 Implications for the neuroscience of arousal states**

#### **5.2.1 Potential mechanisms for propofol-induced unconsciousness**

These findings suggest an overarching theory for how propofol produces unconsciousness. In Chapter 2, I show that propofol induces slow wave activity that fragments cortical networks, disrupting long-range cortical communication and producing unconsciousness (Figure 5.1). Chapter 3 then uses causal manipulations to demonstrate that activation of thalamic reticular nucleus induces slow wave activity and decreases arousal. Modulation of thalamic activity may therefore form part of the mechanism by which propofol causes unconsciousness. Propofol directly enhances TRN's inhibitory effects on thalamus by potentiating and/or activating GABA-A receptors on thalamocortical neurons (Ying and Goldstein, 2005). In addition, propofol could act to disinhibit TRN by enhancing GABAergic projections from the ventrolateral

preoptic area of the hypothalamus onto the brainstem cholinergic centers (peduncolopontine and laterodorsal tegmental nuclei), thus decreasing inhibitory cholinergic inputs to TRN (Brown et al., 2011). Overall, TRN-mediated thalamic inhibition is a compelling candidate for a circuit mechanism through which propofol could induce unconsciousness. However, this circuit can explain only one component of propofol's effects, as the beta oscillations associated with sedation and the alpha oscillations associated with unconsciousness are likely to be mediated by inhibitory actions directly on cortical neurons as well (Ching et al., 2010; McCarthy et al., 2008). Future studies could examine the impact of these dynamics on cortical function and explore potential mechanisms.

These findings are compatible with several current theories for how anesthetics could produce unconsciousness (reviewed in section 1.6.) Overall, I suggest that consciousness is impaired through slow wave activity that disrupts large-scale cortical networks. These findings directly contradict the theory that loss of consciousness is due to decreased cortical firing rates, as we find that cortical neurons can sometimes maintain their firing rates at baseline levels during the unconscious period. Instead, these results are consistent with the theory that thalamus can serve as a control switch for arousal state, as its projections to cortex are key for modulating cortical state (Alkire et al., 2000; Schiff, 2008).

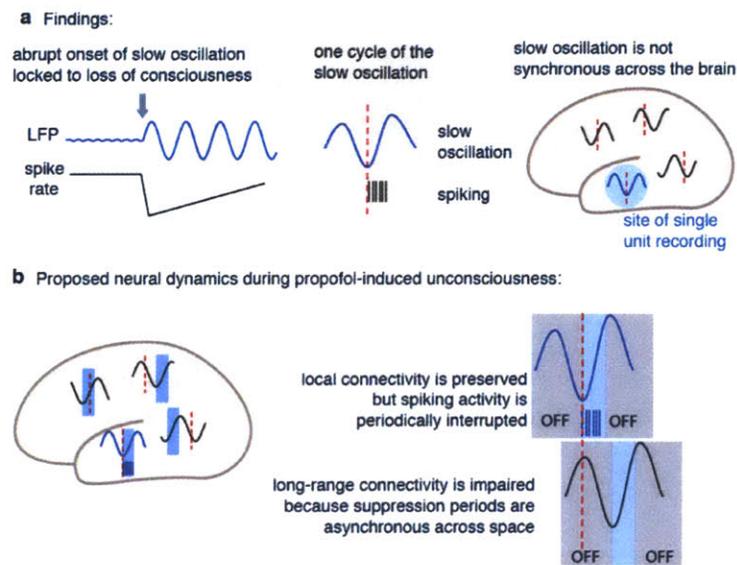
In addition, our findings are compatible with other theories, while not directly supporting their hypotheses. Firstly, we neither support nor refute the theory that unconsciousness is due to loss of gamma synchrony (Mashour, 2004). The onset of slow waves does disrupt gamma synchrony, by preventing simultaneous gamma oscillations across distant cortical sites, but we have not shown that gamma itself plays any role in unconsciousness. However, future studies could test this theory by specifically disrupting gamma rather than inducing a complete OFF period, perhaps through optogenetic manipulation.

Our results are also consistent with the theory that consciousness requires integrated information, as fragmentation of cortical networks into isolated subregions impacts the 'integration' component of this theory (Tononi, 2004). However, it is not clear what inference can be made from this approach; we conclude that integrity of

cortical networks is essential for consciousness but do not necessarily validate the idea that consciousness can be derived from integrated information.

### **5.2.2 Local dynamics in a global brain**

While the results presented here are compatible with several theories, I propose that the key aspect of each brain state considered here is the presence of local, isolated cortical dynamics. Sleep and states of decreased arousal have classically been thought to involve global dynamics across all of cortex, and this assumption is reflected in the use of the term ‘synchronized’ to describe EEG and LFP patterns during these states. Our studies demonstrate that while neurons within a local region are synchronized, producing large-amplitude oscillations due to their phase-locked activity, dynamics across broader cortical regions are often strongly dissimilar. This asynchronicity provides an explanation for loss of consciousness during general anesthesia, as it is expected to disrupt long-range cortical communication. In addition, it provides important information about the circuitry underlying states of decreased arousal, as it suggests that an underlying circuit mechanism must be spatially organized in order to produce these localized dynamics. In the case of unconsciousness, I show that the local circuitry is provided by the corticotopic organization of thalamic reticular nucleus, whereas in burst suppression it reflected local control of cerebral metabolism. The presence of localized dynamics during TRN stimulation could reflect dual-use of this circuitry both to produce global inactivation during sleep, but also to induce localized suppression of specific cortical areas during the awake period. This function would be useful when mediating attention to a single sensory modality while ignoring others, consistent with the concept of thalamic reticular nucleus as an attentional searchlight (Crick, 1984). However, rather than simply limiting the transmission of information from thalamus to cortex, TRN could also withdraw a given cortical region from corticocortical interactions by inducing a sleep-like state in that local area. This local impairment could serve as both a mechanism underlying loss of consciousness when present throughout cortex (Chapter 2), or contribute to cognitive impairment when present in limited cortical regions, as is seen during sleep deprivation (Vyasovskiy et al., 2011).

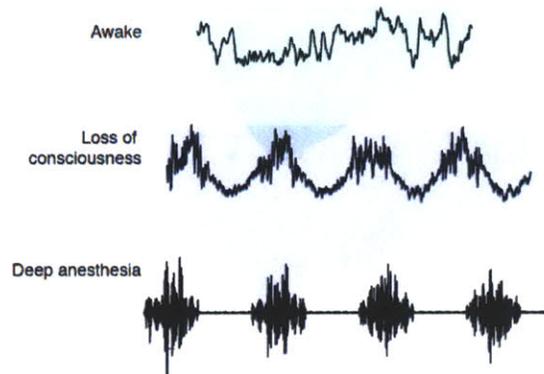


**Figure 5.1: The effect of slow wave dynamics on cortical processing.** A) Findings in Chapter 2 demonstrate that slow oscillations occur abruptly at the onset of unconsciousness during propofol general anesthesia. The slow oscillations are asynchronous across cortex, and local neuronal activity is limited to the active phase of the oscillations. B) These dynamics suggest that cortical processing is disrupted during slow oscillations. Although local neuronal interactions are preserved during a period of firing, distant areas are inactivated and therefore long-range communication is impaired.

### 5.2.3 Nesting of brain states across different levels of anesthesia

Our studies of burst suppression, at higher levels of propofol, highlighted the qualitatively different neurophysiological mechanisms that are engaged at different dosages. During the initial period of loss of consciousness, slow waves alternate between silence (OFF periods) and brief intervals that resemble the normal waking state (Fig. 2.9). Propofol-induced unconsciousness therefore consists of a periodic suppression of the ongoing cortical activity. This ON/OFF rhythm is further nested in burst suppression: the rhythmic slow waves are themselves interrupted by a prolonged isoelectric suppression. Burst suppression therefore appears to act through distinct mechanism, as loss of cerebral metabolism imposes a new, slower rhythm that contains the dynamics of lighter general anesthesia nested within it (Figure 5.2). This finding is a clear illustration that separate mechanisms engaged by a single drug can combine to produce complex dynamics, and

suggests that successful studies of anesthesia will need to study different dosage levels separately. In addition, a useful approach could be to perform local infusions of drug to assess the role of specific regions in generating these dynamics, and use selective agonists or antagonists of specific channels to explore their role in each brain state.



**Figure 5.2: Nested rhythms in propofol general anesthesia.** The ON states of the slow rhythm contain some dynamics present in the awake state (as well as some new ones, such as frontal alpha oscillations), and the bursts during deep general anesthesia contain the slow waves seen at lighter levels.

#### 5.2.4 Slow-wave activity and its mechanisms

A key finding of this thesis is that slow wave activity is not only correlated with unconsciousness, it marks a disruption of long-range cortical networks that is expected to impair cognitive function (Fig. 5.1, Chapter 2). Furthermore, we identified a central circuit that generates slow-wave activity, and found that it also plays a causal role in decreasing an animal's arousal state (Chapter 3). Slow wave activity is found across many states of decreased awareness, including drowsiness, sedation, general anesthesia, certain types of coma, and absence seizures. Slow waves may therefore be a general way that consciousness breaks down across multiple types of decreased arousal. However, as each of these conditions is caused by different mechanisms, this finding also suggests that there are many pathways that result in slow waves. It is possible that slow wave activity is an idling state that corticothalamic loops oscillate in whenever tonic levels of inhibition increase, allowing multiple sources of inhibition to produce the same pattern of cortical activity and the same impairment of cognitive function.

While this thesis identifies a circuit mechanism that can generate slow wave activity, it does not identify the source of rhythmicity: that is, why does tonic activation of TRN result in a 1-4 Hz oscillation? In *in vitro* studies, both thalamus and cortex are capable of inherently generating delta rhythms (Carracedo et al., 2013; McCormick and Pape, 1990). However, *in vivo* it is challenging if not impossible to determine which region is the source of the oscillation, and it may be that the *in vivo* rhythm is truly a joint oscillation rather than being generated by a single structure. In addition, slow wave activity is sometimes thought to be composed of two separate rhythms: a slow oscillation (0.1-1 Hz) and the delta rhythm (1-4 Hz) (Steriade et al., 1993). This thesis did not differentiate between the two rhythms, and it is not clear whether they are truly distinct in the intact brain, as slow oscillations are tightly linked to delta waves rather than occurring as a dissociable phenomenon. For both of these issues, oscillations that are clearly distinct with identifiable cellular-level generators *in vitro* may not directly translate to the *in vivo* case, as the oscillations that arise from the system as a whole are likely to have entirely different properties than those seen in isolated thalamic or cortical slabs. Given the involvement of both cortex and thalamus in generating delta rhythms, I suggest that the delta observed here results from an interaction of the two structures rather than being driven by a single cell type, but future studies (perhaps with *in vivo* intracellular recordings) will be needed to accurately determine the true cellular mechanisms.

### **5.3 Implications for clinical care**

The results presented here suggest specific clinical studies that could translate this information into recommendations for clinical care. First, the observation of a strong correlation between slow wave activity and unconsciousness (Chapter 2) suggests that monitoring patient EEG during general anesthesia could help prevent intraoperative awareness. Future clinical studies should examine whether titrating anesthetic dosage to levels that show consistent slow wave activity leads to a reduction in intraoperative awareness. If so, these results could provide a new patient monitoring method that could help prevent the ~26,000 annual cases of intraoperative awareness in the United States (Sebel et al., 2004).

Second, our findings that engaging the thalamic reticular nucleus can induce cortical sleep states (Chapter 3) suggests new targets for drug development, as targeting this nucleus could aid in clinical induction of either sleep or general anesthesia. Insomnia is an extremely prevalent illness, affecting between 6-10% of the general population (Roth, 2007). Similarly, millions of patients undergo general anesthesia each year, and many experience side effects such as vomiting, respiratory problems, and post-operative cognitive dysfunction. Our results suggest that targeting thalamic reticular nucleus, perhaps in combination with additional arousal control centers such as the ventrolateral preoptic area, could provide a selective way to induce sleep and/or unconsciousness, highlighting a promising avenue for either optogenetic or pharmacological manipulation of arousal states.

Third, our characterization of cortical function during burst suppression (Chapter 4) has substantial clinical implications for a broad range of conditions, as burst suppression is both a symptom of neuropathology and is induced as a treatment for many neurological conditions. Our findings suggest specific monitoring systems that could be useful in clinical practice: first, analyzing the spectral content of bursts could reveal the underlying brain state that would be present if not for burst suppression. This information could be useful for assessing when to lift a medically induced coma, such as in treatment of status epilepticus or traumatic brain injury. Second, analyzing the spatial distribution of bursts could improve measurements of the burst suppression ratio, a commonly used metric for assessing depth of anesthesia, thereby improving clinical control over brain state during medically induced coma. Finally, both the spectral and temporal methods could be used to explore whether specific characteristics of burst suppression can help with diagnosis and prognosis of patients with burst suppression coma due to neurological insult.

## **5.4 Conclusion and future directions**

### **5.4.1 Sedation and emergence**

This thesis characterized the dynamics underlying loss of consciousness and deep general anesthesia, but many other states remain to be explored. First, the sedation that

occurs during low doses of propofol is a fascinating and unexplained state. The mechanisms that produce both the EEG dynamics and remarkable behavioural effects of low-dose propofol, which can range from drowsiness to agitation, remain unknown. The approach used in this thesis could be applied to study sedation and paradoxical excitation, as the combination of human studies and optogenetic manipulations in mice would easily translate to studying this altered state of consciousness. Repeating these experiments with lower doses of propofol may therefore provide insight into the brain states underlying sedation and paradoxical excitation.

In addition, this thesis focuses on induction and maintenance of general anesthesia, but emergence may involve a different set of dynamics. After cessation of anesthetic drug, arousal state gradually increases, but how neural systems recover and re-engage consciousness remains unknown. In addition, children and the elderly are vulnerable to emergence delirium, a period of agitation that occurs as the anesthetic wears off. Future work could examine the cortical dynamics associated with these cognitive disturbances, and explore whether changing the protocols for anesthetic delivery could reduce the incidence of emergence delirium.

#### **5.4.2 Approaches for studying arousal states**

This thesis used two approaches to the study of arousal states that enabled us to overcome previous limitations in the field. First, an essential aspect of these experiments was that they examined the temporal evolution of brain states rather than averaging across long time periods. By analyzing how neural dynamics changed simultaneously with loss of consciousness, and by examining the link between TRN-induced delta waves and behavioral state, we identified a circuit mechanism that directly controls arousal state. This approach has significant advantages over the standard method of simply comparing the awake state to a stably anesthetized state. Even a single anesthetic drug can induce a wide range of different neurophysiological and cognitive effects, ranging from beta oscillations during sedation, slow waves during unconsciousness, and burst suppression during deep anesthesia. Analyzing these states separately is essential in order to understand the functional impact of the observed neurophysiology.

Second, we find that it is useful to integrate both animal and human studies when studying loss of consciousness. The phenomenology of consciousness has some fundamentally human elements, and we can interrogate a person's arousal state far more precisely than an animal's, which may alter its behaviour in different ways during sedation and loss of consciousness. Furthermore, any effort to obtain clinically relevant findings must include human studies, as the EEG and the relationship between neural dynamics and cognitive function will have substantial differences across species. However, sophisticated manipulation and recording of specific neural circuits is only possible in animal models, and transgenic mouse lines have opened up a new world of possibilities for experimental manipulations. Linking human and animal studies therefore has the potential to make major new advances in the neuroscience of arousal states, by identifying key phenomenology in the human brain but then exploring its underlying mechanisms in animal models.

#### **5.4.3 Conclusions**

In conclusion, this thesis characterizes the neural dynamics underlying anesthetic-induced loss of consciousness and coma, and identifies neurophysiological mechanisms that produce these states. These results suggest that disruption of long-range cortical networks may be a key element of unconsciousness, and point to the thalamus as a central modulator of arousal state. Furthermore, these findings highlight the importance of directly linking neurophysiology to cognitive function, rather than classifying states according to drug dosage, as a single drug can induce qualitatively different brain states through multiple mechanisms. In addition, I demonstrate the efficacy of linking human and animal studies to identify circuit mechanisms that control arousal state, and suggest clinical studies that could improve patient care.

## 5.5 References

- Alkire, M., Haier, R., and Fallon, J. (2000). Toward a unified theory of narcosis: brain imaging evidence for a thalamocortical switch as the neurophysiologic basis of anesthetic-induced unconsciousness. *Consciousness and Cognition* 9, 370–386.
- Brown, E.N., Purdon, P.L., and Van Dort, C.J. (2011). General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual Review of Neuroscience* 34, 601–628.
- Carracedo, L.M., Kjeldsen, H., Cunnington, L., Jenkins, A., Schofield, I., Cunningham, M.O., Davies, C.H., Traub, R.D., and Whittington, M.A. (2013). A neocortical delta rhythm facilitates reciprocal interlaminar interactions via nested theta rhythms. *Journal of Neuroscience* 33, 10750–10761.
- Ching, S., Cimenser, A., Purdon, P.L., Brown, E.N., and Kopell, N.J. (2010). Thalamocortical model for a propofol-induced alpha-rhythm associated with loss of consciousness. *Proceedings of the National Academy of Sciences* 107, 22665–22670.
- Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences* 81, 4586–4590.
- Mashour, G.A. (2004). Consciousness unbound: toward a paradigm of general anesthesia. *Anesthesiology* 100, 428–433.
- McCarthy, M.M., Brown, E.N., and Kopell, N. (2008). Potential network mechanisms mediating electroencephalographic beta rhythm changes during propofol-induced paradoxical excitation. *Journal of Neuroscience* 28, 13488–13504.
- McCormick, D.A., and Pape, H.C. (1990). Properties of a hyperpolarization-activated cation current and its role in rhythmic oscillation in thalamic relay neurones. *The Journal of Physiology* 431, 291–318.
- Roth, T. (2007). Insomnia: definition, prevalence, etiology, and consequences. *Journal of Clinical Sleep Medicine* 3, S7–S10.
- Schiff, N.D. (2008). Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Annals of the New York Academy of Sciences* 1129, 105–118.
- Sebel, P.S., Bowdle, T.A., Ghoneim, M.M., Rampil, I.J., Padilla, R.E., Gan, T.J., and

- Domino, K.B. (2004). The Incidence of Awareness During Anesthesia: A Multicenter United States Study. *Anesthesia & Analgesia* 99, 833–839.
- Steriade, M., Nunez, A., and Amzica, F. (1993). Intracellular analysis of relations between the slow (<1 Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. *Journal of Neuroscience* 13, 3266.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience* 5, 42.
- Vyazovskiy, V.V., Olcese, U., Hanlon, E.C., Nir, Y., Cirelli, C., and Tononi, G. (2011). Local sleep in awake rats. *Nature* 472, 443–447.
- Ying, S.-W., and Goldstein, P.A. (2005). Propofol-block of SK channels in reticular thalamic neurons enhances GABAergic inhibition in relay neurons. *Journal of Neurophysiology* 93, 1935–1948.