Pathogenesis of Obstructive and Central Sleep Apnea

David P. White
Division of Sleep Medicine, Harvard Medical School and Brigham and Women’s Hospital, Boston, Massachusetts

Considerable progress has been made over the last several decades in our understanding of the pathophysiology of both central and obstructive sleep apnea. Central sleep apnea, in its various forms, is generally the product of an unstable ventilatory control system (high loop gain) with increased controller gain (high hypercapnic responsiveness) generally being the cause. High plant gain can contribute under certain circumstances (hypercapnic patients). On the other hand, obstructive sleep apnea can develop as the result of a variety of physiologic characteristics. The combinations of these may vary considerably between patients. Most obstructive apnea patients have an anatomically small upper airway with augmented pharyngeal dilator muscle activation maintaining airway patency awake, but not asleep. However, individual variability in several phenotypic characteristics may ultimately determine who develops apnea and how severe the apnea will be. These include: (1) upper airway anatomy, (2) the ability of upper airway dilator muscles to respond to rising intrapharyngeal negative pressure and increasing Co2 during sleep, (3) arousal threshold in response to respiratory stimulation, and (4) loop gain (ventilatory control instability). As a result, patients may respond to different therapeutic approaches based on the predominant abnormality leading to the sleep-disordered breathing.

Keywords: Apnea, sleep, pathogenesis; central sleep apnea; obstructive sleep apnea; upper airway

CONTENTS
The Control of Upper Airway Patency
  Collapsing Forces on the Pharyngeal Airway
  Dilating Forces on the Pharyngeal Airway
  Ventilatory Control Stability
Central Sleep Apnea Pathogenesis
Obstructive Sleep Apnea Pathogenesis
  Pharyngeal Anatomy
  Pharyngeal Dilator Muscle Activation
  Loop Gain
  Pharyngeal Nerve/Muscle Damage
Conclusions

That abnormal patterns of ventilation can emerge during sleep has been increasingly recognized over the last two decades, a product primarily of rising awareness of the adverse neurocognitive and cardiovascular consequences of such disordered breathing (1, 2). We have also steadily improved our understanding of the mechanisms leading to this disordered breathing. With most such work being focused on two principal areas: (1) the mechanisms that influence upper airway patency awake and asleep and (2) ventilatory control stability during sleep (3, 4). Although problems with maintenance of upper airway patency result primarily in obstructive sleep apnea, ventilatory control instability can lead to either central or obstructive apnea depending on the collapsibility of the individual airway. Thus, these two general scientific areas (the upper airway and ventilatory control) will be addressed first followed by a discussion of how abnormalities or individual differences in either can lead to sleep apnea.

THE CONTROL OF UPPER AIRWAY PATENCY

In humans, the upper airway, from the posterior end of the nasal septum to the epiglottis, has relatively little bony or rigid support. This being the case, there are anatomic and physiologic influences that tend to collapse this position of the airway that must be offset by dilating forces if airway patency is to be maintained (Figure 1) (5). The two primary forces tending to collapse the airway are the intraluminal negative pressure generated by the diaphragm during inspiration and the extraluminal tissue pressure (that pressure resulting from tissue and bony structures surrounding the airway). These influences must be offset primarily by the action of pharyngeal dilator muscles, although longitudinal traction on the airway resulting from lung inflation likely contributes as well (6–9).

Collapsing Forces on the Pharyngeal Airway

Intraluminal negative pressure. Intraluminal negative pressure in a collapsible tube will inherently reduce the airway area. Thus, during each inspiration, the diaphragmatically generated negative pressure would diminish airway size depending on the compliance of the airway walls and opposing dilating forces. That being said, the airway pressure required to collapse the pharyngeal airway has been best described by the critical closing pressure (Pcrit), a concept developed and evolved by Schwartz and colleagues (10, 11). This closing pressure is dependent on many variables, most of which are addressed below. However, in terms of intraluminal pressure, Pcrit is not a product of hypopharyngeal pressure but rather the pressure upstream to the collapsing segment. Thus, the negative pressure generated by respiratory pump muscles can reduce airway size, but will generally not collapse the airway. This concept is addressed in detail by Schwartz and colleagues (10, 11).

Airway anatomy. Airway anatomy can also importantly influence pharyngeal patency. In normal, nonobese individuals, when muscle activity is completely inhibited (passive condition), the airway generally remains patent and requires about approximately 5 cm H2O to collapse (12). As a result, the extraluminal tissue pressure (the pressure in the soft tissue surrounding the airway) in these individuals must be 0 cm H2O, negative, or not sufficiently positive to overcome the elastance of the airway wall. Thus, the quantity of soft tissue located in the bony compartment created by the mandible and spinal column relative to the size of the
compartment is sufficiently small that it does not apply a collapsing pressure on the pharyngeal airway (Figure 2) (13). In patients with obstructive sleep apnea, this is not the case. When total muscle paralysis is induced (passive condition), the airway collapses and positive pressure is required to open the airway (12). As a result, the extraluminal tissue pressure must be positive and sufficiently high to overcome the elastance of the airway walls. Thus, in the patient with apnea, the quantity of tissue in the bony compartment relative to the size of the compartment is sufficient to generate a positive pressure on the airway, which partially or completely collapses it in the passive condition. One must assume that there is a continuum in humans from the completely open passive airway (0 or negative tissue pressure) to the firmly collapsed one (positive tissue pressure). Increasing tissue pressure (Figure 2) is likely a product of either fat deposition (extra soft tissue in a normal-sized bony compartment) or a crowding of normal pharyngeal structures into a smaller bony compartment (14). Physical structures that fill (partially or completely) the airway lumen can also increase collapsibility. An example would be tonsils or adenoids (15). Thus, abnormal anatomy can have an important effect on airway patency (16, 17).

Several additional factors may also affect anatomy or airway size. These include vascular perfusion, the posture of the individual (supine vs. lateral), airway secretions, and tissue microstructure. Of these mechanisms, posture is likely the most important, with it potentially having a strong influence on tissue pressure due primarily to the effects of gravity on airway tissues. In the supine posture, tongue and palatal structures would move posteriorly due to gravitational effects if not offset by other forces, thus generating a more positive tissue pressure. These mechanisms and others have been addressed in detail by Olson and colleagues (18) and Fouke and Strohl (19), and we refer the reader to these articles for more detail.

**Dilating Forces on the Pharyngeal Airway**

**Pharyngeal dilator muscle activation.** Pharyngeal dilator muscle activation is the primary process that counteracts the collapsing forces described above. There are a number of such dilator muscles that must behave in a complex and coordinated manner to maintain airway patency. These muscles are addressed in detail by van Lunteren and Strohl (20). These muscles may activate during inspiration (inspiratory phasic pattern) with less activity during expiration (21) or have a similar level of activity across the respiratory cycle (tonic pattern). The control of these muscles is complex and likely varies between muscles (3). The genioglossus, an inspiratory phasic muscle, is the best studied such muscle (16) and will be addressed primarily here, although the tensor palatini, a tonic muscle, will be discussed as well.

There are three primary neural inputs controlling the genioglossus muscle (Figure 3). First, negative pressure in the airway reflexively activates mechanoreceptors located mainly in the larynx, leading to superior laryngeal nerve afferent activity and ultimately increased hypoglossal output to the genioglossal muscle (22–24). Thus, any event that threatens airway patency will lead to increased negative airway pressure and subsequent genioglossal activation to counter the threat (25–29). As a result, airway patency is protected. Second, the respiratory pattern generating neurons in the medulla also influences genioglossal activation. This is most clearly demonstrated by the activation of the genioglossus muscle about 50 to 100 ms before diaphragmatic activation or the development of negative pressure in the airway (30). This activity, therefore, is not driven by mechanoreceptors but by respiratory neurons (31). Thus, during inspiration, both respiratory neurons and the negative-pressure reflex modulate...
muscle activity. Third, neurons that modulate arousal (active awake, less active, or inactive asleep), such as serotonergic or noradrenergic neurons, have a tonic excitatory influence on upper airway motoneurons such as hypoglossal motoneurons (32, 33). This has been called the “wakefulness stimulus” and generally increases muscle activity (34, 35). With these three inputs, pharyngeal muscle activity is linked to respiration, local conditions in the airway (negative pressure), and arousal state (wake vs. sleep) (36).

With the onset of sleep, the control of these muscles changes importantly. The negative-pressure reflex is substantially reduced during non-REM sleep and further during REM sleep (26, 37–40). It is not lost and the muscles can still respond to negative pressure, but not as effectively or as quickly as when a person is awake. The “wakeful” input to these muscles is diminished during sleep as well, which may explain loss of tonic (expiratory) activity in phasic muscles such as the genioglossus and the large decrement in the activation of purely tonic muscles such as the tensor palatini (41). Finally, although minimally studied, the respiratory input to these muscles is likely maintained or reduced minimally during sleep. Thus, there is a general fall in pharyngeal muscle activity at sleep onset due to loss of “wakeful” neuronal input and a reduced ability to defend the airway due to diminished negative-pressure responsiveness (32). Thus, the airway becomes vulnerable. The neural processes in the brainstem driving these events during sleep are becoming increasingly well understood but are beyond the scope of this discussion (42).

Changes in lung volume. Changes in lung volume can also importantly influence pharyngeal patency. Lung inflation applies a caudal traction on the trachea and larynx, thereby inducing a longitudinal tension on the pharyngeal airway (7, 8). Because of the redistribution of tissue at higher lung volumes, tissue pressure may change as well. This caudal force tends to stiffen the airway and reduces collapsibility. This has been convincingly demonstrated in animals using a variety of approaches and more recently in humans as well (7–9). Thus, decrements in lung volume, which can occur with changes in posture (upright to supine) or transitions from wakefulness to sleep, result in less tension on the airway walls (43). As a result, the extraluminal tissue pressure required to collapse the airway is importantly reduced at lower lung volumes.

VENTILATORY CONTROL STABILITY

The quantity and pattern of ventilation in humans is tightly regulated to both maintain oxygen and carbon dioxide levels within narrow limits and to minimize the work required to accomplish this. This is a product of multiple feedback loops, including the chemoreceptors (O₂ and CO₂), intrapulmonary receptors, and respiratory muscle afferents, among others. A change in Pco₂ will therefore initiate a series of events to quickly correct the problem and return Pco₂ to the desired level. This is generally the case during both wakefulness and sleep, although during sleep, the set point for CO₂ rises and several ventilatory control mechanisms deteriorate (44, 45). For instance, the ability to respond to increments in the work of breathing (resistive loading) is substantially reduced during all stages of sleep, as are the chemoresponses during REM sleep (46). However, despite these changes, ventilation during sleep is largely regulated by the same mechanisms that govern breathing while awake, although most behavioral influences on breathing are lost during sleep.

Any mechanical system that is regulated by feedback loops, such as the respiratory system, has the potential to become unstable. This is best explained in the context of loop gain. Loop gain is an engineering term used to describe the overall gain of any system controlled by feedback loops (4, 47–50). A high-gain system responds quickly and vigorously to a perturbation, whereas a low-gain system responds more slowly and weakly. The two primary variables influencing loop gain are what are known as controller gain and plant gain, and both are important in ventilatory stability. Controller gain is synonymous with chemoresponsivity or the hypoxic and hypercapnic ventilatory responses. Thus, a high controller gain is generally due to brisk hypercapnic responsiveness. Plant gain largely reflects the effectiveness of a given level of ventilation to eliminate CO₂. Thus, high plant gain would occur if a small change in ventilation produced a large change in Pco₂. Conditions causing a high plant gain include low functional residual capacity, low dead space, low metabolic rate, low cardiac output, and a high Pco₂.

For a system to become unstable (waxing and waning ventilation), two conditions must be met. First, there must be a phase delay between the effector portion of the system (the lungs) and the sensor for the system (CO₂ detection in the carotid body and brainstem). This is always the case for the respiratory system as there is an inherent delay between blood gas changes in the
Figure 4. The ventilatory response to an apnea (first disturbance in both figures) is demonstrated for (A) an individual with a loop gain (LG) of 0.5 and (B) an individual with an LG of 1. In A, ventilation quickly returns to a regular pattern, whereas in B, a sustained oscillation is established.

Figure 5. Diagrammatic representation of the relationship between alveolar ventilation (VA) and alveolar PCO₂ (PACO₂) at a fixed (resting) VCO₂ (250 ml/min; PₐCO₂ = [VCO₂/VA] × K). See text for full explanation. Reprinted by permission from Reference 51.
altitude is high enough (53). This is a form of ventilatory instability produced by ambient hypoxia (low barometric pressure). Hypoxia at altitude leads to high controller gain, resulting in hyperventilation and hypocapnia. In this situation, the increased control gain is sufficient to overcome the somewhat reduced plant gain (low PCO₂) and ventilation becomes unstable. As a result, ventilation waxes and wanes between apnea and hyperpnea. For this to occur, there must be an adequate hypoxic ventilatory response to drive the hyperventilation (high loop gain) resulting in hypocapnia (54). Thus, the periodic breathing at altitude occurs more commonly in individuals with high controller gain (53). However, most such individuals do not have periodic breathing at sea level. Thus, an inherently high loop gain that is augmented further by altitude (hypoxia) is required for this respiratory pattern to emerge. Over time at altitude (other than at extreme elevations), the periodic breathing resolves due primarily to further decrements in plant gain that cannot be offset by a high controller gain.

Idiopathic central sleep apnea is a relatively uncommon disorder seen at sea level in individuals with high controller gain, generally an elevated hypcapnic ventilatory response. PCO₂ drives ventilation during sleep at sea level as it does at altitude, and patients with idiopathic central sleep apnea tend to have low PCO₂ levels, even during wakefulness (55, 56). Figure 5A would suggest that an individual with high baseline respiratory output (high VA and low Pₐ₄C₇O) would require a large change in ventilation to reach the apnea threshold (low plant gain). However, as demonstrated in Figure 5B, a sufficiently high controller gain can close this gap and render ventilation unstable (48). This disordered breathing occurs primarily during non-REM sleep because chemosensitivity (controller gain) is reduced during REM sleep and is not sufficient to drive the cycling respiration. As expected, these central apneas can be abolished with inhaled CO₂ (57).

A third form of central sleep apnea called Cheyne-Stokes respiration is generally seen in patients with congestive heart failure and is a product of high controller gain (increased CO₂ responsiveness), hypocapnia resulting from lung edema (high filling pressures), and a long circulation time (1, 58–60). This combination of traits is particularly destabilizing to ventilation and yields a characteristic crescendo–decrescendo pattern of breathing, with a cycle time of approximately 1 min. As expected in a high loop gain disorder, CO₂ administration can regularize ventilation (61). As in idiopathic central sleep apnea, Cheyne-Stokes respiration occurs primarily during non-REM sleep, although it can be detected during wakefulness if carefully sought. However, it is uncommon during REM sleep, again likely secondary to decreased controller gain. Because both traits described above (high controller gain and long circulation time) are generally required to manifest Cheyne-Stokes respiration, it is not seen in all patients with congestive heart failure, even those with severe congestive heart failure (62).

Finally, patients with waking hypcapnia primarily due to ventilatory control abnormalities (obesity hypoventilation syndrome, central alveolar hypventilation) or neuromuscular disease may have central apneas during sleep as well. This is likely a product of the high plant gain in such patients (Figure 5A) or an absence of ventilatory drive during sleep when respiration is largely dependent on these chemical control mechanisms (63, 64).

**Obstructive Sleep Apnea Pathogenesis**

On the basis of the information provided above, it would seem likely that obstruction of the pharyngeal airway could occur during sleep in response to a number of possible physiologic traits or combinations of traits, which are discussed in the following sections.

**Pharyngeal Anatomy**

As stated above, most patients with obstructive apnea have an anatomically small pharyngeal airway likely due to either increased soft tissue surrounding the airway or a small bony compartment in which the airway is enclosed (16, 65–68). During wakefulness, pharyngeal patency is maintained primarily by reflex-driven, augmented pharyngeal dilator muscle activity, which offsets the positive extraluminal tissue pressure collapsing the airway (Figure 1) (29, 69). Thus, normal ventilation is maintained. At sleep onset and/or during REM sleep, reflex muscle activation is reduced as is arousal (wakefulness)-modulated excitatory output to the upper airway musculature (38). Lung volume falls as well (43). If the airway anatomy is quite deficient, these events alone will likely lead to substantial or complete airflow obstruction, yielding a hypopnea or apnea. As a result, hypoxia and hypocapnia develop, ventilation is stimulated, and often arousal from sleep in response to respiratory activation is required to reestablish airway patency to allow a recovery of ventilation (70–72). With the next sleep onset, this cycle will repeat.

**Pharyngeal Dilator Muscle Activation**

The scenario leading to pharyngeal obstruction described above makes one fundamental assumption. It assumes that the loss of pharyngeal dilator muscle activation associated with sleep and leading to airway obstruction cannot be recovered without arousal. This is not likely to be completely correct (73). Clearly, pharyngeal dilator muscles can respond to mechanical (negative-pressure) and respiratory (CO₂) stimulation during sleep (74, 75). A number of studies have shown that the upper airway muscles in normal subjects can be recruited during sleep in response to resistive loading and rising CO₂ (74). Patients with apnea, over the course of an apnea, also demonstrate a considerable increment in pharyngeal muscle activation that likely results from similar mechanisms (71). These responses are not as brisk or as rapid asleep as awake, but they are present.

The question then becomes whether this increasing muscle activity can overcome the inherently deficient pharyngeal anatomy and collapsing negative inspiratory airway pressure during sleep. Although this has not been formally studied, most evidence suggests that this is often the case. The most supportive piece of evidence for this is the observation that virtually all patients with obstructive sleep apnea have normal rhythmic ventilation at least part of the time during sleep (73). This most often occurs during stable non-REM sleep (i.e., stages 2, 3, and 4). Assuming no change in sleeping position, one has to assume that the upper airway dilator muscles have been sufficiently recruited to maintain pharyngeal patency. This concept is further supported by the fact that it generally requires a relatively long period of time (at least 1–2 min) to reach stages 3–4 sleep from wakefulness. Thus, the patient with apnea must stay asleep for several minutes to achieve slow-wave sleep. This would likely be sufficient time to adequately recruit upper airway dilator muscles that support pharyngeal patency, thereby allowing stable sleep and breathing. This assumes, however, that the muscles can be recruited during sleep and that the individual can stay asleep long enough to allow such recruitment to occur. There is likely considerable variability in both, as is addressed below.

A number of studies suggest that the ventilatory effort (esophageal pressure) required to induce arousal from a respiratory stimulus is quite variable both in normal subjects (Figure 6) and in patients with apnea (70, 76). One would speculate therefore that an individual with a low arousal threshold would never reach stable sleep (77). If such an individual fell asleep and upper airway resistance rose substantially, there would not be adequate...
time to recruit upper airways muscles before arousal (4). Thus, the individual would wax and wane between sleep and wakefulness and between hypopnea and hyperpnea. As a result, individual variability in arousal threshold could determine whether one reaches stable sleep with activated upper airway muscles or arouses before adequate recruitment of the pharyngeal musculature.

Similar principles probably apply to the activation of the upper airways muscles during sleep as well. Although there has not been a systematic study assessing individual variability in muscle responsiveness during sleep to a collapsing upper airway, several observations suggest such variability exists. First, Younes (73) observed that “compensatory effectiveness” accounted for a considerable portion of the variability in apnea severity in a cohort of patients with apnea. This compensatory effectiveness represents the ability of an individual to augment inspiratory flow during a “drop down” in continuous positive airway pressure. The decreased continuous positive airway pressure leads to a flow limited condition. Over time, in some individuals, substantially improved air movement develops before arousal. This improvement in flow is almost certainly a product of increased pharyngeal muscle recruitment. In the flow-limited condition, increased ventilatory effort does not lead to augmented airflow. Thus, muscle recruitment must occur in some individuals during sleep. Second, studies that have examined the passive $P_{cm}$ demonstrate considerable overlap between normal control subjects, snorers, and individuals with frequent apneas and hypopneas (10, 78). Thus, two individuals may have the same $P_{cm}$ (probably our best measure of passive anatomy) yet one has obstructive sleep apnea while the other has little if any disordered breathing (79). This suggests remarkably different abilities of the upper airway muscles to compensate for deficient anatomy during sleep, as we have observed (Figure 7). As a result, individual variability in both arousal threshold and muscle responsiveness during sleep may importantly influence whether a given person develops sleep apnea or not.

**Loop Gain**

As stated above, loop gain is a measure of the stability or instability of a system controlled by feedback loops, with a high gain indicating inherent instability. The question becomes what role does loop gain play in the development of obstructive sleep apnea. A number of observations suggest that loop gain can contribute to obstructive apnea. First, the upper airways muscles are quite responsive to the respiratory system with their activity increasing or decreasing substantially depending on respiratory drive. Thus, if respiratory drive is waxing and waning (unstable ventilatory control), then the activity of the pharyngeal musculature will do so as well (80). This can promote upper airway collapse at the nadir of such respiratory cycling. Second, it has been demonstrated that airway collapse (partial or complete) is common during a central apnea (81). Thus, if respiratory drive is eliminated (central apnea), the upper airways muscles are turned off as well and the pharynx may collapse depending on its passive characteristics. Third, hypoxia-induced periodic breathing can lead to obstructive hypopneas and apneas in individuals in whom only snoring was observed under normoxic conditions (82, 83). Thus, destabilizing ventilatory control in someone with a moderately collapsible airway (a snorer) can lead to apneas. These tend to be mixed apneas or at least apneas with quite variable ventilatory effort. All of these observations suggest a role for ventilatory control instability in the pathogenesis of obstructive sleep apnea.

This concept is supported by several more direct assessments as well. Hudgel and colleagues (50), using pseudorandom binary CO$_2$ stimulation during wakefulness (another measure of loop gain), reported higher loop gain in patients with apnea than in control subjects. Younes and coworkers (4) observed that patients with severe obstructive sleep apnea have higher loop gain during non-REM sleep than patients with mild apnea, although direct comparisons with normal control subjects were not attempted. Finally, Wellman and colleagues (79) measured both loop gain and passive $P_{cm}$ in a group of patients with apnea during stable, supine non-REM sleep. Although loop gain correlated loosely with apnea severity for the entire group ($r = 0.36$, $p = 0.07$), the relationship was much tighter ($r = 0.88$, $p < 0.01$) in patients with a moderately collapsible airway ($P_{cm} > -1$, but $< +1$). Patients with a highly collapsible airway ($P_{cm} > +1$) may have apnea regardless of loop gain and those with a less collapsible airway ($P_{cm} < -1$) may not be adequately sensitive to respiratory-induced changes in pharyngeal dilator muscle...
activity. However, all of these studies suggest that increased loop gain may play a role in apnea pathogenesis in some individuals.

Pharyngeal Nerve/Muscle Damage

There is an evolving literature suggesting that other variables may play a role in apnea pathogenesis or propagation as well. Several studies suggest that there may be inflammation and trauma to the upper airway due to snoring, vibration, extrinsic contraction, or fatigue (84–87). This may lead to a reduction in sensory mechanisms in the pharyngeal airway (ability to detect negative pressure), derervation of pharyngeal dilator muscles, or actual damage to the muscles themselves. Any event that renders the upper airway muscles less able to generate force or to respond to negative airway pressure certainly has the potential to cause or worsen sleep apnea. However, at this point, it is unclear how important these mechanisms are in apnea pathophysiology.

CONCLUSIONS

Most evidence would suggest that there are a number of phenotypic traits that predispose an individual to the development of sleep apnea. In the case of central apnea, this generally relates to loop gain and circulation time. For obstructive apnea, pharyngeal anatomy, upper airway muscle responsiveness during sleep, arousal threshold, and loop gain may all contribute to apnea presence and severity. The relative contribution of each may vary between patients. It is unclear at this time whether defining these traits in patients with apnea would have therapeutic implications, although this seems possible. Thus, as always, more information is needed.

Conflict of Interest Statement: D.P.W. is a consultant for the following companies: Respironics, Inc.; Alfred E. Mann Foundation; Aspire Medical; and Itamar Medical. He receives grants from Respironics, Alfred E. Mann Foundation, and Cephalon. He has received lecture fees from Orphan Medical and Cephalon.

References


