Historical Perspective

The history of sleep medicine and sleep research is a history of remarkable progress and remarkable ignorance. In the 1940s and 1950s, sleep had been in the forefront of neuroscience, and then again in the late 1990s, there has been a resurgence of our understanding of the neurobiology of sleep. Sleeping and waking brain cycles can now be studied by sophisticated neuroimaging techniques which have shown remarkable progress by mapping different areas of the brain during sleep states and stages. Electrophysiological research has shown that even a single neuron sleeps as evidenced by the electrophysiological correlates of sleep-waking at the cellular (single cell) level. Despite recent progress, we are still groping for answers to two fundamental questions: What is sleep? and why do we sleep? Sleep is not simply an absence of wakefulness and perception nor is it just a suspension of sensorial processes but is a result of a combination of a passive withdrawal of afferent stimuli to the brain and functional activation of certain neurons in selective brain areas. The more important question, however, is “How do we stay awake?” In the mother’s womb, we were all asleep, and we wake up the moment we are born. But even in the newborn, sleep occupies 16 out of 24 h. Therefore, the fundamental inquiry should be directed at the mechanism of wakefulness.

Since the dawn of civilization, the mysteries of sleep have intrigued poets, artists, philosophers, and mythologists. The fascination with sleep is reflected in the literature, folklore, religion, and medicine. Upanishad [1, 2] (circa 1000 B.C.), the ancient Indian textbook of philosophy, sought to divide human existence into four states: the waking, the dreaming, the deep dreamless sleep, and the superconscious (“the very self”). This is reminiscent of modern classification of three states of existence (see later). One finds the description of pathologic sleepiness (possibly a case of Kleine–Levin syndrome) in the mythologic character Kumbhakarna in the great Indian epic Ramayana [3, 4] (circa 1000 B.C.). Kumbhakarna would sleep for months at a time, then get up to eat and drink voraciously before falling asleep again (Fig. 2.1). The ancient Chinese believed in two basic principles of life: yang, the active, light, and positive; and yin, the passive, dark, and negative. The yin–yang concept, originated with Fu Hsi (circa 2900 B.C.), has since become a symbol for sleep and wakefulness [5].

Throughout the literature, a close relationship between sleep and death has been perceived, but the rapid reversibility of sleep episodes differentiates sleep from coma and death. There are myriad references to sleep, death, and dream in poetic and religious writings, including the following quotations: “The deepest sleep resembles death” (The Bible, I Samuel 26:12); “sleep and death are similar … sleep is one-sixtieth [i.e., one piece] of death” (The Talmud, Berachoth 576); “There she [Aphrodite] met sleep, the brother of death” (Homer’s Iliad, circa 700 B.C.); “To sleep perchance to dream … For in that sleep of death what dreams may come?” (Shakespeare’s Hamlet); “How wonderful is death; Death and his brother sleep” (Shelly’s “Queen Mab”).

The 3 major behavioral states in human—wakefulness, non-rapid eye movement (NREM), and rapid eye movement (REM) sleep—are three basic biological processes that have independent functions and controls. The reader should consult Borbely’s monograph Secrets of Sleep [2] for an interesting historical introduction to sleep.
What is the origin of sleep? The words *sleep* and *somnolence* are derived from the Latin word somnus; the German words *sleps*, *siaf*, or *schlaf*; and the Greek word *hypnos*. Hippocrates, the father of medicine, postulated a humoral mechanism for sleep and asserted that sleep was caused by the retreat of blood and warmth into the inner regions of the body, whereas the Greek philosopher Aristotle thought sleep was related to food, which generates heat and causes sleepiness. Paracelsus, a sixteenth-century physician, wrote that “natural” sleep lasted 6 h, eliminating tiredness and refreshing the sleeper. He also suggested that people not sleep too much or too little, but awake when the sun rises and go to bed at sunset. This advice from Paracelsus is strikingly similar to modern thinking about sleep. Views about sleep in the seventeenth and eighteenth centuries were expressed by Alexander Stuart, the British physician and physiologist, and by the Swiss physician, Albrecht von Haller. According to Stuart, sleep was due to a deficit of the “animal spirits”; von Haller wrote that the flow of the “spirits” to the nerves was cut off by the thickened blood in the heart, resulting in sleep. Nineteenth-century scientists used principles of physiology and chemistry to explain sleep. Both Humboldt and Pflüger thought that sleep resulted from a reduction or lack of oxygen in the brain [2].

Ideas about sleep were not based on solid scientific experiments until the twentieth century. Ishimori [6] in 1909, and Legendre and Pieron [7] in 1913, observed sleep-promoting substances in the cerebrospinal fluid of animals during prolonged wakefulness. The discovery of the EEG waves in dogs by the English physician Caton [8] in 1875 and of the alpha waves from the surface of the human brain by the German physician Berger [9] in 1929 provided the framework for contemporary sleep research. It is interesting to note that Kohlschutter, a nineteenth-century German physiologist, thought sleep was deepest in the first few hours and became lighter as time went on [2]. Modern sleep laboratory studies have generally confirmed these observations.


**Fig. 2.1** Waking up of the giant Kumbhakarna, brother of Ravana in the great Indian epic Ramayana (circa 1000 B.C.) by hitting him with weapons and clubs, and shouting in his ear. This is a historical example of pathologic sleepiness resembling contemporary Kleine-Levin syndrome.
scoring technique) [14] which remained the gold standard until now. Recently, the American Academy of Sleep Medicine (AASM) published the AASM manual for the scoring of sleep and associated events [15] which modified the R–K technique and extended the scoring rules. The other significant milestone in the history of sleep medicine was the discovery of the site of obstruction in the upper airway in obstructive sleep apnea syndrome (OSAS) independently by Gastaut et al. [16] in France as well as Jung and Kuhlo [17] in Germany in 1965, followed by the polygraphic observations in the same year by Lugaresi et al. [18] of obstructive central and mixed apnea in these patients associated with periodic fall of blood pressure (BP) during apnea and rise above the baseline on resumption of breathing. The next milestone was the demonstration of dramatic relief of symptoms in these patients following tracheostomy (which bypasses the upper airway obstruction) by Kuhlo et al. [19]. In 1969 in a brief polygraphic report, Chokroverty et al. [20] made two important observations in patients with obesity hypoventilation syndrome (Pickwickian syndrome): Oxygen inhalation produced more prolonged and frequent episodes of apneas–hypopneas indicating the importance of peripheral chemoreceptor-driven hypoxemia causing respiratory stimulation and arousal in the presence of chronic daytime hypercapnia (these findings were later confirmed by other investigators [21]); the other observation is that following weight loss of 100–150-pound patients’ symptoms improved, daytime arterial carbon dioxide normalized but apneas–hypopneas persisted, though these were less frequent than before weight loss. Subsequently, numerous papers were published by Guilleminault et al. [22] who coined the term sleep apnea syndrome. Then came the seminal paper by Sullivan and associates in 1981 [23] of continuous positive airway pressure (CPAP) titration to eliminate such obstruction as the standard treatment modality for moderate-to-severe OSAS. Finally, identification of 2 neuropeptides, hypocretin 1 and 2 (orexin A and B) in the lateral hypothalamus and perifornical regions [24, 25] followed by an animal model of a human narcolepsy phenotype in dogs by mutation of hypocretin 2 receptors (HCTR2) by Lin et al. [26], the creation of similar phenotype in pre-prohypocretin knockout mice [27] and transgenic mice, [28] documentation of decreased hypocretin 1 in the cerebrospinal fluid in humans, [29] and decreased hypocretin neurons in the hypothalamus at autopsy [30, 31] in human narcolepsy patients opened a new and exciting era of sleep research. The definition of sleep and a description of its functions have always bafﬂed scientists. Moruzzi [32] while describing the historical development of the deafferentation hypothesis of sleep quoted the concept Lucretius articulated 2000 years ago—that sleep is the absence of wakefulness. A variation of the same concept was expressed by Hartley [33] in 1749, and again in 1830 by Macnish [34] who deﬁned sleep as suspension of sensorial power, in which the voluntary functions are in abeyance, but the involuntary powers, such as circulation or respiration, remain intact. It is easy to comprehend what sleep is if one asks oneself that question as one is trying to get to sleep. Human sleep can be deﬁned as an altered state in which there is impaired conscious awareness of the external world with different controls, rhythms, emotions, and dreams. It is a transient natural, periodic, physiologic phenomenon which is reversible, thus differentiating it from irreversible coma and death. Consciousness requires two components: awareness (function of cerebral cortex) and arousal (function of ascending reticular activating system). Sleep differs from unconscious state or coma (a pathological state) in the following manner: Sleep besides being a reversible physiological state also shows differences from coma in terms of brain metabolism and circulation which show marked depression and impairment in coma but slight alterations in sleep. Persistent vegetative state, minimally conscious state, and coma are distinctly different from sleep state although superficially may resemble those other states (see Chap. 43). Modern sleep researchers deﬁne sleep on the basis of both behavior of the person while asleep and the related physiologic changes that occur to the waking brain’s electrical rhythm in sleep [35–38]. The behavioral criteria (Table 2.1) include lack of mobility or slight mobility, closed eyes, a characteristic species-speciﬁc sleeping posture, reduced response to external stimulation, quiescence, increased reaction time, elevated arousal threshold, impaired cognitive function, and a reversible unconscious state. Sleep is an active anabolic state (e.g., promoting growth, stimulating immune system) and is observed in all mammals, birds, reptiles, amphibians, and fish. The physiologic criteria (see Sleep Architecture and Sleep Proﬁle) based on the ﬁndings from electroencephalography (EEG), electrocorticography (EOG), and electromyography (EMG) as well as other physiological changes in ventilation and circulation. While trying to deﬁne the process of falling asleep, we must differentiate sleepiness from fatigue or tiredness. Fatigue (see also Chap. 41) can be deﬁned as a state of sustained lack of energy coupled with a lack of motivation and drive but does not require the behavioral criteria of sleepiness such as heaviness and drooping of the eyelids, sagging or nodding of the head, yawning, and an ability to nap given the opportunity to fall asleep. On the other hand, fatigue is often a secondary consequence of sleepiness.

Definition of Sleep

Sleep is “…great nature’s second course, chief nourisher in life’s feast” (Wm. Shakespeare).
The Moment of Sleep Onset and Offset

There is no exact moment of sleep onset—there are gradual changes in many behavioral and physiological characteristics including EEG rhythms, cognition, and mental processing including reaction time. Sleepiness begins at sleep onset even before reaching stage 1 NREM sleep (as defined later) with heaviness and drooping of the eyelids; clouding of the sensorium; and inability to see, hear, or perceive things in a rational or logical manner. At this point, an individual trying to get to sleep is now entering into another world in which the person has no control and the brain cannot respond logically and adequately. This is the stage coined by McDonald Critchley as the “Pre-Dormitum” [39] who also mentioned that Gowers used “sleepening” for this stage, and this is opposite of “awakening.” Slow rolling eye movements (SEMs) begin at sleep onset and continue through stage 1 NREM sleep. At sleep onset, there is a progressive decline in thinking process, and sometimes there may be hypnagogic imagery and hypnic myoclonus [40].

Similar to sleep onset, the moment of awakening or sleep offset (post-dormitum) is also a gradual process from the fully established sleep stages. This period is sometimes described as manifesting sleep inertia or sleep drunkenness. There is a gradual return to a state of alertness or wakefulness.

Sleep Architecture and Sleep Profile

Based on three physiologic measurements (EEG, EOG, and EMG), sleep is divided into two states [41] with independent functions and controls: NREM sleep and REM sleep. Table 2.2 lists the physiologic criteria of wakefulness and sleep and Table 2.3 summaries NREM and REM sleep states. In an ideal situation (which may not be seen in all

### Table 2.1 Behavioral criteria of wakefulness and sleep

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Awake</th>
<th>Non-rapid eye movement sleep</th>
<th>Rapid eye movement sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posture</td>
<td>Erect, sitting, or recumbent</td>
<td>Recumbent</td>
<td>Recumbent</td>
</tr>
<tr>
<td>Mobility</td>
<td>Normal</td>
<td>Slightly reduced or immobile; postural shifts</td>
<td>Moderately reduced or immobile; myoclonic jerks</td>
</tr>
<tr>
<td>Response to stimulation</td>
<td>Normal</td>
<td>Mildly to moderately reduced</td>
<td>Moderately reduced to no response</td>
</tr>
<tr>
<td>Level of alertness</td>
<td>Alert</td>
<td>Unconscious but reversible</td>
<td>Unconscious but reversible</td>
</tr>
<tr>
<td>Eyelids</td>
<td>Open</td>
<td>Closed</td>
<td>Closed</td>
</tr>
<tr>
<td>Eye movements</td>
<td>Waking eye movements</td>
<td>Slow rolling eye movements</td>
<td>Rapid eye movements</td>
</tr>
</tbody>
</table>

### Table 2.2 Physiologic criteria of wakefulness and sleep

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Awake</th>
<th>Non-rapid eye movement sleep</th>
<th>Rapid eye movement sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electroencephalography</td>
<td>Alpha waves; desynchronized</td>
<td>Synchronized</td>
<td>Theta or saw tooth waves; desynchronized</td>
</tr>
<tr>
<td>Electromyography (muscle tone)</td>
<td>Normal</td>
<td>Mildly reduced</td>
<td>Moderately to severely reduced or absent</td>
</tr>
<tr>
<td>Electrooculography</td>
<td>Waking eye movements</td>
<td>Slow rolling eye movements</td>
<td>Rapid eye movements</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sleep state</th>
<th>Sleep time (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NREM sleep</td>
<td>75–80</td>
</tr>
<tr>
<td>1</td>
<td>3–8</td>
</tr>
<tr>
<td>2</td>
<td>45–55</td>
</tr>
<tr>
<td>3</td>
<td>15–23</td>
</tr>
<tr>
<td>REM sleep</td>
<td>20–25</td>
</tr>
<tr>
<td>Tonic stage</td>
<td>–</td>
</tr>
<tr>
<td>Phasic stage</td>
<td>–</td>
</tr>
</tbody>
</table>

NREM: Non-rapid eye movement; REM: Rapid eye movement
normal individuals), NREM and REM alternate in a cyclic manner, each cycle lasting on an average from 90 to 110 mins. During a normal sleep period in adults, 4–6 such cycles are noted. The first two cycles are dominated by slow-wave sleep (SWS) (R–K stages 3 and 4 NREM and AASM stage N3 sleep); subsequent cycles contain less SWS, and sometimes SWS does not occur at all. In contrast, the REM sleep cycle increases from the first to the last cycle, and the longest REM sleep episode toward the end of the night may last for an hour. Thus, in human adult sleep, the first third is dominated by the SWS and the last third is dominated by REM sleep. It is important to be aware of these facts because certain abnormal motor activities are characteristically associated with SWS and REM sleep.

**Non-rapid Eye Movement (NREM) Sleep**—NREM sleep accounts for 75–80 % of sleep time in an adult human. According to R–K scoring manual, [14] NREM sleep is further divided into four stages (stages 1–4), and according to the recent AASM scoring manual, [15] this is subdivided into three stages (N1, N2, and N3), primarily on the basis of EEG criteria. Stage 1 NREM (N1) sleep occupies 3–8 % of sleep time; stage 2 (N2) comprises 45–55 % of sleep time; and stage N3 or SWS makes up 15–23 % of total sleep time. The dominant rhythm during adult human wakefulness consists of the alpha rhythm (8–13 Hz) noted predominantly in the posterior region intermixed with small amount of beta rhythm (>13 Hz) seen mainly in the anterior head regions (Fig. 24.1). This state called stage W may be accompanied by conjugate waking eye movements (WEMs) which may comprise vertical, horizontal, oblique, slow, or fast eye movements. In stage 1 NREM sleep (stage N1), alpha rhythm diminishes to less than 50 % in an epoch (i.e., a 30-second segment of the polysomnographic tracing with the monitor screen speed of 10 mm/s) intermixed with slower theta rhythms (4–7 Hz) and beta waves (Fig. 24.2). Electromyographic activity decreases slightly, and slow eye movements (SEMs) appear. Toward the end of this stage, vertex sharp waves are noted. Stage 2 NREM (stage N2) begins after approximately 10–12 min of stage 1. Sleep spindles (11–16 Hz, mostly 12–14 Hz) and K-complexes intermixed with vertex sharp waves herald the onset of stage N2 sleep (Fig. 24.3). Sleep spindles could be divided into two types: fast spindles (13–15 Hz) seen predominantly in the centroparietal region and slow spindles (11 ≤ 13 Hz) seen mostly in the frontal region. EEG at this stage also shows theta waves and slow waves (0.5–2 Hz) that occupy less than 20 % of the epoch. After about 30–60 mins of stage 2 NREM sleep (stage N2), stage 3 sleep begins and slow waves comprise 20–100 % of the epoch (Fig. 24.4). As stated above, R–K stages 3 and 4 NREM are grouped together as SWS and are replaced by stage N3 in the new AASM scoring manual. Body movements often are recorded as artifacts in PSG recordings toward the end of SWS as sleep is lightening. Stage N3 is briefly interrupted by stage 2 NREM (stage N2), which is followed by the first REM sleep approximately 60–90 mins after sleep onset.

**Rapid Eye Movement (REM) Sleep**—REM sleep accounts for 20–25 % of total sleep time. Based on EEG, EMG, and EOG characteristics, REM can be subdivided into two stages (tonic and phasic). This subdivision is not recognized in the recent AASM scoring manual. A desynchronized EEG, hypotonia or atonia of major muscle groups with the exception of the diaphragm and the oculomotor muscles, and depression of monosynaptic and polysynaptic reflexes are characteristics of tonic REM sleep. This tonic stage persists throughout the REM sleep, whereas the phasic stage is discontinuous and superimposed on the tonic stage. Phasic REM sleep is characterized by bursts of rapid eye movements (REMs) in all directions, in singlets or clusters. Phasic swings in blood pressure and heart rate, irregular respiration, spontaneous middle ear muscle activity (MEMA), periorbital integrated potentials (PIPs) [42], and myoclonic twitching of the facial and limb muscle and tongue movements [43] are all characteristics of phasic REM sleep (Fig. 2.2). These and other phasic phenomena of REM sleep are listed in Box 2.1. A few periods of apnea or hypopnea may occur during REM sleep. Electroencephalographic tracing during REM sleep consists of low-amplitude fast pattern in the beta frequency range mixed with a small amount of theta rhythms, some of which may have a “saw tooth” appearance (Fig. 2.2). Saw tooth waves are trains of sharply contoured, often serrated, 2–6 Hz waves usually with rapid ascent and slow descent seen maximally over the central regions and are thought to be the gateway to REM sleep, often preceding a burst of REMs. PIPs are seen during REMs (Fig. 2.2) but not all REMs are accompanied by PIPs. During REM sleep, there may be some intermittent intrusions of alpha rhythms in the EEG lasting for a few seconds. The first REM sleep lasts only a few minutes. Sleep then progresses to stage 2 followed by stage N3 before the second REM sleep begins.

**Box 2.1. Lists of Phasic Events of REM Sleep**

- Rapid eye movements (REMs)
- Phasic muscle bursts (myoclonic or transient muscle bursts)
- Phasic tongue movements
- Periorbital integrated potentials (PIPS)
- Middle ear muscle activity (MEMA)
- Ponto-geniculo-occipital waves (PGO or P waves) [rats and cats; also reported in human during corticography while performing epilepsy surgery]
- Phasic alterations of breathing (brady-tachypnea)
- Phasic swings of BP (up and down)
- Phasic swings of heart rate (brady-tachycardia)
- Hippocampal theta waves (rhythms) [animal study]
Saw tooth waves in the EEG

Alpha bursts in the EEG during REM sleep

Phasic increase in brain intracellular firing rates during REM sleep

Penile erections in men and clitoral tumescence in women in REM sleep

Phasic increase in myocardial oxygen demand in REM sleep

Phasic vivid dreaming in REM sleep

Phasic suppression of REM muscle atonia

Phasic pupillary dilation and constriction

Phasic fractionations of diaphragmatic activity (pauses of 40–80 ms occurring in clusters) correlating with PGO waves which are phasic events of REM sleep

In summary, during normal sleep in adults, there is an orderly progression from wakefulness to sleep onset to NREM and then to REM sleep. Relaxed wakefulness is characterized by behavioral state of quietness and physiological state of alpha and beta frequency in the EEG, WEMs, and increased muscle tone. NREM sleep is characterized by progressively decreased responsiveness to external stimulation accompanied by SEMs, followed by electroencephalographic slow-wave activity associated with spindle and K-complexes, and decreased muscle tone. REMs markedly reduced or absent muscle tone and low voltage fast electroencephalographic activity mixed with distinctive saw tooth waves and PIPs characterize REM sleep.

The R–K system addresses normal adult sleep and macrostructure of sleep. In patients with sleep disorders such as sleep apnea, parasomnias, or sleep-related seizures, it may be difficult to score sleep according to R–K criteria. Furthermore, the R–K staging system does not address the microstructure of sleep. The recent AASM sleep scoring criteria with a brief reference to R and K system have been outlined in Chap. 24. The macrostructure of sleep is summarized in Box 2.2. There are several endogenous and exogenous factors, which will modify sleep macrostructure (Box 2.3).

**Box 2.2: Sleep Macrostructure**

<table>
<thead>
<tr>
<th>Sleep states and stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep cycles</td>
</tr>
<tr>
<td>Sleep latency</td>
</tr>
<tr>
<td>Sleep efficiency (the ratio of total sleep time to total time in bed expressed as a percentage)</td>
</tr>
<tr>
<td>Wake after sleep onset</td>
</tr>
</tbody>
</table>

**Box 2.3: Factors Modifying Sleep Macrostructure**

- Exogenous
  - Noise
  - Exercise
  - Ambient temperature
  - Drugs and alcohol
Endogenous
- Age
- Prior sleep–wakefulness
- Circadian phase
- Sleep pathologies

Sleep Microstructure—Sleep microstructure includes momentary dynamic phenomena such as arousals, which have been operationally defined by the Task Force of the American Sleep Disorders Association (now called American Academy of Sleep Medicine) [44] which has remained essentially unchanged in the recent AASM scoring manual, [15] and the cyclic alternating pattern (CAP), which has been defined and described in various publications by Termano and co-investigators [45–47]. Other components of microstructure include K-complexes and sleep spindles (Box 2.4).

Box 2.4: Sleep Microstructure

- Arousals
- Cyclic alternating pattern
- Sleep spindles
- K-complexes

Arousals are transient phenomena resulting in fragmented sleep without behavioral awakening. An arousal is scored during sleep stages N1, N2, N3 (or REM sleep) if there is an abrupt shift in EEG frequency lasting from 3 to 14 s (Fig. 2.3) and including alpha, beta, or theta activities but not spindles or delta waves. Before an arousal can be scored, the subject must be asleep for 10 consecutive seconds. In REM sleep, arousals are scored only when accompanied by concurrent increase in segmental EMG amplitude. K-complexes, delta waves, artifacts, and only increased segmental EMG activities are not counted as arousals unless these are accompanied by EEG frequency shifts. Arousals can be expressed as number per hour of sleep (an arousal index), and up to 10 arousal index can be considered normal.

The Cyclic Alternating Pattern—The CAP (Fig. 2.4) indicates sleep instability, whereas frequent arousals with or without stage shifts signify sleep fragmentation [47]. A fragmentation index (number of arousals and stage shifts per hour) can also be calculated to indicate sleep instability. Sleep microstructure is best understood by the CAP, wherein an EEG pattern that repeats in a cyclical manner is noted mainly during NREM sleep. This is a promising technique in evaluating both normal and abnormal sleep, as well as in

Fig. 2.3 Polysomnographic recording shows two brief periods of arousals out of stage N2 sleep in the left- and right-hand segments of the recording, lasting for 5.58 and 6.40 s and separated by more than 10 s of sleep. Note delta waves followed by approximately 10-Hz alpha activities during brief arousals. Reprinted from Chap. 2, 3rd edition
understanding the neurophysiologic and neurochemical basis of sleep. A CAP cycle [48] consists of an unstable phase (Phase A) and relatively stable phase (Phase B) each lasting between 2 and 60 s. Phase A of CAP is marked by an increase in EEG potentials with contributions from both synchronous high amplitude and slow, and desynchronized fast rhythms in the EEG recording standing out from a relatively low-amplitude slow background. The A phase is associated with an increase in heart rate, respiration, blood pressure, and muscle tone. CAP rate (total CAP time during NREM sleep) and arousals both increase in older individuals and in a variety of sleep disorders including both diurnal and nocturnal movement disorders. Non-CAP (sleep period without CAP) is thought to indicate a state of sustained stability.

In summary, sleep macrostructure is based on cyclic patterns of NREM and REM states, whereas sleep microstructure mainly consists of arousals, periods of CAP, and periods without CAP. An understanding of sleep macrostructure and microstructure is important because emergence of abnormal motor activity during sleep may be related to disturbed macrostructure and microstructure of sleep.

The Ontogeny of Sleep

Evolution of EEG and sleep states (see also Chap. 52) from the fetus, preterm, and term infant, early childhood, adolescence to adulthood follows in an orderly manner depending upon the maturation of the central nervous system (CNS) [49–52]. Neurological, environmental, and genetic factors as well as comorbid medical or neurological conditions will have significant effects on such ontogenetic changes. Sleep requirements change dramatically from infancy to old age. Newborns have a polyphasic sleep pattern, with 16 h of sleep per day. This sleep requirement decreases to approximately 10 h per day by 3–5 years of age. In preschool children, sleep assumes a biphasic pattern.
Adults exhibit a monophasic sleep pattern, with an average duration from 7.5 to 8 h per night. This returns to a biphasic pattern in old age.

Upon falling asleep, a newborn baby goes immediately into REM sleep, or active sleep, which is accompanied by restless movements of the arms, legs, and facial muscles. In premature babies, it is often difficult to differentiate REM sleep from wakefulness. Sleep spindles appear from 6 to 8 weeks and are well formed by 3 months (may be asynchronous during the first year and by age 2 are synchronous). K-complexes are seen at 6 months but begin to appear at over 4 months. Hypnagogic hypersynchrony characterized by transient bursts of high amplitude waves in the slower frequencies appears at 5–6 months and is prominent at one year. By 3 months of age, the NREM/REM cyclic pattern of adult sleep is established. However, the REM/NREM cycle duration is shorter in infants, lasting for approximately 45–50 mins and increasing to 60–70 mins by 5–10 years and the normal adult cyclic pattern of 90–100 mins by age of 10 years. A weak circadian rhythm is probably present at birth, but by 6–8 weeks, it is established. Gradually, the nighttime sleep increases and daytime sleep decreases and the number of naps decreases. By 8 months, the majority take two naps (late morning and early afternoon). The first 3 months is a critical period of CNS reorganization, and striking changes occur in many physiological responses. In newborns, total sleep time is about 16 h. The total sleep time decreases to 14 h at 4 months and to 13 h at 6–8 months. By 3–6 months, major concentration of sleep occurs at night.

Sleep onset in the newborn occurs through REM sleep. During the first three months, sleep onset REM begins to change. In the newborn, active sleep (REM) occurs 50 % of the total sleep time. This decreases during the first 6 months of age. By 9–12 months, REM sleep occupies 30–35 % of sleep and by 5–6 years, REM sleep decreases to adult levels of 20–25 %. The napping frequency continues to decline, and by age 4–6 years, most children stop daytime naps. Nighttime sleep patterns become regular gradually, and by age 6, nighttime sleep is consolidated with fewer awakenings. At 9–10 years of age, most children sleep for 10 h at night. Pre-adolescents are highly alert during the day with multiple sleep latency test (MSLT) showing a mean sleep latency of 17–18 min. In summary, the multiphasic sleep pattern of newborns and infants gradually changes to biphasic sleep in preschool children, and finally to monophasic sleep in adults [49, 50]. Sleep reverts to biphasic pattern in the elderly.

There are two other important changes that occur in the sleep pattern in old age: repeated awakenings throughout the night, including early morning awakenings that prematurely terminate the night sleep, and a marked reduction of the amplitude of the slow waves resulting in a decreased percentage of slow-wave sleep (SWS) in this age group. The percentage of REM sleep in normal elderly individuals remains relatively constant, and the total duration of sleep time within 24 h is also no different from that of young adults; however, elderly individuals often nap during the daytime compensating for lost sleep during the night. Figure 2.5 shows schematically the evolution of sleep step distribution in newborns, infants, children, adults, and elderly adults. Night sleep histograms of children, young adults, and of elderly adults are shown in Fig. 2.6.

The significant evolutionary changes in the respiratory and cardiovascular functions [51, 53]. Respiratory controllers are immature and not fully developed at birth. Respiratory mechanics and upper airway anatomy are different in newborns than in adults contributing to breathing problems, particularly during sleep in newborn infants. Brief periods of respiratory pauses or apneas lasting for 3 s or longer, periodic breathing and irregular breathing may be noted in newborns, especially during active (REM) sleep. According to the National Institutes of Health, Consensus Development Conference on Infantile Apnea, [54] the term periodic breathing refers to respiratory pauses of at least 3 s with less than 20 s of normal breathing in between the pauses. Cheyne–Stokes breathing is periodic waxing and waning of respiration accompanied by central apneas and may be noted in preterm infants. Periodic breathing and occasional central apneas of up to 15 s duration in newborns may be noted without any clinical relevance unless accompanied by bradycardia or cyanosis. These breathing events gradually disappear during the first few weeks of life. The respiratory rate also gradually slows during the first few years of life. Another important finding in the newborn, particularly during active sleep is paradoxical inward motion of the rib cage. This occurs because of

![Fig. 2.5](image)
high compliance of the rib cage in newborns, circular rather than elliptical thorax and decreased tone of the intercostal and accessory muscles of respiration. This paradoxical breathing causes hypoxia and reduced diaphragmatic efficiency. Similar breathing in adults occurs during diaphragmatic weakness. At term, posterior cricoarytenoid muscles which assist in maintaining upper airway patency are not adequately coordinated with diaphragmatic activity causing a few periods of obstructive apneas especially during active sleep. Ventilatory responses to hypoxia are also different in newborns than in adults. In quiet sleep, hypoxia stimulates breathing as in adults, but in active sleep after the initial period of stimulation, there is ventilatory depression. Laryngeal stimulation in adults causes arousal, but in infants, this may cause an apnea. Breathing becomes regular, and the respiratory control is adequately developed by the end of the first year. Changes in cardiovascular function indicate changes in the autonomic nervous system (ANS) during infancy and early childhood. There is greater parasympathetic control for children than infants (as assessed by heart rate low-frequency [LF] and high-frequency [HF] analysis (see also Chap. 11): 0.15–0.5 Hz [HF] indicates parasympathetic and 0.04–0.15 Hz [LF] indicates sympathetic activity). The better parasympathetic control for children than infants indicates ANS maturity. Respiratory heart rate modulation is variable in newborns as assessed by LF and HF heart rate spectral analysis. In active sleep, most of the power is in LF. In older infants and children, there is significant respiratory heart rate modulation termed normal sinus arrhythmia. Respiratory rate during quiet sleep decreases, and the respiratory variability decreases with age.

**Sleep Habits**

Sleep specialists sometimes divide people into two groups: “evening types” (owls) and “morning types” (larks). The morning types wake up early feeling rested and refreshed and work efficiently in the morning. These people get tired and go to bed early in the evening. In contrast, evening types have difficulty getting up early and they feel tired in the morning; they feel fresh and energetic toward the end of the day. These people perform best in the evening. They go to sleep late at night and wake up late in the morning. The body temperature rhythm takes on different curves in these two types of people. The body temperature reaches the evening peak an hour earlier in morning types than in evening types. What determines a morning or evening type is not known, but heredity may play a role. Katzenberg et al. [55] using the 19-item Horne–Ostberg questionnaire to determine morningness/eveningness in human circadian rhythms discovered a clock gene polymorphism associated with human diurnal preference. One of two H clock alleles (3111C) is associated with eveningness. These findings have been contradicted by later studies [56]. There is a third type (intermediate type) not conforming to either morning or evening type [57].

**Sleep Need and Requirement**

Sleep requirement or sleep need is defined as the optimum amount of sleep required to remain alert and fully awake and to function adequately throughout the day. Sleep debt is defined as the difference between the ideal sleep requirement and the actual duration of sleep obtained. There are two divergent views: Harrison and Horner [58] concluded that society is not sleep deprived, whereas Bonnet and Arand [59] stated that modern society is chronically sleep deprived. Between 1910 and 1963, there was a mean reduction of 1.5 h of sleep in adolescents aged 8–17 years [60].
However, there may be a significant sampling error in this survey. A study by Bl wired and associates [61] in healthy adults aged 50–65 years showed a reduction of about one hour of sleep between 1959 and 1980 surveys.

It has been traditionally stated that women need more sleep than men, but this has been questioned in a recent field study [62]. There is also a general perception based on questionnaire, actigraphy, and PSG studies that sleep duration decreases with increasing age [63, 64]. This relationship, however, remains controversial. Older adults take naps, and these naps may compensate for nighttime sleep duration curtailment. Sleep is regulated by homeostasis (increasing sleep drive during continued wakefulness) and circadian factors (the sleep drive varying with time of the day). The influence of these factors is reduced in older adults but is still present. Older adults are also phase advanced (e.g., internal clock set earlier yielding early bedtime and early morning awakenings).

Sleep requirement for an average adult is approximately 7½–8 h regardless of environmental or cultural differences [65]. Most probably whether a person is a long or a short sleeper and sleep need are determined by heredity rather than by different personality traits or other psychological factors. Sleep behavior is regulated by genetics. Sleep duration is influenced by the gene DEC2, mutation of which in some people may reduce sleep duration. Social (e.g., occupational) or biological (e.g., illness) factors may also play a role. Sleep need is genetically determined, but its physiologic mechanism is unknown. Slow-wave activity (SWA) in sleep EEG depends on sleep need and homeostatic drive. Adenosine, a peptide, seems to have a direct role in homeostasis. Prolonged wakefulness causes increased accumulation of adenosine which decreases during sleep. SWA increases after sleep loss. Long sleepers spend more time asleep but have less SWS [66] and more stage 2 NREM sleep than do short sleepers [67]. An early important epidemiological study [68] found that the chances of death from coronary arterial disease, cancer, or stroke are greater for adults who sleep less than 4 h or more than 9 h when compared to those who sleep an average of 7½–8 h. There have been some more recent studies by Kripke et al. [69] and others [70–72] confirming these observations. In later studies, other factors such as sleeping medications may have confounded these issues. There is no clear-cut conclusion yet.

There is controversy whether a person can extend sleep beyond the average requirement. Early studies by Taub and Berger [73, 74] showed that sleep extension beyond the average hours may cause exhaustion and irritability with detriment of sleep efficiency. The authors refer to this as the “Rip Van Winkle” effect [74]. Sleep extension studies in the past reported conflicting results regarding MSLT scores, vigilance, and mood ratings [75]. When subjects are challenged to maximum sleep extension, there is substantial improvement in daytime alertness, reaction time, and mood [75]. Most individuals carry a large sleep debt and as extra sleep reduces carryover sleep debt, it is then no longer possible to obtain extra sleep [76].

### Sleep and Dreams

Freud [77] called dreams the “Royal Road to the Unconscious” in his seminal book, The Interpretation of Dreams, published in 1900. The Freudian theory postulated that repressed feelings are psychologically suppressed or hidden in the unconscious mind and often manifested in dreams. Sometimes those feelings are expressed as mental disorders or other psychologically determined physical ailments, according to this psychoanalytic theory. In Freud’s view, most of the repressed feelings are determined by repressed sexual desires and appear in dreams or symbols representing sexual organs. In recent times, Freudian theory has fallen in disrepute. The modern sleep scientists try to interpret dreams in anatomic and physiologic terms. Nevertheless, we still cannot precisely define what is “dream” and why we dream.

The field of dream research took a new direction since the existence of REM sleep was first observed by Aserinsky and Kleitman [11] in 1953. It is postulated that approximately 80% of dreams occur during REM sleep and 20% occur during NREM sleep [78]. It is easier to recall REM dreams than NREM dreams. It is also easier to recall dreams if awakened immediately after the onset of REM dreams rather than trying to remember them the next morning upon getting out of bed. REM dreams are often vivid, highly emotionally charged, unrealistic, complex, and bizarre. In contrast, dream recall which sometimes may partially occur upon awakening from the NREM dream state is more realistic. People are generally oriented when awakening from REM sleep but are somewhat disoriented and confused when awakened from NREM sleep. Most of all, dreams take place in natural color, rather than black and white. In our dreams, we employ all five senses. In general, we use mostly the visual sensations, followed by auditory sensation. Tactile, smell, and taste sensation are represented least. Dreams can be pleasant, unpleasant, frightening, or sad. They generally reflect one’s day-to-day activities. Fear, anxiety, and apprehension are incorporated into our dreams. In addition, stressful events of past or present may occupy our dreams. The dream scenes or events are rarely rational but often occur in an irrational manner with rapid change of scene, place, or people or a bizarre mixture of these elements. Sometimes, lucid dreams may arise in which the dreamer seems to realize vividly that he or she is actually dreaming [79].

The neurobiologic significance of dreams remains unknown. Sleep scientists try to explain dreams in the terms...
of anatomical and physiological interpretation of REM sleep. During this state, the synapses, nerve cells, and nerve fibers connecting various groups of nerve cells in the brain become activated. This activation begins in the brainstem, and the cerebral hemisphere then synthesizes these signals and creates colorful or black and white images giving rise to dreams. Similarly, signals sometimes become converted into auditory, tactile, or other sensations to cause dream imagery. Why the nerve circuits are stimulated to cause dreaming is not clearly understood. Some suggestions to explain significance of dreams include activation of the neural networks in the brain, [80] restructuring and reinterpretation of data stored in memory [81]. This resembles Jouvet’s hypothesis of a relationship between REM sleep and recently acquired information [82]. According to molecular biologist and Nobel laureate Crick and Mitchison [83], the function of dreaming is to unlearn, that is, to remove unnecessary and useless information from the brain. Some have also suggested that memory consolidation takes place during the dream stage of sleep (see Chap. 13). In addition, stories abound regarding artists, writers, and scientists who develop innovative ideas about their art, literature, and scientific projects during dreams. Dream-enacting behavior associated with abnormal movement during sleep (REM behavior disorder) and frightening dreams called nightmares or dream anxiety attacks constitute two important REM parasomnias.

**Phylogeny of Sleep**

Studies have been conducted to find out whether, like humans, other mammals have sleep stages [1, 84–87]. The EEG recordings of mammals show similarities to those of humans. Both REM and NREM sleep stages can be differentiated by EEG, EMG, and EOG in animals. Dolphins and whales are the only groups of mammals showing no REM sleep on recordings [1, 88, 89]. Although initially thought to have no REM sleep, [90] some recent evidence suggests that Australian spiny anteaters (the monotremes, or egg-laying mammals, *echidna*) do have REM sleep [91–93]. Siegel and colleagues [94] suggest that the echidna combines REM and NREM aspects of sleep in a single sleep state. These authors further suggest that REM and NREM sleep evolved from a single, phylogenetically older sleep state.

Like humans, mammals can be short or long sleepers. There are considerable similarities between sleep duration and length of sleep cycles in small and large animals. Small animals with a high metabolic rate have a shorter life span and sleep longer than larger animals with lower metabolic rates [95]. Smaller animals also have a shorter REM–NREM cycle than larger animals. The larger the animal, the less it sleeps; e.g., elephants sleep 4–5 h and giraffes sleep even less than that.

A striking finding in dolphins is that during sleep, half the brain shows the characteristic EEG features of sleep while the other half shows the EEG features of waking [96]. Each sleep episode lasts approximately 30–60 mins; then, the roles of the two halves of the brain reverse. Similar uni-hemispheric sleep episodes with eye closure contralateral to the sleeping hemisphere are known to occur in the pilot whale and porpoise [35]. It is of interest to note that there are indications from computerized coherence analysis of such interhemispheric SWS asymmetry even in human with severe OSA during apneic arousals [97, 98].

Both vertebrates and invertebrates display sleep and wakefulness [99]. Most animals show the basic rest–activity rhythms during a 24-h period. There is behavioral and EEG evidence of sleep in birds but the avian REM–NREM cycles are very short [99, 100]. Although birds are thought to have evolved from reptiles, the question of the existence of REM sleep in reptiles remains somewhat controversial [99]. The absence of REM sleep in reptiles and the presence of NREM and REM sleep in both birds and mammals would be in favor of REM sleep being a more recent development in the phylogenetic history of land-dwelling organisms [99]. Sleep has also been noted in invertebrates, such as insects, scorpions, and worms, based on behavioral criteria [100].

In conclusion, the purpose of studying the phylogeny of sleep is to understand the neurophysiologic and neuroanatomic correlates of sleep as one ascends the ladder of phylogeny from inframammalian to mammalian species. Tobler [35] concluded that sleep is homeostatically regulated, in a strikingly similar manner, in a broad range of mammalian species. These similarities in sleep and its regulation among mammals suggest common underlying mechanisms that have been preserved in the evolutionary process.

**Circadian Sleep–Wake Rhythm**

The existence of circadian rhythms has been recognized since the eighteenth century, when the French astronomer de Mairan [101] noted a 24-h rhythm in heliotrope plants. The plants closed their leaves at sunset and opened them at sunrise, even when they were kept in darkness, shielded from direct sunlight. The discovery of 24-h rhythm in the movements of plant leaves suggested to de Mairan an “internal clock” in the plant. Experiments by chronobiologists Pittendrigh [102] and Aschoff [103] in 1960 clearly proved the existence of 24-h rhythms in animals.

The term *circadian rhythm*, coined by chronobiologist Halberg [104], is derived from the Latin circa, which means *about*, and *diē*, which means *day*. Experimental isolation from all environmental time cues (German *Zeitgeber*) has clearly demonstrated the existence of a circadian rhythm in humans independent of environmental stimuli [105, 106].
Earlier investigators suggested that the circadian cycle is closer to 25 h than 24 h of a day–night cycle; [1, 107–109] however, recent research points to a cycle near 24 h (approximately 24.2 h). Ordinarily, environmental cues of light and darkness synchronize or entrain the rhythms to the night–day cycle; however, the existence of environment-independent, autonomous rhythm suggests that the human body also has an internal biological clock [1, 105–108].

The experiments in rats in 1972 by Stephan and Zucker [110] and Moore and Eichler [111] clearly identified the site of the biological clock, located in the paired suprachiasmatic nuclei in the hypothalamus, above the optic chiasm. Experimental stimulation, ablation, and lesion of these nuclei altered circadian rhythms. The existence of a circadian pacemaker in the suprachiasmatic nuclei (SCN) in humans was confirmed by Lydic and colleagues [112]. There has been clear demonstration of the neuroanatomic connection between the retina and the suprachiasmatic nuclei—the retinohypothalamic pathway [113]—that sends the environmental cues of light to the SCN. The SCN serves as a pacemaker, and the neurons in the SCN are responsible for generating the circadian rhythms [106, 114–117]. The master circadian clock in the SCN receives afferent information from the retinohypothalamic tract which sends signals to multiple synaptic pathways in other parts of the hypothalamus, plus superior cervical ganglion and pineal gland where melatonin is released. The SCN contains melatonin receptors, and there is a feedback loop from the pineal gland to the SCN. Several neurotransmitters have been located within terminals of the SCN afferents and interneurons, including serotonin, neuropeptide Y, vasoactive intestinal peptide, and γ-aminobutyric acid [106, 116, 118].

Time isolation experiments have clearly shown the presence of daily rhythms coordinated by the master clock, the SCN in many physiological processes such as the sleep–wake cycle, body temperature, and neuroendocrine secretion. Body temperature rhythm is sinusoidal, whereas cortisol and growth hormone secretion rhythms are pulsatile. It is well known that plasma levels of prolactin, growth hormone, and testosterone are all increased during sleep at night (see Chap. 11). Melatonin, the hormone synthesized by the pineal gland (see Chap. 11), is secreted maximally during night and may be an important modulator of human circadian rhythm entrainment by the light–dark cycle. Sleep decreases body temperature, whereas activity and wakefulness increase it. It should be noted that internal desynchronization occurs during free-running experiments, and the rhythm of body temperature dissociates from the sleep rhythm as a result of that desynchronization [2, 106–108] (see Fig. 11.10). This raises the question of whether there is more than one circadian (or internal) clock or circadian oscillator [1]. The existence of two oscillators was postulated by Kronauer and colleagues [119]. They suggested that a 25-h rhythm exists for temperature, cortisol, and REM sleep and that the second oscillator is somewhat labile and consists of the sleep–wake rhythm. Some authors, [120] however, have suggested that one oscillator could explain both phenomena. Recent development in circadian rhythm research has clearly shown the existence of multiple circadian oscillators (peripheral clocks) functioning independently from the SCN [121–124] (Fig. 2.7).

The molecular basis of the mammalian circadian clock has been the focus of much recent circadian rhythm research [125–128] (see Chap. 6). The paired SCN are controlled by a total of 8 or more genes (e.g., Clock, Bmal, Per, Cry, and Tim), and their protein products and regulatory enzymes (e.g., casein kinase 1 epsilon and casein kinase 2 delta), and this knowledge is still evolving. By employing “forward genetics” approach, remarkable progress has been made in a few years in identifying key components of the circadian clock in both the fruit flies (Drosophila) and mammals [125–127]. It has been established that the circadian clock gene of sleep–wake cycle is independent of other circadian rhythm functions. There is clear anatomical and physiological evidence to suggest a close interaction between the SCN and the regions regulating sleep–wake states [128, 129]. There are projections from SCN to wake promoting hypocretin (orexin) neurons (indirectly via dorsomedial hypothalamus) and locus coeruleus as well as to sleep-promoting neurons in ventrolateral and median preoptic neurons. Physiological evidence of increased firing rates in single neuron recordings from the appropriate regions during wakefulness or REM sleep and decreased neuronal firing rates during NREM sleep complement anatomical evidence of such interaction between SCN and sleep–wake regulating systems [129, 130]. Based on the studies in mice (e.g., knockout mice lacking core clock genes and mice with mutant clock genes), it has also been suggested that circadian clock genes may affect sleep regulation and sleep homeostasis independent of circadian rhythm generation in other process [131].

Molecular mechanisms applying gene sequencing techniques have been found to play a critical role in uncovering the importance of clock genes, at least in two human circadian rhythm sleep disorders. Mutation of hPer 2 gene (a human homologue of period gene in Drosophila) causing advancing of the clock (alteration of the circadian timing of sleep propensity) and polymorphism in some familial advanced sleep phase state (ASPS) [132–134] and polymorphism in hPer 3 genes in some subjects with delayed sleep phase state (DSPS) [135, 136] suggest genetic control of the circadian timing of the sleep–wake rhythm. Kolker et al. [137, 138] have shown reduced 24-h expression of BMAL1 and clock genes in the SCN of old golden hamsters pointing to a possible role for the molecular mechanism in understanding age-related changes in the circadian clock.
Kondratov et al. [139] reported that mice deficient in the circadian transcription factor BMAL1 have reduced life span and display a premature aging phenotype. Later, Antoch et al. [140] corroborated these findings by showing that clock mutant mice respond to low-dose irradiation by accelerating this aging process and developing phenotype that are reminiscent of those in BMAL1-deficient mice. It is important to be aware of circadian rhythms, because several other sleep disturbances are related to alteration in them, such as those associated with shift work and jet lag.

**Fig. 2.7** The retinal ganglion cells transmit information about “time of day” via the retinohypothalamic tract to the suprachiasmatic nuclei (SCN), which conveys timing information to the brain including the paraventricular (PVN), arcuate (ARC) nuclei subparaventricular zone (SPZ), medial preoptic area (MPOA), intergeniculate leaflet (IGL), and paraventricular nucleus of the thalamus. The SCN also relays timing cues to peripheral tissues (peripheral oscillators) via the autonomic nervous system and hormones, which together organize complex behaviors. Reproduced with permission from Arble and Sandoval [239]
Chromobiology, Chronopharmacology, and Chronotherapy

Sleep specialists are becoming aware of the importance of chronobiology, chronopharmacology, and chronotherapy; however, studies are spares in these fields. Chromobiology refers to the study of the body’s biological responses to time-related events. All biological functions of the cells, organs, and the entire body have circadian (approximately 24 h), ultradian (<24 h), or infradian (>24 h) rhythms. It is important, therefore, to understand how the body responds to treatment at different times throughout the circadian cycle and that circadian timing may alter the pathophysiologic responses in various disease states (e.g., exacerbation of bronchial asthma at night and a high incidence of stroke late at night and myocardial infarction early in the morning; see Chap. 47).

Biological responses to medications may also depend on the circadian timing of administration of the drugs. Potential differences of responses of antibiotics to bacteria, or of cancer cells to chemotherapy or radiotherapy, depending on the time of administration, illustrate the importance of chronopharmacology, which refers to pharmacokinetic or pharcodynamic interactions in relation to the timing of the day.

Circadian rhythms can be manipulated to treat certain disorders, a technique called chronotherapy. Examples of this are phase advance or phase delay of sleep rhythms and application of bright light at certain periods of the evening and morning.

Cytokines, Immune System, and Sleep Factors

Cytokines are proteins produced by leukocytes and other cells functioning as intercellular mediators that may play an important role in immune and sleep regulation (see also Chap. 12). Several cytokines such as interleukin (IL), interferon alpha (IF-α), and tumor necrosis factor (TNF) have been shown to promote sleep. There are other sleep-promoting substances called sleep factors which increase in concentration during prolonged wakefulness or during infection and enhance sleep. These other factors include delta sleep-inducing peptides, muramyl peptides, cholecystokinin, arginine vasotocin, vasoactive intestinal peptide, growth hormone-releasing hormone (GHRH), somatostatin, prostaglandin D2 nitric oxide (NO), and adenosine. The role of these various sleep factors in maintaining homeostasis has not been clearly established [148]. It has been shown that adenosine in the basal forebrain can fulfill the major criteria for the neural sleep factor that mediates these somnogenic effects of prolonged wakefulness by acting through A1 and A2a receptors [157, 158].

The cytokines play a role in the cellular and immune changes noted during sleep deprivation [148, 149, 159–163]. The precise nature of the immune response after sleep deprivation has, however, remained controversial, and the results of studies on the subject have been inconsistent. These inconsistencies may reflect different stress reactions of subjects and different circadian factors (e.g., timing of drawing of blood for estimation of plasma levels) [148, 159, 164].

Infection (bacterial, viral, and fungal) enhances NREM sleep but suppresses REM sleep. It has been postulated that sleep acts as a host defense against infection and facilitates the healing process [148, 159, 163, 165–168]. It is also believed that sleep deprivation may increase vulnerability to infection [169]. The results of experiments with animals suggest that sleep deprivation alters immune function [1, 159, 160, 165]. Sleep thus has a profound impact on the immune system. It has been shown [170] that following sleep deprivation (e.g., sleeping four hours per night for four nights), there was a delay in immune response 10 days after flu vaccination.

There is evidence that cytokines play an important role in the pathogenesis of excessive daytime sleepiness in a variety of sleep disorders and in sleep deprivation [170–173]. Sleep deprivation causing excessive sleepiness has been associated with increased production of proinflammatory cytokines IL-6 and TNF-α [172, 173]. Viral or bacterial infections causing excessive somnolence and increased NREM sleep are associated with increased production of TNF-α and IL-B [174–176]. In other inflammatory disorders such as HIV infection and rheumatoid arthritis, increased sleepiness and disturbed sleep are associated with increased amount of circulating TNF-α [177–180]. Several authors suggested that excessive sleepiness, in OSAS, narcolepsy, insomnia, or idiopathic hypersomnia may be mediated by cytokines such as IL-6 and TNF-α [181–187]. In a recent review, Kapsimalis et al. [170] concluded that cytokines are mediators of sleepiness and implicated in the pathogenesis of symptoms of OSAS, narcolepsy, sleep deprivation, and insomnia and indirectly play an important role in the pathogenesis of the cardiovascular complications of OSAS.

Theories of the Function of Sleep

The function of sleep remains the greatest biologic mystery of all times. Several theories of the function of sleep have been proposed (Box 2.5). None of the theories are satisfactory to explain the exact biologic functions of sleep. Sleep deprivation experiments in animals have clearly shown that sleep is necessary for survival, but from a practical point of view, complete sleep deprivation for a prolonged period cannot be conducted in humans. Sleep deprivation studies in humans have shown an impairment of performance which demonstrates the need for sleep (see also Chap. 3). The
performance impairment of prolonged sleep deprivation results from a decreased motivation and frequent “micro-sleep.” Overall, human sleep deprivation experiments have proven that sleep deprivation causes sleepiness and impairment of performance, vigilance, attention, concentration, and memory. Sleep deprivation may also cause some metabolic, hormonal, and immunologic affects. Sleep deprivation causes immune suppression, and even partial sleep deprivation reduces cellular immune responses. Studies by Van Cauter’s group [188, 189] include a clearly documented elevation of cortisol level following even partial sleep loss suggesting an alteration in the hypothalamic–pituitary–adrenal (HPA) axis function. This has been confirmed even in chronic sleep deprivation which causes impairment of glucose tolerance. Glucose intolerance may contribute to memory impairment as a result of decreased hippocampal function. Chronic sleep deprivation may also cause a detriment of thyrotropin concentration, increased evening cortisol level, and sympathetic hyperactivity which may serve as risk factors for obesity, hypertension, and diabetes mellitus. It should, however, be noted that in all of these sleep deprivation experiments, stress has been a confounding factor, raising a question about whether all these undesirable consequences relate to sleep loss only or a combination of stress and sleeplessness.

Box 2.5: Theories of Sleep Function

- Restorative theory
- Energy conservation theory
- Adaptive theory
- Instinctive theory
- Memory consolidation and reinforcement theory
- Synaptic and neuronal network integrity theory
- Thermoregulatory theory
- Immune and endocrine regulation
- Metabolic homeostasis theory

Restorative Theory
Proponents of the restorative theory ascribe body tissue restoration to NREM sleep and brain tissue restoration to REM sleep [190–193]. The findings of increased secretion of anabolic hormones [194–196] (e.g., growth hormone, prolactin, testosterone, luteinizing hormone) and decreased levels of catabolic hormones [197] (e.g., cortisol) during sleep, along with the subjective feeling of being refreshed after sleep, may support such a contention. Increased SWS (rebound) after sleep deprivation [1] further supports the role of NREM sleep as restorative. The critical role of REM sleep for the development of the CNS of young organisms is cited as evidence of restoration of brain functions by REM sleep [198]. Several studies of brain basal metabolism suggest an enhanced synthesis of macromolecules such as nucleic acids and proteins in the brain during sleep, [199] but the data remain scarce and controversial. Protein synthesis in the brain is increased during SWS [200]. Confirmation of such cerebral anabolic processes would provide an outstanding argument in favor of the restorative theory of sleep. Recent work in animals suggests formation of new neurons during sleep in adult animals and this neurogenesis in dentate gyrus may be blocked after total sleep deprivation [201].

Energy Conservation Theory
Zepelin and Rechtschaffen [95] found that animals with a high metabolic rate sleep longer than those with a slower metabolism, suggesting that energy is conserved during sleep. There is an inverse relationship between body mass and metabolic rate. Small animals (e.g., rats and opossum) with high metabolic rate sleep for 18 h per day, whereas large animals (e.g., elephants and giraffes) with low metabolic rates sleep only for 3–4 h. It has been suggested that high metabolic rates cause increased oxidative stress and injury to self-leading to the hypothesis [202] that higher metabolic rates in the brain require longer sleep time to counteract the cell damage by free radicals, and facilitate synthesis of molecules protecting brain cells from this oxidative stress. During NREM sleep, brain energy metabolism and cerebral blood flow decrease, whereas during REM sleep, the level of metabolism is similar to that of wakefulness and the cerebral blood flow increases. In summary, sleep-related reduction in general metabolism including metabolic heat production, lowering of core body temperature, and certain behavioral signs (e.g., immobile posture minimizing heat exchange) conserve energy. Although these findings might suggest that NREM sleep helps conserve energy, the fact that only 120 calories are conserved in 8 h of sleep makes the energy conservation theory less than satisfactory. Considering that humans spend one-third of their lives sleeping, [203] one would expect far more calories to be conserved during an 8-h period if energy conservation was the function of sleep.

Adaptive Theory
In both animals and humans, sleep is an adaptive behavior that allows the creature to survive under a variety of environmental conditions [204, 205].

Instinctive Theory
The instinctive theory views sleep as an instinct [190, 206], which relates to the theory of adaptation and energy conservation.

Memory Consolidation and Reinforcement Theory
The sleep–memory consolidation hypothesis is a hotly debated issue with both proponents and opponents with proponents
outnumbering the opponents. In fact, McGaugh et al. [207] earlier suggested that sleep- and waking-related fluctuations of hormones and neurotransmitters may modulate memory processes. Crick and Mitchison [83] suggested that REM sleep removes undesirable data from the memory. In a later report, these authors hypothesized that the facts that REM deprivation produces a large rebound and that REM sleep occurs in almost all mammals make it probable that REM sleep has some important biological function [208].

The theory that memory reinforcement and consolidation take place during REM sleep has been strengthened by scientific data provided by Karni et al. [209]. These authors conducted selective REM and SWS deprivation in six young adults. They found that perceptual learning during REM deprivation was significantly less compared with perceptual learning during SWS deprivation. In addition, SWS deprivation had a significant detrimental effect on a task that was already learned. These data suggest that REM deprivation affected the consolidation of the recent perceptual experience, thus supporting the theory of long-term consolidation during REM sleep. Recent studies by Stickgold and Walker [210] and Walker and Stickgold [211] strongly supported the theory of sleep–memory consolidation (see Chap. 13). There is further suggestion by Hu et al. [212] that the facilitation of memory for emotionally salient information may preferentially develop during sleep. Stickgold’s group concluded that unique neurobiological processes within sleep actively promote declarative memories [213]. Several studies in the past decade have provided evidence to support the role of sleep in sleep-dependent memory processing which include memory encoding, memory consolidation and reconsolidation, and brain plasticity (see review by Kalia) [214]. Hormung et al. [215] using a paired associate word list to test declarative memory and mirror-tracking tasks to test procedural learning in 107 healthy older adults aged 60–82 years concluded that REM sleep plays a role in procedural memory consolidation. Walker’s group [216] concluded after sleep deprivation experiments that sleep before learning is critical for human memory consolidation. Born et al. [217] concluded that hippocampus-dependent memories (declarative memories) benefit primarily from SWS. They further suggested that the different patterns of neurotransmitters and neurohormone secretion between sleep stages may be responsible for this function. Backhaus and Junghanns [218] randomly assigned 34 young healthy subjects to a nap or wake condition of about 45 mins in the early afternoon after learning procedural and declarative memory tasks. They noted that naps significantly improved procedural but not declarative memory, and therefore, a short nap is favorable for consolidation of procedural memory. Goder et al. [219] tested the role of different aspects of sleep for memory performance in 42 consecutive patients with non-restorative sleep. They used Rey–Osterrieth Complex Figure Test and the paired associative word list for declarative memory function and mirror-tracking tasks for procedural learning assessment. The results supported the contention that visual declarative memory performance is significantly associated with total sleep time, sleep efficiency, duration of NREM sleep, and the number of NREM–REM sleep cycles but not with specific measures of REM sleep or slow-wave sleep. In contrast to all these studies, Vertes and Siegel [220–223] took the opposing position contending that REM sleep is not involved in memory consolidation or at least not in humans citing several lines of evidence. Vertes and Siegel [220] cited the work of Smith et al. [224, 225] that REM sleep is not involved with memory consolidation. Schabus et al. [226] agreed that declarative material learning is not affected by sleep. In their study, subjects showed no difference in the percentage of word pairs correctly recalled before and after 8 h of sleep. The strongest evidence cited by Vertes and Siegel [220] includes examples of individuals with brain stem lesions with elimination of REM sleep [227] or those on antidepressant medications suppressing REM sleep exhibiting no apparent cognitive deficits. Vertes and Siegel [220] concluded that REM sleep is not involved in declarative memory and REM sleep is not critical for cognitive processing in sleep.

In summary, based on recent studies in our understanding of molecular mechanisms of memory, it can be stated reasonably that sleep strengthens new memories, i.e., sleep can rescue memories that are lost during wakefulness requiring synthesis of new proteins and ribonucleic acid (RNA) by the neurons. Memory consolidation and reconsolidation require both NREM (sleep spindles and SWS) and REM sleep.

**Synaptic and Neuronal Network Integrity Theory**

There is a new theory emerging that suggests the primary function of sleep is the maintenance of synaptic and neuronal network integrity [203, 228–230]. According to this theory, sleep is important for the maintenance of synapses that have been insufficiently stimulated during wakefulness. Intermittent stimulation of the neural network is necessary to preserve CNS function. This theory further suggests that NREM and REM sleep serve the same function of synaptic reorganization [228]. This emerging concept of the “dynamical stabilization” (i.e., repetitive activations of brain synapses and neural circuitry) theory of sleep suggests that REM sleep maintains motor circuits, whereas NREM sleep maintains nonmotor activities [228, 230]. Gene expression [231] studies using the DNA microarray technique identified sleep- and wakefulness-related genes (brain transcripts) subserving different functions (e.g., energy metabolism, synaptic excitation, long-term potentiation, and response to cellular stress during wakefulness; and protein synthesis, memory consolidation, and synaptic downscaling during sleep).
Thermoregulatory Function Theory

The thermoregulatory function theory is based on the observation that thermoregulatory homeostasis is maintained during sleep, whereas severe thermoregulatory abnormalities follow total sleep deprivation [232]. The preoptic anterior hypothalamic neurons participate in thermoregulation and NREM sleep (see also Chap. 11). These two processes are closely linked by preoptic anterior hypothalamic neurons but are clearly separate. Thermoregulation is maintained during NREM but suspended during REM sleep. Thermoregulatory responses such as shivering, piloerection, panting, and sweating are impaired during REM sleep. There is a loss of thermosensitivity in the preoptic anterior hypothalamic neurons during REM sleep.

Role of Sleep in Immune and Endocrinal Regulation

This has been briefly outlined above.

Metabolic Homeostasis Theory

Recent discovery of a new metabolic waste clearing pathway (equivalent to the lymphatic system in the body) in the CNS of mice directs our attention to a new understanding of the function of sleep. This also is beginning to answer a long-standing question of mechanism of clearing of metabolic byproducts of the brain in the absence of a lymphatic system. This pathway was termed “the glymphatic system” by the original researchers led by Nedergaard [233–235] because of its dependence on glial cells (astrocytes) performing a “lymphatic-like” cleansing of the brain interstitial fluid in the perivascular space between the brain blood vessels and leptomeningeal sheaths surrounding these vessels. These researchers have shown that this extracellular space expanded by about 60% along with shrinkage of the glial cells in the sleeping brain as compared with that in the waking brain to promote clearance of interstitial waste products [234] (Fig. 2.8). Sleep, therefore, restores the function of the brain by promoting glymphatic clearance of neural metabolic waste products accumulated in the waking brain. The investigators have further shown the ability of this system to remove misfolded or aggregated proteins by experimental injection of labeled beta amyloid proteins into the brains of sleeping and awake mice [236, 237]. Cerebrospinal fluid cleared these proteins outside of the cells twice as fast during sleep as during wakefulness. These findings may revolutionize our understanding of sleep dysfunction in many neurodegenerative diseases resulting from proteinopathies (e.g., Alzheimer’s disease, Parkinson’s disease, and others). Once these findings are replicated in humans using sophisticated neuroimaging techniques, this discovery will electrify the scientific world of sleep toward drug development to prevent or halt the progression of these neurodegenerative diseases.

Fig. 2.8 Physiological differences in cerebrospinal fluid (CSF) flow between the awake and the sleeping states in mice. The tissues perfused by CSF including blood vessels are indicated in red (arteries) and blue (veins). The interstitial space in the mouse cerebral cortex, through which cerebrospinal fluid moves, increases from 14% in the awake to 23% in the sleeping animal, an increase that allows the faster clearance of metabolic waste products. Reproduced with permission from Herculano-Houzel [240]
References

6. Ishimori K (1909) True causes of sleep—a hypnogenic substance as evidenced in the brain of sleep-deprived animals. Igakkaizasshi (Tokyo) 23:429
33. Hartley D (1749) Observations on man, his frame, his duty, and his expectations. Leake and Frederick, London
34. Macnish R (1830) The philosophy of sleep E. M’Phun, Glasgow
44. Chokroverty S. Sleep disorders atlas task force of the American sleep disorders association (preliminary report): EEG arousals, scoring rules and examples. Sleep 15:174
52. Scher MS (2008) Ontogeny of EEG sleep from neonatal through infancy periods. Sleep Med 9:615
73. Taub JM, Berger RJ (1976) Effects of acute sleep pattern alteration depend upon sleep duration. Physiol Psychol 4:412
89. Lyamin OL, Mukhametov IM, Siegel JM et al (2002) Unihemispheric slow wave sleep and the state of the eyes in a white whale. Behav Brain Res 129:125
95. Lyamin OL, Mukhametov IM, Siegel JM et al (2002) Unihemispheric slow wave sleep and the state of the eyes in a white whale. Behav Brain Res 129:125
96. Lyamin OL, Mukhametov IM, Siegel JM et al (2002) Unihemispheric slow wave sleep and the state of the eyes in a white whale. Behav Brain Res 129:125
190. Moruzzi G (1972) The sleep-waking cycle. Ergeb Physiol 64:1
Overview of Normal Sleep


220. Vertes R, Siegel JM (2005) Time for the sleep community to take a critical look at the purported role of sleep in memory processing. Sleep 28:1228


Sleep Disorders Medicine
Basic Science, Technical Considerations and Clinical Aspects
Chokroverty, MD, S. (Ed.)
2017, XIX, 1269 p. 340 illus., 156 illus. in color. In 2 volumes, not available separately. Hardcover