Selective REM sleep deprivation in humans: effects on sleep and sleep EEG

TAKURO ENDO, CORINNE ROTH, HANS-PETER LANDOLT, ESTHER WERTH, DANIEL AESCHBACH, PETER ACHERMANN, AND ALEXANDER A. BORBELY
Institute of Pharmacology, University of Zürich, CH-8057 Zürich, Switzerland

Endo, Takuro, Corinne Roth, Hans-Peter Landolt, Esther Werth, Daniel Aeschbach, Peter Achermann, and Alexander A. Borbély. Selective REM sleep deprivation in humans: effects on sleep and sleep EEG. Am. J. Physiol. 274 (Regulatory Integrative Comp. Physiol. 43): R1186–R1194, 1998.—To investigate rapid eye movement (REM) sleep regulation, eight healthy young men were deprived of REM sleep for three consecutive nights. In a three-night control sleep deprivation (CD) session 2 wk later, the subjects were repeatedly awakened from non-REM sleep in an attempt to match the awakenings during the REM sleep deprivation (RD) nights. During the RD nights the number of sleep interruptions required to prevent REM sleep increased within and across consecutive nights. REM sleep was reduced to 9.2% of baseline (CD nights: 80.7%) and rose to 140.1% in the first recovery night. RD gave rise to changes in the EEG power spectra of REM sleep. Power in the 8.25- to 11-Hz range was reduced in the first recovery night, an effect that gradually subsided but was still present in the third recovery night. The rising REM sleep propensity, as reflected by the increase of interventions within and across RD nights, and the moderate REM sleep rebound during recovery can be accounted for by a compensatory response that serves REM sleep homeostasis. The changes in the electroencephalogram power spectra, which were observed during enhanced REM sleep propensity, may be a sign of an altered quality of REM sleep.

rapid eye movement sleep, non-rapid eye movement sleep; sleep homeostasis; rectal temperature

Sleep consists of the two substates non-rapid eye movement (NREM) sleep and REM sleep, which cyclically alternate with a periodicity of 90–100 min (20). The ratio of NREM sleep to REM sleep within the sleep cycle undergoes a prominent circadian modulation. REM sleep episodes are longest at or shortly after the circadian phase of minimal core body temperature (17, 22, 24, 40).

Advances have been made in elucidating the dynamics of NREM sleep regulation. An important feature is that not only the duration of this sleep state but also its intensity can vary. The response to a sleep deficit consists in an intensification rather than a prolongation of NREM sleep. Electroencephalogram (EEG) slow-wave activity (SWA; spectral power in the 0.75- to 4.5-Hz range) is a reliable marker of NREM sleep intensity. Thus SWA in recovery sleep after prolonged waking is increased (12, 21), whereas in sleep preceded by a daytime nap it is decreased (26, 43). Because the level of SWA is determined by the duration of prior waking and sleep, it can serve to delineate the homeostatic process regulating NREM sleep. The dynamics of SWA in various experimental protocols have been successfully simulated (1, 43).

Sleep propensity is not only determined by prior waking and sleep but also by a circadian process. In the two-process model of sleep regulation, the variations of sleep propensity during waking, NREM sleep intensity during sleep, as well as the timing of sleep and waking are specified by the interaction of the sleep-waking-dependent homeostatic process S and the sleep-waking-independent circadian process C (9, 10, 18). Major tenets of the model, such as the dependence of SWA on the history of sleep and waking rather than on the circadian phase and the control of the REM sleep fraction of the sleep cycle by the combined action of a circadian and a sleep-dependent process, were recently confirmed by the analysis of experimental data obtained in a forced desynchrony protocol (22).

In contrast to NREM sleep, the regulation of REM sleep has not yet been extensively described by mathematical models. The NREM-REM sleep cycle was simulated on the basis of Lottka-Volterra and limit cycle oscillators (34, 35). REM sleep has been also introduced into elaborate versions of the two-process model as a data-derived external parameter (1). However, the homeostatic facet of REM sleep regulation was not included in any of these models. The reason is not only our insufficient understanding of the underlying mechanisms but also the limitations and inconsistencies in the experimental data. The complexity of the regulatory processes is evident from the observation that a selective REM sleep deprivation is not followed invariably by an REM sleep rebound (16). Likewise, a moderate extension of waking does not cause an increase in REM sleep, and only after a prolonged total sleep deprivation >-60 h in humans: 64 h (44), 108 h (8), 205 h (31); >-12 h in the rat (11, 13, 25, 28, 33, 39) is an REM sleep rebound elicited. The lack of response to shorter deprivation periods seems to be partly a consequence of the inhibitory action of the enhanced NREM sleep propensity (i.e., increased level of process S). Partial sleep deprivation protocols in humans, which induce only a minor NREM sleep deficit (i.e., sleep restriction to the first 4 h of the habitual sleep episode), result in an immediate and consistent REM sleep rebound (14, 15). Moreover, if REM sleep is prevented during the first 5 h of sleep, a time interval in which NREM sleep intensity can largely dissipate, an REM sleep rebound is observed in the final, undisturbed segment of sleep (5). Taken together, the results indicate that REM sleep is indeed regulated by sensitive homeostatic mechanisms, but that the interference by other factors renders their investigation difficult.
One of the most effective challenges to REM sleep regulation is the selective deprivation of this sleep state. Classical REM sleep deprivation studies in humans (2, 19, 30, 37) as well as old and recent studies in animals (7, 11, 25) have shown that the rising "REM sleep pressure" is reflected in the increasing frequency of interventions (within and across nights) required to prevent the occurrence of REM sleep episodes. The aim of the present experiment was to make use of this powerful protocol for gaining insights into the regulatory mechanisms. In the selective REM sleep deprivation nights, care was taken to prevent REM sleep as thoroughly as possible by intervening at the earliest REM sleep signs. The subjects were kept awake for a sufficient time to avoid an immediate relapse into REM sleep while keeping the waking episodes short enough to allow frequent interventions. To control for the effect of the awakenings per se, the protocol included a separate control sleep deprivation session in which the subjects were repeatedly awakened from NREM sleep. The present paper focuses on the time course of the interventions as well as on the effect of selective sleep deprivation on sleep stages and EEG spectra. Future reports will deal more specifically with the NREM sleep deprivation and the periodic occurrence of REM sleep. The first NREM sleep episode (i.e., the interval with the highest amount of SWA) was left undisturbed. The awakenings were started in the second episode after at least 2 min of stage 2 and continued in the subsequent NREM sleep episodes. After each intervention the subjects were allowed to have at least 2 min of stage 2 before the next intervention. If an REM sleep episode was expected (i.e., >60 min of NREM sleep or a reduction of the tonic EMG amplitude), the time between interventions was increased.

Involuntary napping or microsleep after the third deprivation night was precluded by keeping the subjects continuously in the company of an experimenter until bedtime of the first recovery night.

Polygraphic Recordings and Rectal Temperature

During sleep episodes, the EEG (derivations: C3A2 and F3, P3, O1, F4, C4, P4, O2 and Oz against C3), submental EMG, electrooculogram (EOG), electrocardiogram (ECG), and rectal temperature were recorded by a polygraphic amplifier (PSA24; Braintronics, Almere, The Netherlands), digitized, and transmitted via fiber-optic cables to a personal computer (PC). The EOG was recorded by a bipolar derivation that even enabled detection of small eye movements. EEG, EMG, and EOG signals were conditioned by the following analog filters: a high-pass filter (−3 dB at 0.16 Hz), a low-pass filter (−3 dB at 102 Hz, −10 Hz at 256 Hz), and a notch filter (50 Hz). Data were sampled with a frequency of 512 Hz, digitally filtered [EEG and EOG: low-pass finite impulse response (FIR) filter, −3 dB at 49 Hz; EMG: band-pass FIR filter, −3 dB points at 15.6 and 54 Hz], and stored on the hard disk of a PC with a resolution of 128 Hz. The EEG data were subjected to a fast Fourier transform routine. The epoch length was 4 s, and

**METHODS**

**Subjects and Protocol**

Eight right-handed male subjects (mean age 24.1 ± 0.6 yr), recruited from the student population of the University of Zürich and the Swiss Federal Institute of Technology, were paid for their participation in the study. Written informed consent from the subjects and approval from the local ethical committee for research on human subjects were obtained before the study. The medical history, current health status, and subjective sleep quality were assessed by questionnaires. Only subjects who reported good health, no serious medical history, no medication intake, and no sleep disturbances were included. A screening night served to exclude subjects with sleep apnea, nocturnal myoclonus, or a sleep efficiency of <80%. All subjects were nonsmokers, and their habitual alcohol intake was less than five drinks per week (mean 1.8 ± 0.5 glasses/wk). They reported a habitual consumption of zero to three caffeinated beverages per day (mean 1.3 ± 0.3 cups/day). In the week preceding the study as well as during the study, the subjects were asked to abstain from alcohol and caffeinated beverages, to schedule their sleep from 2300 to 0700, and to refrain from napping. Compliance with these instructions was verified by continuous recording of wrist activity, by assaying caffeine in saliva samples, and measuring breath-ethanol concentration before bedtime.

The study consisted of two sessions of nine consecutive nights. The first session contained three REM sleep deprivation (RD) nights and the second session contained three control sleep deprivation (CD) nights. The two sessions were at least 2 wk apart, and the subjects were unaware of the type of deprivation. Each session consisted of an adaptation night, two baseline nights (B1, B2), three RD or CD nights (D1, D2, D3), and three recovery nights (R1, R2, R3). During these 18 nights, the subjects slept from 2300 to 0700 in the completely dark bedrooms of the sleep laboratory and were exposed to a continuous background level of white noise (38 ± 0.5 dB) to mask any sporadic noise.

During the RD nights, the subjects were awakened at the first sign of REM sleep by switching on a dim red light (<5 lx) and entering their room. The criteria for interventions were a 20-s epoch of desynchronized EEG without spindles or K complexes and the concomitant reduction of the tonic electromyogram (EMG) amplitude, regardless of the occurrence of REMs. REMs could not be used as a criterion because the scoring rules of Rechtschaffen and Kales (36) necessitate the retroactive scoring of REM sleep on their occurrence. This would have precluded an effective RD. After awakening from REM sleep, the subjects were required to sit up in bed for 2 min and complete a questionnaire and visual analog scales (VAS) and perform a letter cancellation task. The questionnaire assessed dreams (present/absent; pleasant/unpleasant) and the 100-mm VAS evaluated alertness (tired-recuperated) and mood (good mood-bad mood).

A 2-min awakening period was selected on the basis of a pilot experiment that showed that with a shorter awakening period the subject tended to sleep immediately into REM sleep. The longer awakening period carried the risk of excessive NREM sleep deprivation. The low level of red light was used to minimize light-induced effects on the circadian system. The CD session served as a control for the repeated awakenings. During the CD nights, the subjects were awakened from NREM sleep and kept awake for 2 min using the same experimental protocol as for RD. An attempt was made to match for each individual and night the number of awakenings as well as the duration of stages with high vigilance [i.e., waking (W) + stage 1 (S1) + movement time (MT)] to the corresponding RD night. It was not feasible to match also the pattern of increasing interventions within the night. In the CD session, care was taken to minimize interference with SWA and the periodic occurrence of REM sleep. The first NREM sleep episode (i.e., the interval with the highest amount of SWA) was left undisturbed. The awakenings were started in the second episode after at least 2 min of stage 2 and continued in the subsequent NREM sleep episodes. After each intervention the subjects were allowed to have at least 2 min of stage 2 before the next intervention. If an REM sleep episode was expected (i.e., >60 min of NREM sleep or a reduction of the tonic EMG amplitude), the time between interventions was increased.

During sleep episodes, the EEG (derivations: C3A2 and F3, P3, O1, F4, C4, P4, O2 and Oz against C3), submental EMG, electrooculogram (EOG), electrocardiogram (ECG), and rectal temperature were recorded by a polygraphic amplifier (PSA24; Braintronics, Almere, The Netherlands), digitized, and transmitted via fiber-optic cables to a personal computer (PC). The EOG was recorded by a bipolar derivation that even enabled detection of small eye movements. EEG, EMG, and EOG signals were conditioned by the following analog filters: a high-pass filter (−3 dB at 0.16 Hz), a low-pass filter (−3 dB at 102 Hz, −10 Hz at 256 Hz), and a notch filter (50 Hz). Data were sampled with a frequency of 512 Hz, digitally filtered [EEG and EOG: low-pass finite impulse response (FIR) filter, −3 dB at 49 Hz; EMG: band-pass FIR filter, −3 dB points at 15.6 and 54 Hz], and stored on the hard disk of a PC with a resolution of 128 Hz. The EEG data were subjected to a fast Fourier transform routine. The epoch length was 4 s, and
a 10% tapered-cosine window was applied. Data reduction of power spectra was achieved by omitting values $>25$ Hz and by collapsing two (0.25–5 Hz) or four (5.25–25 Hz) adjacent 0.25-Hz bins. Five consecutive spectra were then averaged to yield a 20-s spectrum after excluding 4-s epochs contaminated by artifacts. Vigilance states were visually scored for 20-s epochs according to standard criteria (36). One value of rectal temperature was stored every 20 s.

REM Density in REM Sleep

REMs in REM sleep were visually scored when an EOG deflection unrelated to an EEG event showed a peak-to-peak amplitude of $\geq 75$ $\mu$V and a slope of $\geq 200$ $\mu$V/s (adapted from Ref. 43). REM density was calculated as the percentage of 4-s epochs with REMs per total number of 4-s epochs in REM sleep.

Subjective Sleep Quality and Assessment of Mood and Alertness During the Day

In the morning, 15 min after lights on, subjective sleep quality and self-rated state were evaluated by a questionnaire and visual analog scales. Throughout the experimental days, subjects rated their subjective mood and alertness at 2-h intervals between 0800 and 2200 h on a 20-point scale of a handheld computer (PSION Organizer II).

Statistics

The data were subjected to an ANOVA for repeated measures (rANOVA) using the SAS general linear model procedure with Greenhouse-Geisser correction. Contrasts were tested by two-tailed paired $t$-tests. If not otherwise mentioned, the significance level is $P < 0.05$. To approximate a normal distribution, sleep latencies and absolute EEG power density values were log transformed, and subjective ratings (mood and alertness) were $z$-transformed before the statistical tests. Sleep variables, as well as EEG power density values of the two baseline nights in each session, were averaged to compare the baseline mean (BL) with deprivation or recovery nights.

The general procedure of performing rANOVA is exemplified on the data of Fig. 4. To test for effects between the two treatments, a three-way ANOVA with the factors “condition” (RD, CD), “night” (D1, D2, D3), and “cycle” (1–4) was used. If the factor condition was significant, a two-way rANOVA was performed with the factors condition (e.g., R1 of RD and CD) and cycle (1–4) for each night, followed by a paired $t$-test for the corresponding cycles.

To test for effects within a condition, a two-way rANOVA with the factors night and cycle (1–4) was performed. The factor night contained BL, R1, R2, and R3 or R1, R2, and R3, depending on whether the effects were compared with baseline or only within the recovery nights. If the factor night was significant, the cycles of corresponding nights were tested by a two-tailed paired $t$-test.

RESULTS

Figure 1 illustrates sleep profiles and SWA of one subject for a BL night, an RD night, and a CD night. In the RD night, REM sleep was successfully reduced by the repeated awakenings. The frequency of interventions, which is evident from the waking epochs, shows an increasing trend during the night. Moreover, distinct clusters of awakenings are separated by intervals with consolidated NREM sleep. In the CD night, sleep was interrupted during NREM sleep. Note that the periodic recurrence of REM sleep was preserved, and SWA was suppressed.

Number of Interventions

Over the three RD nights, the number of interventions increased (Fig. 2, numeric values are indicated). A
three-way rANOVA revealed significant effects (P < 0.01) for night (D1–D3) and "2-h interval" (1–4), but not for condition (RD, CD) (P > 0.05). Within each RD night a significant increasing trend was present when the consecutive 2-h intervals were compared (P < 0.05). In the third RD night, the interventions in all four 2-h intervals were significantly more frequent than in the first RD night. Quantification of the rising trend by a linear regression analysis revealed that the slope of the regression line was significantly steeper (P < 0.001) in D2 (2.46 interventions/h, r² = 0.82) and D3 (2.61, r² = 0.80) than in D1 (1.64, r² = 0.58), whereas the intercepts did not differ significantly.

**Effectiveness of Selective REM Sleep Deprivation**

Table 1 summarizes the distribution of the vigilance states during the baseline, deprivation, and recovery nights. The selective RD protocol was successful, because REM sleep was reduced from 103.7 to 9.5 min (mean of baseline and deprivation nights, respectively). In contrast, CD induced only a moderate reduction of REM sleep (from 107.9 to 87.1 min). Whereas the number of awakenings did not differ significantly between the two conditions, the time spent in stages with high vigilance (i.e., W + S1 + MT) was higher in RD.

**Sleep Variables Derived from Visual Scoring**

Baseline nights. None of the variables differed significantly between B1 and B2 (comparisons within and between conditions) (Table 1).

Deprivation nights. Total sleep time was longer and waking after sleep onset was shorter for CD than RD. Significant condition effects were present for stages 3 and 4 separately and combined [i.e., slow wave sleep (SWS); data not shown], as well as for REM sleep (2-way rANOVA with factors condition and night). Contrasts revealed for some CD nights significantly less stages 3 and 4 and for all CD nights more REM sleep than in RD.

Table 1. Sleep variables and REM density of baseline, deprivation, and recovery nights in the RD and CD sessions

<table>
<thead>
<tr>
<th>Condition</th>
<th>B1</th>
<th>B2</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency to stage 1</td>
<td>RD</td>
<td>6.5 ± 2.0</td>
<td>5.2 ± 1.1</td>
<td>9.2 ± 2.4</td>
<td>5.0 ± 1.1</td>
<td>3.8 ± 1.1</td>
<td>3.8 ± 1.1</td>
<td>4.0 ± 0.7</td>
</tr>
<tr>
<td>Latency to stage 2</td>
<td>RD</td>
<td>4.6 ± 1.5</td>
<td>7.2 ± 2.3</td>
<td>8.3 ± 2.8</td>
<td>3.5 ± 0.9</td>
<td>3.2 ± 1.1</td>
<td>4.1 ± 2.0</td>
<td>4.2 ± 1.5</td>
</tr>
<tr>
<td>REM sleep latency</td>
<td>CD</td>
<td>9.4 ± 3.1</td>
<td>7.9 ± 1.4</td>
<td>12.7 ± 2.4</td>
<td>7.0 ± 1.3</td>
<td>7.0 ± 1.8</td>
<td>6.4 ± 1.1</td>
<td>7.2 ± 1.1</td>
</tr>
<tr>
<td>Total sleep time</td>
<td>RD</td>
<td>7.6 ± 1.6</td>
<td>10.0 ± 2.3</td>
<td>11.3 ± 3.2</td>
<td>6.0 ± 1.1</td>
<td>5.3 ± 1.2</td>
<td>6.4 ± 2.2</td>
<td>7.6 ± 1.8</td>
</tr>
<tr>
<td>Total sleep time</td>
<td>CD</td>
<td>71.4 ± 6.5</td>
<td>68.0 ± 4.0</td>
<td>70.0 ± 4.2</td>
<td>61.1 ± 4.4</td>
<td>62.9 ± 3.9</td>
<td>51.8 ± 7.7</td>
<td>52.4 ± 8.0</td>
</tr>
<tr>
<td>Total sleep time</td>
<td>CD</td>
<td>71.2 ± 10.3</td>
<td>67.9 ± 4.5</td>
<td>63.1 ± 7.3</td>
<td>83.4 ± 18.5</td>
<td>89.7 ± 19.8</td>
<td>64.5 ± 4.6</td>
<td>67.4 ± 4.4</td>
</tr>
</tbody>
</table>

Values are means ± SE (n = 8) in minutes for sleep variables and percent for rapid eye movement (REM) density. Latencies to stages 1 and 2 are from lights off. REM sleep latency, latency of REM sleep from stage 1; B1, B2, baseline nights; D1, D2, D3, deprivation nights; R1, R2, R3, recovery nights; RD, REM sleep deprivation; CD, control sleep deprivation. *Significant difference from baseline value; †significant difference from corresponding D2 value (P < 0.05, 2-tailed paired t-tests).

Fig. 2. Number of interventions required in RD nights (A) and CD nights (B). Each bar represents mean values (+SE; n = 8) of consecutive 2-h intervals. Numeric values at bottom indicate the average number of sleep interventions per night (SE in parentheses; n = 8). * Significant difference from corresponding D1 value; †significant difference from corresponding D2 value (P < 0.05, 2-tailed paired t-tests).
The sleep variables showed the expected changes in the deprivation nights. Noteworthy is the marked increase of stage 1 in RD, an effect that was not present in CD. REM density did not change in CD.

Recovery nights. The two significant condition effects were the higher amount of REM sleep in the first recovery night of RD and the lower amount of stage 2 in R1 and R3. Both variables in R1 differed also significantly from baseline. It should be noted that the REM sleep rebound after RD was limited to the first recovery night, and that the latency to REM sleep was not affected in either condition. REM density in R1 was below baseline for both conditions. The increase in stage 4 was evident for both conditions but reached significance only for CD.

EEG Power Spectra in NREM Sleep and REM Sleep

The deprivation protocols affected EEG power during recovery sleep. The power density spectra in Fig. 3 are plotted relative to baseline (100% level). In NREM sleep, power in the low-frequency range was significantly enhanced in the first recovery night for both RD and CD (Fig. 3A). A significant increase was still present in some frequency bins in the second recovery night of CD. Power in the range of sleep spindles (13.25–15 Hz) was reduced in R1 for both conditions, but significance was reached only in one bin for RD. Comparison of power in corresponding bins revealed no significant condition effect.

Figure 3B shows the power spectra in REM sleep. In the RD condition, power within the α-band (8.25–11 Hz) was significantly reduced in all three recovery nights. This effect was most prominent in R1 and then appeared to gradually subside. In R2, activity at 9.25–10 Hz tends to be higher than in R1 (P = 0.064), and in R3 the values were significantly higher than in R1 for all three bins. A significant reduction of power was also present in the β-range (16.25–20 Hz). In CD, the reduction of power in the α-band was limited to R1. Significant changes were present at the low and high end of the R1 spectrum, and a rise in the upper δ-band was seen in R1 and R2. During the RD nights, REM sleep episodes were too short for performing a meaningful spectral analysis, and during the CD nights there were no significant spectral changes from baseline (data not shown).

Because major changes of power were seen in the α-band, they were examined in more detail by plotting them for each sleep cycle (Fig. 4). Condition, night, and

**Fig. 3.** Mean relative power density spectra (n = 8) in NREM sleep (A) and REM sleep (B) of recovery nights (R1–R3) after RD and CD. Values in each frequency bin are expressed as a percentage of the corresponding values in baseline (horizontal line at 100%). Diamonds plotted at the upper limit of the frequency bins (gray level corresponding to line style) indicate significant differences from baseline (P < 0.05; two-tailed paired t-test on log-transformed absolute values); t-tests were only performed when a repeated-measure ANOVA (rANOVA) with factor night (BL, R1, R2, R3) revealed a significant effect.

**Fig. 4.** Effect of selective sleep deprivation on α-activity in REM sleep in the first 4 NREM-REM sleep cycles of the 3 recovery nights (R1–R3). Power density in the 8.25–11 Hz range (3 1-Hz bins) for recovery nights after RD (A) and CD (B). Bars represent mean deviation (n = 8) from baseline expressed as percentage of baseline value (100%). *Significant difference between corresponding cycles of RD and CD (P < 0.05; bottom asterisks, 2-way rANOVA with factors condition and cycle performed for each night; asterisks below bars, 2-tailed paired t-test on log-transformed absolute values). The following significant differences from baseline are not indicated. RD, R1 in cycles 1–4; R2 in cycles 1, 2 and 4; R3 in cycle 4. CD, R1 in cycles 3 and 4.
cycle effects were significant (P < 0.01; 3-way rANOVA). Significant condition effects were then observed in R1 and R2. In addition to the overall effect of α in the first two recovery nights, correspondingly significant condition effects were observed. In comparison to baseline, α-activity in all four cycles of R1 and in various cycles of the subsequent nights was significantly reduced after RD, whereas after CD only power in the last two cycles of R1 was decreased.

Rectal Temperature

Figure 5 illustrates for both conditions the time course of rectal temperature in the baseline nights and in the first recovery nights. The hypothermic response during sleep was enhanced in the second part of the first CD recovery night. Significant differences were present relative to baseline as well as to the corresponding RD values.

Self-Rated Alertness and Mood During the Day

Table 2 shows the daily mean values of the self-ratings for alertness and mood. Alertness was reduced for both conditions during the days after D2 and D3 and for RD also on the day after D1. Alertness returned to the BL level after the first recovery night. There was no significant condition effect. The mood ratings after deprivation and recovery nights did not deviate significantly from baseline. However, the overall scores in CD were lower than in RD [P < 0.01; 3-way ANOVA with factors condition (RD, CD), 2-h intervals (1–8), and day (1–6)], which may constitute an order effect.

DISCUSSION

The selective REM sleep deprivation procedure was effective in reducing REM sleep for three nights to a low residual level. A control for nonspecific effects of the awakenings was achieved by a matching number of interventions in NREM sleep during the control deprivation protocol. It did not appear reasonable to mimic the time course of interventions, because this would have disrupted the normal sleep architecture. The control deprivation paradigm did not interfere substantially with the NREM-REM sleep cycle (see Fig. 1). However, it could not be avoided that total sleep time was longer in the control deprivation schedule because the subjects returned to sleep faster after having been awakened from NREM sleep in comparison to REM sleep. Nevertheless, the sleep deprivation effects were comparable for the two conditions. Sleep propensity is known to be associated with SWA and inversely related to spindle frequency activity and REM density (4). The control sleep deprivation protocol can therefore be considered an adequate control for nonspecific sleep deprivation effects.

Attenuation of α-Activity

Whereas the two deprivation procedures had comparable effects on NREM sleep spectra in the recovery nights, this was not the case for the REM sleep spectra. Power in the α-band was attenuated after RD in all three recovery nights. The effect was most pronounced in the first recovery night and declined in the following nights. The analysis by cycles clearly showed that the effect was much less prominent after CD than after RD. Because a mild REM sleep deprivation prevailed also during CD, the reduction of α-activity could have been due to this factor. A progressive and persistent attenuation of α-activity in REM sleep has been observed previously in a 4-day partial sleep deprivation protocol that induced a preferential deficit in REM sleep (14). The interpretation of those findings that the attenuation of α-activity was due to the increased REM sleep pressure, is corroborated by the present data. The present results are also consistent with the observations made in a forced desynchrony study, where the circadian variation of REM sleep propensity (i.e., REM sleep per total sleep time) was inversely related to α-activity in REM sleep (23).
Is there an association between $\alpha$-activity, which is prevalent in the waking EEG, and the $\alpha$-component in the REM sleep spectrum? Two observations suggest that this might be the case. First, a recent topographic analysis of the sleep EEG revealed that $\alpha$ in REM sleep is more prominent in the parietooccipital derivation than in more anterior derivations (41, 42). At NREM-REM sleep transitions, power in the $\alpha$-band exhibited an anteroposterior shift. Thus the well-known occipital predominance of $\alpha$-activity in waking is present also in REM sleep. Second, the progressive suppression of $\alpha$-activity in the repeated partial sleep deprivation study was observed not only in REM sleep but also in waking (14). Taken together, the results suggest that the accumulated REM sleep deficit is reflected by an EEG variable that is present in both wakefulness and REM sleep. This is reminiscent of the situation in the rat where $\theta$-activity is a salient feature of both REM sleep and active wakefulness. In the course of sleep deprivation, power in the $\theta$-band was progressively increased in the waking EEG, and during recovery sleep this increase was present in the REM sleep EEG (11, 28). In both the human and rat, the changes in the REM sleep EEG outlasted the REM sleep rebound. This suggests that an REM sleep deprivation causes prolonged alterations in a state-specific process that contributes to the generation of the REM sleep EEG.

Implications for the Regulation of REM Sleep

In terms of characterizing the regulatory processes of REM sleep, apparently discrepant observations must be reconciled. On one hand, the dramatic rise in the number of interventions during the REM sleep deprivation nights points to a rapid increase in REM sleep propensity. On the other hand, there is a more modest rise in the number of interventions from night to night and merely a 40% REM sleep rebound in the first recovery night, which is far less than the amount of REM sleep lost. The fact that sleep episodes were limited to 8 h may have obscured a rebound that would have appeared during extended sleep. However, permitting unlimited sleep during the first recovery night would have led to large individual variations in sleep duration, which would have rendered it difficult to assess the changes during the second and third night. In the following, two interpretations will be considered that differ with respect of the strength of the homeostatic drive.

Hypothesis 1: Strong homeostatic drive. The rapid rise in the number of awakenings in the course of the sleep episode may be evidence for a strong homeostatic drive of REM sleep. However, the high REM sleep propensity at the end of the night must decline during the subsequent waking episode, because the number of awakenings in the initial part of the following night is much lower. Does this decline occur because the waking state as a whole or a particular substrate of waking functionally substitutes for REM sleep? Common features in the waking and REM sleep EEG (i.e., $\alpha$ in humans, $\theta$ in the rat) could be regarded as evidence for common brain mechanisms. Still, waking does not fully counteract the rise in REM sleep propensity. In the present as well as in previous selective REM sleep deprivation studies (2, 19, 30, 37), the number of interventions increased from night to night. Moreover, prolonged total sleep deprivation induces a massive REM sleep rebound (>60 h in humans (8, 31, 44) and >12 h in the rat (11, 13, 25, 28, 33, 39)). Thus not only a dissipation but also a buildup of REM sleep propensity may occur during waking.

Hypothesis 2: Weak homeostatic drive. The rising trend in the number of interventions during selective REM sleep deprivation may be due to circadian factors and a sleep-dependent disinhibition of REM sleep propensity rather than to REM sleep homeostasis. The REM sleep fraction of total sleep time is known to exhibit a marked circadian rhythm whose maximum coincides closely with the minimum of the circadian core body temperature rhythm (17, 22, 24, 40). REM sleep exhibits typically a rising trend during the night. This trend is seen only during the habitual nighttime hours and is not present when sleep is displaced to other phases of the 24-h cycle (3). In addition to the circadian factor, REM sleep is under the influence of a sleep-dependent component. The latter has been postulated in a theoretical paper (9) and was recently demonstrated in a forced desynchrony protocol where a gradual rise in the amount of REM sleep occurred irrespective of the circadian phase (22). It is therefore conceivable that the combined effect of the circadian and sleep-dependent factors is responsible for the rapid rise in REM sleep propensity and that REM sleep homeostasis plays a minor role. A weak homeostatic drive would be responsible for the moderate night-to-night increase in the number of interventions and for the minor and short-lasting REM sleep rebound. There would be no need to postulate a decay of REM sleep propensity during waking.

Selective REM sleep deprivation during sleep episodes scheduled in the late morning at a phase opposite to that of habitual sleep will allow for the testing of the two hypotheses. Although these hypotheses are not mutually exclusive, they highlight the basic question of the potency of REM sleep homeostasis. Data from animal experiments seem to be relevant to this problem. Thus selective REM sleep deprivation with (11, 25) or without (7) prior total sleep deprivation was performed in the rat. In all these studies, the number of interventions required to prevent REM sleep rose dramatically within the first 2-h period. In view of this short time interval, a circadian influence can be practically ruled out. A sleep-dependent disinhibition of REM sleep could be involved, because the intensity of NREM sleep as reflected by SWA declined during the deprivation period. An alternative proposition advanced by Benington and Heller (6) is that REM sleep is not disinhibited but actively promoted by NREM sleep. Although this hypothesis can account for a number of experimental results, the increase in REM sleep propensity during prolonged total sleep deprivation can be
explained only by assuming that NREM sleep-like states occur during prolonged waking. For human studies this possibility is still conjectural, and there is no evidence that signs of NREM sleep occur in the waking EEG to any significant extent. In an effort to resolve these problems, Franken (27) proposed, in the context of simulating sleep in the rat under various sleep deprivation schedules, that the initiation and maintenance of REM sleep could be controlled by separate processes. According to this hypothesis, the timing of NREM-REM sleep transitions depends on an hourglass process whose level rises during NREM sleep episodes and declines during REM sleep, whereas the amount (i.e., expression) of REM sleep is controlled by a process whose level increases during both NREM sleep and waking and declines during REM sleep. The former process may be perceived as an NREM sleep-dependent "sensitization" to the initiation of REM sleep (cf. aborted transitions to REM sleep in rats (6)), which would account for the time course of interventions during the deprivation night and, assuming a decay during waking, for the increasing trend across nights. The latter process (corresponding to the weak homeostatic drive in hypothesis 2) would be responsible for the modest REM sleep rebound and the gradual normalization of the REM sleep EEG during the recovery nights.

Change of Rectal Temperature

The hypothermia in the first recovery night after the control deprivation was an unexpected observation, which could indicate a change in thermoregulation during sleep. Rapid sleep-related changes of brain temperature have been documented in animals (e.g., Ref. 28), and large 24-h variations of intracranial temperature are known to prevail in humans (32). The present findings could be related to the prolonged hypothermic response of cortical temperature after total sleep deprivation in the rat (28), an effect that could reflect an altered thermoregulatory set point. A similar change could have been induced by the repeated awakenings from NREM sleep, a sleep state during which brain temperature in animals exhibits a steady decline.

Perspectives

The properties of the homeostatic process underlying NREM sleep regulation are relatively well defined and can be used for simulating the timing and architecture of sleep in humans and animals (1, 18, 29). Comparable data are scarce for REM sleep, and the present study was designed to gain insights into the regulatory processes of this sleep state. Although the emerging picture is still complex, the results may contribute to conceptualizing the problem. The analysis of NREM sleep was heavily based on specific EEG measures. A major finding of the present study is that the REM sleep EEG is not only a state marker but also an indicator of REM sleep propensity. This aspect is not evident from visual inspection and necessitates a quantitative analysis. An obvious question is whether the attenuation of α-activity is also present during waking and, if so, whether the changes occur in parallel in the two states. Regardless of this problem, REM sleep clearly emerges as a state with qualitative attributes that need to be taken into consideration in its further analysis. Related to this aspect is the possibility of other states to functionally substitute for REM sleep. A prime candidate in this respect is stage 1, which shares with REM sleep so many EEG features that the latter was referred to for some time as "emergent stage 1." Is the dramatic rise of stage 1 during the RD nights (Table 1) a sign that it is a functional equivalent of REM sleep? One of the difficulties in specifying the regulatory characteristics may be the "brittleness" of the state concept. In contrast to NREM sleep, the classical definition of REM sleep uses the two non-EEG measures tonic EMG activity and phasic EOG activity. The fact that either of the two can dissociate from an REM sleep EEG raises the problem whether one is still dealing with the same state. The typical changes in autonomic nervous activity can be regarded as further non-EEG state indicators whose periodic occurrence during sleep may persist in the absence of other REM sleep signs (38).

In conclusion, the present study may help to place REM sleep regulation in a different perspective by specifying a functionally relevant EEG component, by contrasting the homeostatic response of different measures, and by considering the possibility that different processes underlie the initiation and maintenance of REM sleep. Almost one-half a century after its discovery we are still far from understanding the paradoxes of this puzzling and fascinating facet of sleep.

We thank Dr. Beatrice Baldinger for help during the study, Evelyne Weber for scoring of rapid eye movements, and Dr. Irene Tobler for comments on the manuscript.

The study was supported by the Swiss National Science Foundation Grant 3100-042950.94 and the Human Frontiers Science Program Grant RG-8196. Dr. Takuro Endo was supported by a fellowship from the Human Frontier Science Program LT 46/93.

Address for reprint requests: A. A. Borbély, Institute of Pharmacology, Univ. of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

Received 12 August 1997; accepted in final form 9 January 1998.

REFERENCES


