

COMMENT

## Debating how REM sleep is regulated (and by what)

JOEL H. BENINGTON

Department of Biology  
St Bonaventure University  
St Bonaventure, NY, 14778  
USA  
Tel.: +1 716 375 2564  
Fax: +1 716 375 7618  
E-mail: jbening@sbu.edu

Paul Franken has published in the March 2002 issue of this journal (pp. 17–28) an interesting article in which he re-analyzes two of his datasets to test a new model of rapid eye movement sleep (REMS) regulation. In Franken's model, REMS timing is regulated in the short term by the accumulation of a REMS propensity during non-rapid eye movement sleep (NREMS). This part of Franken's model builds on a publication of Benington and Heller (1994), which in turn was based on a number of studies showing that REMS timing is a NREMS dependent process (reviewed by Benington and Heller 1994). Additionally, Franken has posited the long-term accumulation of another form of REMS propensity in both waking and NREMS, which he feels is necessary to account for REMS rebounds following total sleep deprivation (TSD). In Franken's words, 'the differences in REMS during recovery cannot be explained by assuming that the need for REMS increases exclusively during NREMS.' This part of Franken's model is in effect rather similar to mainstream views of REMS regulation, according to which REMS expression is controlled by accumulation of a REMS propensity during waking.

In this commentary, I will discuss the strengths and weaknesses of interpreting the effects of TSD in terms either of a REMS propensity accumulating during waking, during NREMS, or as in Franken's model during both waking and NREMS. In my opinion, none of these models accounts for the available data in a decisive and entirely unproblematic way, yet all are at least broadly consistent with the available data and so remain viable models.

However, we should not restrict our consideration of this question to abstract analyses of such models. REMS, NREMS and waking are physiological states of the brain and body. The homeostatic regulation of REMS presumably involves physiological changes as a propensity for REMS accumulates and is discharged. As I will discuss later in this commentary, my own support for the idea that REMS is regulated by the accumulation of a REMS propensity in NREMS derives primarily from a consideration of the neurophysiological characteristics of waking, NREMS and REMS, not from purely phenomenological analyses of arousal state distributions.

### FRANKEN'S MODEL OF LONG-TERM REMS REGULATION

The finding that TSD produces a REMS rebound is generally interpreted as evidence that REMS is regulated by the accumulation of a REMS propensity during waking (reviewed by Benington 1992). The natural assumption is that REMS expression increases during the recovery period to compensate for the REMS lost during the deprivation period. As the deprivation period consists largely of waking, it is likewise natural to assume that the propensity for REMS accumulates during waking. In Franken's model, the long-term form of REMS propensity accumulates 'in the absence of REMS' – that is, in both waking and NREMS. As periods of TSD consist largely of waking, this model generates predictions that are similar but not identical to those of the more common waking-related model.

The strength of these two models is that they account in a straightforward and intuitive way for the occurrence of REMS rebounds following TSD. However, the REMS rebounds that do occur do not fully compensate (on a minute-for-minute basis) for the REMS lost during sleep deprivation. This in itself is not necessarily a problem, as the same is true also of NREMS rebounds following TSD. NREMS following sleep loss is thought to be 'deeper' and more intense than normal NREMS; hence, more efficient in discharging NREMS propensity, and increases in EEG slow-wave activity in recovery NREMS are a measure of the increased efficiency of recovery sleep.

Franken has correctly observed that REMS rebounds following TSD are, 'proportional to the loss incurred by that deprivation.' However, and this is an important point, REMS recovery is more 'complete' the longer the TSD is. In other words, short TSDs are followed by no REMS rebounds or REMS rebounds that recover only a very small fraction of the REMS lost. By contrast, longer TSDs produce REMS rebounds that recover a greater fraction of the REMS lost (reviewed by Benington and Heller 1994). These findings are not what one would expect given the hypothesis that REMS rebounds are homeostatic responses to the loss of REMS during TSD. Generally speaking, homeostatic responses

depart from linearity (producing less 'complete' rebounds) more markedly when the system is more vigorously perturbed, and this is, in fact, what happens in the case of NREMS rebounds following TSD.

To account for the nature of REMS rebounds as we have NREMS rebounds, we would have to suppose that (1) recovery REMS is more intense and therefore more efficient than baseline REMS and (2) the efficiency of the REMS recovery process *decreases* following longer TSDs. This latter supposition would be inconsistent with the general concept of homeostatic regulation, and I for one cannot think of any examples from nature of such a homeostatic control system. So while REMS rebounds following TSD are indeed, 'proportional to the loss incurred by that deprivation', the *nature* of that proportionality is difficult to reconcile with the idea that REMS propensity accumulates during the TSD period, either in waking or in both waking and NREMS.

### NREMS-RELATED MODELS OF REMS REGULATION

Benington and Heller have suggested that both REMS homeostasis and the timing of REMS episodes within sleep periods can be neatly explained in terms of accumulation of REMS propensity in NREMS (Benington and Heller 1994). One objection that we anticipated was the observation that TSD produces REMS rebounds, an observation that (as noted above) is commonly taken as evidence for accumulation of REMS propensity in waking. We explained this observation by noting that both NREMS and NREMS-like neuronal activity in waking commonly occur during TSD, and that this should lead to some NREMS-dependent accumulation of REMS propensity during the deprivation period that would not occur were the sleep deprivation truly 'total'. This hypothesis explains the observation that REMS rebounds in both humans and experimental animals are negligible or absent following shorter TSDs, and become more marked the longer the TSD gets, as it is more difficult to produce a truly 'total' TSD as the experimental subject becomes increasingly sleepy.

Franken challenges this explanation by noting that the amplitude of electroencephalogram (EEG) slow waves in drowsy waking during TSD is low compared with that of the EEG slow waves that occur in normal NREMS. If we were to assume that the accumulation of REMS propensity during NREMS is always a direct function of the magnitude of EEG slow-wave activity, then we would indeed predict the accumulation of very little REMS propensity during drowsy waking in the deprivation period. But the existing evidence argues *against* a more rapid accumulation of REMS propensity in 'deeper' NREMS, in association with higher levels of EEG slow-wave activity (reviewed by Benington and Heller 1994). Moreover, the evidence for a quantitative link between EEG slow-wave activity and the rate of the NREMS recovery process (and perhaps also the rate of accumulation of REMS propensity) applies only to normal, undisturbed NREMS. The administration of benzodiazepine hypnotics, for example, reduces

EEG slow-wave activity in NREMS and yet appears to produce normal or supranormal discharge of sleep propensity, suggesting that when the system is experimentally perturbed the level of EEG slow-wave activity is no longer a reliable measure of the rate of the NREMS recovery process (Achermann and Borbely 1987; Borbely *et al.* 1983, 1985). So measured levels of EEG slow-wave activity during drowsy waking may underestimate the degree to which this NREMS-like activity in waking is functionally equivalent to normal NREMS.

Franken considers this possibility, but finding that EEG-based measures do not enable us, 'to quantitatively predict the different rebounds in REMS within the framework of the hypothesis that REMS need only accumulates during NREMS', he concludes that, 'the most parsimonious hypothesis to explain the differences in REMS during recovery is to assume that REMS need increases in its absence and that this 'need' and the subsequent compensatory response during recovery is therefore higher after 24-h SD as compared with 12-h SD.' I agree that a good quantitative measure of the rate of accumulation and discharge of REMS propensity associated with NREMS and NREMS-like activity would further strengthen the case for a NREMS-related model of REMS regulation. But the absence of such a measure is not in itself an argument against that hypothesis. The history of biology is replete with examples of hypotheses that could not at first be verified on a physiological level but that nevertheless proved to be correct.

Moreover, this aspect of our hypothesis is in fact testable. If the REMS rebound following TSD results not from waking but from short bouts of NREMS and NREMS-like activity, then a TSD protocol that maintains a higher level of vigilance in the experimental animals should produce a smaller REMS rebound than one that permits more drowsy waking. One could compare the effects of TSD produced by constant locomotor activity vs. a sleep deprivation paradigm that deliberately permits short bouts of NREMS before animals are aroused. According to Franken's model there should be no difference in REMS rebound between these two conditions. If REMS propensity accumulates only during waking, then one would expect if anything a slightly smaller REMS rebound in the latter condition. Our model would on the other hand predict a greater REMS rebound when some amount of NREMS is allowed to occur during the deprivation period.

### PHYSIOLOGICAL IMPLICATIONS

In sum, my contention is that the effects of TSD on REMS expression can be accounted for by any of the three models of REMS regulation under consideration. Certain assumptions must be made in order to reconcile any of the three models with all available observations. The issue is not therefore whether phenomena *can* be explained by a given model, but whether the explanations of one model are more intuitively satisfying than the explanations of another.

However, this conclusion is based on a purely abstract analysis of these models. If we are going to seriously investigate the functions of NREMS and REMS, we should not treat them as abstract entities but should consider what they are representative of on the cellular level. I will briefly address this issue in the remainder of the commentary. For a more detailed consideration, and for references to the primary literature, consult Benington and Heller (1994).

Waking and REMS are neurophysiologically quite similar, while NREMS and REMS are much more different. In both waking and REMS, release of acetylcholine throughout the forebrain produces an activated state in which neurons are tonically depolarized and thus highly responsive. In NREMS, lower levels of acetylcholine release result in hyperpolarization and produce synchronized rhythmic bursting in neurons in many brain regions.

The one noteworthy difference between CNS activity in waking vs. REMS is that serotonin and norepinephrine release is high in waking and virtually absent in REMS. One waking-related hypothesis of REMS function has been based on this difference (Siegel and Rogawski 1988). Otherwise, waking-related models of REMS regulation have neglected to identify the specific neurophysiological attributes of REMS as distinct from waking that cause a propensity to accumulate in waking and be discharged in REMS.

By contrast, we have hypothesized that REMS reverses some consequence of the synchronized bursting activity that occurs in NREMS (Benington and Heller 1994). We have argued that, 'the ubiquity of REMS in eutherian mammals suggests that the restorative properties of NREMS directly produce accumulation of REM-sleep propensity.' In other words, if the restorative properties of NREMS are in fact linked to synchronized bursting activity (still an open question, by the way), then the accumulation of REMS propensity in NREMS should likewise be so linked. In another publication, Benington *et al.* (1995) further speculated that accumulation of REMS propensity in NREMS may be tied to increased neuronal potassium conductance associated with the synchronized bursting firing mode. Increased potassium conductance is of course just one concomitant of the distinctive pattern of neuronal activity in NREMS, but this hypothesis is an example of how specific we need to be in seeking the physiological basis of what now we can only abstractly refer to as 'REMS propensity'.

When neurophysiological considerations do not enter into hypothesis formation, and arousal states and 'propensities' are treated as mere abstract entities, conclusions are arrived at that are in my opinion physiologically implausible. Franken has based his model on the assumption that REMS propensity accumulates simply in the absence of REMS. Abstractly, this may indeed seem a plausible and even conservative assumption, but it is hard to imagine how two states as different physiologically as waking and NREMS could equally result in the accumulation of a propensity for REMS. The more common assumption that REMS propensity accumulates during waking is equally incomprehensible unless one focuses

on release of norepinephrine or serotonin, as Siegel and Rogawski have, or until someone articulates another plausible physiological difference between REMS and waking that could make sense of this assumption.

## CONCLUSIONS

Ultimately, these issues will not be resolved until the function of REMS is determined. At that point, we will know the physiological substrate of REMS propensity, and there will no longer be any question whether that propensity accumulates during waking, during NREMS, or merely in the absence of REMS. The current debate is of value primarily insofar as it provides suggestive guidance in our search for the specific function of REMS.

This issue is no mere intellectual exercise. A NREMS-related model of REMS regulation naturally suggests a different range of possible functions for REMS than does a waking-related model or Franken's model. In rejecting any one of these models, we thereby dismiss from our thinking any of the potential functions for REMS that entail such a model. My own conviction that REMS is regulated in relation to NREMS necessarily limits the hypothetical functions for REMS that I am likely to give serious consideration to. If I am correct, then I have adroitly focused my search – if not, then I have seriously handicapped myself. I therefore counsel readers to be circumspect as they attempt to resolve this matter in their own thinking, or at least to keep an open mind.

## REFERENCES

- Achermann, P. and Borbely, A. A. Dynamics of EEG slow wave activity during physiological sleep and after administration of benzodiazepine hypnotics. *Human Neurobiol.*, 1987, 6: 203–210.
- Benington, J. H. *REM-sleep Homeostasis in the Rat*. Department of Biological Sciences, Stanford University, Stanford, CA, 1992.
- Benington, J. H. and Heller, H. C. Does the function of REM sleep concern non-REM sleep or waking? *Prog. Neurobiol.*, 1994, 44: 433–449.
- Benington, J. H., Woudenberg, M. C. and Heller, H. C. Apamin, a selective SK potassium channel blocker, suppresses REM sleep without a compensatory rebound. *Brain Res.*, 1995, 692: 86–92.
- Borbely, A. A., Mattmann, P., Loepfe, M., Fellmann, I., Gerne, M., Strauch, I. and Lehmann, D. A single dose of benzodiazepine hypnotics alters the sleep EEG in the subsequent drug-free night. *Eur. J. Pharmacol.*, 1983, 89: 157–161.
- Borbely, A. A., Mattmann, P., Loepfe, M., Strauch, I. and Lehmann, D. Effect of benzodiazepine hypnotics on all-night sleep EEG spectra. *Human Neurobiol.*, 1985, 4: 189–194.
- Siegel, J. M. and Rogawski, M. A. A function for REM sleep: regulation of noradrenergic receptor sensitivity. *Brain Res. Rev.*, 1988, 13: 213–233.

## RESPONSE BY PAUL FRANKEN

The objective of my article was to make the case that two processes regulating REMS should be considered. It was an effort to quantify their dynamics in the rat and to model sleep architecture before and after a sleep deprivation based on how

these two processes might interact. The intent was not to disprove the Benington and Heller (1994b) hypothesis. However, as the concepts underlying those two processes are not novel and have long been recognized (they may indeed be referred to as 'mainstream' and 'conservative'), my interpretations, by definition, are inconsistent with the interpretations the alternative Benington and Heller model offer. Benington's rebuttal places me in the awkward situation of having to defend a mainstream view on REMS regulation against a hypothesis that did not gain much, if any, momentum in terms of supportive data or advocates since its inception almost 8 years ago. Part of this controversy regarding the regulation of REM sleep has already been addressed by the journal *SLEEP* (Rechtschaffen and Bergmann 1999 vs. Benington and Heller 1999). I will not repeat the issues raised there and only address the main points Benington raises in his rebuttal to my manuscript.

Benington states that 'the issue is not whether phenomena *can* be explained by a given model, but whether the explanations of one model are more intuitively satisfying than the explanations of another'. 'Intuitively satisfying' is a criterion that works best in the absence of data that clearly support or refute a model. In my view, his model does not even meet this criterion. According to the Benington and Heller model, pressure for REMS accumulates exclusively in NREMS. However, to account for the first and most obvious sign of a homeostatic regulation of REMS, i.e. the REMS rebound after a sleep deprivation, an exception already has to be made; now REMS pressure does increase also during wakefulness. Furthermore, the central part of his thesis that 'REMS reverses some consequence of the synchronized bursting activity that occurs in NREMS' is counter-intuitive as variations in synchronized activity in NREMS, quantified as EEG delta power, do not correlate (or correlate negatively) with the expression of REMS. So in contrast to Benington, I clearly have a more 'conservative' point-of-view and put more emphasis on the 'phenomenology'. Any model concerning REMS regulation should first address and explain the observed data before any inferences can be made on its neuro-physiological substrate or even function.

Two earlier studies that were cited in my manuscript tested the predictions of Benington and Heller hypothesis in the rat. Endo *et al.* (1997) measured the rebound in REMS after four different protocols. All four protocols were preceded by a 12-h total sleep deprivation (SD) that was followed either by (1) recovery sleep, (2) an additional 4-h total SD (abolishing both NREMS and REMS), (3) a 4-h slow-wave SD (mainly affecting NREMS), or (4) a 4-h selective REMS SD (only marginally affecting NREMS and almost abolishing REMS). The different amounts of NREMS obtained in the respective protocols could not predict the REMS rebound during recovery sleep and the absence of REMS during the protocol seemed to be the crucial factor. Similarly, Ocampo-Garcés *et al.* (2000) studied sleep in rats under three different protocols. In all three, REMS was deprived for 3 h, but the protocols differed in the amount of NREMS that was allowed

within these 3 h. Again the rebound in REMS during recovery sleep did not depend on the amount of NREMS. The conclusion of both studies was that the build-up of REMS pressure and the subsequent rebound in REMS was primarily related to the loss of REMS. Together with the data presented in my manuscript, these results clearly show that REMS expression can notably dissociate from the expression of NREMS even in shorter-term SDs in which a large NREMS propensity does not accumulate and therefore rule out a potential contribution of 'drowsy' wakefulness to REMS propensity. The notion that REMS can dissociate from NREMS and thus is not a neuro-physiological consequence of NREMS is further underscored by genetic studies of sleep where different genetic factors seem to underlie NREMS and REMS expression (e.g. Tafti *et al.* 1997), by pharmacological studies of sleep that only affect one sleep state but not the other (e.g. Landolt *et al.* 2001), or neuro-anatomical studies where specific lesions affect only NREMS expression but not that of REMS (e.g. Lu *et al.* 2000).

Benington states that short total SDs are followed by no or very small REMS rebounds whereas longer total SDs produce more complete rebounds and concludes that these findings are difficult to reconcile with the idea that REMS propensity accumulates during wakefulness and NREMS. He refers to the homeostatic regulation of NREMS for comparison. First, short SDs *can* elicit a substantial REMS rebound. Ocampo-Garcés *et al.* (2000) showed that of the REMS lost during a 3-h SD, 51% was already compensated within the 3 h for which recovery sleep was measured. Second, it is plausible to assume that under undisturbed laboratory conditions not all REMS expressed is 'needed' (obligate vs. facultative). Therefore, to elicit a significant homeostatic response, a larger REMS deficit might be required. Third, one cannot study REMS in isolation as NREMS and REMS can compete for expression. In the rat the dynamics of NREMS propensity seem to be faster than those of the long-term REMS process. Therefore, a shorter SD can already result in a (relatively) large NREMS pressure that can suppress or delay REMS recovery. Fourth, one cannot compare the dynamics of the regulation of NREMS with those of REMS. The homeostatic regulation of NREMS is best reflected by EEG delta power, an intensity measure. Clearly, there are neuro-physiological limits to how much delta power can be produced per unit of time in response to a SD. For the rat this limit seems to be reached after SDs between 12 and 24 h. Unlike NREMS there is no intensity marker for REMS and lost REMS is primarily compensated by increased REMS duration. Limits to the increase in REMS duration are not yet reached even after 4 days of SD (Rechtschaffen *et al.* 1999).

Benington finds it 'hard to imagine how two states as different physiologically as waking and NREMS could equally result in the accumulation of a propensity for REMS.' This is not a convincing argument, which only makes sense if *a priori* it is assumed that REMS propensity increases as a function of NREMS associated synchronized burst-pause firing patterns in thalamocortical neurons. There is more to sleep than

thalamocortical activity. Based on the available data, others and I argue that a need for REMS increases in the absence of REMS from which it follows that the important function(s) that REMS subserves, can only be fulfilled during REMS. For comparison, propensity for feeding behavior (i.e. hunger), which is also influenced by a strong homeostatic component, can also be expected to increase irrespective of behavioral state (although the increase rates might vary, an aspect that has not yet been addressed for the long-term regulation of REMS).

Finally, I want to emphasize that the Benington and Heller model was fully incorporated into my model to describe the short-term REMS process. There is ample evidence suggesting that the timing of REMS during a sleep episode (i.e. the NREM-REMS cycle) is regulated by a NREMS-dependent process (reviewed in Benington and Heller 1994b). Vivaldi *et al.* (1994) referred to this process as a short-term homeostatic process, a term I adopted here. Based on the analysis of sleep architecture Benington initially concluded that it was the timing of REMS that was controlled by a NREMS dependent accumulation of REMS propensity (Benington *et al.* 1994a). It is, however, hard to envision how one and the same process can drive a cycle in the range of minutes and at the same time can regulate the amount of REMS expressed on a daily basis. Nevertheless, the analyses I presented was built on Benington's findings. His analytical tools to identify REMS attempts (NRTs or PREs) were instrumental in determining how the long- and short-term processes interact to shape sleep architecture. Perhaps we should not limit our thinking by trying to identify the function of REMS. It is still possible that for the short-term NREMS dependent process REMS is needed to reverse some effects of 'synchronized burst-pause firing' and the mere presence of REMS might suffice. However, the amount of REMS is clearly regulated independent of NREMS and is therefore likely to depend on other factors.

Paul Franken, PhD  
 Department of Biological Sciences  
 Stanford University

## REFERENCES

- Benington, J. H. and Heller, H. C. REMS timing is controlled homeostatically by accumulation of REMS propensity in non-REMS. *Am. J. Physiol.*, 1994a, 266: R1992-R2000.
- Benington, J. H. and Heller, H. C. Does the function of REMS concern non-REMS or waking? *Prog. Neurobiol.*, 1994b, 44: 433-449.
- Benington, J. H. and Heller, H. C. Implications of sleep deprivation experiments for our understanding of sleep homeostasis. *Sleep*, 1999, 22: 1033-1037.
- Endo, T., Schwierin, B., Borbély, A. A. and Tobler, I. Selective and total sleep deprivation: effect on the sleep EEG in the rat. *Psychiatry Res.*, 1997, 66: 97-110.
- Landolt, H. P., Raimo, E. B., Schnierow, B. J., Kelsoe, J. R., Rapaport, M. H. and Gillin J. C. Sleep and sleep electroencephalogram (EEG) in depressed patients treated with phenelzine. *Arch. Gen. Psychiatry*, 2001, 58: 268-276.
- Lu, J., Greco, M. A., Shiromani, P. and Saper, C. B. Effect of lesions of the ventrolateral preoptic nucleus on NREM and REM sleep. *J. Neurosci.*, 2000, 20: 3830-3842.
- Ocampo-Garcés, A., Molina, E., Rodríguez, A. and Vivaldi, E. A. Homeostasis of REMS after total and selective sleep deprivation in the rat. *J. Neurophysiol.*, 2000, 84: 2699-2702.
- Rechtschaffen, A. and Bergmann, B. M. Sleep rebounds and their implications for sleep stage substrates: a response to Benington and Heller. *Sleep*, 1999, 22: 1038-1043.
- Rechtschaffen, A., Bergmann, B. M., Gilliland, M. A. and Bauer, K. Effects of method, duration, and sleep stage on rebounds from sleep deprivation in the rat. *Sleep*, 1999, 22: 11-31.
- Tafti, M., Franken, P., Kitahama, K., Malafosse, A., Jouvét, M. and Valatx J. L. Localization of candidate genomic regions influencing paradoxical sleep in mice. *Neuroreport*, 1997, 8: 3755-3758.
- Vivaldi, E. A., Ocampo, A., Wyneken, U., Roncagliolo, M. and Zapata, A. M. Short-term homeostasis of active sleep and the architecture of sleep in the rat. *J. Neurophysiol.*, 1994, 72: 1745-1755.