Controversies in Neuroscience V: Persistent pain: Neuronal mechanisms and clinical implications: Introduction

Pain is not a single entity but is instead a collection of sensory experiences commonly associated with tissue damage. There is growing recognition that not all pains are equivalent, that pains and pathologies are not related in a simple manner, and that acute pains differ in many respects from persistent pains. Great strides have been made in improving our understanding of the neuronal mechanisms responsible for acute pain, but the studies leading to these advances have also led to the realization that a bewildering array of processes are interposed between tissue damage and sensations of pain, especially in persistent pains. Persistent pains often seem unrelated or disproportionate to identifiable pathology, and they are modulated by a multitude of factors. This complexity in such a vital function serves as a challenge both to scientists seeking fundamental understanding and to clinicians faced with the immediate need to treat patients with painful disorders.

The understanding of persistent pain poses special challenges as compared to acute induced pain. In persistent pain, the responses of primary afferent nociceptors change in response to sustained sensory input. The sensitivity of these primary afferent neurons increases or decreases over time; they express and transport different modulators and receptors; and their activity contributes and responds to peripheral inflammation, which can, in turn, produce changes in central excitability and cause neurotoxicity. Central and peripheral modulatory processes and hormones have significant roles in persistent pains. The sympathetic nervous system modulates inflammatory and sensory processes and supraspinal systems can also play important roles depending on such factors as sensory input, prior history, expectation, fear, hormones, and reinforcement. Thus, persistent pain is influenced by a large number of factors with potentially complex interactions.

Controversies in the understanding of persistent pains were the focus of a symposium entitled "Controversies in Neuroscience V: Persistent Pain: Neuronal Mechanisms and Clinical Implications" upon which this issue of *BBS* is based. The symposium, hosted by the Robert S. Dow Neurological Sciences Institute in August 1994, brought together scientists and clinicians for the purpose of discussing unresolved, important issues related to persistent pains. The formats of both the symposium and the updated, refereed versions of the papers that appear in this issue of *Behavioral and Brain Sciences* were designed to encourage the expression of diverse views, stimulate debate, and catalyze further efforts to understand and better manage persistent pains. Six target articles by the primary speakers are presented, followed by commentaries and the author's responses to the commentaries.

The target articles by **BERKLEY** and **MCMAHON** address sex differences in pain and differences between visceral and somatic afferents related to pain, respectively. **BERKLEY** reviews reported sex differences in response to acute induced pain and the influence of factors such as menstrual status and other physiological and hormonal conditions that contribute to these differences. Most of her target article deals with sex differences in endogenous pains and the complexities and reported contradictions therein. She reviews anatomical, hormonal (specifically hormonal influences in GABA-ergic, opioidergic, and non-opioidergic inhibitory systems and growth factors), and sympathetic nervous system differences related to sex. This section, relevant commentaries, and Berkley's response to these commentaries illustrate clearly the complexities of this area of research.

MCMAHON points out the importance of pains related to the viscera and the relative paucity of information about the primary visceral afferents. He provides an extensive review of the features of visceral pains and how visceral and somatic pain differ, including anatomical, physiological, neurochemical and trophic factors, with a primary focus on peripheral processes. This review makes the point that pains should not all be considered equivalent, and that diagnosis and management of a particular pain syndrome must take into account the unique features of that pain.

The next three target articles by **DICKENSON, CODERRE & KATZ,** and **WIESENFELD-HALLIN et al.** deal with issues related to plasticity in nociceptive systems and the effects of plasticity on peripheral and central processing of nociceptive information. **DICKENSON** focuses primarily on modulatory processes in the central nervous system, particularly opioids, their diverse central actions, and plasticity in those systems. He also explores other substrates related to central hypersensitivity, including peptides, excitatory amino acids, and nitric oxide. Clinical implications and therapeutic strategies are discussed relative to each of these substrates.

CODERRE & KATZ review clinical and experimental evidence for peripheral and central hyperexcitability in persistent pains. Neuronal hyperexcitability due to injury or disease is an area of intense study in pain research, in part because of its clinical importance. Controversies about the importance of central versus peripheral components of hyperexcitability and about interactions between the two are numerous, as demonstrated by the commentaries appearing in this issue. Arguments are presented regarding different therapeutic strategies for dealing with persistent pains directed toward central or peripheral processes. Coderre & Katz acknowledge that peripheral processes are important, but they stress that central neuroplasticity is critical to pathological persistent pain states.

WIESENFELD-HALLIN et al. provide an insightful review of dysfunctions in central inhibitory systems. Experimental evidence for changes in GABA-ergic systems after peripheral nerve injury or spinal cord ischemia is reviewed and symptomatic parallels between the experimental animals and injured humans are noted. These authors also review experimental and clinical evidence for pains that are insensitive to opiates and they suggest that upregulation of cholecystokinin in primary afferent neurons following nerve injury is likely to antagonize the actions of endogenously or exogenously administered opioids.

In the last paper, **BLUMBERG et al.** reappraise the involvement of the sympathetic nervous system in pain syndromes in humans. These authors describe three symptomatically distinct types of syndromes with potential sympathetic involvement and provide evidence for different underlying mechanisms in each syndrome. This target article and the commentaries it evoked illustrate the difficulties of studying complex pain syndromes in humans. The diversity of opinions regarding sympathetically mediated pain syndromes reminds us that, in research, what is observed and reported is often constrained by the conceptual frameworks of the observers. Informed public debate, like that offered here, is an important process in advancing our understanding of persistent pain syndromes.

Editors:
Bill Roberts
Paul Cordo
Stevan Harnad

 0140-525X/97 $9.00+.10

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 371–380
Printed in the United States of America

Sex differences in pain

Karen J. Berkley
Program in Neuroscience, Department of Psychology, Florida State University, Tallahassee, FL 32306–1051
Electronic mail: *kberkley@neuro.fsu.edu*

Abstract: Are there sex differences in pain? For experimentally delivered somatic stimuli, females have lower thresholds, greater ability to discriminate, higher pain ratings, and less tolerance of noxious stimuli than males. These differences, however, are small, exist only for certain forms of stimulation and are affected by many situational variables such as presence of disease, experimental setting, and even nutritive status. For endogenous pains, women report more multiple pains in more body regions than men. With no obvious underlying rationale, some painful diseases are more prevalent among females, others among males and, for many diseases, symptoms differ between females and males. Sex differences in attitudes exist that affect not only reporting, coping, and responses to treatment, but also measurement and treatment. So many variables are operative, however, that the most striking feature of sex differences in reported pain experience is the apparent overall lack of them. On the other hand, deduction from known biological sex differences suggests that these are powerful sex differences in the operation of pain mechanisms. First, the vaginal canal provides an additional route in women for internal trauma and invasion by pathological agents that puts them at greater risk for developing hyperalgesia in multiple body regions. Second, sex differences in temporal patterns are likely to give rise to sex differences in how pain is "learned" and stimuli are interpreted, a situation that could lead to a greater variability and wider range of pains without obvious peripheral pathology among females. Third, sex differences in the actions of sex hormones suggest pain-relevant differences in the operation of many neuroactive agents, opiate and nonopiate systems, nerve growth factor, and the sympathetic system. Thus, while inductive analysis of existing data demonstrate more similarities than differences in pain experience between females and males, deductive analysis suggests important operational sex differences in its production.

Keywords: γ-aminobutyric acid (GABA); female; male; nerve growth factor; neuroactive peptides; sex hormones; sympathetic nervous system

Are there sex differences in pain? Ask the opinion of almost anyone and the answer will usually be yes (Bendelow 1993; McCaffery & Ferrell 1992). In fact, when carefully reviewed, evidence can be found for sex differences in virtually every sensory realm (Velle 1987). Consensus disappears, however, on what the differences are.

1. Assessments of somatic stimuli delivered under experimentally controlled circumstances

A number of psychophysical studies in humans have been carried out over the years on sex differences in the attribution of intense somatic stimulation as painful (e.g., see reviews in Ellermeier & Westphal 1995; Feine et al. 1991; Fillingim & Maixner 1995; Goolkasian 1985; Lander et al. 1990; Lautenbacher & Rollman 1993; Maixner & Humphrey 1993; Procacci et al. 1972; Rollman 1993; Velle 1987). When differences are observed under these carefully controlled experimental circumstances, it is often the case that women have lower thresholds, rate similar stimuli as more painful, or have less tolerance for intense stimuli.

Of interest is that certain parameters may be important for observing such differences. One such parameter is the type of stimulus (Lautenbacher & Rollman 1993). Thus, ratings of electrical and pressure stimuli are more reliably observed to exhibit sex differences than are ratings of thermal stimuli. The timing characteristics of the stimulus are also important, sex differences being more readily observed with less temporal summation (Hogeweg et al. 1992; Lautenbacher & Rollman 1993). Spatial aspects may also be important, including the size or bodily locus of the stimulus (Hogeweg et al. 1992; Lautenbacher & Rollman 1993; Lee & Essick 1993; Lipman et al. 1990).

Factors other than stimulus characteristics are also relevant. Situation variables are important. For example, the sex of the experimenter can affect sex differences in pain estimates when this factor is exaggerated by selecting "attractive" experimenters (e.g., compare results of Levine & DeSimone 1991 with Feine et al. 1991). The setting is also relevant. Somatic stimuli that are delivered in clinical settings (such as venipuncture and postoperative incisional cleaning) fail to show sex differences (Lander et al. 1990).

Other variables impact differently on males and females in a manner that would affect conclusions about the nature of any sex differences in pain. Thus, the presence of other disease conditions can affect pain ratings. In dysmenorrheic women, pressure-pain estimates are decreased during the follicular phase of the menstrual cycle (Hapidou & DeCatanzaro 1988), whereas abdominal muscle becomes hyperalgesic throughout the cycle (i.e., pain thresholds are decreased), with maximum hyperalgesia appearing premenstrually (Giamberardino et al. 1995). Pressure-pain thresholds are elevated in women suffering from bulimia nervosa (Faris et al. 1992).

Menstrual phase and reproductive status also affect pain ratings (Cogan & Spinnato 1986; Gintzler 1980; Goolkasian 1980; 1985; Hapidou & DeCatanzaro 1988; Procacci 1993; Tedford et al. 1977; Whipple et al. 1990), although the actual nature of the variations is inconsistent and variations

 0140-525X/97 $9.00+.10

are not always observed, especially if thermal or ischemic stimuli are used (e.g., Amodei & Nelson-Gray 1989; Dunbar et al. 1988). Sex differences in sugar and fat consumption are also likely to impact pain ratings (Frye et al. 1994; Krahn et al. 1994). And finally, willingness to report is also a likely factor in pain ratings, but the issue is difficult to resolve because of the difficulty in separating sensory from response factors (see discussions in Chapman 1977; Clark 1994; Ellermeier & Westphal 1995; Lander et al. 1990; Rollman 1977; Rollman & Harris 1987).

Data on sex differences in nociception in animals also exist (i.e., in their responses to stimuli that produce or threaten to produce injury). As reviewed by Bodnar et al. (1988), there are sex differences in basal nociceptive thresholds in rodents and these differences can be hormonally modulated and shown to vary with estrous condition, a result also recently observed by others (Forman et al. 1989; Frye et al. 1992; 1993; Gintzler & Bohan 1990; Martinez-Gomez et al. 1994). Similarly, chronic and acute sucrose consumption affects nociceptive responses differently in female rodents (Frye et al. 1992; 1993).

Thus, overall, experimental studies using exogenously delivered somatic stimuli indicate that sex differences in pain reports do exist, with females generally reporting lower thresholds, higher ratings, and less tolerance. However, the differences are small and inconsistently observed, being most prominent for pressure or electrical stimuli. Furthermore, even under these rigidly controlled experimental circumstances, the presence and direction of the differences are influenced by situational, health related, hormonal, motivational, and nutritional factors.

2. Assessments of endogenous pains

The issue of sex differences becomes even more complex for clinical situations when one considers reports by humans of their own endogenous pains (i.e., pains not evoked by others via experimentally delivered noxious stimuli) or of the endogenous pains of others. The literature is immense (see review in Unruh 1996), in large part because there is a wider range and less control over the research conditions. In addition, relevant information is often buried in the results sections of studies directed at other questions.

Many variables enter into any attempts to make sense of the available data (Mendelson 1991; Turk & Melzack 1992). These variables include *what* is being measured, *where* the measures are being taken, *who* is being studied, *when* they are studied, and *how* they are studied. As examples, for "what," clearly relevant to the results is whether it is physiological or psychological measurements that are being made. For "where," the setting in which measurements are taken is of clear importance (e.g., hospital, home, college classroom, survey in clinic waiting room, etc.). For "who," clearly significant is the age (Bodnar et al. 1988; Von Korff et al. 1988), ethnicity (Bates 1987), and health condition (Basoglu 1992) of the subjects. Also significant is the population from which the subjects are drawn; for example, the general population or a clinical one (Crook et al. 1989). For "when," time of day is of importance, as is the interval (e.g., days, weeks, months, etc.) over which the pain is being studied, as is whether subjects are being asked to estimate current or previous pains (e.g., Von Korff 1992). And, finally, for "how," pain diaries clearly differ from various scales such as visual analogue scales (VAS), from brain imaging techniques, from verbal reports, from questionnaires, and so on. A good example is the current debate about not only the value of but also what the best approaches are to assessing "pain behavior" (Keefe & Dunsmore 1992a; 1992b).

Given all of these interacting variables, it is at present possible to derive only a few general conclusions on the nature of sex differences in endogenous pain. Health surveys in both North America and Europe show that women's overall morbidity and use of health care is higher than that of men (Gijsbers van Wijk et al. 1991). Of importance, however, is that much of the higher morbidity in women can be accounted for by specific gynecological or obstetrical problems (Gijsbers van Wijk et al. 1992). Another important factor may be a willingness on the part of women to perceive and report physical symptoms as indicative of illness (Gijsbers van Wijk et al. 1991; Grove & Hughes 1979), although, as discussed above, it is difficult to separate the response and sensory factors.

This same sex difference appears to hold true, but considerably less so for pain complaints, many of which are musculoskeletal or visceral, and persistent, episodic, or chronic, unlike the acute somatic stimuli studied experimentally. In general, when large-scale studies of pain prevalence are carried out, chronic pain prevalence is not always higher in women than men (Andersson et al. 1993). However, women do report more multiple or recurrent pains than men, particularly more in certain body regions and at certain ages (Andersson et al. 1993; Eggen 1993; Ektor-Andersen et al. 1993; James et al. 1991; Klonoff et al. 1993; Von Korff et al. 1988; 1990), although, as noted in some of these studies, some of the differences appear to be related specifically to gynecological or obstetric problems (see review in Unruh 1996).

Studies on specific disorders in which pain is a prominent component provide additional information on the nature of sex differences in pain. The recently published *Classification of chronic pain* (Merskey & Bogduk 1994) describes about 500 such disorders and provides information about sex prevalence for about 85 of them. This information is summarized in Table 1.

As discussed in Berkley (1993), the data on sex prevalence as shown in Table 1 are currently derived from such disparate, uncertain, and often anecdotal sources that it is unclear at this time how to interpret them. Bush and colleagues (1993), in a detailed study of temporomandibular disorder, provide a strong argument that virtually all of the female predominance in that disorder can be explained by a "greater health awareness or interest in symptoms by women than by men." They go on to suggest that this explanation might also apply to other disorders. As discussed above, this explanation has logical problems, and, in any case, is difficult to apply to all of the conditions listed in Table 1, particularly those with a male predominance. Other tentative interpretations of the sex prevalence data are considered in the section 3 of this target article ("A switch from induction to deduction").

A factor that further complicates this issue are sex differences within individual disorders in the nature of the symptoms that are reported. For example, the diagnosis of *irritable bowel syndrome* (IBS) can be one of exclusion; that is, "chronic abdominal pain with no apparent cause associated with alteration of bowel habit" (Merskey &

Table 1. *Sex prevalence of various painful disorders*

Female Prevalence	Male Prevalence	No Sex Prevalence
migraine headache with aura	migraine without aura	acute tension headache
chronic tension headache	cluster headache	cluster-tic syndrome
post-dural puncture headache	post traumatic headache	"jabs" and "jolts" syndrome
hemicrania continua	SUNCT syndrome	secondary trigeminal neuralgia
cervicogenic headache	Raeder's paratrigeminal syndrome	neuralgia of nervus intermedius
tic douloureux	Pancoast tumor	painful ophthalmoplegia
temporomandibular joint disorder	thromboangiitis obliterans	maxillary sinusitis
occipital neuralgia	brachial plexus avulsion	toothache due to dentinoenamel defects
periapical periodontitis & abscess	pancreatic disease	toothache due to pulpitis
atypical odontalgia	duodenal ulcer	cracked tooth syndrome
burning tongue	abdominal migraine	dry socket
carotidynia	lateral femoral cutaneous neuropathy	vagus nerve neuralgia
chronic paroxysmal hemicrania	post herpetic neuralgia	stylohyoid process syndrome
temporal arteritis	hemophilic arthropathy	thoracic outlet syndrome
carpal tunnel syndrome	ankylosing spondylitis	brachial plexus tumors
Raynaud's disease		esophageal motility disorders
chilblains		chronic gastric ulcer
causalgia		Chron's disease
reflex sympathetic dystrophy		diverticular disease of colon
hemicrania continua		carcinoma of the colon
chronic venous insufficiency		familial Mediterranean fever
fibromyalgia syndrome		hereditary coproporphyria
esophagitis		acute herpes zoster
reflux esophagitis with peptic ulcer		burns
slipping rib syndrome		
twelfth rib syndrome		
gallbladder disease		
post-cholecystectomy syndrome		
irritable bowel syndrome		
interstitial cystitis		
acute intermittent porphyria		
proctalgia fugax		
chronic constipation		
pyriformis syndrome		
peroneal muscular atrophy		
multiple sclerosis		
rheumatoid arthritis		
pain of psychological origin		

Age Dependent Sex Differences	
Female Prevalence	Male Prevalence
gout (after age 60)	gout (before age 60)
osteoarthritis (after age 45)	osteoarthritis (before age 45)
livedo reticularis (after age 40)	coronary artery disease (before age 65)
	erythromelalgia (over age 50)

Source: Compiled from Merskey & Bogduk (1994) and others (see text).

Bogduk 1994). To help reduce the number of investigative procedures and referrals, many gastroenterologists have followed a differential diagnostic process using as criteria for IBS the existence of some proportion of a set of six specific gastric symptoms described by Manning et al. in 1978. It has been shown, however, that these specific criteria are of diagnostic value only in women (correlation was 0.47 for women; 0.16 for men; Smith et al. 1991), which is one of the factors that led to the adoption of a different criteria set (Rome criteria; see Drossman 1994).

Similarly, for *acute appendicitis,* while men and women share some similar symptoms (tenderness, rigidity, guarding, leucocytosis, location of pain at diagnosis), other symptoms differ considerably. For men, but not women, significant predictors include previous abdominal surgery, rectal digital tenderness, rebound, and elevated body temperature, whereas for women, but not men, the absence of renal tenderness is a good predictor (Eskelinen et al. 1994). *Migraine headaches* provide another example. Women show a prevalence of migraine without aura twice that of migraine with aura, while the opposite is true for men (Rasmussen et al. 1992). These two types of migraine are now considered to be different diseases (Merskey & Bogduk 1994). In *coronary artery disease,* chest pain is a much poorer predictor in women of abnormal angiography or positive thallium-20 scans (Garber et al. 1992; Sullivan et al. 1994). Risk factors also differ, diabetes being more prevalent in women than men with cardiac disease, while external triggers (e.g., exceptional stress) or high weight are of higher prevalence in men than women with cardiac

disease (Behar et al. 1993; Seeman et al. 1993; Sullivan et al. 1994).

An even further complication involves differences in the temporal aspects of pain between males and females. Females undergo large changes in hormonal status and other aspects of their lives as a function of changing reproductive status, including puberty, menstrual cycling, pregnancy, and menopause. Hormonal and other conditions for males change as well, but differently. Males reach puberty at a different age and undergo more gradual changes in hormonal conditions with age. These differences have an as yet unclear impact on pain mechanisms (see below for hypotheses), but the incidence of some clinical pains are well known to vary with menstrual stage (e.g., migraine; Marcus 1995), and others (reviewed in Berkley 1993) vary in their incidence, disappearance, and prevalence as a function of puberty, pregnancy, menopause, and age.

Another major factor that enters into this issue comprises attitudes towards pain that can exert effects not only on subjects' reports of their own pains but also on the experimenters or clinicians who carry out the assessments (Unruh 1996). It is evident that cultural, religious, cognitive, and sociological variables interact to give rise to attitudes about pain that affect the decision on the part of any given individual at any given moment to report the existence of endogenous pain (e.g., Bates 1987; Mendelson 1991; Roy 1992; Strong et al. 1992). The possible differential impact of these variables on reports of pain by males and females has not been extensively studied directly, but when it has been it is no surprise that sex differences in attitudes towards pain exist (Bendelow 1993; Strong et al. 1992) and that such differing attitudes can give rise to sex differences in reports of endogenous pain. For example, one recent study (Bendelow 1993) showed clearly that as a result of differences in attitude of the particular group of individuals she studied (those living in a inner-city area of North London), women were more likely to give "holistic, integrated" reports of their pain, whereas men were more reluctant to classify pain associated with emotional suffering as "real" pain.

On the other hand, sex differences in attitudes might not always affect pain ratings. For example, Fowler-Kerry and Lander (1991) found that although female and male children and adolescents differ in their estimates of how painful an impending venipuncture might be (females overestimated; males underestimated), there were no sex differences in their ratings of the venipuncture pain that was actually produced. Whatever the circumstances with respect to ratings of pain, however, it has been shown that attitudes towards pain can affect coping behaviors and responses to treatment (e.g., see reviews in Edwards et al. 1992; Unruh 1996). Relevant here are recent data showing that in certain specific situations women are more likely than men to benefit from behaviorally based treatments (Hapidou 1994; Jensen et al. 1994).

In addition to the impact of sex differences in attitudes towards pain on an individual's pain reports, coping behaviors, and responses to treatment, attitudinal differences are likely to have an impact on the experimenters who carry out the assessments, the type of assessments that are made, and the treatments that are provided. Some even argue, for example, that the current major positive advance from a Cartesian biomedical model of illness towards a biopsychosocial one is due to the inclusion of a feminist point of view in science and health care systems in many parts of the world (Johnson 1991; Rose 1994).

Thus, another layer of factors that enters into the issue comprises sex differences in attitudes about how gender influences pain. In a recent study carried out by two female nurses (McCaffery & Ferrell 1992), it was found that the majority of nurses (mostly female) have opinions about sex differences in pain that result in differential assessments of pain in female and male patients. It is not illogical to conclude from these results that sex differences in these opinions could give rise to a complex set of sex differences in the diagnosis and treatment of pain disorders in females and males (Unruh 1996). There is in fact some evidence for this sex bias in the general health care system. For example, a number of studies have found evidence for a "less aggressive" management of acute nontraumatic chest pain in women than men (e.g., Heston & Lewis 1992; Hsia 1993), although this difference has been debated (Mark et al. 1994; Vacek et al. 1995). Such differences could relate to the sex of the physician (predominantly male in the case of cardiologists) as has been shown, but not always (Furman et al. 1993) in other realms of health care (Lurie et al. 1993).

In summary, studies of assessments of endogenous pain provide very little concrete information on the nature of sex differences, if any, in pain, be it acute, persistent, or chronic. Women report more pain in more body regions and of more varied types than men. In addition, there is a female or male sex predominance for many painful diseases and the symptoms reported for them sometimes differ between the sexes. All of these differences, however, are subject to situational, temporal, attitudinal, and social factors. The net result is that any sex differences that might exist in the amount of pain experienced or its impact on the life of the individual, or in response to treatment, represent only a small part of the vast array of other factors that impact on pain.

Thus, as discussed at length in the literature on sex differences in cognition (e.g., Halpern 1992), it must be remembered that the differences observed so far are statistical and small. For the most part it is the similarities between the sexes that should be emphasized, particularly when it comes to using the data to form strategies for treatment. It is therefore inappropriate at this time to use data from a single limited experiment as a basis for recommending different overall treatment regimens for females and males. For example, in a recent paper published in the journal *Pain* (Jensen et al. 1994), the authors found significant sex differences in coping strategies in their sample of 139 chronic pain subjects. From these differences, the authors recommend that "the management of each gender should be distinctive, focusing on observed differences of the determinants of outcomes and coping strategies. The message is of importance to primary care practitioners, rehabilitation specialists, clinical psychologists, and occupational physicians" (p. 171). Although these authors go on to warn that "our findings could be easily over-interpreted and much caution is in order," such general recommendations are very clearly unwarranted by their data on a limited sample of patients. Furthermore, as pointed out clearly by McCaffery and Ferrell (1992), such recommendations, if acted upon in a clinical setting, are dangerous because they could result in inadequate or inappropriate treatment.

3. A switch from induction to deduction

Although the discussion above indicates that it is premature and indeed may be impossible to form definitive conclusions on the nature of sex differences in pain from an inductive analysis of the literature, this conclusion does not mean that the analysis of sex differences is irrelevant to improving our understanding of the mechanisms of and best array of treatments for various aspects of pain. Another strategy is to take a deductive approach. In other words, in addition to trying to determine the nature of sex differences in pain by analyzing the troublesome literature as above, it may prove useful as well to develop hypotheses from what is already clearly known about sex differences in general.

Few would disagree that females and males differ in many components of their socialization, psychology, and biology. Although the actual differences are most often controversial or statistical, females and males do differ virtually absolutely and unarguably in three aspects of their reproductive biology. Their pelvic reproductive organs differ and their hormonal conditions differ both temporally and compositionally. How then might these three biological differences operate to affect pain?

4. Sex differences in reproductive organs

Although many aspects of body structure, such as height, weight, muscle composition, fat distribution, and so forth differ in females and males, those differences are statistical, not absolute. With the exception of a number of relatively infrequent intersex states, females and males differ absolutely in the characteristics of their reproductive organs. Females have a vagina, clitoris, cervix, uterus, Fallopian tubes, and ovaries; males have a penis, vas deferens, epididymis, prostate organ, seminal vesicles, and testes.

These differences create major differences in the organization of the entire pelvic region, including muscles, skeletal structures, and the relationship between reproductive, urinary, and alimentary tract organs. The impact of these structural sex differences on the functioning of nonreproductive pelvic organs (e.g., bladder) in females and males is only just beginning to be appreciated (Bavendam 1992; Chalker & Whitmore 1990). However, except for the fact that some of the increased use of health care and morbidity in women compared to men can be accounted for by gynecological and obstetrical problems (Gijsbers van Wijk et al. 1992), little is known about the possible impact of structural differences in pelvic organs on the mechanisms of pain.

As discussed by Slocumb (1984), one factor that could contribute to sex differences in pain derives from the fact that the vagina and cervix provide ready access to internal pelvic structures. Such access is of course important for sperm, but, like the mouth (and anus), that access also provides a route for the entrance of viral and other pathological agents. Susceptibility is high in the vagina because of its continual invasion by potentially damaging objects such as the penis during copulation, tampons during menstruation, and various instruments during gynecological and obstetrical procedures. Phrased in another way, the vaginal canal and cervix increase the vulnerability in women of the T10–L1 (innervates uterus and cervix) and S2–S4 (innervates vagina and cervix) segments to morbidity (Bonica 1990).

The consequences of such increased vulnerability and morbidity could be far reaching. To understand these potential consequences, it is necessary to consider five sets of evidence from the animal research literature. *First,* recent evidence in animals has shown that (1) peripheral pathological events can alter response characteristics of neurons within the dorsal horn of the spinal cord to produce a long-lasting state of hyperexcitability or "central sensitization" of those neurons and (2) that this state is probably part of the mechanism underlying certain chronic pain conditions that outlast their initial pathology (Coderre et al. 1993; McMahon 1992; McMahon et al. 1993; Woolf 1984). *Second,* input from C-fibers (which are the predominant type of fiber innervating the vaginal canal and cervix; Berkley et al. 1993) is particularly efficient at producing such states and probably does so not only by increased electrical activity but also by changes in the axonal transport of neuroactive substances induced by pathological events. (See discussions in Donnerer et al. 1992; Kitchener et al. 1994; Lewin et al. 1992; McMahon et al. 1993; Wall 1989; Woolf et al. 1994.) *Third,* viral agents have access via axonal transport in peripheral C-fibers not only to their spinal cord segment of entry but also transneuronally sequentially across several synapses to other parts of the spinal cord and brain (Card et al. 1990). *Fourth,* virtually all spinal dorsal horn neurons receive convergent input from local and distant sources, the input from distant sources arising both from diverging long-range peripheral afferent fibers (Wall & Shortland 1991) and from multiple intraspinal linkages (see discussion in Berkley & Hubscher 1994; Wall et al. 1993). Indeed, neurons in the C1–C2 segments of the spinal cord have recently been shown to receive input from and then provide a possible major influence back on neurons located throughout the entire length of the spinal cord (reviewed by Foreman 1994; 1995). *Fifth,* dorsal horn neurons that have been provoked into a hyperexcitable state by peripheral pathological events appear to be able in turn to produce inflammatory changes and sensitize peripheral afferents back out in the periphery by a variety of efferent mechanisms both direct and indirect via sympathetic fibers (see discussions in Dubner & Basbaum 1994; Levine & Taiwo 1994; McLachlan et al. 1993; Rees et al. 1994; Sluka et al. 1994).

Taken together, these actions – central sensitization induced by alterations in C-fiber activity and transport, viral access via peripheral afferents across synapses in spinal cord and brain, intraspinal divergence and convergence, and, finally, "retrograde" inflammation and hyperexcitability – produce a situation in which an initial noxious or pathological event at one locus could initiate a sequence of events that would eventually give rise to morbid conditions referred to multiple regions remote from the original site. Certain aspects of this scenario have indeed been postulated as the basis of referred hyperalgesia that occurs in association with many visceral pathological conditions (Vecchiet et al. 1993), sometimes to more than one locus (Slocumb 1984; 1990), and as the basis for certain aspects of fibromyalgia (Goldenberg 1993). The most common region of referral is to somatic loci within the same segments that receive input from the organs initially (but not necessarily currently) involved. However, multiple and long-range divergence–convergence patterns within the spinal cord and transneuronal transfer of viral agents clearly indicate the opportunity for a cascade of referrals to more distant ones, including the C1–C2 segments.

It is thus quite possible that a woman's vaginal canal provides her with an additional bodily entrance that puts her at increased risk relative to that of men for multiple sites of referred pain and hyperalgesia. Examples include fibromyalgia (because of multiple convergence patterns within the spinal cord; see Goldenberg 1993) as well as certain types of headache and other facial/trigeminal pains as listed in Table 1 (because of the possible involvement of C1–C2). The clinical implications of this possibility are obviously significant.

What is important to remember about the scenario described above is that it would apply to males as well as females. In other words, although a woman's increased vulnerability could be at least part of the basis for the huge female predominance of certain forms of pain and for the fact that, overall, women report more multiple pains than men, the lesson to be derived from the analysis above applies to both sexes. That is, as discussed in the example of chest pain at the end of this article, diagnostic power and appropriate treatment in both sexes would be advanced by a more persistent pursuit and analysis of patterns of multiple symptomology (e.g., Facchinetti et al. 1993) and past history of pathological conditions (e.g., Laws 1993).

5. Sex differences in temporal features of hormonal action

Because of the importance of all three sex hormones (i.e., estrogen, progesterone, and testosterone) to many aspects of life in both sexes (McEwen 1991), how sex differences in those hormones affect pain is obviously unlikely to be straightforward. One obvious difference between the sexes, however, is in the temporal characteristics of estrogen, progesterone, and testosterone levels (Goodman 1994). All three change cyclically on a monthly cycle throughout most of a female's reproductive lifetime and episodically at puberty, during and after pregnancy, and at menopause as well as gradually afterwards. Levels of these hormones in males also exhibit chronobiological changes, particularly with aging, but overall they are more stable than they are in females. Regardless of the compositional changes (to be discussed below), how might simply the temporal differences affect pain?

Much of the research on the impact of temporal hormonal characteristics on nonreproductive functions has focused on menstrual or estrous cyclicity, although other biological rhythms obviously exist (e.g., Binkley 1992; Moore-Ede et al. 1982; Procacci et al. 1972; Sothern et al. 1993). In women, numerous conditions can be demonstrated to vary with menstrual cycle. Twelve examples include the following: memory (Phillips & Sherwin 1992), creativity (Krug et al. 1994), mood disorders (Endicott 1993), affective state (Johnston & Wang 1991), thermoregulation (Frascarolo et al. 1992; Kolka & Stephenson 1989; Sothern et al. 1993), CO_2 sensitivity (Dutton et al. 1989), caffeine elimination time (Lane et al. 1992), gastrointestinal transit time (Wald et al. 1981), wrist activity (Binkley 1992), epilepsy (Newmark & Penry 1980), the tics of Tourette syndrome (Schwabe & Konkol 1992), and the symptoms of multiple sclerosis (Giesser et al. 1991; Smith & Studd 1992).

Similar temporal variations are observed in female rodents. Nine examples include: binding of mu opioid receptors in the hypothalamus and $GABA_B$ receptors in the cerebral cortex (Al-Dahan et al. 1994; Maggi et al. 1993); norepinephrine levels in cerebral cortex (Parada et al. 1991); cannabinoid receptors in the hypothalamus (Rodríguez de Fonesca et al. 1994); galinin-like immunoreactivity in the spinal cord (Newton 1992); glial characteristics in dentate gyrus (Luquin et al. 1993); seizure susceptibility for certain types of seizures (Thomas 1990; Woolley & Timiras 1962); susceptibility to viral inoculation (Teepe et al. 1990); cutaneous receptive field sizes of ganglion cells and peripheral nerves (Adler et al. 1977; Bereiter & Barker 1980), and potency of central morphine analgesia (Kepler et al. 1989).

It is important to note that it is also sometimes the case that temporal variations are not observed. For example, menstrual variations were not observed in cardiovascular responses to mild stressors (Stoney et al. 1990) or in certain verbal or spatial cognitive tasks (Gordon & Lee 1993), and estrous variations were not observed in seizure susceptibility to certain kindling procedures (Wahnschaffe & Löscher 1992) or in various parameters of hippocampal electrical activity in struggling or immobilized awake rats (Mead & Vanderwolf 1992). As discussed earlier in this article, the same sort of conflict in evidence demonstrating or failing to demonstrate menstrual or estrous variations exists for various pain conditions. It is thus necessary to conclude that while menstrual and estrous periodicities clearly exist for many conditions, it is also likely to be the case that certain subsets of individuals are more vulnerable to the temporal aspects of hormonal status than others and that those aspects are relevant only under certain circumstances. What will be important in the future is to characterize these subsets and circumstances.

In the meantime, despite the seemingly confused state of current knowledge, some conclusions can in fact be derived from these wide-ranging data. One conclusion is that the regular cyclical changes that females undergo throughout much of their lives have the potential to affect their overall pain conditions in at least two important ways.

First, a possible consequence of menstrual/estrous periodicity could be the production of periodic experiences that by associative learning eventually do not require the same exogenous stimuli that originally provoked them. For example, *time* alone could come to serve as a discriminitive stimulus for pain, so that simply the passage of one month could give rise to certain "appropriate" pains, as it does in some women after menopause who continue to suffer from "dysmenorrhea." Another example is that animal data have shown that the *hormones* estradiol, progesterone, and testosterone possess distinct stimulus properties and can therefore serve as discriminative stimuli for conditioning (Heinsbroek et al. 1987; Peeters et al. 1992). It is unknown whether these hormones can similarly serve as discriminative stimuli in humans; if they can, if a particular noxious stimulus is paired for several menstrual cycles with, say, high progesterone and estrogen just before menses (Ferin et al. 1993) or with high testosterone and estrogen just before ovulation (Ferin et al. 1993; Vermeulen & Verdonck 1976), then, by associative learning, the presence of these hormones alone would give rise to experiences similar to the ones evoked by the original noxious stimulus; that is, pain. The net result of both of these associative learning cues (time, hormones) would be a greater number of pain

conditions without obvious peripheral pathology in females than males.

Second, in addition to the possible consequences of menstrual/estrous cyclicity on "learned pain," the fact that such cycles also regularly subject females to a regularly fluctuating affective drive (Johnston & Wang 1991) might give rise to regular fluctuations in their responses to either nociceptive events or treatments, or in their interpretation of experiences as being painful or benign. The fluctuations could produce a situation in which some females develop a greater tolerance for various painful states and others develop an enhanced response to various stimulus events or experiences. The end result would be a greater variability in pain behaviors in females than males. Such a situation could be part of the explanation, for example, for the seemingly conflicting data on coronary disease, which indicate on the one hand that many women with chest pains wait too long before seeking health care (Moser & Dracup 1993) while many others have chest pains without evidence of cardiac pathology (Hsia 1993; Sullivan et al. 1994).

Although one message from these considerations is that menstrual cyclicity could put women at more risk than men for developing a wider range of pains without obvious peripheral pathology, the most important message is that chronobiological factors in general (that is, all types of temporal factors in addition to menstrual ones) are likely to be much more important for pain in both sexes than has been realized (Berkley 1993; Lautenbacher & Rollman 1993; Procacci 1993). In other words, more active consideration of the impact of when and how regularly either nociceptive events have occurred or pain is reported as having been experienced on whether or not any given set of circumstances or treatments will produce or affect pain is needed both in research and treatment.

6. Sex differences in sex hormones

All three main sex hormones – estradiol, progesterone, and testosterone – are functionally active in both sexes (Goodman 1994). What varies are their relative concentrations in various tissues, the receptors for them in various tissues, certain aspects of their metabolism, and, as discussed above, their temporal fluctuations. It is thus inappropriate to consider estrogen and progesterone as purely "female hormones" and testosterone as purely a "male hormone." It is also inappropriate to assume that all sex differences and all functions that vary with estrous or menstrual stage are due to the action of hormones (Reisert & Pilgrim 1991).

Nevertheless, given the large changes that occur in estrogen and progesterone levels over the ovarian cycle as well as during puberty, pregnancy, and menopause, it is not unreasonable to hypothesize that these hormones are functionally involved when an activity fluctuates with estrous/menstrual stage or as a function of puberty, pregnancy, or menopause (Van Goozen et al. 1995). Similarly, given the large differences in testosterone and, at certain times, estrogen and progesterone between the sexes, it is not unreasonable to hypothesize that these hormones might be functionally involved in activities that exhibit sex differences. And finally, conversely, because of these changes and differences, if one of these hormones is demonstrated to be important for some function, then it is not unreasonable to hypothesize that there would be sex differences in that function.

Given this logic, a number of intriguing possibilities applicable to possible sex differences in pain mechanisms in humans present themselves from the recent animal literature. Four of these possibilities will be considered here, on the following topics: the action of γ-aminobutyric acid (GABA) and other neuroactive agents, mechanisms of opioid and nonopioid analgesia, mechanisms of nerve growth factor (NGF) operation, and sympathetic nervous system function.

6.1. GABA and other neuroactive agents

Numerous studies have now clearly demonstrated a very definite association between sex hormones and the action of GABA, not only in epilepsy but also in many other aspects of neural function (Carey et al. 1992; Grattan & Selmanoff 1993; Jussofie 1993; Lambert & Peters 1989; Maggi & Perez 1986; McCarthy 1995; Perez et al. 1986; Smith 1991; Weiland 1992), including pain (Frye & Duncan 1994; McCarthy et al. 1990; Schwartz-Giblin et al. 1989). How this association could produce operational differences in females and males is likely to prove complex but should be studied. For example, Westerling et al. (1991) showed that estrous stage influences the potentiation by barbiturate but not by benzodiazepine of primary afferent depolarization in slices of cuneate nucleus. One clear area is in the use of various pharmaceutical agents that act on GABAergic mechanisms whose effects on pain as well as other functions would then vary as a function of various hormonal states.

In addition to GABA, sex differences exist in the action of a number of other neuroactive agents such as serotonin (Kojima & Sano 1984; Mendelson 1992), dopamine (Beyer et al. 1992; Zanin & Takahashi 1994), thyrotropin-releasing hormone (Deshpande et al. 1987), calcium-dependent nitric oxide (Weiner et al. 1994), and various peptides (Micevych et al. 1988; Newton et al. 1990). Most of these sex differences (Biegon et al. 1980; Demotes-Mainard et al. 1993; Fernández-Ruiz et al. 1991; Fischette et al. 1984; Peris et al. 1991; Weiner et al. 1994), but not all of them (Beyer et al. 1992; Ovtscharoff et al. 1992), appear to be related to hormonal action.

Because all of these substances are variously involved in pain mechanisms, the sex differences in their actions together with those associated with GABA are likely to be of importance for overall sex differences in the mechanisms of pain and its control. As highlighted in an elegant review of the mechanisms of migraine (Marcus 1995), this important possibility currently cries loudly to be addressed.

6.2. Opioid and nonopioid analgesia

One arena where the sex and hormonally associated differences in the action of neuroactive agents is of clear relevance is analgesia. Recently, a substantial number of animal studies have emerged demonstrating sex differences in many aspects of analgesia. For example, female rats show less analgesia to morphine than do males (Bodnar et al. 1988). As thoroughly reviewed by Fillingim and Maixner (1995), these differences have been shown to depend upon how the analgesia is induced and to involve a number of opioid and nonopioid mechanisms, many of which are modulated by the action of estrogen and other steroids

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

(Aloisi et al. 1994; Baamonde et al. 1989; Bodnar et al. 1988; Dawson-Basoa & Gintzler 1993; 1996; Islam et al. 1993; Kavaliers & Colwell 1991; Kavaliers & Innes 1993; Kepler et al. 1989; 1991; Mogil et al. 1993; Ratka & Simpkins 1991; Romero & Bodnar 1986; Ryan & Maier 1988) as well as by various reproductive conditions such as multiparity (Mann & Bridges 1992).

As discussed in a recent issue of the *Journal of NIH Research* (Touchette 1993), although it is clear that sex differences in pain modulatory mechanisms exist in animals, the mechanisms at work to produce these differences are not yet understood. The potential for this difference being of importance in humans is obvious, where very few clinical studies on this issue exist (e.g., DeKock & Scholtes 1991) and where currently only brief mentions are made of sexual, hormonal, or chronobiological factors that might affect anesthetic or other drug usage in humans (Collins 1993; Hrushesky 1994).

6.3. NGF

In addition to its action on nerve development and sprouting, nerve growth factor (NGF) has recently been found to be actively involved in many aspects of nociception, including inflammation, hyperalgesia, and the regulation of afferent activity (Donnerer et al. 1992; Fitzgerald et al. 1985; Lewin & Mendell 1993; Lewin et al. 1992; McMahon 1992; Woolf et al. 1994). In certain circumstances, for example, in female rat vocalizations and rejection behaviors during mating, NGF involvement has been shown to depend on estrogen and progesterone conditions (Gibbs et al. 1993). Sex hormones are potently involved in inflammatory mechanisms as well as immune system functions important for inflammatory effects in diseases such as rheumatoid arthritis (e.g., DaSilva & Hall 1992; DaSilva et al. 1993; Lahita 1992; Szekeres-Bartho 1992). Furthermore, estrogen receptors are prevalent throughout the nervous system where they develop in a sexually dimorphic manner in some regions (Kornack et al. 1991) and colocalize with receptors for NGF, including parts of the spinal cord relevant to pain (Keefer et al. 1973; MacClusky et al. 1987; Morrell et al. 1982; Toran-Allerand et al. 1992a; 1992b; Urschel & Hulsebosch 1992). And, finally, estrogen regulates NGF receptor mRNAs in sensory neurons and GAP-43 mRNAs in various parts of the central nervous system (CNS) (Shughrue & Dorsa 1993; Sohrabji et al. 1994).

Taken together, these facts suggest that NGF may operate differently in females and males in circumstances important for both peripheral and central plastic changes associated with pain. It is difficult to predict from these fast emerging and sometimes confusing data (e.g., NGF can both alleviate and increase pain under different circumstances) what the sex differences might be, but they could have important implications for the mechanisms of persistent pain (e.g., an increased vulnerability in women for the development of sympathetically maintained pain; see below), for temporal aspects of surgery (e.g., certain menstrual conditions might promote postoperative recovery better than others; Emerson et al. 1993) and for the use of pharmacological agents now under development that act on NGF (e.g., Apfel et al. 1992; some of these agents might act differently in males and females under different reproductive conditions).

6.4. Sympathetic nervous system function

It is evident from a perusal of Table 1 that a large proportion of the disorders that show differences in their sex prevalence characteristics involve the cardiovascular system and visceral organs, particularly gastrointestinal structures. As discussed above, further analysis of some of these disorders shows that their pain components are influenced by menstrual stage, reproductive status, or hormonal treatments. These sex differences could therefore be due in part to sex differences in the operation of the autonomic nervous system.

6.4.1. Structural differences. Numerous studies have provided evidence for sex differences in the structural organization of many groups of neurons throughout the CNS, mainly in brain regions directly associated with reproduction. These differences occur as a result of both hormonal and nonhormonal influences during development (reviewed in Breedlove 1994; Kelley 1986; Reisert & Pilgrim 1991; Tobet & Fox 1992). Structural sex differences, however, also appear to be produced during development in neural systems associated more generally with other physiological functions. Thus, Calaresu and Henry (1971; Henry & Calaresu 1972) showed that there were many fewer sympathetic preganglionic neurons in the spinal cords of female cats than male cats, with a female-to-male ratio of 0.78.

The possibility of an overall sex difference in the structural features of the sympathetic supply to internal structures has a number of implications important for pain. Differences in afferent input from internal structures to the CNS via sympathetic nerves could not only produce different visceral pains in females and males (such as those referred to earlier in this article for various painful diseases), but, assuming the importance of visceral input for the perception of emotions (James 1884; Lange & James 1922/1967), sex differences in visceral afferent input could also result in different emotional consequences of pain experiences. Differences in the efferent supply of different viscera through sympathetic nerves as demonstrated by Calaresu and Henry would affect the responses of these organs to various stimuli that would in turn affect the interpretation of those stimuli through afferent mechanisms.

6.4.2. Functional differences. Considerable evidence supports the possibility of important sex differences in the functional organization and operation of the sympathetic system. In humans at rest, overall integrated levels of resting sympathetic nerve activity to skeletal muscles through the right peroneal nerve are lower in women than men (Ng et al. 1993). On the other hand, sympathetic output to the skin is higher in women than men; this appears to account for lower basic levels of skin blood flow and perfusion in women (Cooke et al. 1990). In response to various stressors, Morrison and Pickford (1971) showed that sympathetic nerve activity was differentially sensitive in female and male cats and dogs to arterial pressure changes induced by angiotensin and noradrenalin. These effects occurred mainly at higher pressures and were attributed by the authors to sex differences in the state of the blood supply to arterial chemoreceptors. Along these lines, Maixner and Humphrey (1993) found clear sex differences in humans in cardiovascular responses to forearm ischemia,

with males having greater blood pressure responses during postexercise ischemia. Interestingly, only in males did pain assessments of the ischemia correlate with the cardiovascular responses. Similarly, sex differences in adrenergic sensitivity leading to different patterns of cardiovascular responses to various stressors have been observed in humans (e.g., Claustre et al. 1980; Girdler et al. 1993) and these differences have been shown in animals to have important consequences for the action of alpha 2 adrenoreceptor pharmacological agents, some of which, such as clonidine, are of value for pain control (e.g., Heal et al. 1989).

Evidence exists in other arenas of autonomic function as well. For example, gastrointestinal transport and absorption are influenced so powerfully by the menstrual cycle (McBurney 1991; Wald et al. 1981) that certain gastrointestinal disorders occur almost always in females (MacDonald 1993). Norepinephrine levels and release in the cerebral cortex are influenced by the estrous cycle in rats (Parada et al. 1991). Of relevance to possible sex differences in sympathetic functions related to cutaneous and muscle pain are results showing that sweating, skin blood flow, and postural vasoconstriction reflexes are powerfully influenced by the menstrual cycle in normal women (Bartelink et al. 1990; Frascarolo et al. 1992; Hassan et al. 1990; Kolka et al. 1989) and there are sex differences in peripheral reflex-induced muscle atrophy produced by bone fractures (Urbancova et al. 1993). And finally, there are sex differences in the CNS, sometimes modulated by hormonal conditions, of responses to traumatic events such as hemorrhage, contusion, and electroconvulsive shock (Emerson et al. 1993; Heal et al. 1989; Iyengar & Laycock 1993; Öztas et al. 1991; 1992; Roof et al. 1992; Stein 1995) that are of relevance to pain (Wiesenfeld-Hallin et al. 1993).

Another important realm in which sex differences in sympathetic function would be important for pain, particularly persistent or chronic pain, relates to plastic changes in sympathetic action induced by injury. Recently, McLachlan et al. (1993) observed large increases in noradrenergic fibers surrounding dorsal root ganglion cells in response to ligation of the sciatic nerve in rats. Although the authors stated that there were no sex differences in this sprouting response, their study did not directly address the issue. In other studies where the issue has been directly addressed, sprouting of sympathetic fibers into the hippocampus induced by neural injury has been shown to be more restricted in male than in female rats and to be affected by neonatal (but not adult) manipulations of testosterone levels (Milner & Loy 1982). Evidence exists as well for the involvement of both estrogen and testosterone in plastic responses to injury under other circumstances (Demotes-Mainard et al. 1993; Jones 1988; Kujawa et al. 1991). And, finally, the associations discussed above between NGF's functions in injury-induced sprouting and inflammation and pain, and the interactions between NGF and sex hormones, together enhance even more the possibility of sex differences in the plastic changes of sympathetic action induced by injury.

A net result of such differences in humans could be that females might be either at a greater risk of developing more potent forms of a condition involving the sympathetic nervous system variously referred to as reflex sympathetic dystrophy, causalgia, or sympathetically maintained pain (Campbell et al. 1992; Jänig 1992; Jänig et al. 1991; Roberts 1986), or their symptoms might vary as a function of these conditions, or they might be affected differently from males by treatments for this condition, especially in certain hormonal conditions.

Very little evidence directly related to this hypothesis exists, but it is certainly the case that there is a large female predominance of subjects reported in most studies of sympathetically maintained pain, causalgia, or reflex sympathetic dystrophy. Although in one study, no sex differences were reported in responses to guanethidine blocks for algodystrophy in humans (Eulry et al. 1991), Abelli et al. (1993) reported that while sympathectomy abolishes corneal lesions produced by neonatal sensory denervation with capsaicin in female rats, it only reduces those lesions in male rats. Similarly, in another study currently underway at Johns Hopkins (e.g., see Shir et al. 1993), the ratio of females to males whose pains are greatly reduced by phentolamine (e.g., those classified by the authors as suffering from sympathetically maintained pain) is greater than the female to male ratio of patients whose pains are only slightly reduced or unaffected (e.g., those classified as suffering from another form of pain; Campbell, personal communication).

Thus, women may have a better response than men to certain forms of therapy that address pathological alterations in sympathetic function or they might respond differently to such therapies under different hormonal conditions. For example, this possibility could relate to the fact that men are more at risk of developing duodenal ulcers than women (Schubert et al. 1993). Clearly, the issue warrants further study.

7. Summary and conclusions

Taken together, inductive analysis of the available literature indicates that for experimentally induced acute somatic (usually skin) stimulation, females often have lower thresholds, greater ability to discriminate, higher pain ratings, and less tolerance of noxious stimuli than males. The evidence also indicates, however, that these differences are inconsistently observed, relatively minor, exist only for certain forms of stimulation, and can be affected by numerous situation variables in daily life such as the presence of disease, the setting of the experiment, the characteristics of the experimenter, and even nutritive status.

The relevance of these minor differences to the clinical situation seems questionable (Berkley 1995), because, in contrast to experimentally induced pain, endogenous pain most often involves pains that are episodic, persistent, or chronic (as well as acute) and that are more likely to involve muscles and internal organs than skin. However, results from studies of endogenous pains do indicate some sex differences. Thus, women report more multiple pains in a greater number of body regions than men. Some painful diseases are more prevalent among females, while some are more prevalent among males, with no basis yet evident for these differences. For many diseases, symptoms differ between females and males, again with no obvious rational basis. Some pains vary with hormonal status of females or males, once again with no clear basis for the variability. And, finally, substantial differences in attitudes exist not only in the individual experiencing pain, thereby affecting her or his reporting and coping behaviors and responses to treatment, but also in the individual taking the measures,

thereby affecting how and what measurements are made and what treatments are carried out.

Taking all of this information together, however, the major consequence of so many diverse factors being operative in both experimental and clinical settings is that the most striking overall feature of sex differences in reported pain experiences is the apparent lack of them.

On the other hand, deductive analysis of the available literature gives rise to a number of possible factors that could operate differently to affect pains in females and males. Three examples are discussed here. First, the existence of a vaginal canal vulnerable to trauma or invasion by pathological agents in females and not males, together with the existence of neural mechanisms of sensitization and divergence that operate to give rise to referred hyperalgesia in regions remote from the initial source of the problem, could put females at greater risk than males of developing hyperalgesia in multiple regions remote from the initial problem. Second, differences in temporal patterns of many aspects of life produced by differences in temporal patterns of sex hormones could give rise to sex differences in how pain is "learned" and how various sensory experiences are interpreted, the net result being a greater variability and wider range of pains without peripheral pathology in women. Third, differences in compositional aspects of hormones give rise to a number of hypotheses associated with the actions of estrogen, progesterone, and testosterone. Thus, there are sex and hormonally dependent differences in the operation of GABA and other neuroactive substances that could produce both sex differences and hormonally dependent differences in pain modulatory situations where those substances are operative and in the actions of their agonist and antagonist pharmacological agents. Along these lines, evidence is emerging on the existence of fundamental sex- and hormone-dependent differences in both opioid and nonopioid mechanisms of analgesia; the potential clinical implications of these differences are substantial but as yet unclear. In another arena, potential sex- and hormone-dependent differences in the organization and operation of the sympathetic nervous system give rise to the possibility of differences in visceral symptomology for various diseases and the cardiovascular and emotional sequelae of noxious events. These differences are also relevant to pain situations in which the sympathetic system may be involved in the production of hyperalgesic states as a result of injury. This relevance is strengthened by evidence suggesting that sex and hormonal factors are likely to be involved in the operation of NGF both centrally and peripherally.

It is not immediately evident how conclusions such as these, derived from inductive and deductive approaches to the literature, might be applied to future research and current treatment of humans and animals. It is clear, however, that when patient A appears in a health care facility to report that she or he is experiencing, say, chest pain, a large number of variables have already contributed to that report that may seem to have only a remote relation to the cause of A's pain. On the other side of the scene, a large number of seemingly remote factors also enter into the response of health care worker B, who is faced with A's report of chest pain. These remote factors operate together to have a large impact on A's overall health.

It is also evident, moreover, that both A's and B's sex is one of the more important factors that must be considered in this scenario. In fact, the issue of sex and hormonal factors in the significance, diagnosis, and treatment of cardiac pain and cardiac disease is currently under intense scrutiny and is on the threshold of important findings (Behar et al. 1993; Cabral et al. 1988; Chae et al. 1993; Dewhurst et al. 1991; Dittrich et al. 1988; Garber et al. 1992; Hamilton & Seidman 1993; Harris & Weissfeld 1991; Heston & Lewis 1992; Hsia 1993; Karlson et al. 1993; Mark et al. 1994; Moser & Dracup 1993; Puntillo & Weiss 1994; Seeman et al. 1993; Sullivan et al. 1994; Vacek et al. 1995). Perhaps, then, it is not an exaggeration when modern philosophers make statements such as the following:

> Sexual difference is one of the major philosophical issues, if not the issue, of our age. . . . Sexual difference is probably the issue in our time which could be our "salvation" if we thought it through. (Irigeray 1993, p. 5)

BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 381–391
Printed in the United States of America

Are there fundamental differences in the peripheral mechanisms of visceral and somatic pain?

Stephen B. McMahon
Department of Physiology, St. Thomas' Hospital Medical School, London, SE1 7EH, England.
Electronic mail: *s.mcmahon@umds.ac.uk*

Abstract: There are some conspicuous differences between the sensibilities of cutaneous and visceral tissues: (1) Direct trauma, which readily produces pain when applied to the skin, is mostly without effect in healthy visceral tissue. (2) Pain that arises from visceral tissues is initially often poorly localised and diffuse. (3) With time, visceral pains are often referred to more superficial structures. (4) The site of referred pain may also show hyperalgesia. (5) In disease states, the afflicted viscera may also become hyperalgesic. In this target article, I consider to what extent differences in the physiology, anatomy, and chemistry of peripheral processing systems explain these different sensibilities. In almost every aspect, there are subtle differences in the properties of the processing mechanisms for cutaneous and visceral information. These may arise because of distinct developmental cues operating in the two domains. Many of the differences between visceral and cutaneous afferents are quantitative rather than qualitative. The quantitative differences, for example in the density of afferent innervation, can be large. The quantitative differences in the numbers of afferents alone may be a sufficient explanation for some aspects of the differential sensibility, for example, the poor localisation of sensation and the apparent insensitivity to focal yet tissue-damaging stimuli. In addition, the few clear qualitative differences apparent in the innervations of the two tissue types may be of special importance. That the encoding of visceral nociceptive events may occur by an intensity mechanism rather than a specificity mechanism could be the key difference in viscerosensory and somatosensory processing.

Keywords: hyperalgesia; nociception; pain; primary sensory neurons; referred pain; somatosensation; viscera; viscerosensation

1. Introduction

It is nearly a century since the publication of Sherrington's influential work, "The integrative action of the nervous system" (1900). Sherrington established a framework of ideas about the organisation and operation of the nervous system that has dominated thinking for much of this century. In relation to pain sensibility, Sherrington proposed the existence of nociceptors, afferent neurones that would detect tissue damage. This idea arose from Sherrington's observations on the reflex responses triggered by the application of strong stimuli to the skin. Indeed, Sherrington wrote that "pain appears the psychical adjunct to protective reflexes" (1900). Several decades after proposing the existence of nociceptors, direct experimental evidence for large numbers of such receptors in skin was obtained. It is difficult, however, to apply Sherrington's concept to deep tissues such as viscera, because many forms of tissue injury (such as neoplastic destruction of the solid organs) are often not painful in man, as discussed more fully in section 2.1. Further, some stimuli that are painful (such as distention of the hollow organs) do not damage, or even threaten to damage, the tissue. It is also not always clear what might constitute an adequate noxious stimulus for some visceral organs. For example, a receptor found to respond to supra-physiological levels of distension of a viscus might also respond to entirely physiological pressure changes in the associated vasculature or in the mesenteric attachments of that viscus. One example can be seen in the case of renal colic. At first it might appear intuitively obvious that the pain arises from the mechanical stimulation offered by the passage of a sometimes large and rough stone through the restricted lumen of a ureter. However, an alternative explanation, for which there is much evidence, is that the stone merely obstructs the ureter, leading to a distension of the entire upper urinary tract. The increased hydrostatic pressure acting on receptors in the pelvis of the kidney may actually be the stimulus that is detected (Bretland 1972).

Persistent pain of visceral origin is undoubtedly a greater clinical problem than pain from skin, but the overwhelming focus of experimental work on pain mechanisms has considered cutaneous systems, and much of what we know of these mechanisms relates specifically to cutaneous sensibilities. Until relatively recently, however, it was often tacitly assumed that these ideas could be transferred more or less whole to the visceral domain. Indeed, there has been such a reliance on the data and ideas arising from the study of somatic tissues that in some cases a visceral "peg" has been made to fit a cutaneous "hole." Arguably, it is true that if the initial experimental studies had been undertaken on visceral tissue, we might now have a completely different general theory of the processing of painful stimuli.

In this target article, I will seek to compare directly the peripheral properties of somatosensory and viscerosensory

 0140-525X/97 $9.00+.10

systems. Such a discussion is timely because of recent interest in viscerosensory processes. This interest comes not only from a basic science perspective but from the growing clinical belief that disorders such as irritable bowel syndrome, previously thought to be primarily the result of motor dysfunctions, may actually be at least partly explained by altered sensory function. I will consider first the nature of sensations from visceral and somatic domains, and then attempt to correlate these sensibilities with the properties of primary afferent neurones innervating the two systems.

2. Nature of superficial and visceral pains

2.1. Effective stimuli

In normal, healthy people, a variety of intense stimuli applied to skin readily produces pain. These include mechanical and thermal stimuli that might be considered noxious (i.e., tissue damaging), as well as nondamaging events such as electrical stimuli and some chemical irritants. When the intensities of cutaneous stimuli are raised above threshold, pain does not usually radiate. On the contrary, pain frequently becomes more focal, as can be experienced if a pencil point is increasingly pushed against the skin. Also note that increasing the area over which a stimulus acts causes a modest increase in perceived pain, but the threshold for pain is not markedly reduced (Price et al. 1989). Deep somatic tissues, such as joint and muscle, are similarly sensitive to direct tissue-threatening stimuli. For example, strong mechanical pressure on a muscle or distortion of a joint beyond its working range causes pain. Irritant chemicals, directly injected into muscles or joints, can induce pain (see Mense 1986).

The sensitivity of visceral tissues is markedly different. Some structures, such as the lung, liver, and the parenchymatous part of the kidney, appear not to give rise to pain with any stimulus, including their gross destruction by malignant growth. In addition, surgeons working in the early part of this century, using only local anaesthesia of the body wall, were surprised to find that a wide variety of traumatising stimuli including crushing, cutting, and burning, very rarely gave rise to any sensation when applied to healthy visceral tissue (Capps 1932; MacKenzie 1909; Morley 1931). There are some exceptions: the mesenteries are said by most authors to be sensitive to tension or clamping and it is recognised that the trigone region of the bladder neck can give rise to pain when probed directly or stimulated by the presence of a stone. Viscera, however, are sensitive to other forms of stimulation. The most widely recognised is distension of the hollow muscular-walled organs. Distension of the gastrointestinal tract from oesophagus to rectum, the urinary tract from kidney pelvis to bladder, and of the gall bladder, all produce pain (Bentley & Smithwick 1940; Bretland 1972; Csendes & Sepulveda 1980; Denny-Brown & Robertson 1933a; Goligher & Hughes 1951; Hertz 1911; Lewis 1942; Nathan 1956; Pollard & Bloomfield 1931; Ray & Neill 1947; Risholm 1954). The severity of distension-induced pain is often only modest in healthy subjects. It arises with a short latency (measured in seconds), suggesting that indirect effects (e.g., ischemia) are not the cause. Active contractions of smooth muscle, around an obstruction, for example, may exacerbate pain, and result in pain that comes in waves, as is so apparent in the case of labour pains. One can readily demonstrate similar contraction-associated pain by voluntarily checking the flow of urine in the middle of micturition. Closure of the urethral outlet causes a large reflex isovolumetric contraction of the bladder that, in most people, is distinctly painful.

One of the problems in transferring the concept of nociception from cutaneous to viscera domains is that the distending pressures associated with pain are not tissue-damaging (e.g., 30–40 mm Hg in the case of bladder distension). Estimates of the threshold pressures producing pain in a particular viscus often vary considerably. One reason is that the area of tissue stimulated may be a crucial determinant of threshold. Unlike skin, spatial summation may drastically reduce the effective threshold for pain. This viewpoint was strongly argued by Goldscheider (1920). Comparisons of different studies in man and animals suggest that spatial summation can appreciably lower the threshold for visceral pain (Lewis 1942; Peterson & Youmans 1945). The existence of appreciable spatial summation of visceral inputs may explain the failure of localised mechanical stimuli, even frankly damaging ones, to produce pain.

Another effective stimulus for visceral pain is ischemia. The best recognised example is ischemic heart disease but it is likely that ischemia of other visceral tissues also produces pain (Lewis 1942; Poole et al. 1987). With coronary occlusion there is the possibility of secondary mechanical effects (e.g., the spasm of arteries; Osler 1910), but it is frequently assumed that an important component of the stimulus is an accumulation of pain-producing chemicals in the ischemic tissue (see Malliani 1986; Ness & Gebhart 1990 for opposing views). A number of well-recognised algogenic chemicals do produce pain when applied to human visceral tissues. The best studied is bradykinin, a naturally occurring agent. This substance produces pain when infused into the abdomen of healthy volunteers (Lim et al. 1967). It is less clear if it is algogenic in the heart (Euchner-Warmser et al. 1994; see also Pagani et al. 1985). Another reason for caution is that bradykinin may exert an indirect action via smooth muscle contraction (Floyd et al. 1977). Synthetic algogens have been shown to induce pain in some viscera, for example, the urinary bladder (Head 1893; Maggi et al. 1989; Nesbit & McLellan 1939).

A final – and clinically important – circumstance where visceral pain may be triggered is in inflammatory states (Head 1893; Wolf 1965). In the urinary and alimentary tracts, inflammation is common and can be painful. In cystitis, for example, the sensations during bladder emptying often become unpleasant and painful (Nesbit & McLellan 1939; Petersen & Franksson 1955).

2.2. Hyperalgesia

In the wake of strong stimuli, the sensitivity of skin changes markedly. Previously innocuous stimuli become capable of evoking pain, and noxious stimuli produce more pain than in normal tissue. This phenomenon is called primary hyperalgesia. There is a wealth of experimental evidence that this hyperalgesia arises at least in part from a sensitization of primary sensory nociceptors. Surrounding the area of damage, skin sometimes becomes more sensitive as well, a process called secondary hyperalgesia. The wide spread of secondary hyperalgesia that is sometimes observed strongly

suggest that the underlying cause lies not in the properties of primary afferents but within the central nervous system; there is indeed considerable experimental support for such a view (see McMahon et al. 1993).

The opportunities to observe such changes in visceral structures are much more limited, and it is not surprising that we have only meagre information on this point. Nonetheless, there are a number of anecdotal reports that visceral structures may become hyperalgesic, particularly in inflammatory states. Kinsella (1940) reported that direct mechanical stimulation of the inflamed – but not the healthy – appendix caused pain. Other reports exist for the ureter, kidney, bladder, ovary, stomach, and oesophagus (Head 1893; Hertz 1911; Hurst 1911; McLellan & Goodell 1943; Petersen & Franksson 1955; Ruffin et al. 1953; Wolf 1965). One example comes from the much-studied patient, Tom, who had a gastric fistula. Tom's gut mucosa was normally insensitive to pinching, but the same stimuli produced pain when the mucosa was inflamed. Quantitative studies of the increased sensitivity of inflamed viscera are few, but some data exist for patients with irritable bowel syndrome and noncardiac chest pain. In one recent study, Trimble et al. (1995) reported lower pain thresholds to distension in irritable bowel syndrome and functional dyspepsia (see also Mayer & Raybould 1990). The altered sensibility of visceral tissue in pathological conditions such as inflammation may indicate the emergence of new neurophysiological processes, a view supported by growing experimental evidence (see sect. 2.3).

2.3. Accuracy of localisation

Pain of cutaneous origin has distinct features. It is usually focal (e.g., with well-defined boundaries) and often has a burning quality. It is well localised, and even if any tactile cues are removed by a block of large diameter afferent fibres, people can localise a noxious stimulus to the skin of the hand within 10–20 mm (Lewis 1942). For visceral pain, two distinct types of localisation have been noted. In some cases, visceral pain may be referred to a distant structure, as described in more detail in section 2.4. In other cases, however, the pain is perceived as being deep within the body. This type of visceral pain, so-called "true" visceral pain (or as early authors such as Ross [1888] called it, "splanchnic pain"), is usually perceived as arising in the midline. The pain may be perceived as anterior or posterior, and occasionally radiates over considerable distances. One example is the initial sensation perceived after myocardial infarction (Proccaci et al. 1986). Another is the early pain of appendicitis, which is initially felt in the midline. "True" visceral pain is usually extensive rather than focal, perceived over an area much larger than that of the stimulus. Deep pain has diffuse boundaries. It is frequently associated with a sense of nausea and ill-being. Autonomic and motor reflexes associated with deep pains are often extreme and prolonged. Muscle rigidity may itself form a new source of pain, although this is contested by some. Only in exceptional circumstances is deep pain well localised.

2.4. Referred pain

In contrast to pain deriving from stimulation of skin, much visceral pain is localised to distant structures, a phenomenon known as referred pain. The area of referral is generally segmental and superficial, that is, to muscle and/or skin innervated by the same spinal nerves as the viscus giving rise to the referred sensation. A classic example is the pain that develops shortly after myocardial infarction. Although the initial pain in these cases may be felt deep within the chest, with time (usually measured in minutes) it is often felt in parietal structures. Here it is still not well localised, but most often perceived as diffuse within the anterior chest and left arm. In some patients, the referred pain becomes even more superficial, involving cutaneous structures as time progresses (Procacci et al. 1986). Another example is the pain of renal colic which is felt in the iliac fossa and scrotum. The general pattern of referral is consistent enough to be of diagnostic use, although confusion can arise from viscera that share a common segmental innervation (i.e., those within a viscerotome), for example, the heart and oesophagus. One notable feature of referred pain is that it masks the original "true" visceral pain.

Descriptive studies on the nature of referred sensations in patients are sometimes confounded by the possibility that the effective stimulus moves from a visceral site to a parietal one. For example, the rupture of an inflamed appendix is associated with the sudden appearance of a pain localised in the lower right quadrant of the abdomen (Silen 1987). Similarly, the growth of a tumour may newly involve nonvisceral tissue. Stimulation of the body wall – and especially its membranous linings – is well recognised as giving rise to poorly localised deep pain in some cases, or a more superficial referred pain in others. It is clear, however, that some examples of referred pain cannot be explained on this basis, such as the pain of renal colic, angina, and in the extreme, the referred pain felt when the splanchnic nerve of conscious humans has been stimulated electrically (Foerster 1933; Leriche 1937).

Another important feature of referred pain is that the site of referral may additionally show hyperalgesia (e.g., Head 1893; Procacci et al. 1986). This is true for the pain referred to muscle and to skin as well. Such tenderness develops slowly, taking many minutes or even hours to become manifest and, equally, persisting for prolonged periods, measured in hours.

2.5. Mechanisms of referral

Figure 1 illustrates a number of possible explanations that have been offered for referred sensations. The first case shown originated from Sinclair et al. (1948). The hypothesis is that some primary sensory neurones have widely bifurcating axons and innervate both somatic and visceral targets, thus obscuring the source of afferent activity, and explaining the segmental nature of referred sensations. In support, Bahr et al. (1981) found that 18% of a relatively small sample of unmyelinated fibres in the lumbar splanchnic nerves could be driven by electrical stimuli applied to segmentally appropriate somatic nerves. Some of these may have been sensory neurones, but no attempts were made to identify receptive fields in peripheral tissues. There have been other positive results reported for pairs of somatic nerves (Pierau et al. 1982; Taylor & Pierau 1982), but these findings have been challenged on technical grounds (Devor et al. 1984). The only positive data for sensory neurones with receptive fields in two tissues comes from a study by Mense et al. (1981) who reported single sensory neurones with both skin and muscle fields innervating the tail of the

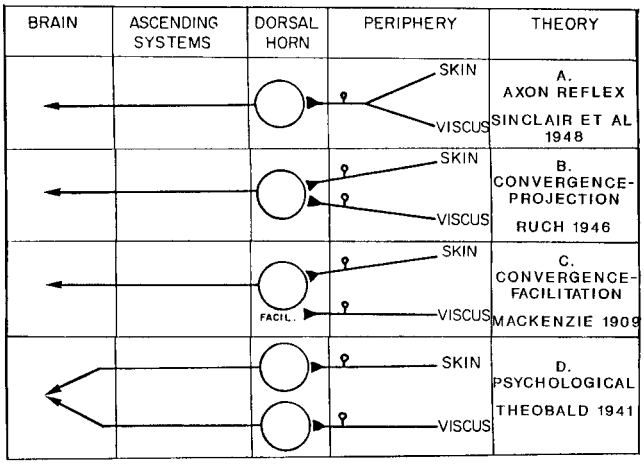


Figure 1. Summary diagram illustrating the various theoretical mechanisms of referred pain (from Morrison 1987).

cat. Recently, Takahashi et al. (1993) provided evidence for bifurcating nociceptors with terminals in intervertebral discs and skin. This first hypothesis does not explain the time delay in the evolution of referred pain however. Nor does it explain the referred hyperalgesia that frequently develops, because antidromic activity (that might invade the distant branch) does not appear capable of inducing a sensitization of peripheral terminals (Reeh et al. 1986).

Another putative mechanism of referred pain is that visceral and somatic primary sensory neurones converge onto common spinal neurones. This is the projection-convergence theory, suggested as such by Ruch (1946), but derived from earlier ideas of Sturge (1883) and Ross (1888). This proposes that the activity in ascending spinal pathways is misconstrued as originating from somatic structures. This theory can explain the segmental nature of referred pain. There is now considerable experimental evidence that somatovisceral convergence is common in spinal neurones (as reviewed in McMahon 1994; Ness & Gebhart 1990), but it should be remembered that many instances of such convergence may relate to the integration of somatic and visceral reflexes, rather than to viscerosensory processing. The theory does not explicitly address the issue of referred hyperalgesia. It is possible that summation of inputs from visceral and cutaneous structures could underlie cutaneous hyperalgesia, but the theory offers no explanation of the slow evolution of referred tenderness.

A variation on the theme of convergence-projection (see Fig. 1c) derives from the ideas of MacKenzie (1909), and is called the convergence-facilitation theory. Because MacKenzie was convinced that the viscera were wholly insensitive and believed therefore that visceral afferent activity itself never gave rise to pain, he proposed instead that this activity was capable of creating an "irritable focus" within the spinal cord. Thus, other, segmentally appropriate somatic inputs could now produce abnormal and referred pain sensations (MacKenzie 1909). His theory did not find general acceptance, in part because it implicitly denied the existence of "true" visceral pain. However, the theory offers an explanation for referred hyperalgesia and, perhaps, the delay in the referral of sensations (allowing for the generation of an "irritable focus"). The concept of an irritable focus has been resurrected with another label – central sensitization, which appears to be of cardinal importance in hyperalgesia from somatic and visceral structures (see Mayer & Gebhart 1995; McMahon 1994).

A final view of referred pain is illustrated in Figure 1d, and suggests that interactions at supraspinal levels lead to the phenomenon (Theobald 1941). Most electrophysiological data we have relating to projections from the spinal cord to brainstem suggest that viscerosomatic convergence is extremely common, and such convergence is contrary to this theory (see McMahon 1994). There is some evidence, however, that a subset of ascending spinal neurones conveys exclusively visceral information (e.g., Akeyson & Schramm, 1994) and there is additional evidence that, whatever the mechanism, some supraspinal structures may functionally respond to visceral information only (see Cechetto & Saper 1987).

In summary, it is clear that the nature of pain from cutaneous and visceral tissues differs in a number of important respects, as summarised in Table 1. There are a number of general reasons why this might arise. First, quantitative differences in the density of innervation of somatic and visceral tissues may have a major impact in the precision of perceived sensations. It is also true that the tissues exhibit very different physical properties, such as their degree of viscoelasticity. Such differences may greatly alter the encoding properties of afferent terminals. Alternatively, the observed differences in sensory capacities may have arisen secondarily to the evolution of other systems. For example, the skin is endowed with a numerically large and highly specialised system of large afferent fibres that can relay information about tactile stimuli with great precision. This system is associated with a large and precise cortical representation of the body surface. It is therefore possible that the high degree of localisation seen for cutaneous pain derives incidently from the existence of the accurate cortical map of the body surface. Finally, visceral and somatic systems may be fundamentally different for good teleological reasons; for instance, that the tissues are normally exposed to very different types of stimuli and participate in different behavioural repertoires.

3. Comparison of cutaneous and visceral afferent primary sensory neurone properties

3.1. Anatomy

The skin of all mammalian species is richly innervated by sensory neurones. In man, perhaps a million or so fibres project from the skin of the body and head to the spinal cord (Holmes & Davenport 1940). In contrast, the abdominal and thoracic viscera receive but a sparse innervation, probably amounting to only 5% or so of the numbers of somatic afferents. This paucity is all the more impressive remembering that many visceral structures are thrown into folds and offer an immense surface area. The innervation density in viscera is therefore only a small fraction of that seen in skin. Afferents to both structures are bundled in peripheral nerves. In the case of visceral afferents, these fibres run with the sympathetic or parasympathetic nerves. There is an exception with the afferent innervation of much of the peripheral vasculature (considered a visceral target because it is innervated by the sympathetic system), which is innervated by afferents running in appropriate somatic nerves. This anatomy has led to some confusion over terminology:

Table 1. *Features of visceral and cutaneous pain compared*

	Visceral	Cutaneous
Effective stimuli	Direct trauma ineffective Distension & ischema effective	Direct trauma effective
Summation	Yes?	No
Localisation	Local–diffuse Often referred	Local–precise Not referred
Primary hyperalgesia	Yes	Yes
Secondary hyperalgesia	Yes, at site of referral	Yes, around site of damage

the sensory neurones in the vagal or pelvic parasympathetic nerves are often called parasympathetic afferents, and similarly, those in the hypogastric nerves, for example, are frequently referred to as sympathetic afferents. The terms sympathetic and parasympathetic strictly refer to efferent systems, but there is an obvious economy in applying them adjectivally to afferent neurones. On the other hand, the use of these terms appears to bestow, a priori, some special property on visceral afferents that may be unwarranted. Some texts claim that only "sympathetic" afferents are responsible for signalling visceral nociceptive events. This generalisation is also an oversimplification. For example, for the urinary bladder (which receives afferents via both sympathetic and parasympathetic nerves), both clinicopathological investigations and studies after surgical interruption of individual nerves have determined that the pain of acute overdistension or cystitis can be signalled by primary afferents in the parasympathetic pelvic nerve (Bors & Comarr 1971; Denny-Brown & Robertson 1933b; Gunterberg et al. 1975; Head 1893; Head & Riddoch 1917; Learmonth 1931; Petersén & Franksson 1955; Ray & Neill 1947; Riddoch 1921; White et al. 1952). Indeed, for this organ, little information exists about the function of the sympathetic afferents, projecting to the thoraco-lumbar spinal cord. Interruption of these pathways (as is often the case today in radical retroperitoneal lymphadectomy for testicular cancer) does not appreciably interfere with bladder sensation. Similarly, the ability of visceral afferents to induce c-fos in central neurones, often tacitly assumed to represent activation of pain-signalling pathways, actually appears much more pronounced for "parasympathetic" than "sympathetic" afferents (e.g., Traub et al. 1994).

Visceral afferents have the same general anatomy as their somatic counterparts, with terminals in both peripheral targets and spinal cord/brainstem, and cell bodies in dorsal root ganglia or ganglia of cranial nerves. Although afferents from skin project to the brainstem and along the entire length of the spinal cord, visceral afferent projections are more restricted. In the spinal cord, the afferents running with sympathetic nerves project to thoracic and upper lumbar segments, whereas those running with parasympathetic nerves project to the second–fourth sacral segments in man.

A major difference in cutaneous and visceral afferents is seen in the size distribution of fibres present. The classification of afferent fibres that is applied to skin derives from the work of Gasser and Grundfast (1939). Three groups are recognised: large myelinated (Aβ) fibres, small myelinated (Aδ) fibres, and unmyelinated (C) fibres. In a typical cutaneous nerve the Aβ fibres amount to some 20–25% of the total, the Aδ's, about 10–15%, and the C's, 60–70%. There is a great deal of functional specialisation amongst the cutaneous afferent fibres. The large myelinated afferents respond to innocuous events such as light touch or limb movement. Many of the smaller diameter somatic afferents, conducting in the Aδ and C velocity range, are nociceptors (see sect. 3.2). In visceral nerves, very few large myelinated fibres are present, numbering only a small percent of the total population. These appear to innervate Pacinian or Pacinian-like corpuscles located mostly in the mesenteries (see Jänig & Morrison 1986). The vast majority of afferents are Aδ and C fibres, and these have to encode both innocuous and noxious stimuli. The ratio of unmyelinated:myelinated fibres, unlike skin, is about 10:1 (see Jänig & Morrison 1986; Willis & Coggeshall 1991).

One final anatomical difference is seen in the preponderance of ventral root afferents. In a variety of species, including rat, cat, and man, it has been reported that some afferent fibres project not in the dorsal root but in the ventral root (see Willis & Coggeshall 1991). These fibres have their cell bodies, as normal, in the dorsal root ganglion, and although some of the ventral root fibres may be loops or branches or dorsal root projections, in some cases the ventral root projection appears to be the only one (Häbler et al. 1990c). These ventral root projections are much more common in spinal segments receiving a visceral projection and a relatively high proportion of visceral afferents have these ventral root projections (Häbler et al. 1990c). However, because the functional role of this group of fibres remains unknown (indeed it is not even clear if these fibres project into the spinal cord or end blindly in the ventral root; see Willis & Coggeshall 1991), the significance of this difference in somatic and visceral afferents is entirely speculative. It is possible that, during development, selective guidance cues operate for afferents innervating different targets, and that the nature of these cues for visceral afferents simply leads to more developmental errors in trajectory.

3.2. Physiology

The area of greatest controversy in viscerosensory processing relates to the encoding properties of visceral afferents. As described above, afferent Aβ fibres are relatively abundant in somatic tissues. These fibres obviously have the highest conduction velocities and they are well suited to the

rapid transmission of precise somatosensory information about the impact of the outside world on the body. The smaller cutaneous afferents, conducting in the Aδ and C velocity range, are mostly nociceptors that respond when stimulus intensities are raised so as to threaten the integrity of the tissue. This distinction between low- and high-threshold afferents provides the main prop for the so-called specificity theory of sensory processing (Fig. 2a). The concept is straightforward. As stimulus intensity rises from liminal levels, a specialised group of low-threshold, tactile afferents is recruited. As stimulus intensities increase through the normal innocuous, range, these afferents increase their discharge rates and then begin to saturate. With yet further increases in stimulus intensity into the noxious range, an entirely new group of afferents, small diameter nociceptors, are recruited.

In visceral nerves, as I have indicated, there are practically no Aβ afferents. The few present are probably incapable of encoding information related to individual viscera (Jänig & Morrison 1986). This strongly suggests that both painful and nonpainful sensations, and the afferent information used to regulate visceral reflexes, must be carried by the small afferent fibres. Therefore, a crucial question is whether the visceral afferents, like somatic afferents, can be divided into separate groups responding to innocuous and noxious events, respectively, conforming again to the specificity theory of pain (Fig. 2a). The problem of transferring the concept of "noxious" from somatic to visceral tissues has already been discussed, and it is not always clear what properties one would expect from a specific visceral nociceptor. The problem is compounded by the fact that the response properties of visceral afferents have been determined in animals. Even when the nature of the effective stimulus is clear, as in the case of distension of the hollow viscera, one must extrapolate across species

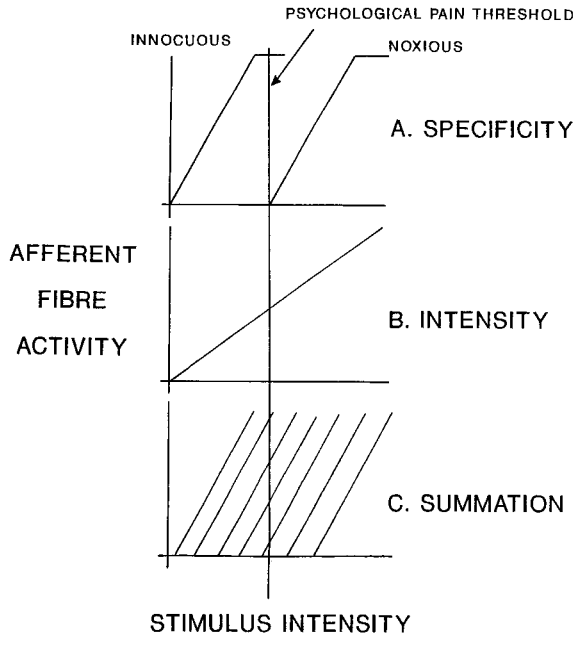


Figure 2. Diagrammatic representation of three possible encoding mechanisms for noxious events by visceral afferent nerve fibres. See text for details. (Adapted from Cervero 1988.)

to determine the levels at which the stimulus becomes painful. The use of pseudoaffective responses (such as increases in blood pressure) as a determinant of nociceptive threshold is not without its problems (see McMahon 1994).

An alternative to the specificity theory is illustrated schematically in Figure 2b. This is the intensity theory, which requires that individual fibres encode physiological, innocuous events and, with higher discharge frequencies, supraphysiological, presumed noxious ones. Clearly these two theories are mutually exclusive because (1) the specificity theory denies any contribution from other than specific nociceptors and (2) the intensity theory requires this contribution. Some workers have reported on visceral afferents that appear to conform to the specificity theory, whilst others find afferents that are clearly signalling events in both the physiological and supraphysiological ranges. It is interesting that, for the most part, the electrophysiological data put forth by proponents of the two theories do not relate to the same viscus.

By way of example, the electrophysiological evidence deriving from two well-studied tissues and representing some of the conflicting interpretations that have been made will be considered in the following sections.

3.2.1. The urinary bladder. Distension of the urinary bladder in healthy humans initially gives rise to a sensation of fullness and eventually pain as volume increases and intravesical pressure exceeds about 25–35 mm Hg (Bors & Comarr 1971; Bors et al. 1956; Denny-Brown & Robertson 1933a; Morrison 1987; Nathan 1956). Primary sensory neurones innervating the normal urinary bladder have been repeatedly and carefully studied (Bahns et al. 1986; 1987; Dmitrieva & McMahon 1996; Floyd et al. 1976; Häbler et al. 1988; 1990a; 1993a; Sengupta & Gebhart 1994b; Wen & Morrison 1994). Nearly all are small myelinated (Aδ) or unmyelinated (C), and travel with sympathetic (hypogastric) or parasympathetic (pelvic) nerves. Some (or many) exhibit a low level of ongoing discharge when the bladder is empty. There appear to be some species differences, as follows:

In the cat, bladder distension excites nearly all thin myelinated afferents, with pressure thresholds corresponding to the values at which humans report the first sensation of fullness. Nearly all mechanosensitive afferents are activated by the intraluminal pressures reached during normal, nonpainful micturition, and they form a homogenous population in terms of their stimulus-response functions. Mechanosensitive afferents respond in a graded fashion to increases in the intravesical pressure throughout the innocuous and into the supraphysiological, noxious, pressure range (Floyd et al. 1976; Häbler et al. 1990a; 1993a). These afferents reflect the magnitude and the temporal profile of intravesical pressure changes with high accuracy.

Also in the cat, the unmyelinated population of afferents projecting through the pelvic nerve differ in their properties. Very few fibres (<2.5%) respond to changes in intraluminal pressure in normal animals, and these differ significantly in their response properties from the population of thin myelinated fibres (Häbler et al. 1990b). They have pressure thresholds of 30–50 mm Hg, outside, or at the top end of, the physiological range. Thus, there are only a few afferents that could be called specific nociceptors in the bladder, and which would signal only painful levels of distension. It is illuminating to estimate the magnitude of

the afferent inflow arriving at the sacral spinal cord by these different afferent populations (McMahon & Koltzenburg 1993). At an intravesical pressure of 50 mmHg (painful in man and beyond the normal physiological pressures), the total afferent discharge in the cat is about 4500 action potentials per second, of which only about 225 (around 5%) are contributed by the unmyelinated fibre population.

In the rat, most afferents in the pelvic nerve appear to have some spontaneous activity (Sengupta & Gebhart 1994b). The clear distinction between Aδ and C fibre properties seen in the cat is much less clear. Fibres that do not respond even to supraphysiological levels of distension are found by all workers (Dmitrieva & McMahon 1996; Sengupta & Gebhart 1994b; Wen & Morrison 1994), but they are less common than in the cat, represent only some of the unmyelinated population, and are not restricted to this population. The mechanosensitive population, as in the cat, form a largely homogenous population with pressure thresholds in the innocuous range, but with responsiveness extending clearly into the supraphysiological range. However, Sengupta and Gebhart (1994b) reported that 20% of mechanosensitive fibres (and therefore perhaps 10–15% of all bladder afferents) had higher pressure thresholds, averaging 34 mm Hg. A similar proportion of bladder afferents (7%) from a smaller sample were found by Dmitrieva and McMahon (1996) to have high pressure thresholds.

The controversy relates to the interpretation of this data. On the one hand, some have argued that the existence of a subgroup of afferents with thresholds for activation in the range that might be associated with pain in humans, supports the specificity theory in relation to viscerosensory processing. An alternative interpretation is that these fibres constitute the "tail end" of a distribution of fibres that are essentially intensity encoding. The latter interpretation is supported by the relative paucity of these high-threshold fibres and the fact that they are not distinguishable from intensity-encoding afferents in other respects (e.g., presence or absence of ongoing activity, sensitivity to bradykinin; cf. Sengupta et al. 1992).

3.2.2. Internal reproductive organs. Kumazawa and coworkers have made extensive studies of the response properties of afferent nerve fibres in the superior spermatic nerve of the dog (see Kumazawa 1986), and Berkley et al. (1993) and Hong et al. (1993) have studied in some detail the innervation of female reproductive organs. In the testes, other than a small population of rapidly adapting, low-threshold mechanoreceptors (numbering about 3%), Kumazawa found that afferent fibres form a homogenous group (of both myelinated and unmyelinated afferents) with polymodal receptors in testis and/or epididymis. They can be excited in a slowly adapting fashion to stimuli applied to one or more sensitive spots, each about a millimetre or so in diameter. The threshold for activation varies over a wide range, but about 80% of afferents respond to mechanical stimuli of less than 17g/mm. The afferents are polymodal in the sense that they respond to stimulation with algesic chemicals and heating as well as mechanical stimuli. Bradykinin and hypertonic saline solutions are effective stimuli for the afferents. Prostaglandins sensitize but do not excite the afferents to other stimuli. Heating the exposed testis excites afferents when stimuli exceed about 45°C.

Kumazawa and colleagues correlated the response of canine testicular afferents with earlier psychophysical studies on the thresholds of testicular compression that cause pain in man (Woollard & Carmichael 1933). The testicular afferents could encode the level of compression up to several kilograms of force but most had thresholds below 50g. Woollard and Carmichael (1933) reported that pain was felt in man with compressive forces over 200g or so. Kumazawa concluded that these afferents, although similar in some respects to the polymodal fibres innervating somatic tissues, could not be considered specific nociceptors because they encoded innocuous as well as noxious stimuli.

A sizeable population of hypogastric afferents supplying the uterus of the cat was studied by Hong et al. (1993). These sensory neurones had many features in common with testicular afferents. Nearly all responded in a slowly adapting fashion to mechanical stimuli. The majority were also polymodal in that they were excited by chemical (bradykinin, capsaicin, or potassium chloride) as well as mechanical stimuli. Their response threshold to stimulation with von Frey hairs varied over two orders of magnitude, with about 20% of the fibres forming a high threshold group activated only by high intensity stimuli. It is perhaps surprising that relatively few of the fibres responded to spontaneous contractions of the uterus, and very high levels of intrauterine distension (about 100 mm Hg) were necessary to activate two afferent neurones.

In female rats, Berkley and coworkers have demonstrated a marked dichotomy between uterine and vaginal sensory systems. They have shown that sensory fibres in the hypogastric nerve and pelvic nerve supply the uterus and vagina, respectively (Berkley et al. 1993). Moreover, they find that the pelvic afferents innervating the vagina encode low and high levels of distension, whereas the uterine afferents respond only to extreme levels of distension that are associated with ischemia. The same group (Berkley et al. 1995) has recently correlated the response properties of these afferents with behavioural responses to distension. For vaginal distension they found that animals exhibited detection thresholds that were lower than escape thresholds, and that both of these were in the range encoded by pelvic nerve afferents. One could propose, therefore, an intensity coding function for this structure. In contrast, however, in the uterus, detection thresholds to distension were no lower than escape thresholds, and escape thresholds were very much greater than for vaginal stimulation. In fact, most animals did not respond to 100% of stimuli at any pressure, and a significant minority failed to respond at all. One interpretation of these data is that hypogastric innervation of the uterus is specialised to respond to supraphysiological mechanical events, perhaps only those associated with ischemia. However, there are a number of factors complicating this interpretation. One is that the uterine pressures producing escape were greater than those ever likely (or unlikely) to occur in the nonpregnant rat. Another factor is that the escape threshold was substantially higher than the threshold necessary to evoke hypogastric nerve afferent activity. The authors conclude that "the results also indicated that activity produced in these [hypogastric afferent] fibres, even by abnormal stimuli, does not inevitably result in behaviour" (Berkley et al. 1995). (A similar conclusion was also reached by Eucher-Wamser et al. [1994] for chemical stimulation of cardiac afferents.) A final complication for the female reproductive organs is that there appears to be a marked variation in the sensitivity of afferent

systems through the estrous cycle (Robbins et al. 1992). Given these ambiguities, one can see how contrary interpretations of the data have been possible.

The properties of afferents innervating some visceral organs are held as demonstrating a clear case for intensity coding: distension of the gall bladder in man, both pathologically (following obstruction of the bile ducts) and experimentally, causes pain when intraluminal pressures exceed about 35–45 mm Hg (Csendes & Sepulveda 1980; Newman & Northrup 1956; Ray & Neill 1947). One study on the properties of afferents innervating the biliary system of the ferret (Cervero 1982a) reports the existence of fibres with high pressure thresholds to distension that might therefore be considered nociceptors. These afferents, travelling via the sympathetic splanchnic nerves, have no ongoing activity and respond to direct tactile stimuli applied to restricted sites in the gall bladder and its ducts. These afferents represent perhaps the clearest case of specific nociceptors in visceral tissue. Yet even here all the data are not consistent with a simple intensity theory for pain, because studies on the spinal representation of biliary information disagree in some respects. Cervero (1982b; 1983) reported the existence of dorsal horn neurones with similarly high thresholds to gall bladder distension, but Ammons & Foreman (1984) and Ammons et al. (1984) found in their studies that pressure thresholds were generally in the range 0–10 mm Hg, well below what might be considered noxious.

It is difficult to reconcile some of the differences reported for visceral afferent encoding properties. One potentially complicating factor is that the stimulus-response functions of individual afferent fibres exhibit a continuum of mechanical thresholds, and so the situation depicted in Figure 2c may actually represent physiological reality in many cases. This form of the intensity theory is compatible with the existence of a subpopulation of afferents that appears to fulfil the criteria of specific nociceptors in a particular viscus (i.e., the stimulus-response functions illustrated in both Figure 2a and 2b could be present amongst the afferents innervating a single viscus).

Another consideration is that views on processing of cutaneous noxious stimuli are changing somewhat. The classic description of specific nociceptors has been largely confirmed by microneurographic studies in man in which the sensations evoked by stimuli have been correlated with the firing of individual afferent fibres. A clear conclusion can be drawn when thermal stimuli are applied. When skin temperatures reach levels that subjects judge painful, nociceptors are recruited. Increasing stimulus temperatures are associated with increasing discharge rates of these nociceptors. It has been known for some time, however, that the discharge frequencies that are associated with a just-painful mechanical stimulus are much higher (van Hees & Gybels 1981). More recently, a careful microneurographic study on the relationship between evoked sensations and afferent nociceptive activity using graded mechanical stimuli has findings equally well suited to the intensity theory of sensations, with a significant degree of "nociceptor" activity at stimulus levels not judged painful (Koltzenburg & Handwerker 1994).

3.2.3. Sensitization and the recruitment of "silent" afferents. It is very well established that the encoding properties of cutaneous afferent nociceptors can change in the wake of a strong stimulus. The classic descriptions relate to a leftward shift in the stimulus-response functions of nociceptors to thermal stimuli after the skin containing the nociceptive terminal is subjected to a mild burn. That is, some of these afferents show a lowered heat threshold (and are activated by what are normally innocuous temperatures), and increases in their responses to suprathreshold heating (see J. N. Campbell et al. 1994 for review). It is now clear that a variety of strong stimuli, mechanical and chemical as well as thermal, can trigger this sensitization process. However, it has been much more difficult to demonstrate peripheral sensitization to mechanical stimuli in these polymodal nociceptors. It is not seen, for example, with the same mild burns of skin. Very damaging stimuli can induce some mechanical sensitization of cutaneous afferents (Woolf & McMahon 1985) but desensitization is just as likely with these stimuli. In the visceral domain, there have been far fewer attempts to demonstrate peripheral sensitization. For the most part these attempts have not considered thermal sensitization, because this is rarely a normal stimulus for visceral afferents. (In the case of testicular afferents, thermal sensitization has been reported.) However, in contrast to cutaneous afferents, mechanical sensitization has been observed repeatedly. For example, with mechanosensitive afferents innervating the urinary bladder, the inflammation induced by intraluminal injection of chemical irritants such as turpentine and mustard oil has been shown to excite and sensitize appreciable numbers of afferents to mechanical stimuli (Häbler et al. 1993b), with a leftward shift of the stimulus-response function to changes of intravesical pressure. Recently, we have observed a rapid onset mechanical sensitization of a very large proportion of vesical afferents to peripheral application of nerve growth factor (Dmitrieva & McMahon 1996). Cervero and Sann (1989) have provided some evidence for a lowering of mechanical threshold of ureteric afferents in some conditions.

Haupt et al. (1983) have reported that colonic afferents subjected to ischaemia can show an increase in their levels of ongoing and contraction-related activity. These changes take considerable time to develop. The results suggest a chemosensitivity of these afferents, and this was directly demonstrated in some cases using bradykinin and potassium.

In most of these cases it is not clear what agent is finally responsible for the sensitization. What is clear, however, is that the process of mechanical sensitization is relatively easily induced in visceral afferents and not in cutaneous ones. This is likely to be the case even where the same intervention/substance is compared in the two tissues. For example, nerve growth factor appears to induce a rapid thermal sensitization of cutaneous nociceptors, and a rapid mechanical sensitization of some visceral afferents (Andreev et al. 1994; Dmitrieva & McMahon 1996). It may be that these different propensities to sensitize depend on different receptors expressed on the two types of afferent, which I discuss in section 3.3.

It has recently become clear that at least some unmyelinated afferent fibres do not respond appreciably to physiological or supraphysiological forms of mechanical stimuli. Some of these afferents respond specifically to chemical stimuli and have been called "silent" or "sleeping" afferents (see McMahon & Koltzenburg 1990). They appear ideally suited to signalling changes occurring in inflammatory states. At the onset of an experimental inflammation, some of these fibres become active and, moreover, develop a

novel mechanosensitivity so that they now respond to events such as distension. The presence of these fibres serves to further reinforce the idea that mechanisms of pain may change dramatically when one moves from normal healthy tissue to diseased pathological states.

These silent afferents have been difficult to study experimentally, because of the obvious problems associated with their identification. Even studies specifically designed to isolate this type of fibre may result in systematic under- or overestimates of their numbers. Nonetheless, there have now been several studies of this type of fibre in skin. The emerging picture is that an appreciable minority of unmyelinated afferents (in the range of 10–20%) may be of this type. In viscera, only the innervation of the urinary bladder has so far been systemically studied for this type of fibre.

In the pelvic nerve of the cat, very few fibres (2.5%) respond to changes in intraluminal pressure in normal animals. The overwhelming majority of unmyelinated pelvic afferents innervating the bladder appear therefore to fall into this "silent" category. In the rat, the incidence of these silent fibres appears lower, but now three independent groups have suggested the existence of significant numbers in the pelvic nerve (Dmitrieva and McMahon 1996; Sengupta and Gebhart 1994b; Wen and Morrison 1994). Most or all of these afferents are likely to have a sensitivity to chemical agents, such as capsaicin (Hu-Tsai et al. 1992) or even the constituents of normal urine (Wen & Morrison 1994).

At the onset of an acute vesical inflammation induced by intraluminal injection of chemical irritants such as turpentine and mustard oil some of these silent afferents are excited (Häbler et al. 1990b). These neurones show an initial burst of activity that settles to a lower level as the chemically induced inflammation progresses. Some of these initially mechanically insensitive afferents also acquire a novel mechanical sensitivity in the biologically relevant pressure range (Häbler et al. 1990b). Compared to the number of unmyelinated afferents responding in the normal animal, some four times as many can be excited by distension at the onset of an acute inflammation. The activation of a numerically significant population of initially unresponsive afferents indicates that peripheral afferent mechanisms encoding pain from pelvic viscera are highly malleable and are strongly affected by the state of the tissue. These peripheral changes are obviously likely to be important for signalling pain and discomfort in inflammatory conditions.

The major difference between skin and visceral afferents in this regard appears to be numerical, although it is not yet known whether the afferent innervation of other visceral structures will be similar to the bladder. Lombardi et al. (1981) made incidental observations suggesting that specific chemoreceptors were uncommon or absent amongst afferents innervating the heart via the sympathetic nerves. Sengupta and Gebhart (1994a) also note that some "silent" afferents can be recruited after chemical stimulation of the colon.

3.3. Neurochemistry

3.3.1. Fast neurotransmitters. The majority of primary sensory neurones appear to use the excitatory amino acids glutamate or aspartate as their principal neurotransmitter (see McMahon et al. 1993 for review). The transmitter is localised to at least 70% of dorsal root ganglion cells. Many of the postsynaptic responses of dorsal horn neurones are blocked by amino-acid antagonists. The cutaneous receptive fields of dorsal horn neurones are abolished by the blockade of glutamate receptors. We have less direct information about visceral afferents, but indirect evidence is consistent with the idea that glutamate is the principal neurotransmitter. Thus, micturition contractions triggered by bladder afferent activity are depressed by glutamate antagonists (Rice & McMahon 1994), and central sensitization induced by activity in visceral afferents is blocked by NMDA (N-methyl-D-aspartate) receptor antagonists. The ability of visceral afferents to induce c-fos expression in dorsal horn neurones is also partially blocked by glutamate antagonists (Birder & de Groat 1992). There is therefore no compelling reason to believe that visceral and cutaneous afferents differ fundamentally in the fast neurotransmitters they use.

3.3.2. Neuropeptides and other primary afferent markers. The cell bodies of afferent neurones innervating somatic tissues can be subdivided into three minimally overlapping populations (Averill et al. 1995) as follows: (1) the classically described large light cells, rich in neurofilaments and stained positively with the antibody RT97. These cells have myelinated axons and mostly innervate tactile low-threshold mechanoreceptors. They amount to 30–40% of all somatic afferents; (2) the peptidergic afferents, marked by the CGRP (calcitonin gene-related peptide). These are small and medium sized cells, with some overlap with the RT97 population. Subsets of these cells contain other neuropeptides, such as substance P. Most cells have unmyelinated or thinly myelinated axons and innervate nociceptors. They constitute about 45% of the total somatic population; (3) the nonpeptidergic small dark population. These cells, too, mostly innervate nociceptors via unmyelinated axons. They do not contain neuropeptides, but are marked in a number of other ways. They react to the antibody LA4, and express the enzyme FRAP (fluorine resistant acid phosphatase). They also react with the lectin IB4, but this shows more overlap with the CGRP population. About 40% of cells form this group.

Visceral afferents exhibit different proportions in these classes. The paucity of large afferent neurones necessarily leads to greater proportions of small cells. It is perhaps not surprising that the number of peptidergic afferents is generally found to be greater than among somatic afferents. There are quantitative differences between viscera. For example, the prevalence of substance P-expressing neurones is reported to vary from about 10% in pancreatic afferents to 40% for stomach afferents in the rat (Dockray & Sharkey 1986). In the afferents innervating the urinary bladder, some 60–70% of cells appear to express CRGP and 25% substance P (see Bennett et al. 1996; de Groat 1986; Hunt et al. 1992). FRAP is also expressed by urinary bladder afferents (McMahon 1986). In these respects, visceral afferents appear only quantitatively different from their cutaneous counterparts. However, there are some very striking qualitative differences. Most notably, the peptide VIP (vasoactive intestinal polypeptide) is apparently expressed in very many visceral afferents (at least in the cat), but many fewer somatic ones (see de Groat 1986). It may therefore be a relatively good marker of visceral

afferents. VIP expression shows another interesting feature. Peripheral axotomy of somatic nerves leads to the induction of VIP in substantial numbers of the damaged afferents, whereas axotomy of visceral nerves leads to a reduction in VIP expression. Another marker showing high levels of target specificity may be somatostatin, which is found in a minority of cutaneous afferents but practically no visceral ones.

The enzyme nitric oxide synthase (NOS), which makes the putative neurotransmitter nitric oxide (NO), also appears to be differentially expressed in somatic and visceral afferents. In the rat, for example, cells in the L4 and L5 dorsal root ganglia innervate somatic targets and very few normally express NOS. In contrast, relatively large numbers of afferents express NOS in spinal ganglia innervating visceral targets (Aimi et al. 1991). By retrograde labelling of the splanchnic nerve we have confirmed that it is visceral afferents that normally express NOS (Train et al., unpublished). Like VIP expression, NOS is induced in many somatic afferents after peripheral axotomy (Fiallos-Estrada et al. 1993) but down-regulated in visceral afferents (unpublished observations). It is less clear if nitric oxide normally functions as a transmitter for visceral afferent neurones. It has recently been reported that NOS inhibitors do not block micturition reflexes, although they do prevent the development of bladder hyper-reflexia after experimental inflammation (Rice 1995).

3.3.3. Capsaicin sensitivity. The pungent extract of chili peppers, capsaicin, has a variety of actions on sensory neurones, apparently via a specific capsaicin receptor. It can strongly activate afferents, but with repeated applications can desensitize them. In developing animals capsaicin can selectively kill small diameter afferent neurones. Capsaicin activates peptidergic afferent neurons and leads to the release of, for example, CGRP and substance P. Capsaicin acts on both cutaneous and visceral afferents: about 25% of cutaneous afferents and about 60% of visceral afferents appear to be sensitive to capsaicin (Hu-Tsai et al. 1992). Because the neuropeptides released from the peripheral terminals of afferents by capsaicin are neuroactive, it is not surprising that so-called neurogenic extravasation is well developed in skin and in at least some visceral structures. Again, the difference between visceral and cutaneous systems appears mostly one of degree.

3.3.4. Neurotrophin receptors. In development, neurons of the peripheral nervous system are dependent for survival on neurotrophic factors, which are produced in target tissues. The most important molecules for survival appear to be the neurotrophins, which in mammals currently comprise nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3), and NT-4/5. These proteins exert their actions via another family of tyrosine kinase receptors, known as trks. The receptor trkA mediates NGF responses, trkB mediates BDNF and NT4/5 responses, and trkC mediates NT-3 responses. The neurotrophins not only regulate survival, but may also determine the phenotype of sensory neurones (Ritter et al. 1991). The neurotrophins may also continue to exert actions on sensory neurone phenotypic properties in the adult animal. For example, the availability of NGF appears to regulate the levels of the sensory neuropeptides CGRP and substance P. It is possible that some or many of the differences between cutaneous and visceral afferents (described above) may arise because of the different types or levels of neurotrophic support that afferents receive. It is interesting, therefore, to ask if there are differences in neurotrophin sensitivity between the afferent groups. We have recently reported on the expression of trk mRNAs in functionally defined groups of sensory neurones (McMahon et al. 1994). We found that about 45% of somatic afferents express trkA (the receptor for NGF). These cells overlap almost completely with the CGRP population (Averill et al. 1995). Smaller numbers of afferents express trkB and trkC (about 25% and 20%, respectively). The trkC-expressing cells were large in diameter, whereas the trkB-expressing cells could be large or small. In marked contrast, 90% of afferents projecting through the pelvic nerve expressed trkA, and 94% expressed trkB. Only 2% were found to express trkC. A recent immunocytochemical study broadly supports these findings (Bennett at al. 1996). Clearly, nearly all pelvic visceral afferents must express both trkA and trkB. The large numbers of visceral afferents sensitive to NGF (i.e., expressing trkA) may explain why more of these cells express neuropeptides. The coexpression of trkA and trkB in these afferents also distinguishes them from cutaneous afferents. A summary of the properties of visceral and cutaneous sensory neurones is given in Table 2.

4. Conclusions

Clinical observations on the nature of visceral and cutaneous sensibilities in health and disease, as discussed in the second section of this article and summarised in Table 1, provided a powerful impetus for suggesting that there may be fundamental differences in the organisation and properties of the respective sensory systems. As discussed in the third section of this article, there are indeed differences in the somatosensory and viscerosensory systems, listed in Table 2. Some of these differences are qualitative, such as the fibre size spectrum of primary sensory neurones innervating the two systems, but many are quantitative, for example, the proportions of afferent neurones expressing particular neuropeptides or receptors. An important question, therefore, is to what extent the observed differences in properties can account for the observed differences in sensibilities. The quantitative differences in the density of innervation of somatic and visceral tissues alone may be a sufficient explanation. The apparent insensitivity of visceral tissue to focal yet tissue-damaging stimuli, combined with the effectiveness of distension and ischemia, is readily explained on such a basis. Further, the diffuse nature and poor localisation of "true" visceral pain may have the same explanation. The spatial summation of visceral stimuli, hypothesised to be a crucial determinant of the pain threshold in the hollow viscera, is at least partly explainable in terms of innervation density, although it is not clear on this basis alone why spatial summation appears not to be relevant in establishing the pain threshold in skin, at least for thermal stimuli. Numerical differences in visceral and cutaneous innervation seem at first less relevant to the question of referred pain. It is possible, however, that one consequence of the development of a large and precise cortical representation for cutaneous – but not visceral – information may have led incidentally to a system whereby visceral sensations are at best poorly localised, and frequently "default" to segmentally appropriate somatic structures.

Table 2. *Comparison of properties of primary afferent fibres innervating visceral and cutaneous targets*

	Visceral afferents	Cutaneous afferents
Axon calibre/velocity	Aδ,C	Aβ,Aδ,C
Stimulus-response function	Frequently intensity encoding	Separate high and low threshold populations
Sensitization	Mechanical sensitization common	Thermal sensitization common, mechanical uncommon
"Silent" afferents	Numerous in some organs (e.g., 90% of cat pelvic C fibres)	Less common (c. 20% of C fibres?)
Neurotransmitters	Probably mostly glutamate	Mostly glutamate
Neuropeptides/ neuromodulators	CRGP & SP common, VIP in some populations Nitric oxide constitutively expressed in many afferents	CGRP & SP common, VIP very unusual. Nitric oxide not normally expressed
Capsaicin sensitivity	Many afferents sensitive	Many small-diamter afferents sensitive
Neurotrophin receptors	trkA and trkB in 90% trkC rare	trkA,B,C more evenly distributed. Many small afferents no known trk
Receptors	5HT, GABAa, bradykinin, opiate	5HT, GABAa, bradykinin, opiate
Ventral root afferents	Relatively common	Uncommon
Neurogenic extravasation	Well-developed	Well-developed

An alternative view would be that the few marked qualitative differences in visceral and cutaneous innervations underpin the observed differences in sensibilities. The most significant, if controversial, difference relates to the hypothesised physiological encoding properties of the two afferent systems. If it is true that much or all visceral pain is signalled by intensity-encoding afferents, rather than a dedicated group of specific nociceptors, then it could be argued that the ineffectiveness of local trauma to produce pain in viscera arises because of insufficient activation of afferents. Another feature of an intensity-encoding mechanism is that it inherently encompasses the idea of summation, also arguably explaining the effectiveness of distension and ischemic stimuli. Encoding by an intensity mechanism does not itself address the issue of referred pain or referred hyperalgesia, which would appear to require a central rather than peripheral explanation. The observed sensitization of visceral afferents, including the sensitization of so-called silent afferents, is likely to contribute to primary hyperalgesia, although intensity-encoding and specific systems have equal explanatory power in this respect.

ACKNOWLEDGMENT

The work of the author reported in this target article was supported by the Medical Research Council (UK).

BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 392–403
Printed in the United States of America

Plasticity: Implications for opioid and other pharmacological interventions in specific pain states

Anthony H. Dickenson
Department of Pharmacology, University College London, London WC1E 6BT, United Kingdom
Electronic mail: *anthony.dickenson@ucl.ac.uk*

Abstract: The spinal mechanisms of action of opioids under normal conditions are reasonably well understood. The spinal effects of opioids can be enhanced or reduced depending on pathology and activity in other segmental and nonsegmental pathways. This plasticity will be considered in relation to the control of different pain states using opioids. The complex and contradictory findings on the supraspinal actions of opioids are explicable in terms of heterogeneous descending pathways to different spinal targets using multiple transmitters and receptors – therefore opioids can both increase and decrease activity in descending pathways. These pathways could exhibit considerable plasticity. There is increasing evidence that delta opioid receptor agonists have the potential to replace morphine as major analgesics with reduced side-effect profiles. The concept of preemptive analgesia, based on preventing the induction of some of the negative plastic influences on opioid controls and the detrimental effects of pain, is sound, but experimental verification in the clinical setting is difficult. For example, a delayed compensatory upregulation of inhibitory systems, particularly in inflammation, may counter persistent painful inputs. Combination therapy with opioids may be beneficial in many pain states where either negative influences are blocked or inhibitory controls are enhanced. Finally, developmental aspects of these systems are discussed in connection with the treatment of pain in young children, where inhibitory systems in the spinal cord are immature.

Keywords: analgesia; cholecystokinin; development of pain; excitatory amino acids; hypersensitivity; nociception; opioids; peptides; spinal cord

1. Introduction: Plasticity in pain

The two major types of clinical pain arise from distinct events in the periphery. Inflammatory pain arises from tissue damage such as that produced by trauma, surgery, childbirth, or invasion of tissue by a tumour. The second type of pain is termed "neuropathic" pain and results from damage to a nerve; trauma, surgery, and cancer can also cause this type of pain. In the case of inflammatory pain, the damage to tissue causes the local production of a number of chemical mediators that sensitize and/or activate the peripheral endings of nociceptive C-fibres (Dray et al. 1994). With neuropathic pain, however, activity is set up in the nerve itself. Regardless of the origin of the pain, impulses are conveyed to the first synapse in the dorsal horn of the spinal cord, where the interplay between excitatory and inhibitory events determines the ascending messages that are transmitted to higher centres. However, descending controls from the brain stem can be triggered and further alter processing in the spinal cord (Besson & Chaouch 1987; Fields et al. 1988). Opioid analgesics can exert controls on these events by direct actions on the spinal cord but can also interact with systems at the origins of the descending controls (Duggan & North 1984; Dickenson 1994a). Thus, at a number of levels there is the potential for alteration in the messages that ultimately give rise to the final sensation of pain. This article is an attempt to bring together some of the interactions between excitatory and inhibitory events to explain some of the different characteristics of inflammatory and neuropathic pain and to investigate how alterations in these systems can give rise to difficulties in treating certain pains, especially neuropathic pains.

It is well established that the repertoire of the adult central nervous system (CNS) is not fixed and immutable. Plasticity, the ability of central nervous function to change in response to internal and external events, can be due to alterations in connectivity (Woolf & Doubell 1994; Dray et al. 1994). Plasticity is of great relevance to the control of these different pain states (McQuay & Dickenson 1990), yet plasticity can also result from a relatively rapid induction and activation of different pharmacological systems under different circumstances (Dickenson 1994a; 1994b; McMahon et al. 1993; Price et al. 1994b; Woolf 1994). This target article will concentrate on interactions between spinal pharmacological systems and opioid analgesia. States of increased pain transmission, central hypersensitive states (Woolf 1983), can result from activation of spinal systems that do not participate in the responses to brief stimuli. In these cases where the level of excitatory transmission is augmented, opioid inhibitions will need to be increased to compensate. Consequently, the actions of opioids are not fixed but highly dependent on activity in other transmitter systems that in turn appear to be influenced by different types of pain. This target article will consider this plasticity in relation to the control of different pain states using

0140-525X/97 $9.00+.10

opioids as a theme. The main emphasis will be on results from models of acute and more prolonged pains such as that arising from inflammation and neuropathy and I will attempt to discuss why the actions of opioids may differ in different pain states. Thus, pain arising from tissue damage (inflammatory nociceptive pain) can respond well to opioids whereas neuropathic pain (arising from peripheral or central nerve damage) and allodynia (where touch is perceived as pain) can show poor opioid sensitivity.

2. Opioid receptors

Opioids act by activating three opioid receptors: the mu, the delta, and the kappa (Kosterlitz 1985). Whereas the endogenous opioid peptides, the natural ligands for the receptors – namely the enkephalins, dynorphins, and endorphin family – are not entirely specific for any one of these receptors, a number of synthetic agents with high selectivity are available to study the individual receptors (Dickenson 1994a; Kosterlitz 1985). Neurones producing the different opioids can be unequivocally identified now using messenger RNA probes for the precursor propeptides. Opioid peptide synthesis can be altered in animal models of different pain states, for example, dynorphin levels in the spinal cord during inflammation are increased enormously due to the switching on of the gene for the synthesis of the parent propeptide (Dubner & Ruda 1992). The biological lifetimes of the endogenous opioids, particularly the enkephalins, are brief due to rapid peptidase degradation. It is now possible to protect the enkephalins from breakdown by the use of peptidase inhibitors, some of which are now active by systemic routes (Roques et al. 1993). The use of these agents together with the synthesis of stable analogues of the endogenous opioids themselves has provided the means for the study of the roles and function of opioids and their receptors with far greater selectivity than with only the endogenous opioids.

2.1. Mechanisms of opioid analgesia. There are three key mechanisms of action for opioids. The underlying events by which opioids interfere with the transmission of pain are the same as the mechanisms by which opioids cause their other actions, including side effects (Dickenson 1994a; Duggan & North 1984). These are:

1. A presynaptic action on the terminals of neurones whereby transmitter release is reduced by activation of opioid receptors. In tissues where relative receptor location has been gauged, the number of presynaptic opioid sites predominates over postsynaptic locations.

2. There are significant numbers of postsynaptic opioid receptors and after activation the resultant hyperpolarisation reduces evoked activity in the neuronal pathways. The postsynaptic effects can be on cell bodies of output neurones, interneurones, or dendrites.

3. An alternative postsynaptic action involves disinhibition: in a circuit of two neurones, where the second cell is held in check by an inhibitory neurone, opioid inhibition of the first neurone allows the second cell to become active.

These actions produce analgesia at a number of sites in the nervous system. The two key sites would appear to be a spinal and a midbrain/brain stem action in normal circumstances, but an additional peripheral site in inflamed tissue can also be induced.

3. Spinal analgesia

Opioid receptors in the spinal cord are a critical site in the production of analgesia. Spinal opioid analgesia demonstrates how basic research in animals can have a rapid and important application to the clinical relief of pain. Opioid inhibition of nociceptive neurones in spinal animals and then evidence for analgesia following epidural and intrathecal opioids in animals was soon followed by clinical usage (Besson & Chaouch 1987; Yaksh & Nouiehed 1985).

3.1. Presynaptic actions. The highest levels of opioid receptors in the spinal cord are around the C-fibre terminal zones in lamina 1 and the substantia gelatinosa with lower levels found in deeper layers. The best current estimates suggest that the mu receptor forms 70%, the delta 24%, and the kappa 6% of the total opioid sites in the rat spinal cord (Besse et al. 1990). The idea that kappa levels are higher in the mouse and guinea pig spinal cord has been put forward. However, studies in species other than the rat have not been carried out with the most selective ligands for the receptors and so may not represent the true relative distribution of the receptors. We lack systematic quantitative studies in a number of species on the relative distribution of the receptors at a variety of CNS sites. Recent studies using probes for the selected sequences of the delta receptor have shown unequivocally that many of these receptors are located presynaptically on afferents and in close apposition to enkephalin-containing cells (Dado et al. 1993). In addition, spinal application of antisense to the delta receptor has shown that this leads to a marked reduction in delta-mediated analgesia without alteration of the effects of morphine (Uhl et al. 1994).

The relative numbers of presynaptic and postsynaptic receptors can be calculated after nerve section and the former predominate. The proportions of presynaptic opioid receptors in the spinal cord varies from 70% to 50%, with over 70% of the total mu receptor sites (Besse et al. 1990), along with large numbers of delta receptors (Dado et al. 1993), on the afferent terminals. Given the large number of receptors, it is not surprising that evidence for a presynaptic action of opioids emerges from studies of opioid inhibition of C-fibre evoked release of transmitters (substance P and glutamate) as well as in vitro and in vivo electrophysiological studies (Dickenson 1994a; Yaksh & Malmberg 1994; Yaksh & Nouiehed 1985). However, other approaches have failed to demonstrate this presynaptic action (e.g., Lang et al. 1991).

Presynaptic actions on transmitter release result from an opening of potassium channels (mu and delta receptors) or a closing of calcium channels (kappa), both of which lead to a reduction in calcium influx into C-fibre terminals – thus diminishing transmitter release (North 1989). C-fibres are believed to release a number of coexisting transmitters including the tachykinins, excitatory amino acids, and a number of excitatory peptides that act on multiple receptors (Besson & Chaouch 1987; Dickenson 1994a; Dray et al. 1994). Consequently, the presynaptic action of opioid's ability to reduce the release of many transmitters will be a highly effective route to analgesia, since it will be equivalent to the block of multiple postsynaptic receptors. It is not probable, therefore, that any single antagonist of one of these postsynaptic receptors will have sufficient efficacy to compete with the opioids as do powerful analgesics in acute and chronic pains. An exception to this is likely to be agents

acting directly, or indirectly, to modulate the N-methyl-D-aspartate (NMDA) receptor in pains where central hypersensitivity, an augmented spinal response to a low or moderate afferent input, is generated.

Section of a peripheral nerve will lead to degeneration of the nerve and the presynaptic receptors synthesized in the dorsal root ganglion will be lost (reducing mu opioid receptor sites in the dorsal horn by 70%; see Besse et al. 1990). It would be beneficial to know if less severe nerve damage impairs the production of functional opioid receptors. However, in animal models where the presynaptic opioid receptors have been removed by prior rhizotomy, whether electrophysiological with systemic dosing in spinal preparations (Lombard & Besson 1989) or behavioural with spinal application (Xu & Wiesenfeld-Hallin 1991), morphine is still effective, although higher doses are required in the former case. Thus when there is a loss of opioid receptors, there will be a reduction in opioid actions that should predictably be overcome by dose escalation (Lombard & Besson 1989). If nerve damage does lead to a loss of some of the spinal opioid receptors, one could predict that opioid delivery to supraspinal sites would target the normal population of supraspinal receptors. In clinical practice there is a reduced systemic opioid effectiveness in neuropathic pains (Arner & Meyerson 1988) that can be overcome by increasing the dose in some circumstances (Jadad et al. 1992; Portenoy et al. 1990). Dose increases may not always be possible, since side effects may become intolerable. When the side effects can be controlled and increase in dose does not overcome the pain, we have to suspect that opioid receptor loss or dysfunction is not the only factor.

3.2. Postsynaptic actions. Evidence for functional opioid actions at postsynaptic receptors is based on electrophysiological and behavioural approaches (Duggan & North 1984; Lombard & Besson 1989). Postsynaptic hyperpolarisations again result via the opening of K^+ channels or the closing of calcium channels (North 1989). These receptors could hyperpolarise the dendrites of projection neurones and interneurones (both would be selective for noxious transmission) as well as the cell body of projection cells that may not be selective for nociceptive inputs, since many but not all neurones in the dorsal horn receive both nociceptive and tactile inputs.

An important indirect postsynaptic action is the opioid disinhibitory effect mediated via GABA and enkephalin neurones in the substantia gelatinosa, which in turn leads to an inhibition of output neurones. Thus in this instance, neurones can be recorded in the substantia gelatinosa that are facilitated by opioids, an action that requires $GABA_A$ receptor function. There is both morphological and electrophysiological evidence to support this action (see Magnuson & Dickenson 1991).

These postsynaptic actions of opioids present some problems of interpretation, since any direct hyperpolarisation of a cell soma would inhibit all responses of the cell including the innocuous inputs onto convergent or multireceptive cells. However, many of the opioid receptors in the substantia gelatinosa could be on the dendrites of the deep cells penetrating into the C-fibre terminal zone; inhibitory effects here would also be selective, as they are likely to be spatially distinct from the large fibre inputs. Another possibility is that the postsynaptic disinhibitory effects of opioids selectively feed onto nociceptive circuitry. When allodynia and some of the hyperalgesias are transmitted through A-fibre afferents, a poor sensitivity to opioids might be found (Yaksh 1989), since the only opioid control of A-fibre inputs is via the relatively small number of postsynaptic receptors on the output neurones (Besson & Chaouch 1987; Duggan & North 1984). Doses of morphine that abolish C-fibre-evoked responses in normal animals have only minor effects on A-fibre activity (Dickenson & Sullivan 1986). In these pain states, novel nonopioid therapy directed at the spinal systems generating the tactile-evoked pain may have to be considered (Yaksh 1989).

4. Alternatives to morphine?

Surprisingly, it is unclear whether different opioids may have slightly different ranges of pharmacological actions that could allow a choice of opioid for different pains. Clinical studies comparing different opioids in different pain states are needed. Morphine at physiological doses probably acts only at the mu receptor. This drug has a high affinity for the mu receptor, a relative affinity of 50 times less for the delta receptor, and a minimal affinity for the kappa receptor (Kosterlitz 1985). Thus relatively nonselective effects could occur with very high doses such as those achieved both in neurochemical and in binding studies where non-mu effects of morphine have been reported. However, in vivo, doses of spinal morphine, which are sufficient to abolish the C-fibre-evoked responses of dorsal horn nociceptive neurones, are probably entirely mu receptor mediated. In support of this, there is no evidence for mu–delta cross tolerance from physiological studies (Kalso et al. 1993). Since morphine is the standard opioid for clinical practice, plasticity that is related to mu receptor mechanisms is of great importance (Dickenson 1994a).

Possible opioid drugs that act on receptors other than the mu receptor for morphine are analgesics with reduced morphine-like side-effect profiles. As is the case with any transmitter system, the greater the number of receptors the greater the chance that the desirable effects can be separated from the unwanted effects. With opioid receptors, a further division of the receptors from the main three – the mu, delta, and kappa – has been proposed. The mu receptor has been suggested to consist of a mu 1 and a mu 2 subtype (Pasternak & Wood 1986); the delta has been also subdivided and the kappa receptor has been divided into three subtypes (Jiang et al. 1991; Traynor 1989). Whether these subtypes have functional consequences remains to be seen: physiological consequences are not yet discerned except for the delta subtypes where there is evidence for differential effects of the two receptors (Jiang et al. 1991). The recent cloning of the opioid receptors (Uhl et al. 1994) will further facilitate this task, because probes based on the receptor sequence will provide unequivocal proof of location of the particular receptor and important insights into the mechanisms of opioid actions and the existence of subtypes. At the present time there is no evidence from the cloning studies for receptor subtypes: the receptors that have been isolated, whether mu, delta or kappa, were single identical species. There may be alternative splicing that produces the subtypes or local neuronal tissue environments that allow the subtypes to be expressed and, in addition, the cell lines used so far may underestimate the variability within the opioid receptor family (Uhl et al. 1994). However, it is known that the rat and mouse opioid

receptors are, as far as can be determined, identical in structure and pharmacology to the human receptors, further verifying the important links between animal studies and clinical practice (Uhl et al. 1994).

4.1. Delta opioids. Antagonists for the opioid receptors have demonstrated the independence of mu, delta, and kappa receptors in terms of antinociception, although there have been problems demonstrating kappa receptor agonist effects in some studies (Millan 1990). The independent analgesic effects following activation of non-mu receptors indicate potential for opioid analgesics that are delta or kappa agonists. Kappa opioids are not always particularly effective analgesics in animals and this appears to be reflected in the initial early clinical studies with these drugs in humans. The delta receptor may well be an important target for novel opioid therapy. Animal studies have shown that opioids selective for the delta receptor can equal the analgesic effects of morphine by actions at both spinal and supraspinal sites in a number of nociceptive tests (Dickenson et al. 1987; Jiang et al. 1991; Sullivan et al. 1989a). One could also predict reduced respiratory depression and gastrointestinal effects of delta as compared to mu opioids. There are now reports of potent and selective nonpeptide delta opioids, which have been tested in a number of paradigms in animals. A highly selective delta opioid, SNC 80, has been produced. It was found to be effective by central and systemic (including oral administration) routes, and analgesic effects were reversed by a number of delta but not mu opioid receptor antagonists. Importantly, in tests of respiratory function, SNC 80 stimulated rather than reduced respiratory rate in sheep (Bilsky et al. 1995). The potential of delta opioids is therefore high and eventually delta opioids may be clinical alternatives to morphine. As will be discussed in sections 5 and 7.1, there are suggestions that some of the reduced effects of morphine in neuropathy may be due to negative effects of cholecystokinin on mu receptors – these effects do not extend to delta-mediated actions. Consequently, it is possible that delta opioids may be better analgesics than mu opioids in the treatment of neuropathic pains.

4.2. Endogenous opioids. What about endogenous opioids? The enkephalins are rapidly degraded by membrane-bound peptidases. The synthesis of peptidase inhibitors has been a successful strategy, so that kelatorphan, a mixed peptidase inhibitor inhibiting at least two of the important breakdown enzymes, affords almost complete protection to the enkephalins (Roques et al. 1993). The spinal application of the inhibitor produces a reduction of nociceptive responses of cells, with the pool of enkephalins protected by the inhibitor likely to be derived from both a segmental release and from descending pathways activated by the stimulus. The inhibitions are reversed by a selective delta antagonist (Dickenson et al. 1987a). The very recent reports of RB 101, a systemically active mixed peptidase inhibitor, constitutes the next stage toward the clinical application of this novel approach to pain relief. In addition, the side-effect profile of RB 101 appears to be unlike that of morphine in terms of physical and psychological dependence (Roques et al. 1993).

4.3. Novel peripheral analgesia. Whereas opioids lack peripheral actions in undamaged tissue, there is now good evidence that the consequences of inflammation can reveal a novel site of opioid action that appears rapidly (Stein 1994). The synthesis of opioid receptors occurs in dorsal root ganglion cells as well as other locations in the body. These receptors are transported in the fine afferent fibres in both directions; the centrally directed receptors become the presynaptic receptors and the peripherally transported receptors somehow become active only following inflammation. The relative effectiveness of mu, delta, and kappa receptor activation to elicit peripheral analgesia varies between models, but in arthritic states all three are active (Stein et al. 1989). Alongside the appearance of the functional opioid receptors on afferent nerves, the arrival of endogenous opioid peptides at the injury site seems to be related to immune cell proliferation. Thus opioids unable to penetrate the CNS and, as a result, devoid of central side effects may be good analgesics in inflammatory states via these peripheral sites. There have now been a number of clinical studies on this effect, the bulk of which have been positive. Thus the local application of morphine into the knee joint in patients has been shown to produce a local analgesic effect. A recent study has also shown that the degree of analgesia can be related to the amount of tissue damage and, presumably, to the degree of inflammation (Stein 1994). Peripherally acting opioids may then have potential analgesic effects in inflammation (Stein 1994; Stein et al. 1989).

4.4. Supraspinal analgesia. The first demonstration of opioid actions within the CNS consisted of analgesia seen following intraventricular morphine. Numerous supraspinal sites of opioid analgesia have been established (Besson & Chaouch 1987). These have now been localized to areas in the medial brain stem around the nucleus raphe magnus and extending rostrally to periaqueductal and periventricular grey and other areas with the monoamines appearing to be critical transmitters in these pathways (Yaksh et al. 1988). The roles of these areas in morphine analgesia have been based on microinjection studies and the ability of naloxone, when applied locally into these areas, to reduce the effects of systemic morphine.

The mechanisms of action in opioids at these supraspinal levels still is unclear, particularly about how they interact with descending inhibitory controls. Whereas opioid induced increases and decreases in descending inhibitory controls have been reported, the roles of these descending pathways in different models of various pain states are unknown. We need more information on the physiological and pharmacological bases for supraspinal analgesia in animal models of persistent pain in order to form a basis for the potential use of manipulation of the monoamines in difficult clinical pains.

Examination of the anatomy and the pharmacology of the descending systems may provide a basis for these disparate results with regard to opioid interactions with descending controls. First, descending controls originate from many different areas of the brain stem and midbrain and, in addition, a complex pharmacology exists in these descending pathways (Yaksh et al. 1988). Noradrenaline, 5HT, enkephalin, and substance P are involved, all coexisting in some neurones projecting from the brain stem and midbrain to the spinal cord. In addition to interactions between these transmitters, there are a number of local transmitter systems (cholinergic, GABAergic, and opioid) in the nuclei where the descending controls originate. Direct opioid inhibitions

or indirect disinhibitions could result from local opioid actions in these supraspinal areas. These opposite effects could themselves be on inhibitory and/or excitatory pathways. A further complexity is that the monoamines have a number of receptors at the spinal terminal sites that when activated could inhibit or excite depending on the receptor – an additional complication is the presence of autoreceptor or heteroreceptor control of the release of transmitter at these terminal sites. Whether the postsynaptic receptors are on excitatory or inhibitory elements is also important. Consequently the problem is not to discern the direction of effect of opioids on these systems but to understand the physiological roles and consequences of the mixed opioid actions on these multiple pathways (Dickenson 1994a).

Consideration of the direction of effect of the spinal monoamine receptors serves to illustrate these points. In the case of noradrenaline, there is a general consensus that, notwithstanding a possible role of the alpha-1 receptors, the predominant spinal targets for the transmitter are alpha-2 receptors located in the spinal cord, postsynaptically to the noradrenergic terminals. In a similar manner to the opioid receptors and the afferent nociceptive fibres, these receptors are located both presynaptically and postsynaptically on spinal sensory circuits; there is ample evidence for alpha-2 agonists being effective analgesics in a number of animal models of acute and more persistent pains. In addition, there is little doubt that alpha-2 agonists synergize with morphine, probably as a result of dual activation of separate receptors with similar locations and effector mechanisms (Dickenson & Sullivan 1993; Yaksh & Malmberg 1994). Relatively little is known about the driving force behind pain-related changes in noradrenergic activity in these models. An exception is a report of increased alpha-2 mediated activity in inflammation, but it does not contribute to enhanced spinal opioid effectiveness (Stanfa & Dickenson 1994a). There may, however, be a supraspinal site of action of noradrenaline in enhancing opioid actions (Hylden et al. 1991).

The effectiveness of tricyclic antidepressants (TADS) for pain relief in humans possibly relates to enhancement of the availability of noradrenaline and serotonin. This is where the problems arise. The number of receptors for serotonin or 5-hydroxytryptamine (5HT) increases on a regular basis. Presently, there are at least seven major receptors with over 20 subtypes. The receptor that underlies descending antinociception at the spinal level is unknown. Thus, increases in 5HT levels in the cord will activate all the receptors irrespective of whether they are excitatory or inhibitory. Knowledge of the particular roles of the monoamine receptors may potentially lead to better therapeutic efficacy by agents (or combinations thereof) acting on particular receptors rather than the indirect indiscriminate activation of multiple receptors produced by the TADS (Max 1994).

Bearing in mind these complexities, how do opioids interact with descending pathways? As many of the sites of opioid actions at supraspinal sites overlap with areas where descending inhibitory controls originate, the simplest situation is that supraspinal opioids increase these descending monoamine inhibitions; in turn these block spinal pain transmission by actions at inhibitory spinal receptors. For opioids to increase descending inhibitions, the mechanism will have to be via disinhibitions (Fields et al. 1988). The clearest demonstration of supraspinal descending inhibitory controls that are increased by morphine is that the spinal induction of c-fos, used as a marker of noxious evoked activity, was found to be very clearly reduced by intraventricular morphine (Gogas et al. 1991).

There are other studies, however, that do not find this direction of effect. Diffuse noxious inhibitory controls (DNICs) are descending controls induced by heterosegmental noxious stimulation and partly involve both opioid and serotoninergic mechanisms. Morphine, either given directly into supraspinal tissues or at low systemic doses without direct spinal actions, reduces these descending controls (Le Bars & Villaneuva 1988). There are considerable difficulties, therefore, in arriving at a simple consensus as to the direction of effect of opioids on descending control systems. However, as discussed above, the multiplicity of these descending controls in terms of their anatomy, their pharmacology, and their spinal projections form a framework in which the various directions of effect of opioids can be incorporated. There is no doubt that whatever the mechanism, supraspinal opioids produce behavioural analgesia (Besson & Chaouch 1987; Yaksh et al. 1988).

It is highly likely that both the level of pain transmission and the effectiveness of opioids in different pain states is determined by alterations in descending control pathways. Other than studies on alpha-2 adrenoceptors, we are ignorant of the extent of plasticity in these systems.

5. Plasticity in opioid controls

Using the previously articulated framework of opioid effects and mechanisms on which to consider plasticity in opioid systems, it is pertinent to consider in this section particular pain states where there is evidence for changed opioid effectiveness. Why are opioids sometimes poorly effective in neuropathic pain states in man and animals (Arner & Meyerson 1988; Jadad et al. 1992; Portenoy et al. 1990)? Yet there is good evidence that in a number of inflammatory models opioids are more effective than in normal animals (Dickenson 1994a). In addition, as there are transmitter systems in the CNS that can reduce opioid effectiveness, preemptive analgesia should stop the induction of these systems and so provide better pain relief: Why has it been so difficult to provide clear and marked clinical benefits for this approach (McQuay 1994)?

The analgesic effects of morphine can vary in different pain states. The mechanisms behind these changes have been elusive, but their identification and eventual manipulation may be of considerable clinical benefit. First, let us consider the pharmacological systems that can interfere with opioid effectiveness, bearing in mind that pathology can also play a role in nerve section and the loss in number of presynaptically located opioid receptors. There appear to be four major pharmacological factors:

1. Interference with mu receptor function by the metabolite of morphine, morphine-3-glucuronide (M3G). This has been proposed on the basis of behavioural studies but is not supported by electrophysiological and clinical studies.

2. Changes in the levels of the nonopioid peptides, FLFQPQRFamide and/or cholecystokinin (CCK), either spinally where there is very strong evidence for CCK as a regulator of morphine analgesia, or supraspinally as a more global negative influence on opioid actions.

3. Increased levels of the opioid peptide, dynorphin, which has been shown to occur after persistent pain. In theory this peptide can reduce mu opioid analgesia, but the physiological role of dynorphin as an opioid modulator is not good.

4. An excess of excitatory activity, so that a spinally generated hypersensitive state is induced, against which opioid controls are insufficiently efficacious. The N-methyl-D-aspartate (NMDA) receptor is a very strong candidate for the final common path for generation of this state, and there is poor opioid sensitivity to a number of electrophysiological and behavioural measures of pain where NMDA receptor activation has been induced.

These four possibilities are not mutually exclusive (Dickenson 1994a). Thus, in a particular pain state where opioids are used to treat the pain, NMDA mediation of spinal transmission may be occurring at the same time as elevated spinal CCK and dynorphin levels with high plasma M3G levels (Dickenson 1991). The evidence for and against these systems altering opioid analgesia will be considered in turn.

5.1. Morphine metabolites. The actions of morphine do not end with metabolism. It is now well established that the glucuronidation of morphine produces two major metabolites, morphine-3-glucuronide and morphine-6-glucuronide, each with remarkably different actions. The 6-glucuronide is more potent than morphine itself, and although the degree of this enhanced action is variable from study to study, it is at least 10 to 30 times more effective in tests of analgesia (Sullivan et al. 1989b). The reasons for this are not obvious; the affinity of morphine-6-glucuronide for the mu receptor is not appreciably greater than morphine itself, although it has more delta and less kappa affinity. However, as discussed earlier with morphine, it is likely that predominant mu activity underlies the analgesia with morphine-like opioids at therapeutic doses. Yet the other metabolite, M3G (morphine-3-glucuronide), has no affinity for the mu receptor and being unable to bind to the receptor has no opioid actions. Results from behaviour after administration of M3G have nevertheless led to the suggestion that M3G is a factor that contributes to reduced opioid sensitivity (Gong et al. 1992; Smith et al. 1990). The metabolite given by the intraventricular route caused marked behavioural agitation that interfered with the tests. In contrast to these studies, there is electrophysiological (where nonspecific effects are less likely to interfere with the results) and behavioural evidence that even with dose ratios of 100:1 (metabolite to morphine), M3G has absolutely no effect on the spinal antinociceptive effects of morphine (Hewett et al. 1993). It is highly unlikely that M3G is an important factor in cases of opioid poorly responsive pain, since (1) the spinal site of action of morphine is a major contributor to systemic analgesia, (2) M3G does not bind to opiate receptors, and (3) in renal insufficiency, where the metabolite will accumulate, opiate effects tend to be enhanced. Thus M3G should not, at present, be used as an excuse not to persevere with or increase the dose of morphine in pain states where opioid responsivity is poor. Patient-controlled analgesia has revealed that neuropathic pain patients can gain relief with morphine – although within this patient group pain control is not as good as in patients with nociceptive pains (Jadad et al. 1992). Dose escalation can also be effective (Portenoy et al. 1990).

5.2. Antiopioid peptides. Among the numerous factors influencing morphine analgesia, accumulating evidence indicates that the nonopioid peptide cholecystokinin (CCK) is an important physiological modulator of analgesic mechanisms. The exogenous spinal application of CCK and another peptide, FLFQPQRFamide, both nonopioid peptides found within intrinsic neurones in the spinal cord, will prevent mu- but not delta-mediated neuronal inhibitions (Baber et al. 1989; Dickenson 1994a), and both reduce intrathecal morphine analgesia in behavioural studies (Xu et al. 1993). Thus in situations where there is a release of these peptides one would expect a reduction in morphine effects without requiring any change in opioid receptor number. In fact, the ability of these peptides to interfere with analgesia is not restricted to the effects of opioids but also includes alpha-adrenoceptor agonist actions.

CCK has been shown to reduce the analgesic effects of morphine at a number of CNS sites and has also been implicated in the development of opioid tolerance (Baber et al. 1989). Spinal and supraspinal delta opioid mediated analgesias are not altered by CCK and, as a result, if there are physiological situations where CCK reduces mu opioid actions, future clinical delta agonists could be effective.

One of the key sites for these interactions is the spinal cord. Negative results with cDNA probes within dorsal root ganglia in the normal rat make it very unlikely that genuine CCK is found in nociceptive C-fibres in normal animals. Surprisingly, induction of the peptide in afferents occurs after pathological damage to the afferents. The consequences of this with regard to neuropathic pain (Xu et al. 1993) are discussed later in this section. Endogenous CCK in the dorsal horn under nonpathological conditions is thought to originate from both intrinsic neurones found in superficial laminae and descending fibres. The receptors are found both presynaptically (approximately 50%–60%) and postsynaptically to the primary afferent fibres, mirroring the mu opiate receptor distribution in the rat spinal cord. The postsynaptic CCK receptors are mainly of the CCK_B type in the rat spinal cord but of the A-type in the primate, whereas the presynaptic receptors are of the CCK_B type in all species. (Ghilardi et al. 1992). Thus CCK_B receptor antagonists will be critical in testing whether CCK influences morphine analgesia in humans (Stanfa et al. 1994).

The mechanism by which CCK attenuates the antinociceptive effect of morphine is not on opioid receptors but via activation of CCK receptors that may then interfere with opioid actions via postreceptor mechanisms. Key sites for these CCK-opioid interactions are likely to be the spinal terminals of C-fibres. Here, one possibility is that CCK mobilizes calcium from intracellular stores. This will counter the opioid suppression of the rise in internal calcium produced by depolarization, the basis for opioid reductions in transmitter release. Again, CCK only reverses the suppression of the induced rise in $[Ca^{2+}]i$ produced by mu but not delta opioid agonists (Wang et al. 1992).

At the same time, there is evidence for another mechanism in the CCK–opioid interaction that involves the endogenous enkephalins acting on the delta opioid receptor. Here, both CCK antagonists and the presence of inflammation (see below, this section) enhance morphine analgesia, an effect that is prevented by delta opioid antagonists. The theory is, then, that CCK inhibits the release of

enkephalins – removal of this control allows the increased levels of the enkephalins to cause a delta-receptor-mediated synergy with the mu receptor (Ossipov et al. 1995; Vanderah et al. 1994). Both theories are not mutually exclusive.

In keeping with CCK reducing opioid analgesia, the ability of morphine to inhibit spinal nociceptive processing is enhanced in the presence of selective CCK_B antagonists, demonstrating physiological antagonism of morphine antinociception by endogenous CCK under conditions of acute nociception. There is now evidence for this interaction in animal models more relevant to clinical situations such as inflammatory and neuropathic pain models (Ossipov et al. 1995; Stanfa et al. 1994).

As discussed earlier, in neuropathies morphine tends to have a reduced effectiveness, whereas after inflammation morphine has enhanced actions. In fact, a few hours after carrageenan inflammation there are mild increases in the potency of delta and kappa opioid effects but marked increases in the effects of morphine (Ossipov et al. 1995; Stanfa et al. 1992). One reason is that due to the novel peripheral action of opioids and systemic dosing, studies will be confounded by this additional site of action. However, spinal morphine is almost 20 times more potent than in normal rats after carrageenan inflammation. The mechanisms for this latter effect must be central and rapidly induced; the increased opioid actions occur within one hour of the inflammation, ruling out receptor upregulation (Stanfa et al. 1992). In this model, exogenous CCK still attenuates the antinociceptive effects of morphine but CCK receptor antagonism no longer produces an enhancement of the antinociceptive effect of morphine. The most likely basis for these results is a decreased availability of CCK within the spinal cord following carrageenan inflammation, either due to a decreased release of CCK or reduced content within the dorsal horn. This reduced functional activity of CCK in inflammation is therefore a major factor in the enhanced potency of spinal morphine seen in these animals (Stanfa & Dickenson 1993). In exactly the same model there is also an increased alpha-2 inhibitory tone in the spinal cord. Yet in this case, antagonist studies have shown that noradrenergic activity is not a factor in the altered opioid sensitivity, although it may well reduce inflammation-induced nociception (Stanfa & Dickenson 1994a).

Neuropathic models reveal that CCK plays an entirely opposite role. In nerve damage, increases in CCK systems have been shown to underlie observed reductions in spinal opioid sensitivity. It has been shown that an increase in spinal CCK (likely to be due to novel synthesis of the peptide in primary afferent fibres; see Xu et al. 1993) leads to a reduction in the potency of spinal morphine in a rat model of neuropathic pain following peripheral nerve injury. If the increased CCK is derived from induction of the peptide in the afferents, the interference would seem to be directed at the presynaptic mu receptor. The opioid responsiveness of this model was restored by CCK_B antagonism. Different pain states may then lead to changes in the levels and synthesis of CCK that can shift opioid sensitivity in either direction (Stanfa et al. 1994). An important point that arises from these studies is that the attenuation of opioid analgesia by CCK is not global but selective for mu but not for delta opioid receptor events. The prediction would be that delta opioids may have efficacy in pain states where morphine is poorly effective due to enhanced CCK levels.

CCK may then be an endogenous "brake" applied to the antinociceptive actions of morphine. In addition to the alterations in CCK induced by different pain states, behavioural studies have suggested that the release of endogenous CCK is even governed by the environment to which an animal is exposed. It is suggested that CCK released in "safe" situations prevents the acute antinociceptive effects of mu agonists and thus reduces the effects of morphine (Wiertelak et al. 1992). Findings such as these may provide a basis to understand how events such as stress and anxiety alter opioid efficacy. Whatever the case, these studies serve to indicate that the release of CCK is not fixed but varies, in both directions, from its normal state according to both external and internal events.

It appears, therefore, that CCK can act to control spinal morphine analgesia. Attenuation of this negative influence leads to augmented spinal opioid controls. This augmentation, along with the novel peripheral actions of opioids in inflammation and enhanced descending controls, could be an adaptation of intrinsic inhibitory systems to balance enhanced nociception during inflammatory states. In contrast to the natural physiological processes of inflammation, the pathological changes in neuropathic pain states result in counterproductive increases in CCK. CCK antagonists currently developed by the pharmaceutical industry may well enhance morphine analgesia in nonpathological pain states and restore morphine analgesia in humans with neuropathic pains. The predicted anxiolytic effects of these antagonists would be a bonus when used as analgesic adjuncts especially in states of chronic pain where anxiety commonly accompanies pain (Stanfa et al. 1994).

5.3. Dynorphin. In inflammatory states there is an increase in the mRNA in the spinal cord for dynorphin and, to a lesser extent, for enkephalin with all the cells increasing dynorphin synthesis having a preceding rise in c-fos, a protooncogene (Dubner & Ruda 1992). Dynorphin can mimic some of the increases in excitability seen after inflammation, such as the increased nociceptive responses of neurones, while inhibiting others (Knox & Dickenson 1987). Kappa opioids can functionally antagonize the mu receptor in the spinal cord, potentially contributing to a *decreased* morphine effectiveness even though opioids are more potent in inflammatory models (Stanfa et al. 1992). Furthermore, increases in spinal dynorphin levels also occur in neuropathic states where opioid actions tend to be reduced. These generally increased dynorphin levels in different physiological pain models, where opioid actions can be increased or decreased depending on the model, make kappa antagonism of morphine unlikely to be of physiological significance. What the functional consequences of the increases in dynorphin mean to the spinal cord is not known (Stanfa & Dickenson 1994b), and since dynorphin can elicit NMDA-receptor-mediated effects as well as opioid actions (see Dubner & Ruda 1992), the picture is complex. Figure 1 depicts some of these opioid interactions with spinal circuitry.

6. Central hypersensitivity

One of the most important new concepts related to pain is the idea that the ascending and propriospinal pain mes-

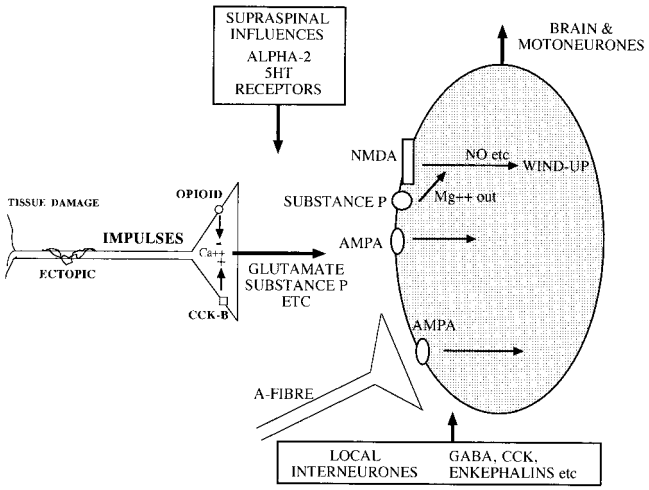


Figure 1. The diagram depicts the interactions between the different pharmacological systems described in section 5. Activity generated in peripheral sensory nerves releases a number of transmitters into the dorsal horn of the spinal cord. Their receptor actions and interactions, subject to control by local and supraspinal pathways, determine the output from dorsal horn projection neurones. Many of these systems are subject to plasticity.

sages from the dorsal horn are not the same under all circumstances. We are nearing explanations for the extreme aberrations of pain transmission – such as phantom limb pains, hyperalgesias, and allodynias – where the relations between the stimulus and the response are markedly perturbed. The basis for this lack of strict concordance between stimulus and response appears to be the generation of central hypersensitivity (Dickenson 1994b; McMahon et al. 1993; Price et al. 1994b; Woolf 1994; Woolf & Thompson 1991).

There are two key observations on this subject. First, high-frequency C-fibre stimuli result in a marked and prolonged increase in the flexion withdrawal reflex in rats recorded from motoneurones in spinal animals (Woolf 1983). Thus, noxious stimuli can enhance spinal excitatory events. Second, the repetition of a constant intensity C-fibre stimulus induces the phenomenon of "wind-up," whereby the responses of certain dorsal horn nociceptive neurones suddenly increase markedly (both in terms of magnitude and duration) despite the constant input into the spinal cord (Dickenson 1994b). Volatile general anaesthesia such as with halothane fails to prevent this type of activity, indicating that the treatment of postoperative pain states needs to take into account potential priming events occurring during the operations. The object of this account is to discuss the possible pharmacological substrates underlying these changes.

6.1. Substrates for central hypersensitivity

6.1.1. Peptides. Historically Substance P (SP) was the first transmitter to be related to the transmission of pain. SP release can be detected in the spinal cord following high- but not low-intensity peripheral stimulation. The use of antibody microprobes to detect the spatial release of SP has shown that it is essentially restricted to the zones where the C-fibres terminate (Duggan et al. 1988). In addition to substance P, the release of neurokinin A and CGRP following C-fibre activation has been demonstrated. However, when CGRP is present the subsequent release of SP is now extended to cover much of the dorsal horn. The interpretation of this finding is that the degradation of SP is reduced by CGRP binding to the peptidase that also cleaves SP, allowing SP to diffuse in the active form over considerable distances (Schaible et al. 1992). The concept of actions at a distance from the release site, so-called volume transmission, has attracted interest as a basis for nonsynaptic transmitter actions. Events such as these may have relevance to pain as the peptides may diffuse to distant receptors, avoiding both peptidases and spatially restricted inhibitory influences. The induction of inflammation is accompanied by enhanced release of these peptides centrally, which may then contribute to the central hypersensitivity (Dray et al. 1994; Sluka et al. 1992; Todd & Spike 1993).

The postsynaptic receptors for the neurokinin family of peptides, substance P, and neurokinins A and B are now well characterized (Otsuka & Yoshioka 1993). Cloning and sequencing have been achieved. Currently it is understood that there are three subclasses of tachykinin receptors: the neurokinin 1, – 2, and – 3 receptors. Early studies on the role of SP in neuronal events in nociception were bedeviled by poorly selective antagonists and nonspecific effects of the drugs. More recent studies have indicated a role of SP at the NK1 receptor in different types of more prolonged nociceptive transmission, including slow excitatory postsynaptic potentials induced by repetitive C-fibre stimulation and C-fibre-induced reflex facilitation. Similarly, the NK1 receptor antagonist RP67580 has only weak actions on acute responses but produces marked inhibitions of the formalin response of the dorsal horn neurones with equivalent effects on the two phases of the response. These recent studies would indicate that the ability of NK1 receptor antagonists to reduce the activation of dorsal horn neurones depends on the type of stimulation used (Otsuka & Yoshioka 1993). A consensus is that the conditions for the release of substance P from the fine afferents include a sufficiently long stimulus at an intensity sufficient to activate C-fibres (Urban et al. 1994). The acute responses of the neurones must therefore include some other transmitter; the evidence implicates glutamate and aspartate. We lack antagonists for the other peptides, but some, such as galanin and neuropeptide Y, are induced in afferents after nerve damage (Todd & Spike 1993; Urban et al. 1994).

6.1.2. Excitatory amino acids. A large proportion of peripheral sensory fibres including both small and large fibres contain glutamate and aspartate (Battaglia & Rustioni 1988). In the case of the C-fibres, the coexistence of glutamate with peptides (Battaglia & Rustioni 1988) makes it highly likely that a noxious stimulus releases both peptides and excitatory amino acids from the afferent nociceptive fibres. Thus in clinical pain states postsynaptic activation of both neurokinin and other peptide receptors together with the receptors for the excitatory amino acids on nociceptive neurones will occur. The development of selective agents for the receptors – the N-methyl-d-aspartate (NMDA), the metabotropic, and the alpha-amino-3-hydroxy-5-methyl-isoxazole (AMPA) receptors – has enabled their roles in spinal processing to be studied.

The metabotropic receptor still has an ill-defined role in pain states but may well contribute by acting to enhance

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

NMDA and AMPA receptor function via intracellular actions. Use of AMPA receptor antagonists indicate that acute noxious but also innocuous stimuli seem to be transmitted via AMPA receptor activation (Dougherty et al. 1992; Neugebauer et al. 1993). The widespread roles of AMPA receptors in CNS function and the lack of nociceptive selectivity mean that the receptor as a therapeutic target looks doubtful. In contrast, the NMDA receptor has become an increasingly important target site as evidence accumulates for a role of the receptor in the enhancement of spinal processing of painful messages (see Dickenson 1990; 1994b; Price et al. 1994a; 1994b) as well as a target site in many long-term events in the brain (Collingridge & Singer 1990; Daw et al. 1993). In the spinal cord, the NMDA receptor may play a similar role, especially in more prolonged pain states involving hypersensitivity where functional alterations in central transmission processes may occur.

The complexity of the NMDA receptor-channel is striking. In order to operate it, certain specific conditions need to be met: The release and binding of the coagonists for the receptor, glycine and glutamate, are needed together with a non-NMDA-induced depolarisation to remove the resting magnesium block of the channel (Dickenson 1994b). C-fibre-induced release of excitatory peptides, either in a restricted spatial zone or via volume transmission, may provide the required depolarisation to remove the block, since neurokinin receptor antagonists can reduce NMDA-mediated responses in the spinal cord (Urban et al. 1994). For these reasons, the NMDA receptor-channel complex is not a participant in "normal" synaptic transmission. Yet when the correct conditions are achieved, the complex will suddenly become activated and add a powerful depolarising or excitatory drive to transmission of pain in the spinal cord, which then appears to lead to enhanced synaptic transmission or hypersensitivity (Dickenson 1990; 1994a; 1994b; Dubner & Ruda 1992; McMahon et al. 1993; Neugebauer et al. 1993; Price et al 1994a; 1994b; Woolf & Thompson 1991). Increased release of afferent peptides in inflammation, for example, could facilitate NMDA transmission by more effective removal of the magnesium block of the receptor channel or by increasing the release of the excitatory amino acids themselves (Kangra & Randic 1990). It is now well established that wind-up and the reflex hypersensitivity are NMDA receptor mediated. Further experiments with formalin indicate that when inflammation is present, "pathological pain" can be distinguished from the acute phase response where there is no damage on the basis of the sensitivity of only the former to NMDA antagonism. Both the induction and the subsequent maintenance of these responses are dependent on NMDA processes (Haley et al. 1990; Neugebauer et al. 1993; Price et al. 1994b).

However, NMDA receptor activation can also influence inhibitory interneurones in the spinal cord; evidence for this appears from carrageenan inflammation where excessive NMDA receptor activation subsequently induces inhibitory influences (Stanfa et al. 1992). Excessive NMDA activation in the CNS is one mechanism behind excitotoxicity and, as a result, elevated NMDA activation may trigger inhibitory systems as an auto-limiting device to prevent over excitation and even cell death. Possibly, the loss of inhibitions to counter NMDA excitatory mechanisms (Woolf & Doubell 1994) leads to some of the problems of neuropathic pain. In this regard, NMDA-mediated allodynia can be induced by a blockade of spinal inhibitory tone in normal animals (Yaksh 1989). Furthermore, failure of inhibitions could underlie the transition from acute to chronic pain.

Other approaches have revealed roles of the NMDA receptor in spinal pain processes including ischaemia and neuropathic pain states, where NMDA antagonists have beneficial effects weeks after induction of the injury against the hyperalgesia and spontaneous pain. Thus, there is evidence for an involvement of the NMDA receptor in inflammatory pain, neuropathic pain, allodynia, and ischaemic pain. Not only can wind-up be demonstrated in elegant psychophysical studies in humans (Price et al. 1994a) but, crucially, recent evidence also has shown an NMDA dependency of allodynias and wind-up pains in controlled clinical studies (Eide et al. 1994).

6.1.3. Nitric oxide and arachidonic acid. Central plasticity can also involve a gas. Nitric oxide (NO), a diffusible gas, is produced in response to NMDA receptor activation and thus may mediate some or all of the consequences of NMDA receptor activation in nociception (Meller & Gebhart 1993). Blockers of nitric oxide synthase (NOS) are effective against inflammatory and neuropathic nociception in animals by spinal actions. There are hints that NO may feed back and enhance the release of the afferent transmitters and as a result set up a positive feedback loop (Sorkin 1993). In addition, an induction of NO in the afferents has been reported after nerve damage. Spinal production of arachidonic acid in response to C-fibre stimulation and NMDA receptor activation may achieve the same end; prevention of this could underlie some of the central analgesic effects of nonsteroidal anti-inflammatory drugs (NSAIDs) (Malmberg & Yaksh 1992).

The evidence for spinal actions of NSAIDs continues to grow. This growth is based not only on the demonstrations of C-fibre-evoked and, more importantly, NMDA-evoked release of prostanoids, but also on the spinal action of NSAIDs that can be shown by intrathecal administration of these agents. In addition, it has been shown that the hyperalgesia produced by NMDA and substance P is reduced by spinal NSAIDs (Malmberg & Yaksh 1992). As there is evidence for prostanoid-, NO-, and NMDA-mediated release of glutamate and substance P, it would appear that the production of novel mediators by NMDA receptor activation underlies retrograde messenger control of transmitter release (Sorkin 1993).

There will be problems with NOS blockers in therapy, since NO is an endothelium-derived relaxing factor and systemic administration may induce analgesia but will be accompanied by severe hypertension. However, it has been demonstrated that neuronal NOS differs from that in the endothelium and thus it may be possible to separate these effects.

7. Consequences of central hypersensitivity

The presently characterized systems operate in the spinal cord, which appears to mediate central hypersensitivity, a state where amplification and prolongation of the afferent barrage occurs. The established roles of these different transmitter systems may offer novel targets for therapy. This may be important since wind-up, the hypersensitized

reflex, and several measures of NMDA-dependent activity in neuropathy and allodynia models can be poorly sensitive to opioids (Dickenson 1994b). As discussed earlier, the reasons may be nociception through channels not controlled by opioids (large fibre-induced allodynias), high levels of excitability (NMDA-mediated amplification), or pathological loss of the opioid receptors.

7.1. Treating opioid poorly responsive pain. The first approach is based on animal studies showing that loss or dysfunction of presynaptic opioid receptors can be overcome by increasing the dose of opioid (Lombard & Besson 1989; Xu & Wiesenfeld-Hallin 1991). In addition, in pains where the NMDA receptor is operating and there is reduced opioid sensitivity (as in some of the models for inflammatory and neuropathic pain), this too can be overcome by dose escalation (Chapman & Dickenson 1992; Yamamoto & Yaksh 1992). The simple augmentation of the dose of morphine should be first tried, although side effects may confound this tactic. Another approach may be the use of high-efficacy opioids such as alfentanyl or sufentanil, but data is lacking on this point. If opioids cannot produce the desired effects, different pharmacological approaches are possible: In the case of the NMDA receptor there are many experimental drugs that effectively block the receptor, the channel, or associated sites. Some of these are in development as potential drugs, but there clearly is a need for agents to be tested now. In fact, ketamine blocks the channel associated with the NMDA receptor and has current use in the relief of pain. Dextrophan and dextromethorphan are also antagonists at this site and are currently used in humans for their antitussive effects. Both have been shown to reduce wind-up itself (Dickenson et al. 1992) as well as to be effective in the Bennett model of neuropathic pain after spinal application (Mao et al. 1993; Tal & Bennett 1993) and in humans (Price et al. 1994a). Recently, the anti-Parkinson drug, memantine, has been shown to be an effective NMDA antagonist. All could be used to test the clinical effectiveness of NMDA blockade in opioid poorly responsive pains. However, the NMDA antagonists would be effective only in reducing hyperalgesia, not in abolishing the pain (Dickenson 1994b). These agents may turn out to be especially useful in the allodynias, which are sensitive to NMDA receptor antagonists but not to opioids. In fact, not only is there psychophysical evidence for wind-up pain in humans being mediated by NMDA receptors, based on studies with dextromethorphan (Price et al. 1994a), but there is also evidence from clinical trials showing that ketamine can reduce allodynias, hyperalgesias, and cause pain relief in circumstances where opioids had poor or restricted efficacy (Eide et al. 1994).

One practical application of the poor opioid responsiveness of NMDA-mediated pains is that the coadministration of morphine with low doses of an NMDA antagonist should be beneficial in these pain states. This is indeed the case; furthermore the combination has been shown to synergize in one study (Chapman & Dickenson 1992) and be additive in another model (Yamamoto & Yaksh 1992). However, both studies have shown that the additional NMDA antagonism restores the opioid sensitivity of the responses. In addition, spinal local anaesthetics synergize with spinal morphine (Akerman et al. 1988), partly due to the ability of the former to reduce NMDA-mediated activity (Fraser et al. 1992). The spinal release of prostaglandins affords another target, and centrally acting NSAIDs have been shown to reduce persistent inflammatory nociception and the behavioural hyperalgesia produced by spinal substance P and NMDA. Predictably, NSAIDs will synergize with opioids (Yaksh & Malmberg 1994).

It has been suggested that once these central hypersensitivity states have been induced, they remain active in the absence of peripheral inputs (Coderre et al. 1990). There is counterevidence from both animal studies and in human pain states where there is clear evidence for central changes that are entirely dependent on peripheral inputs for maintenance (Dickenson & Sullivan 1987; Gracely et al. 1993). Thus it would seem that the central pain hypersensitivity generators are continually triggered by afferent activity. Consequently, there is a place for peripheral local anaesthetics although the symptoms may well reappear once the block wears off. In addition, there is evidence that systemic local anaesthetics may have selective effects on ectopic foci in a damaged peripheral nerve at doses that do not alter conduction in the nerve (Devor et al. 1992). Finally, as stated earlier, spinal sites of action of local anaesthetics include a reduction in wind-up (Fraser et al. 1992).

In all of these studies on opioid poorly responsive pain, the emphasis has been on mu opioids, especially morphine; yet mu and delta opioids and alpha-2 agonists all have similar effects on wind-up, reducing the initial responses of the cells but with wind-up breaking through the inhibitions as the stimulation continues and restoring the cell responses (Dickenson 1991; 1994a). It is unlikely, therefore, that these three systems would have differential effects on NMDA-receptor-mediated events, making drugs such as clonidine unlikely to be alternatives to morphine, at least with regard to spinal events where the NMDA receptor is implicated. In addition, the negative effects of CCK on analgesia are not only directed against mu opioids but also alpha-2 adrenoceptors. However, it is possible that in cases where presynaptic opioid receptors are reduced, such as in cases of peripheral nerve pathology, alpha-2 receptors may persist at postsynaptic sites and thus provide a therapeutic target. Relatedly, there may well be sympathetic blocking effects of clonidine via systemic routes as well as the spinal route, which will be of importance in sympathetically maintained pains. Yet again, alpha-2 adrenoceptor agonists produce powerful potentiations of opioid analgesia (Dickenson & Sullivan 1993; Yaksh & Malmberg 1994).

7.2. Preemptive analgesia. The idea of preemptive analgesia has arisen due to the potential for induction of hypersensitivity, genes, and negative influences on opioid controls in addition to the well-established detrimental effects of the stress and hormonal responses to pain. Thus, treating pain before it arises rather than waiting for it to develop appears to have a rational basis (Woolf 1994). Animal studies lend support to this: several measures of central hypersensitivity are less sensitive to opioids given as a post-treatment as compared to preemptive administration (Chapman et al. 1994). Restricting comments to the use of opioids, there are several reasons why the clinical studies on preemptive analgesia have been either negative or have showed relatively weak benefits (Dahl 1994; McQuay 1994). Most have used postoperative pain measures and invariably operative procedures will induce inflammation.

The key issues in comparing pain relief pre- and post-

operatively are the exact timing of administration and that baseline opioid sensitivity remains the same (McQuay 1994). With the former, it may be that a pretreatment is the same as an early posttreatment, as they both will preempt late-developing central hypersensitivity. Studies with varied timing of opioids on the formalin response support this idea (Chapman et al. 1994). With opioid sensitivity, it is well established in animal models of carrageenan inflammation and arthritis that not only is novel peripheral opioid receptor mediated analgesia rapidly revealed but also spinal opioid sensitivity is enhanced, just as rapidly (Stanfa et al. 1992; Stein 1994). Thus post-treatment with opioids will impinge upon enhanced opioid systems and comparisons with the effectiveness of pretreatments will be biased. In neuropathic pains, where in general opioid sensitivity is reduced, inhibitions are lost, and central and sympathetic processing is aberrant, the impact of pretreatment is far more obvious.

7.3. The roles of inhibitions. When NMDA-mediated central events leading to hypersensitivity are active in the shorter-term models (formalin and the hypersensitive reflex), there is a reduced sensitivity to opioids; whereas once several hours have elapsed (carrageenan and arthritic inflammation), opioid sensitivity is now found to be increased. In the example of the former models, dose-escalation can overcome the reduced opioid sensitivity (Chapman et al. 1994). The reasons for these differences could reside in the unchecked NMDA receptor mediation of activity in the shorter-term models, for example, formalin, prior to the induction of the slower-developing inhibitory changes, overcomes opioid inhibitions. The compensatory increases in spinal opioid sensitivity (via altered CCK) and other inhibitions means that the longer term acute pains respond well to opioids. This is partly due to the enhanced opioid effectiveness per se but also because the increased nonopioid (GABA and alpha-2 adrenoceptor mediated) inhibitory events will reduce the NMDA-driven level of excitability (Castro-Lopes et al. 1994; Stanfa et al. 1994). A good example of the potential controlling influence of inhibitions is seen in the formalin response. NMDA receptor activation in the spinal cord amplifies a low level of C-fibre input (Heapy et al. 1987) to generate the characteristic response. The resultant behavioural and neurophysiological responses to the peripheral injection of formalin last for one hour, but the C-fibre inputs continue for over two hours and the peripheral inflammation for even longer periods (Porro & Cavazzuti 1993). The central responses must surely be curtailed by inhibitory controls. If the plastic changes are leading to compensatory increased central inhibitions such as these, then preemptive approaches may prevent both this beneficial plasticity as well as the target of central hypersensitivity mechanisms. If this is the case, as the preemptive agent wears off, the pain may return in the absence of compensatory inhibitions and thus with greater intensity.

Whereas the roles of inhibitions have generally received less attention than the excitatory systems, in an animal model of allodynia, NMDA antagonists unlike morphine are effective against tactile-evoked nociception (Yaksh 1989). In this model the NMDA-mediated allodynia is induced by a blockade of spinal GABA or glycine inhibitory tone in normal animals. This has bearing on the previously made point that inhibitions are important in controlling and limiting the extent of NMDA receptor participation in nociceptive processing in the spinal cord. Increasing GABA function by the administration of benzodiazepines reduces NMDA-mediated hyperalgesia after ischaemia (Cartmell & Mitchell 1993); another intriguing example of this is the enhanced NMDA-mediated nociception seen when spinal glycine inhibitions are blocked in a neuropathic model in the rat (Seltzer et al. 1991). In contrast, GABA upregulation may be an intrinsic compensatory mechanism in longer-term inflammation (Castro-Lopes et al. 1994).

Failure of these and other inhibitory controls conceivably may lead to chronicity of pain. In fact there is strong evidence that the hyperalgesia seen in neuropathic models is as much a consequence of loss of inhibitions and reorganization as excess excitations (Dray et al. 1994; Woolf & Doubell 1994), which could be due to a destructive loss of inhibitory interneurones – itself exacerbated by pharmacological block of these inhibitory systems (Sugimoto et al 1990).

8. Developmental aspects

Examination of events in the neonatal spinal cord are of great relevance to the role of modulatory systems in controlling excitation, since the maturation of the inhibitory systems is slow. Paediatric pain control has to take into account the findings that the development of the nervous system is accompanied by marked changes in many transmitters and receptors over time. A number of studies on the anatomical and functional development of the excitatory and inhibitory pharmacology of the rat and human spinal cords suggest that the rodent provides a good model for investigating clinical questions (Fitzgerald 1991).

Studies using the neonatal spinal cord have shown that all the excitatory and inhibitory receptors covered in this account are functional at day one in the rat; many of these neuropharmacological systems are present even before birth. Full maturation of the endogenous transmitter pathways and connections is much slower (Marti et al. 1987), particularly the local inhibitory and long descending controls. The levels of excitatory transmitters, such as substance P and excitatory amino acids including glutamate, tend to increase during development. At the same time NK and NMDA receptor location in young animals is far more exuberant than in adults, and their numbers decline with time as the receptors shrink back to assume the discrete adult form (Charlton & Helke 1986). Dendritic development, interneurones (Bicknell & Beal 1984), and descending inhibitory controls (Fitzgerald & Koltzenburg 1986) are slowly developing so that modulation of excitability is delayed. The conclusion from these studies is that the transmission of pain in the spinal cord of the young is likely to be exaggerated compared to the adult as a result of excess early development of excitation and delayed maturity of the intrinsic controlling inhibitory systems.

This difference between the young animal and the adult can be seen once more by using the formalin response, but now in neonatal rats (Guy & Abbot 1992). In the first week of life, the response to peripheral algogen is dramatic and disruptive to normal behaviour. As development proceeds, response declines – until at about three weeks, when inhibitory systems have matured and response resembles the more discrete adult form.

Opioids and their three receptors also change with time: the relative arrival of the mu, delta, and kappa receptors differs as does the development of the endogenous opioids. Although the opioid receptor affinities resemble the adult at very early stages, the numbers of receptors decline over time (Attali et al. 1990; Sales et al. 1989). The early ontogeny of the opioid receptors means that there is a substrate present for the production of analgesia by exogenous opioids, whereas the controlling influences of endogenous opioid neuronal systems may only appear later in development.

Thus the systems in the adult that generate spinal hypersensitivity may be more effective in the young nervous system, due to less inhibitory influences as a result of immaturity of these systems. Exogenous activation of opioid receptors allows pain in the young to be controlled by adequate analgesia either with opioids alone or with some of the combinations discussed previously. Because activity-dependent plasticity is important in determining maturation of connectivity in the developing nervous system, it has to be considered that uncontrolled pain in the young may have long-term consequences for the neurobiology of pain. In this respect, fears regarding the consequences of opioid use in young children, which are probably unfounded as suggested by the lack of dramatic effect from early opioid exposure on later function (Bardo et al. 1982), pale into insignificance when compared to the possible permanent alterations in sensory processing that uncontrolled pain in the young could produce.

9. Conclusions

Many questions remain to be answered about events in different pain states that can alter, in either direction, the analgesic effects of opioids. As has been discussed, this recently accumulating knowledge of plasticity provides a rational basis for combination therapy (Dickenson 1994b; Dickenson & Sullivan 1993; Yaksh & Malmberg 1994). Examples given here fall into two categories: first, the combination of opioids with other inhibitory agents such as alpha-2 adrenoceptor agonists and, second, dual therapy where the nonopioid acts to reduce excitability or to control interfering systems (CCK and NMDA receptor antagonists, NSAIDs, local anaesthetics). Dual therapy such as this has been shown to result in restoration, additivity, or potentiation of opioid analgesia (Dickenson & Sullivan 1993; Yaksh & Malmberg 1994). It is also becoming increasingly clear that particular pain states have different plasticities; despite some similarities (e.g., NMDA mediation of hypersensitivity), inflammatory and neuropathic pains are not only very different from each other but cannot be viewed as uniform syndromes. For example, in neuropathic pains the human and animal studies on opioid sensitivity do not reach a consensus; animal studies suggest both good (Attal et al. 1991; Yamamoto & Yaksh 1992) and poor opioid sensitivity – and the clinical studies are not totally in agreement on this point either (Arner & Meyerson 1988; Jadad et al. 1992; Portenoy et al. 1990). Particular monotherapies and combination therapies will likely be appropriate for different states within a pain syndrome (McQuay & Dickenson 1990). Other key points that need investigation are the following: Does a repeated noxious insult alter the plasticity? Knowledge is building up for the neonate, but what happens in the ageing nervous system? The answers to these and other questions relating to plasticity have important consequences for the clinical treatment of pain.

ACKNOWLEDGMENTS

The funding of research for the author's laboratory is from the Medical Research Council and The Wellcome Trust. Major contributions to this account, both experimental and conceptual, have come from Louise Stanfa, Vicky Chapman, Ann Sullivan, Alison Reeve, Henry McQuay, Eija Kalso, Maria Fitzgerald, and Jane Haley. Their enthusiasm, dedication, and assistance is gratefully acknowledged.

BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 404–419
Printed in the United States of America

Peripheral and central hyperexcitability: Differential signs and symptoms in persistent pain

Terence J. Coderre[a]
[a]*Pain Mechanisms Laboratory, Clinical Research Institute of Montreal; Centre de recherches en sciences neurologiques et département de médecine, Université de Montréal; Department of Psychology, McGill University, Montreal, Quebec, Canada, H2W IR7*
Electronic mail: *coderrt@ircm.umontreal.ca*

Joel Katz[b]
Department of Psychology and Acute Pain Research Unit, The Toronto Hospital; Departments of Behavioral Science and Anaesthesia, University of Toronto, Toronto, Ontario, Canada M5S 2C4
Electronic mail: *j.katz@utoronto.ca*

Abstract: This target article examines the clinical and experimental evidence for a role of peripheral and central hyperexcitability in persistent pain in four key areas: cutaneous hyperalgesia, referred pain, neuropathic pain, and postoperative pain. Each suggests that persistent pain depends not only on central sensitization, but also on inputs from damaged peripheral tissue. It is instructive to think of central sensitization as comprised of both an initial central sensitization and an ongoing central sensitization driven by inputs from peripheral sources. Each of these factors, initial sensitization, ongoing central sensitization, and inputs from peripheral sources, contributes to the net activity in dorsal horn neurons and thus influences the expression of persistent pain or hyperalgesia. Since each factor, peripheral inputs and central sensitization (initial or ongoing), can contribute to both the initiation and maintenance of persistent pain, therapies should target both peripheral and central sources of pathology.

Keywords: hyperalgesia; neurogenic inflammation; neuropathic pain; nociception; phantom limb pain; plasticity; postoperative pain; pre-emptive analgesia; referred pain; sensitization

1. Introduction

Pain is a normal reaction of the somatosensory system to noxious stimulation which alerts the individual to actual or potential tissue damage. It serves a protective function, informing us of injury or disease, and usually remits when healing is complete or the condition is cured. However, in some cases, peripheral tissue damage or nerve injury leads to a pathological state characterized by one or more of the following: pain in the absence of a noxious stimulus (spontaneous pain), increased duration of response to brief stimulation (ongoing pain or hyperpathia), reduced pain threshold (allodynia), increased responsiveness to suprathreshold stimulation (hyperalgesia), and spread of pain and hyperalgesia to uninjured tissue (referred pain and secondary hyperalgesia). For more than a century there has been a heated debate over the role of peripheral and central neural mechanisms in the initiation and maintenance of these pathological conditions. Although the debate has a long history (see Bonica 1992 for a review), most of the empirical evidence in support of one side or the other is relatively new (also see Ruda & Dubner 1992; Willis 1994; Woolf 1992). Perhaps even newer is a growing realization that peripheral and central neural mechanisms are not mutually exclusive, and interact extensively to reinforce the pathological changes that contribute to chronic pain. In this paper we examine both clinical and experimental evidence for peripheral and central neural contributions to pathological pain. In this context, we will review the current state of knowledge concerning the proposed neural mechanisms contributing to the initiation and maintenance of four types of painful conditions, including: (1) hyperalgesia after cutaneous injury; (2) referred pain and hyperalgesia after deep tissue injury; (3) neuropathic pain; and (4) postoperative pain.

2. Cutaneous hyperalgesia

After a cutaneous injury both the injured skin and the uninjured skin adjacent to the injury become more sensitive to specific types of sensory stimulation. In particular, the injured skin becomes more sensitive to non-noxious heating or stroking (thermal and mechanical allodynia), as well as to noxious heating or punctate stimulation (thermal and mechanical hyperalgesia). In contrast to the injured skin, the

0140-525X/97 $9.00+.10

adjacent uninjured skin appears to become more sensitive to mechanical, but not thermal stimuli, with hyperalgesia to punctate stimulation spreading further and lasting much longer (13–24 hrs) than allodynia to stroking (1–2 hrs) (Meyer et al. 1994). For descriptive purposes, many investigators have adopted the terminology first proposed by Hardy et al. (1950) in which hyperalgesia at the site of injury is termed *primary hyperalgesia,* and hyperalgesia in the adjacent uninjured skin is termed *secondary hyperalgesia*. Given a general prevalence of mechanical hyperalgesia over thermal hyperalgesia in this review, subsequent usage of the term hyperalgesia refers to *mechanical* hyperalgesia, unless otherwise specified.

2.1. Nociceptor sensitization. Most would agree that the thermal allodynia and hyperalgesia present in the injured region is due to peripheral sensitization of nociceptors. Following tissue injury there is an increase in the excitability of primary afferent nociceptors. Nociceptor sensitization is reflected by one or more of the following: decreased threshold, increased impulse frequency to the same stimulus, decreased latency of the first impulse, after discharge following extended or intense stimulation, and the appearance of spontaneous firing (Beitel & Dubner 1976). Repeated heat stimulation produces nociceptor sensitization which develops within 1 minute and last for hours (Perl 1976). Following heating of the skin, sensitization to further heat stimuli has been demonstrated in C-fiber polymodal nociceptors in the rat (Lynn & Carpenter 1982), rabbit (Perl et al. 1974), cat (Bessou & Perl 1969), monkey (Beitel & Dubner 1976) and man (Torebjörk et al. 1984). Sensitization after heat injury has also been found in the heat responses of A-delta fiber, high-threshold mechanoreceptor units in the rabbit and cat (Fitzgerald & Lynn 1977) and monkeys (Meyer & Campbell 1981), as well as in the paradoxical responses of cold receptors in monkeys (Dubner et al. 1975).

Whether nociceptor sensitization can account for primary mechanical hyperalgesia is less clear. Thus, some investigators have found that both polymodal nociceptors (Bessou & Perl 1969) and high threshold mechanoreceptors (Fitzgerald & Lynn 1977) become sensitized to mechanical stimuli following heat injury. However, others have found that, within their normal receptive fields, the thresholds of C- and A-fiber mechanoheat sensitive nociceptors are not altered by heat or mechanical injury (J. N. Campbell et al. 1988a; Thalhammer & LaMotte 1982). It has been alternatively suggested that the expansion of receptive fields of nociceptors into an adjacent area of injury may account for primary hyperalgesia to mechanical stimulation, as this occurs after heat (Thalhammer & LaMotte 1982) and mechanical (Reeh et al. 1987) injuries.

Recent studies have looked for a correlation between nociceptor sensitization and reports of primary hyperalgesia following a cutaneous injury. Initially this was performed by comparing magnitude estimations of hyperalgesia in man with neurophysiological recordings in nerve fibers of monkeys (LaMotte et al. 1982; Meyer & Campbell 1981). More recent studies have examined the correlation between human sensory judgements and evoked neural responses in the same subjects using percutaneous recording techniques (Ochoa & Torebjörk 1989; Torebjörk et al. 1984). The results of these studies have been controversial. While Meyer and Campbell (1981) reported that primary hyperalgesia is associated with a sensitization of A-fibers and a desensitization of C-fibers, LaMotte et al. (1982) and Torebjörk et al. (1984) suggested that primary hyperalgesia is related to a sensitization of C-fibers and not A-fibers. It has been proposed that this discrepancy depends on either the type of skin that is injured or the intensity of the stimulus producing the injury.

2.2. Peripheral neurogenic mechanisms. The spread of hyperalgesia to uninjured tissue is probably not due to nociceptor sensitization, but may involve either a neurogenic axon reflex or a sensitization of central neurons. Lewis (1936; 1937) was the first to perform an extensive examination of the spread of cutaneous hyperalgesia into uninjured tissue. According to Lewis, the spread of hyperalgesia to uninjured tissue was due to a peripheral neural mechanism which involved antidromic activity in peripheral nerves leading to the release of a substance which contributed to the development of both hyperalgesia and vasodilatation or flare responses in the skin. In support of his hypothesis, Lewis presented evidence that cutaneous hyperalgesia in response to skin crush does not develop in anesthetized skin and does not spread across an anesthetized strip of skin, until after the anesthesia wears off. Lewis also showed that cutaneous hyperalgesia which occurred in response to electrical stimulation of nerves through the skin (faradic stimulation), was prevented by a local anesthetic nerve block distal to the electrical stimulus. Conversely, when the nerve block was proximal to the faradic stimulation, hyperalgesia developed normally, but only after the anesthesia wore off.

Early studies provided supported for Lewis's mechanism of spreading hyperalgesia. Perl et al. (1974) showed that an extensive skin injury produced a sensitization of C-fiber polymodal units whose receptive fields were removed from the injured region. Fitzgerald (1979) recorded activity in C-fiber nociceptors in the skin near an injury and found that nociceptors in the uninjured tissue were more sensitive to heat following an injury than when there was no injury. The spread of the effect of the injury was induced by nerve impulses, since a local injection of lignocaine anesthetic blocked the spread of sensitization. Fitzgerald also found that there was a spread of nociceptor sensitization following antidromic stimulation of the rabbit sural nerve at C-fiber strength. The effect was independent of the CNS (central nervous system) since it occurred even when the nerve was cut central to the stimulation point. Chahl and Ladd (1976) demonstrated that antidromic stimulation of the rat saphenous nerve produced inflammation and an increased excitability in sensory nerve fibers when the stimulation was of C-fiber, but not A-fiber strength.

More recent studies provide evidence against Lewis's theory of spreading peripheral sensitization. Thus, antidromic stimulation of nociceptive fibers in either the monkey (Meyer et al. 1988) or the rat (Reeh et al. 1986) was not found to produce nociceptor sensitization. In addition, Thalhammer and LaMotte (1982) found that a heat injury in one half of a cutaneous nociceptor's receptive field did not produce heat sensitization in the other half, despite the fact that hyperalgesia spread into this area. Indeed, mechanical and chemical injuries produce extensive, spreading hyperalgesia (LaMotte et al. 1992) without producing the same degree of spreading sensitization of primary afferent nociceptors in monkeys (J. N. Campbell et al.

1988a; Baumann et al. 1991) or humans (LaMotte et al. 1992). Typically, nociceptor sensitization associated with injury is restricted to about 5–10 mm of the site of injury (J. N. Campbell et al. 1984; Fitzgerald 1979), while cutaneous hyperalgesia spreads as far as 10–20 cm beyond the site of injury (Hardy et al. 1950; LaMotte et al. 1991; 1992; Lewis 1936; 1937). Furthermore, the zone of secondary hyperalgesia is typically found to be larger than the flare produced by tissue injury (Koltzenburg et al. 1992; LaMotte et al. 1991; Raja et al. 1984). In fact, as noted by LaMotte et al. (1991), a flare can be produced (by histamine injection) without even inducing secondary hyperalgesia, and secondary hyperalgesia can occur in the absence of a flare response. Finally, secondary hyperalgesia after cutaneous injury typically does not spread beyond the body's midline, whereas flare responses do (LaMotte et al. 1991).

2.3. Central sensitization. In contrast to Lewis, Hardy et al. (1950) proposed that while primary hyperalgesia was mediated by peripheral mechanisms, secondary hyperalgesia was produced by central sensitization. Hardy et al. (1950) confirmed Lewis's finding that cutaneous hyperalgesia (in this case in response to burn injury) did not develop in anesthetized skin until after the anesthesia wore off. However, in contrast to Lewis, Hardy et al. reported that hyperalgesia after faradic stimulation was unaffected by a distal nerve block, but was significantly delayed by a proximal nerve block.

More recent evidence supports the view that hyperalgesia depends, in part, on central sensitization (Guilbaud et al. 1992b; LaMotte 1992; Torebjörk 1992). Hyperalgesia to punctate mechanical stimuli, which develops after intradermal injection of capsaicin, is maintained even after anesthetizing the region where capsaicin was injected (LaMotte et al. 1991). However, if the skin region is anesthetized prior to capsaicin injection, cutaneous hyperalgesia does not develop. Furthermore, hyperalgesic responses to capsaicin can be prevented if the area of skin where the injection is made is rendered anesthetic by a proximal anesthetic block of the peripheral nerve which innervates it. Thus, for hyperalgesia to develop it is critical that initial inputs from the injury reach the CNS. However, once hyperalgesia is established, it does not need to be maintained by inputs from the injured peripheral tissue. In support of this, Torebjörk et al. (1992) have shown that pain thresholds to intraneural electrical stimulation of afferent fibers are dramatically reduced following intradermal capsaicin injection in the skin from which the stimulated nerve emanates: neural stimulation which was felt as tactile before administration of capsaicin, was painful after capsaicin. Importantly, this reduced pain threshold is evident even when the sensory projected field of the afferent nerve is anesthetized after the capsaicin injection. Again, a state of central sensitization is indicated since once they have established their effects, inputs from the injured region are not required to maintain the lowered threshold.

Recent experimental data from animal studies also provide support for Hardy et al.'s central mechanism of secondary hyperalgesia since peripheral injuries typically produce a sensitization of neurons in central nervous system (CNS). Thus, dorsal horn neurons fire with increasing frequency in response to repeated application of a noxious heat stimulus (Kenshalo et al. 1979; Perl 1976). Sensitization of dorsal horn neurons occurs after various types of tissue damage including thermal injury (Kenshalo et al. 1982; Price et al. 1978), chemical injury (Dougherty & Willis 1992; Simone et al. 1991), and polyarthritis (Calvino et al. 1987; Menétrey & Besson 1982), or after stimulation of C-fiber afferents (Chung et al. 1979). Tissue injury/inflammation or electrical nerve stimulation also produces sensitization in spinal motoneurons (Woolf 1983), thalamus (Guilbaud et al. 1986), and somatosensory cortex (Lamour et al. 1983). Repeated C-fiber afferent stimulation also produces a sequential increase in dorsal horn activity resulting in a prolonged discharge of the cell (wind-up), which lasts from seconds to minutes post-stimulation (Mendell 1966; Schouenbourg & Dickenson 1985).

In addition to the sensitization and wind-up of dorsal horn cells, noxious stimulation associated with tissue injury also produces an expansion of the receptive fields of dorsal horn neurons. Neurons in the dorsal horn of the spinal cord with receptive fields adjacent to a cutaneous heat injury expand their receptive fields to incorporate the site of injury (McMahon & Wall 1984). Similar receptive field expansions have been observed in spinal cord following mechanical (Cervero et al. 1988), chemical (Hoheisel & Mense 1989; Woolf & King 1990), and inflammatory (Hylden et al. 1989) injuries, as well as following the induction of polyarthritis (Calvino et al. 1987; Menétry & Besson 1982), and in response to electrical nerve stimulation (Cook et al. 1987). Inflammatory lesions also produce an expansion of receptive fields of cells in the ventrobasal thalamus (Guilbaud et al. 1986). Injury-induced receptive field expansions may contribute to enhanced pain by recruiting primary afferent fibers within the newly expanded field, thus increasing the magnitude of the ascending signal into the CNS, or by modality convergence and activation of previously ineffective synapses (Devor 1989; Dubner et al. 1987). It should be noted, however, that the degree of expansion of neuronal receptive fields is not necessarily related to pain sensation, and varies greatly with the level of anesthesia.

Behavioral and physiological studies in animals also demonstrate hyperalgesia or an increase in the excitability of flexor efferent responses to stimulation of body regions which are at a distance from a cutaneous or deep tissue injury. Woolf (1984) found that localized thermal and chemical injuries cause reductions in flexion reflex thresholds to noxious mechanical and thermal stimulation in the limb contralateral as well as ipsilateral to the injury. Cutaneous (Woolf 1983) and deep (Woolf & McMahon 1985) tissue injury, as well as noxious electrical stimulation of cutaneous and muscle afferent nerves (Wall & Woolf 1984) also produce an increase in the excitability of the ipsilateral and contralateral flexor efferent nerves in response to noxious mechanical stimulation of the hindpaw. Since the increased excitability in the contralateral flexor efferent nerve is maintained even after inputs from the injured paw are blocked by local anesthesia, the results suggest that central, not peripheral, changes underlie this effect. In this way, cutaneous hyperalgesia may depend on central sensitization which is produced by inputs from a peripheral injury, but does not need to be maintained by them. Behavioral studies of thermal withdrawal latencies indicate that the spread of hyperalgesia to the hindpaw contralateral to the paw that received a thermal injury is unaffected by either deafferentation or anesthetic blocks of the injured hindpaw following the injury, but is prevented if deafferen-

tation or anesthetic block precedes the injury (Coderre & Melzack 1985; 1987). The similarity between the physiological and the behavioral data is quite striking considering that the flexor efferent reflex measures were obtained in decerebrate/spinalized animals, while withdrawal latencies were obtained in intact, awake animals. These data provide further evidence that peripheral injury can produce central changes which are maintained even after the inputs from the injury are removed.

2.4. Separate contribution of peripheral and central sensitization to hyperalgesia. The great debate between Lewis and Hardy et al. began chiefly because of their very different interpretations of similar results they each obtained from experiments comparing the effects of proximal and distal nerve blocks on the development of spreading hyperalgesia after faradic stimulation of the skin over peripheral nerves. Lewis concluded that spreading hyperalgesia was mediated by a peripheral neural mechanism because the hyperalgesia was blocked by distal but not proximal nerve blocks. Hardy et al. concluded that secondary hyperalgesia was mediated by a central neural mechanism because the hyperalgesia was unaffected by distal nerve blocks, but was delayed by proximal nerve blocks. While the conclusions based on their apparently different findings were opposite, in reality their experimental findings were very similar (see Fig. 1). First, both Lewis and Hardy et al. found that hyperalgesia to faradic stimulation develops fully after proximal nerve blocks wear off. Lewis reported that hyperalgesia extended throughout the anesthetized region after the nerve block wore off (between 15 and 60 min after stimulation) (Fig. 1B). In Hardy et al.'s experiment, hyperalgesia extended throughout the anesthetized region by 60 min after stimulation (Fig. 1A). Thus, the delay in the spread of hyperalgesia in Hardy et al.'s study was not greater than the variability in the length of anesthesia reported by Lewis.

In the case of distal nerve blocks, there is again no real difference in their data. Although Lewis concluded that hyperalgesia does not spread into the anesthetized region, it is evident from his figures (Lewis 1936, Fig. 6, or Fig. 1D here) that in some cases hyperalgesia did spread up to 2 cm into the anesthetized region. Hardy et al. concluded that hyperalgesia does spread into the anesthetized region after distal nerve blocks. However, when Hardy et al. stimulated nerve trunks at the same distance (3 cm) from the distal nerve block as did Lewis, they also found that hyperalgesia spread only 2–3 cm into the anesthetized region (Hardy et al. 1950, Fig. 8B or Fig. 1C here). Given all the variability inherent in assessing the borders of secondary hyperalgesia, performing nerve blocks, and stimulating nerve trunks through the skin, as well as the normal variability in the development of hyperalgesia seen between different subjects receiving the same stimulus, the similarities of their findings are more impressive than the differences. Thus, it appears that Lewis and Hardy et al. came to completely opposite conclusions with nearly identical data.

What conclusion do we come to from Lewis's and Hardy et al.'s data? With distal nerve blocks, Hardy et al., like Lewis, found very little spread of hyperalgesia into the anesthetized zone, suggesting that peripheral neural mechanisms are critical to secondary hyperalgesia. As for proximal blocks, the fact that both Lewis and Hardy et al. found that the full extent of hyperalgesia to faradic stimulation

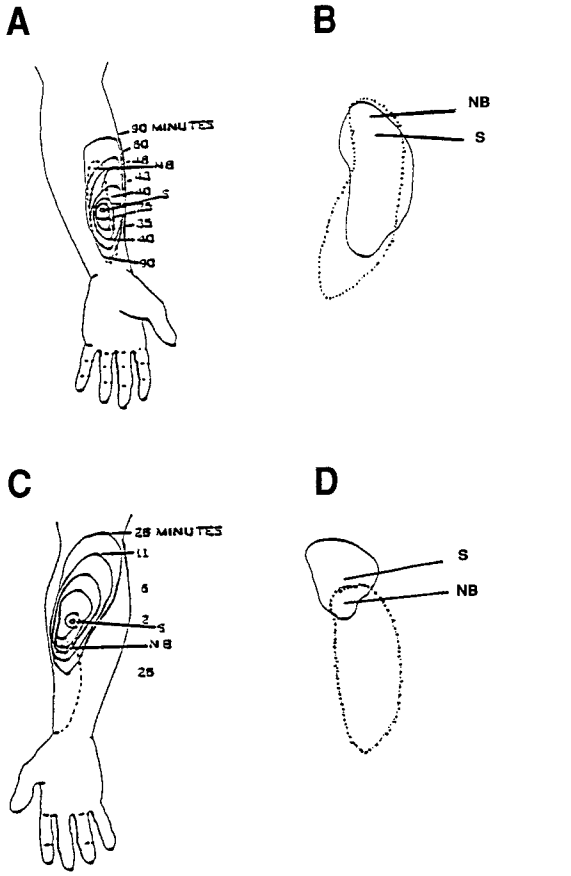


Figure 1. Similarities in the effects of proximal (A & B) and distal (C & D) nerve blocks on hyperalgesia produced by faradic stimulation in studies by Hardy et al. (1950; left side) and Lewis (1936; right side). Nerve blocks (NB) and stimulation (S) were performed where indicated; zones of anesthetic skin and hyperalgesia are indicated by dashed and solid lines, respectively. Lewis's diagrams have been flipped vertically, reduced in scale, and relabeled to enable an easier comparison with those of Hardy et al. Both Hardy et al. (A) and Lewis (B) found extensive spread of hyperalgesia throughout the previously anesthetic zone within 60 min of a proximal nerve block. In contrast, both Hardy et al. (C) and Lewis (D) observed a small, but similar, degree of hyperalgesia spreading into the anesthetic zone after a distal nerve block. Modified from Hardy et al. (1950) and Lewis (1936), with permission.

develops after the block wears off, could be taken as evidence, as Lewis suggests, that central neural mechanisms do not contribute to secondary hyperalgesia. Although the proximal nerve block would prevent neural impulses from reaching the central nervous system during the faradic stimulation, it is highly likely that the intense electrical stimulation required to penetrate the skin and activate high threshold nerve fibers would produce considerable tissue injury (including that produced by peripheral neurogenic and non-neurogenic proceses). Consequently, it is also possible that peripheral tissue injury associated with the faradic stimulation could, after the proximal nerve block wears off, produce central neural changes which contribute to the development of secondary hyperalgesia.

There is no doubt that the faradic stimulation used by Lewis and Hardy et al. did produce extensive tissue injury. In many experiments the skin was stimulated for 5 min with

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

a current at the pain tolerance level. Furthermore, when the skin was infiltrated with a local anesthetic prior to electrical stimulation, there was a delay in the development of secondary hyperalgesia until after the anesthesia wore off. Since the skin was anesthetized during the stimulation, thus preventing axon reflexes and central transmission, it is clear that the electrical stimulation itself did not induce the hyperalgesia. Recently, Dahl et al. (1993) have shown that there is a similar delay in the development of secondary hyperalgesia in response to burn injury (50°C for 6 min) performed within an anesthetized patch of skin. Thus, it appears that an injury performed in anesthetized skin will produce hyperalgesia which appears after the anesthesia has waned, provided the injury is capable of producing non-neurogenic tissue injury.

In contrast to the effects of faradic stimulation or burn injury, the hyperalgesia produced by capsaicin probably involves minor non-neurogenic tissue injury. Subcutaneous injection of capsaicin produces a small bleb at the injection site, which is no larger than that produced by its vehicle, and typically does not even produce a weal (LaMotte et al. 1991). Furthermore, when capsaicin is injected into anesthetized skin, unlike the effects of electrical and burn injuries of the skin, hyperalgesia does not develop at all (LaMotte et al. 1991). For this reason, we suggest that capsaicin is a better stimulus to use when one wishes to assess the effects of brief sensitization, independent of lingering peripheral inputs. Importantly, it has also been found that the secondary hyperalgesia which occurs following burn injury typically does not last as long as the primary hyperalgesia at the site of injury. Moinche et al. (1993) reported that primary hyperalgesia after heat injury persisted for 48 hours, while secondary hyperalgesia to punctate mechanical stimuli did not extend beyond 24 hours. In contrast, LaMotte et al. (1991) found that localized thermal hyperalgesia after capsaicin injury persisted for only 1–2 hrs, while secondary hyperalgesia to punctate mechanical stimuli lasted 13–24 hrs.

It is significant that of the four stimuli that commonly have been used to produce secondary hyperalgesia (faradic stimulation, skin crush, burn, and capsaicin injection), it is only after capsaicin injection that secondary hyperalgesia is prevented by *prior* proximal nerve block or local anesthesia of the skin; the other stimuli invariably lead to a full blown secondary hyperalgesia, which is delayed until the anesthetic wears off. Based on these findings, we hypothesize that prior anesthetic nerve or skin blocks will prevent subsequent hyperalgesia only if there is minimal tissue damage and no continued activation of primary afferents. If this is true, then it would be very difficult to find support for a contribution of central neural mechanisms of hyperalgesia by assessing the effects of prior anesthetic blocks with stimuli which produce extensive tissue injury or continued afferent input.

Since peripheral injury interferes with the ability to assess the contribution of central neural mechanisms of hyperalgesia using prior anesthetic blocks, perhaps more useful information can be gained using postinjury blocks. Lewis (1936) found that hyperalgesia was completely unaffected by local anesthesia of the skin previously subjected to faradic stimulation. Hardy et al. (1950) found the hyperalgesia produced by faradic stimulation could be blocked by subsequent local anesthesia, but only with deep, as opposed to superficial, anesthetization. Dahl et al. (1993) demonstrated that postinjury local anesthesia of a burned region reduced, but did not eliminate, cutaneous hyperalgesia. In the case of capsaicin, LaMotte et al. (1991) found that hyperalgesia to stroking, but not the hyperalgesia to punctate stimuli, was blocked by local anesthesia or cooling of the skin at the site of injury, after hyperalgesia has fully developed. Thus, postinjury blocks have been found either to not affect, to completely block or to partially block hyperalgesia associated with various stimuli.

2.5. Contribution of initial and ongoing central sensitization to hyperalgesia. From the above discussion of cutaneous hyperalgesia it appears that both peripheral and central neural mechanisms may contribute to secondary hyperalgesia, but that any determination of their separate roles is highly controversial. From the above two paragraphs we recognize that the ability to demonstrate a central contribution to secondary hyperalgesia is very much dependent on the degree of tissue injury produced by the stimulus. However, it is likely that both peripheral and central sensitization become more obvious with a greater degree of peripheral injury. We propose that rather than emphasizing the separate roles of peripheral and central neural mechanisms to cutaneous hyperalgesia, the existing data can be better explained by a hypothesis that involves an interactive contribution of both peripheral and central sensitization (see Fig. 2). In addition to this basic interaction, it is important to conceive of central sensitization as composed of two components, initial central sensitization and ongoing central sensitization, which is influenced by peripheral sensitization.

We know that both a brief afferent barrage produced by C-fiber stimulation and the persistent inputs associated with peripheral tissue injury will each sensitize dorsal horn neurons. Thus, when studying the role of central sensitization in hyperalgesia, it is relevant to differentiate between the *initial* central sensitization produced by an injury barrage and the *ongoing* central sensitization associated with lasting tissue injury. In this way, the sensitized state can depend on either one of two processes: an initial intense barrage, or an ongoing lower-level peripheral input. Initial sensitization could be considered an autonomous state that exists without an ongoing or maintaining input, that is, the sensitization which persists after the initial stimulation. In contrast, ongoing sensitization is a more labile state that exists only if there is ongoing peripheral input to maintain it. The separation of central sensitization into these two components may explain many of the differences obtained in studies which assess the effects of pre- or postinjury local anesthesia blockade on secondary hyperalgesia. We hypothesize that with severe tissue injury the afferent input is so intense at the time of testing, that hyperalgesia will be evident regardless of whether or not preinjury anesthetic block was performed. Thus, ongoing central sensitization maintained by the inputs from damaged peripheral tissue overrides much of the benefit of the preinjury block. However, with less extensive tissue injury a preinjury block will reduce or prevent the initial central sensitization which would be produced by the injury barrage. Since there is little ongoing central sensitization associated with minor injury, the secondary hyperalgesia is prevented. Thus, preinjury anesthetic blocks are typically used to provide information about initial central sensitization; however, if there is sufficient ongoing central sensitization, as in the case of

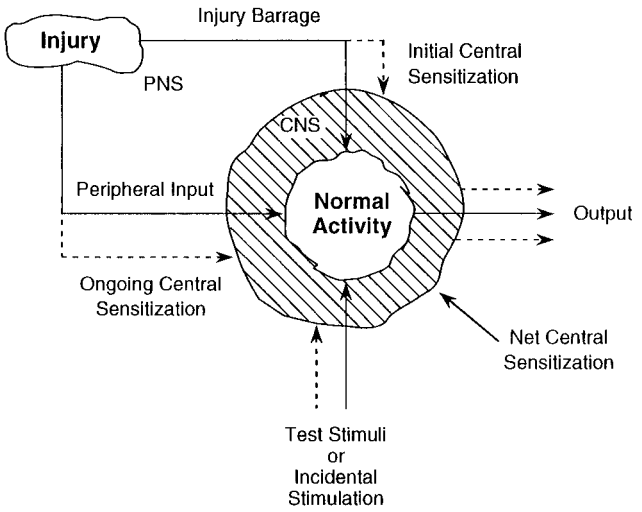


Figure 2. Schematic model of the proposed contribution of both initial and ongoing central sensitization to the output of central nervous system (CNS) neurons that are involved in the processing of nociceptive information. An injury produces both an injury barrage, which underlies initial sensitization, and continued peripheral inputs, which underlie ongoing sensitization of central nociceptive neurons. The combined influence of both initial and ongoing sensitization results in a significant decrease in the threshold of central neurons to further stimulation and is represented by the depiction of a larger area of activation for net central sensitization (diagonal hatched-line fill), as compared with normal activity (open fill). Pain sensitivity (hyperalgesia) and spontaneous pain are influenced by the output of the central nociceptive neurons. The output of central nociceptive neurons is directly influenced by the degree of net central sensitization and by the degree of peripheral inputs which activate afferent fibers that impinge on sensitized neurons. Importantly, peripheral inputs may originate from the injury itself, or from adjacent or remote uninjured regions whose afferent fibers converge directly (monosynaptically) or indirectly (polysynaptically) onto the sensitized neurons. In this way, the output of sensitized neurons reflects not only the degree of input from the injured region, but also the input from test stimuli in an area of secondary hyperalgesia, or incidental stimulation of referred zones (after visceral or deep tissue injury) or focal points (after nerve injury). The dashed-line arrows represent the enhanced output which is associated with an enhanced sensitivity to peripheral inputs, and an enhanced activation of the sensitized central neurons, as compared to the solid-line arrows which represent normal sensitivity, activation, and output of the central nociceptive neurons.

extensive tissue injury, information about initial sensitization is better obtained using *postinjury* anesthetic blocks in order to eliminate the confounding effects of ongoing central sensitization associated with the peripheral injury. In these cases, a role of initial central sensitization is suggested if hyperalgesia persists despite the block. However, hyperalgesia which is dependent on ongoing rather than initial central sensitization is reduced or eliminated by post-injury block.

We also propose that initial and ongoing central sensitization, although temporally separated, are interdependent. Thus, the degree of hyperalgesia is related to the net central sensitization, which derives both from initial and ongoing central sensitization. If the initial central sensitization is great enough, then hyperalgesia may be less dependent on ongoing peripheral inputs. If the ongoing central sensitization is great enough, then hyperalgesia is less dependent on the initial central sensitization. Accordingly, very severe injuries such as faradic stimulation, or burns, that elicit both an intense afferent barrage, as well as subsequent tissue injury, would produce hyperalgesia that involves both a strong initial central sensitization and significant ongoing central sensitization. Since the net central sensitization and resulting hyperalgesia reflect both initial and ongoing central sensitization, in such cases pre-injury anesthetic blocks would only delay hyperalgesia (i.e., it would develop later due to ongoing central sensitization). Similarly, by eliminating ongoing central sensitization, postinjury anesthetic blocks would either not affect, or only partially reduce, hyperalgesia, because the initial sensitization is sufficient to maintain a level of net central sensitization required to produce hyperalgesia.

In the case of a capsaicin injection, where the afferent barrage due to C-fiber activation is high, but the tissue injury is considerably less, the net central sensitization and hyperalgesia depend more on initial sensitization and less on ongoing central sensitization. Thus, preinjury anesthetic blocks are very effective because they reduce the initial central sensitization, but postinjury anesthetic blocks are less effective because the ongoing sensitization is less critical. In contrast, recent evidence (Koltzenburg et al. 1994) suggests that hyperalgesia produced by the topical application of mustard oil may depend more on ongoing central sensitization. Thus, it was established that after mustard oil treatment the degree of brush-evoked secondary hyperalgesia was highly correlated with the degree of ongoing burning pain from the site of injury.

In a recent article in the *IASP Newsletter,* Niv and Devor (1993) raised an important question about the time constant of central sensitization. It was suggested that since secondary stroking hyperalgesia produced by chemical irritation of the skin is diminished by cooling or anesthetizing the injured area, the time constant of central sensitization is short-lived. It was implied that because local anesthetic blockade temporarily relieved secondary hyperalgesia, the central sensitization that underlies the hyperalgesia was eliminated – at least until it was reinitiated after the blocks wore off. We would recommend caution in equating central sensitization and secondary hyperalgesia. While central sensitization may contribute to secondary hyperalgesia, the presence of secondary hyperalgesia does not automatically imply that central sensitization is present, and its absence does not necessarily imply that central sensitization has been eliminated. Importantly, peripheral inputs from chemically irritated skin will retrigger secondary hyperalgesia after a post-treatment anesthetic nerve block, but produce substantially reduced or no secondary hyperalgesia if the nerve block was present at the time of the chemical irritation (Koltzenburg et al. 1994; LaMotte et al. 1991). It is also important to distinguish the concepts of reinitiating versus retriggering of central sensitization. Reinitiating implies that the sensitization has disappeared and must be initiated once again by a similar peripheral input, while retriggering implies the central sensitization did not disappear, but was latent, and is re-established when a necessary threshold is met by peripheral inputs. As pointed out by Gracely et al. (1992), retriggering is more likely than reinitiation, since when first initiated secondary hyperalgesia and allodynia gradually expands outward from the injury, yet when returning after a local anesthetic block of

the injured area, the hyperalgesia and allodynia rapidly expands over its previously existing area.

LaMotte at al.'s (1991) data with capsaicin not only point to the importance of the interaction of initial and ongoing central sensitization, but also attest to the importance of the inputs produced by the test stimulus, since postinjury local anesthetic blockade of the skin more effectively alleviated hyperalgesia to stroking than hyperalgesia to punctate stimuli. The fact that hyperalgesia to stroking is reduced by postinjury blockade, suggests that stroking hyperalgesia is maintained by ongoing central sensitization of the injury. In contrast, the fact that hyperalgesia to punctate stimuli is not alleviated by postinjury blocks, suggests that hyperalgesia to punctate stimuli is not maintained by ongoing central sensitization. We propose that since hyperalgesia ultimately depends on the net activity of dorsal horn neurons, the stroking stimulus may not provide enough input to produce pain when ongoing central sensitization has been abolished by the postinjury local anesthetic blockade. In contrast, the greater sensory input produced by pricking the skin will be sufficient to activate dorsal horn neurons, which have been sensitized by the capsaicin injury barrage (i.e., initial central sensitization), despite the elimination of ongoing central sensitization by the postinjury local anesthetic. In this way, ongoing central sensitization combines with initial sensitization to enhance incoming inputs which must meet or exceed a specific threshold before pain or hyperalgesia is experienced. The elimination of one of these components does not necessarily eliminate hyperalgesia if inputs from one of the other two components is strong enough. It is possible, however, that it is not exclusively sensory intensity that is important, but rather whether high or low threshold afferents are excited, so that high threshold input associated with punctate stimuli is able to overcome the effects of the anesthetic block, while low threshold input associated with stroking is not.

Two phenomena have not adequately been explained by the model in Figure 2: (1) Heat hyperalgesia does not spread into uninjured tissue to the same extent as mechanical hyperalgesia, and (2) mechanical hyperalgesia does not spread across an anesthetized strip of skin. To explain these phenomena LaMotte et al. (1991) proposed that, in addition to mechanoheat sensitive nociceptors and low threshold mechanoreceptors which have less extensive branching, there is a population of chemospecific afferent fibers which branch extensively in the skin; a theory which is now supported experimentally (Schmelz et al. 1994). According to LaMotte et al., capsaicin injury activates the chemospecific afferents which release a neuromodulator from their central terminals. This neuromodulator sensitizes dorsal horn neurons that receive input from myelinated low or high threshold mechanoreceptive afferents. These dorsal horn neurons in turn facilitate the responses of wide dynamic range neurons and high threshold spinothalamic tract (STT) neurons to mechanical stimulation of the skin outside the area of injury, while separate dorsal horn neurons facilitate STT neuron responses to heat stimulation inside the area of injury. According to LaMotte et al., this explains why there is remote hyperalgesia to mechanical but not heat stimuli. LaMotte's model is also proposed to explain the effects of a local anesthetic strip since the anesthesia blocks neural conduction in the lateral branches of the widely branching chemospecific neurons. Although LaMotte's model has been endorsed and expanded upon by some investigators (Meyer et al. 1992), others have been skeptical (Lynn 1992; Wall 1993). In particular, Wall (1993) has suggested that the anesthetic skin strip findings could be explained by central effects, rather than indicating there are widely branched chemospecific afferents. Wall (1993) proposed that a strip of local anesthetic applied to the skin produces a strip of spinal cord cells which are unable to refer to the periphery, and thus blocks the spread of central changes from the original focus to neighbouring cells. Other explanations attribute the spread of mechanical hyperalgesia, despite the absence of spreading heat hyperalgesia, to the sensitization of afferents which are initially insensitive to mechanical stimulation. These afferent fibers develop a sensitivity to mechanical stimulation after exposure to inflammatory mediators released in injured skin (Davis et al. 1993; Kress et al. 1992), as occurs for a group of mechanically insensitive neurons (so-called silent nociceptors) in the inflamed knee joint of rats (Schaible & Schmidt 1988), as well as in skin sensitized by capsaicin or mustard oil in humans (Schmelz et al. 1994; Schmidt et al. 1995).

3. Referred pain and hyperalgesia

First described by Martyn (1864), referred pain is a condition in which pain is not localized within the injured region, but at an adjacent or distant site. Referred pains often occur following injury of deep tissue such as muscle, joints or viscera. Indeed, referred pain and hyperalgesia are often reported in the muscle and skin within the same spinal cord dermatome as injured organs, and are commonly used in the diagnosis of conditions such as appendicitis and angina pectoris. It has been shown that the distribution of referred pain increases with the intensity (Kellgren 1937; 1938; McLelland & Goodell 1943; Vecchiet et al. 1993) and duration (McAuliffe et al. 1943; McLelland & Goodell 1943) of the noxious stimulation from the injured deep tissue. Typically referred pain is restricted to the same spinal segment, however in some cases it has been found to extend great distances and beyond segmental boundaries (Lewis 1942; Livingston 1943). Along with referred pain there is often a development of tenderness in the referred area (i.e., referred hyperalgesia) (Head 1893; Procacci et al. 1986). Importantly, referred pain and hyperalgesia typically develop slowly after injury of deep tissues (Kellgren 1939), similar to the development of secondary hyperalgesia after cutaneous injury (Lewis 1936; Hardy et al. 1950).

3.1. Theories of referred pain. Although several theories have been advanced to account for referred pain, there is general agreement that referred pain depends on neural mechanisms, since local anesthesia of the injured region blocks its expression (Robertson et al. 1947; Vecchiet et al. 1993). Aside from this single point of agreement, the various theories of referred pain have little in common. One theory proposes that referred pain depends on impulses arising from the injured deep tissue region producing a sensitization of the referred area by means of an axon reflex mechanism (Penfield 1925). Another theory suggests that while referred hyperalgesia is dependent on an axon reflex mechanism, referred pain depends on the misinterpretation of inputs from an injured region whose axons also branch to the uninjured referred area (Sinclair et al. 1948). A third, the convergence-projection theory (Ruch 1947), suggests that axons from the injured and referred regions

converge on the same cells in the spinal cord and there is a misinterpretation as to the source of the stimulation. A fourth theory suggests that impulses within axons from the injured region produce a facilitation of cells in the spinal cord at which axons from the referred area also terminate (MacKenzie 1893). This convergence-facilitation theory of referred pain was inspired by earlier suggestions of Sturge (1883) and Ross (1888) that referred pain depended on the development of a "commotion" or "irritable focus" in spinal cord neurons. Other hypotheses for referred pain contend that it is due to a summation of inputs from the injured and referred area within neurons of the brain, rather than the spinal cord (Cohen 1947; Theobald 1941). The major difference between these theories is their reliance on either a peripheral or a central mechanism. The peripheral mechanism is dependent on axon reflexes and peripheral sensitization, and the central mechanism is dependent on convergence in the central nervous system with or without central sensitization.

3.2. Evidence supporting peripheral or central sensitization in referred pain. There is some experimental support for the idea that referred pain may rely on an axon reflex-like response in branched afferent nerves, as proposed in the theories of Penfield (1925) and Sinclair et al. (1948). It has been shown that 18% of unmyelinated lumbar splanchnic nerve fibers can be activated by electrical stimulation of somatic nerves (Bahr et al. 1981). Mense et al. (1981) have also reported that there are sensory neurons with bifurcating axons which innervate both skin and muscle in the cat's tail. However, as indicated by McMahon (1994), it is generally agreed that these types of neurons are rare, if they exist at all. Furthermore, it is unlikely that this branched nerve mechanism could explain the delayed onset of referred pain after deep tissue injury, and it fails to explain pain which has been found to be referred to deafferented areas (Brown 1942; 1948; Kellgren 1938; Livingston 1943). As described in the section on cutaneous hyperalgesia, there is also a debate as to whether antidromic stimulation can produce a spreading sensitization of nociceptors.

There is considerably more evidence to support the idea that referred pain relies on a convergence of inputs at the spinal cord level, as proposed in the theories of Ruch (1947) and MacKenzie (1893). In an extensive review of the literature on the studies of visceral afferent activity Ness and Gebhart (1990) listed over 60 experiments demonstrating the presence of spinal neurons which receive both visceral and somatic input. The percentage of neurons which received convergent input from viscera and somatic inputs ranged from 6 to 100%, but was above 90% in the vast majority of these experiments. While these studies provide significant support for convergence-projection mechanisms of referred pain, they do not adequately explain either the slow development of referred pain or the appearance of tenderness in the referred area. Furthermore, there are many examples in which referred pain is eliminated or significantly reduced by local anesthesia of the referred area (see below).

The slow development of referred pain and hyperalgesia and the reduction of referred pain and hyperalgesia by local anesthesia of the referred area provide more support for a convergence-facilitation theory, since the sensitization of central neurons would most likely take some time to develop, and the resultant sensations would rely on a summation of inputs from the injured and referred areas, and not simply a misinterpretation of the origin of the inputs.

Furthermore, a role of central mechanisms in referred pain is suggested by the observation that phrenic nerve stimulation causes referred shoulder pain even after sectioning all cutaneous nerves from the painful region of the shoulder (Doran & Ratcliffe 1954), and by the finding that the injection of hypertonic saline into intraspinous ligaments produces pain referred to a phantom arm (Harman 1948). If referred pain could be explained exclusively by convergence, then such pains would not provide clear evidence of central sensitization. However, evidence that referred pain is also, in part, dependent on CNS changes is provided by findings that referred pain and hyperalgesia spread to areas which do not share the same dermatome (Lewis 1942; Livingston 1943). For example, it has been shown that pain of cardiac origin is referred to sites as distant as the patient's ear (Brylin & Hindfelt 1984). That pain and hyperalgesia can spread to areas far removed from the injured region implies that central changes and facilitation, as opposed to convergence, are involved in the spread of hyperalgesia.

Referred pain has also been found to spread specifically to sites of a previous injury. Henry and Montuschi (1978) described a case where the pain of an angina attack was referred to the site of an old vertebral fracture, while Cohen (1947) showed that angina brought on by exertion results in pain referred to a prior blister injury of the right elbow or the right mammary region, or to the site of an injury produced by injection of 5% saline into muscles of the back. In each of these cases, no angina pain was referred to these areas before the injuries, and the pain of the injuries had subsided prior to the angina attack which resulted in pain referred to these sites. Furthermore, Hutchins and Reynolds (1947) discovered that alterations in barometric pressure during high-altitude flights caused many of their patients to complain of pain localized to teeth which had been the site of previous painful stimulation (e.g., fillings, caries, and extractions), in many cases years earlier. Reynolds and Hutchins (1948) were able to replicate this finding under controlled conditions. One week after damaged teeth were filled or extracted, pinprick of the nasal mucosa produced pain referred to the previously treated teeth. This phenomenon occurred among patients who had been treated under general anesthesia, but not under the influence of a local anesthetic block. Furthermore, in patients who had bilateral dental treatment without a local anesthetic, subsequent blocks applied to one side permanently abolished the referred pain ipsilateral, but not contralateral, to the anesthetized side.

Theories which propose that central sensitization contributes to referred pain have also received recent support. Studies have shown a sensitization, or expansions of the receptive fields of dorsal horn neurons following inflammatory injury of various visceral tissues such as the urinary bladder (McMahon 1988), colon (Ness & Gebhart 1990) and esophagus (Garrison et al. 1993), or following acute joint inflammation (Dougherty et al. 1992b; Schaible et al. 1987) or electrical stimulation of muscle afferents (Cook et al. 1987). Receptive field expansions have also been observed in trigeminal brainstem neurons following chemical stimulation of deep craniofacial afferents (Hu et al. 1992). Following inflammatory lesions of the rat knee joint, spinal dorsal horn (Neugebauer & Schaible 1990) and thalamic

(Guilbaud et al. 1986) neurons exhibit an enhanced responsiveness not only to mechanical stimulation of the inflamed joint, but also to stimulation of the muscles in the thigh and lower regions of both the ipsilateral and contralateral legs. These findings are consistent with clinical observations that hyperalgesia develops in body regions distant from a deep tissue injury (Hardy et al. 1950), and that flexion reflex thresholds are reduced in patients following gynecological surgery (Dahl et al. 1992a).

3.3. Interactions of peripheral and central mechanisms in referred pain. Perhaps the most controversial and interesting phenomena associated with referred pain relate to the effects of local anesthesia of the area of reference. Some early reports demonstrated that injection of a local anesthetic agent into the referred area reduces or eliminates the referred pain (Lemaire 1926; Morley 1931; Weiss & Davis 1928). Others reported that referred pain was unaffected by local anesthesia of the referred area (Lewis 1942; McClelland & Goodell 1943; Woollard et al. 1932). It may be that the equivocal findings in this area depend to a large extent on the intensity of the inputs from the injury. For example, Theobald (1941) showed that weak faradic stimulation of the cervix produced mild referred pain in the abdominal wall that was eliminated by local anesthesia of the referred area. Alternatively, intense stimulation produced referred pain which was unaffected by local anesthesia of the referred area. In support of this, Doran and Ratcliffe (1954) demonstrated that referred pain in the shoulder after phrenic nerve stimulation could be eliminated by local anesthesia of the referred zone. However, the effect of the anesthetic blockade could be counteracted by increasing the intensity of the stimulation. In addition, Bonica (1967) reported that referred pains of mild intensity associated with the early first stage of labour are virtually eliminated by local anesthesia of the lower abdominal wall, while referred pains of much greater intensity during the late first stage of labour are not affected by local anesthesia of the referred zone. Cohen (1947) has also described patients with left arm amputation who developed pain referred to the phantom arm during an attack of angina. In one patient, local anesthetic block of the brachial plexus eliminated the pain of angina referred to the phantom arm. In a second patient, pain in the phantom arm was reliably brought on by physical exercise. Before local anesthesia of the brachial plexus, pain in the phantom arm developed after the patient walked 120 to 150 yards. After a brachial plexus block the patient could walk up to 600 yards before pain re-appeared in the phantom arm.

The fact that referred pain can sometimes be reduced or eliminated by local anesthesia of the referred zone suggests that referred pain depends on a convergence of visceral and somatic inputs in the spinal cord. However, that local anesthesia is not effective when the visceral stimulus is very intense suggests that when visceral inputs are strong enough, a state of central sensitization develops, over-riding any requirement of input from the referred zone for maintenance of the pain. Thus, as in the case of secondary hyperalgesia, there appears to be an interaction between peripheral and central neural mechanisms underlying referred pain, and in some, but not all, cases afferent input is necessary for its maintenance. Using the terms we previously generated for the model of secondary hyperalgesia in Figure 2, we would argue that referred pain relies predominantly on *ongoing* central sensitization, but that the net output of dorsal horn neurons which results in pain experience is also influenced by inputs from the referred area. If the tissue injury in the visceral organs is extensive enough to produce intense ongoing central sensitization, then inputs from the referred area are not required to produce referred pain. With less extensive injury, there is less ongoing central sensitization and local anesthesia of the referred area will alleviate the referred pain. The interdependence of referred and visceral pain on both the ongoing sensitization from the injured region and the inputs from the referred area are further exemplified by observations that patients with angina can be induced to suffer an angina attack not only by stressing the heart with exertion, but also by irritating the area of reference by producing ischemia in the left arm with a tourniquet (Cohen 1947). The influence of *initial* central sensitization on referred pain is indicated in cases where pain is referred specifically to the site of a previous injury.

4. Neuropathic pain

Pain that occurs as a result of nerve injury is by far the most complex somatosensory phenomenon that we know. Theories of neuropathic pain are probably as numerous as the conditions that lead to them. Symptoms of neuropathic pain include spontaneous pain, paroxysmal pain (episodic, shock-like pain), hyperpathia (pain which is delayed and exaggerated), hyperalgesia and allodynia, as well as extensive secondary hyperalgesia and the development of referred pains and focus points (see Bennett 1994 or Devor 1994 for a recent review). In many instances certain symptoms are closely associated with specific nerve pathologies. However, the diversity of symptoms and pathological changes, and the variability in the relationship between symptoms and pathology across patients allow for considerable debate about the etiology of neuropathic pain.

4.1. Peripheral neural mechanisms. Peripheral factors that contribute to neuropathic pain include an abnormal sensitization of nociceptors (Cline et al. 1989; Culp et al. 1989; Ochoa 1986), the development of abnormal adrenergic sensitivity in nociceptors (J. N. Campbell et al. 1988b; Wallin et al. 1976), the development of ectopic activity in damaged nerves, or in dorsal root ganglion (DRG) cells of damaged nerves (Devor et al. 1994; Kajander et al. 1992; Xie et al. 1995), the formation of ephaptic connections in demyelinated axons (Jänig 1988), and abnormal sensitivity after collateral sprouting of primary afferent neurons (Inbal et al. 1987), among others.

Particularly important in many neuropathic pains is the involvement of the sympathetic nervous system, either through development of abnormal sympathetic function (Hoffert et al. 1984) or through its effects on abnormally functioning afferent nerves (J. N. Campbell et al. 1994). It is clear from animal data, that in certain instances following nerve injury either nociceptors (Hu & Zhu 1989; Sato & Perl 1991), regenerating fibers within a neuroma (Wall & Gutnik 1974; Devor & Jänig 1981), or the somata of injured nerves (Devor et al. 1994; Kajander et al. 1992; Xie et al. 1995) develop an abnormal adrenergic sensitivity. It has also been observed following nerve injury in rats that rapidly adapting mechanoreceptors are modified such that

they respond with low and irregular static discharges during a maintained mechanical stimulus (Na et al. 1993). Importantly, these modified responses appear to depend on sympathetic efferent function, since they are blocked by intravenous administration of the adrenergic antagonist phentolamine. This type of sympathetic-sensory coupling mechanism results in high levels of afferent input in response to post-ganglionic sympathetic output. Pain syndromes such as causalgia and reflex sympathetic dystrophy are often relieved by sympathetic ganglion blocks (Bonica 1979). The importance of sympathetic nervous system involvement in many neuropathic pains has lead to the development of a new taxonomy of neuropathic pain. Current thinking distinguishes between sympathetically maintained pain (SMP) and sympathetically independent pain (SIP) (J. N. Campbell et al. 1994; Roberts 1986). SMP is defined as "pain attributable to sympathetic efferent function in peripheral tissues" (Campbell et al. 1994). SMP is, by definition, abolished when the sympathetic supply to the painful region is blocked. In contrast, SIP is not dependent upon sympathetic efferent function, so that maneuvers that are directed at blocking peripheral sympathetic activity do not affect the pain. One of the major advances of this taxonomy is to dissociate the presence of pain from signs of sympathetic disregulation (e.g., altered temperature, excessive sweating, trophic changes), so that evidence of abnormal sympathetic activity need not accompany SMP.

It is important to recognize that persistent pain and hyperalgesia are not linked exclusively with a single source of peripheral pathology. In some cases, hyperalgesia associated with nerve injury is alleviated by A-fiber nerve blocks (Meyer et al. 1985; Ochoa 1982; Wallin et al. 1976), while in other cases it is resistant to A-fiber block and is likely influenced by C-fibers (Ochoa & Torebjörk 1989). In some instances, hyperalgesia associated with nerve injury is alleviated by sympathetic blocks, while in other cases it is either not relieved or even exacerbated by sympathetic blocks (Bonica 1979; Ochoa & Marchettini 1993). Ochoa and Marchettini (1993) also point out that there are subsets of neuropathic pain patients which can be divided into "hot" and "cold" patients, based on the temperature of their limbs and the effects of temperature on their sensory symptoms. In "hot" patients, the affected skin is hot and pain is elicited by warming, while cooling provides relief of pain and hyperalgesia. In "cold" patients, the affected skin is cold, and pain is provoked by cooling, and may be relieved by warming.

4.2. Central neural mechanisms in neuropathic pain. As pointed out earlier, the development of post-injury adrenergic sensitivity in nociceptors and damaged nerves suggests that SMP can be explained entirely by peripheral mechanisms. However, several investigators have proposed that there is an interaction between peripheral and central neural factors which underlies SMP. Thus, Roberts (1986), who coined the term SMP, suggested that after peripheral nerve injury, an abnormal central state develops in which activity in sympathetic efferents stimulates low-threshold mechanoreceptors, which in turn induce pain by activating sensitized dorsal horn neurons. A similar type of interactive peripheral/central model for SMP has also been proposed by J. N. Campbell et al. (1994). According to their model, following injury, spontaneous pain results from asympathetic-sensory coupling mechanism in which nociceptors upregulate alpha-adrenergic receptors and respond to noradrenaline released from sympathetic terminals in the affected region. Sympathetically generated nociceptor activity produces a dynamically maintained state of central sensitization so that activity in low-threshold mechanoreceptors, which normally is not painful, now evokes allodynia in response to light touch. Blocking the sympathetic supply to the injured region interrupts the sympathetic-sensory coupling mechanism and restores the central neurons to a desensitized state, thus relieving both ongoing pain and allodynia.

Models such as these that propose an interactive role of sympathetic-sensory coupling in the periphery and central neural changes are extensions of an early proposal by Livingston (1943). Livingston argued that afferent activity generated from a peripheral nerve injury elicits an abnormal firing pattern within the spinal cord. A disturbance ensued in an internuncial pool of dorsal horn interneurons which resulted in reverberatory activity that eventually spread to other parts of the spinal cord. The spread of activity to the lateral spinal dorsal horns would have the effect of increasing sympathetic efferent activity, causing a disruption in vasoregulation, trophic changes, and hypersensitivity of peripheral tissue. The resultant increased sensory input, driven by sympathetic outflow, acts to maintain the abnormal firing in the spinal cord, creating Livingston's "vicious circle" of peripheral-central activity.

Recent animal data support the notion that nerve injury produces changes in central neural function. Nerve section leads to the development of increased neuronal activity at various levels of the somatosensory system. In addition to spontaneous activity generated from the neuroma (Wall & Gutnik 1974), peripheral neurectomy also leads to increased spontaneous activity in the DRG (Burchiel 1984; Wall & Devor 1983), dorsal spinal roots (Howe et al. 1977; Wiesenfeld & Lindblom 1980), and spinal cord (Asada et al. 1990; David & Aguayo 1980). Furthermore, after dorsal rhizotomy, there are increases in spontaneous neural activity in the dorsal horn (Basbaum & Wall 1976; Loeser & Ward 1967), the spinal trigeminal nucleus (Anderson et al. 1971; Macon 1979) and the thalamus (Albe-Fessard & Lombard 1983; Lombard et al. 1979). There is also a lowered threshold for evoked activity in thalamic (Guilbaud et al. 1990) and cortical (Guildbaud et al. 1992a) neurons of rats with chronic constriction injuries of the sciatic nerve. These findings parallel those reported in the clinical literature in which patients with neuropathic pain after deafferenting lesions exhibit increased spontaneous neural activity in the somatosensory thalamus (Gorecki et al. 1989; Hirayama et al. 1989). Nerve section also produces expansions of the receptive fields of the neurons adjacent to those which are denervated by peripheral nerve sections (Devor & Wall 1978). These receptive field expansions have significant implications for the development of persistent pain after nerve sections. Thus, Markus et al. (1984) have demonstrated that the development of hypersensitivity in a rat's hindpaw following sciatic nerve section occurs concurrently with the expansion of the saphenous nerve's somatotopic projection in the spinal cord.

It is possible that receptive field expansions and spontaneous activity generated in the CNS following peripheral nerve injury are, in part, mediated by alterations in normal inhibitory processes in the dorsal horn. After peripheral

nerve section, there is a reduction in the dorsal root potential, and the presynaptic inhibition it is assumed to represent (Wall & Devor 1981). Nerve section also induces a reduction in the inhibitory effect of A-fiber stimulation on activity in dorsal horn neurons (Woolf & Wall 1982). Furthermore, nerve injury affects descending inhibitory controls from brainstem nuclei. In the intact nervous system, stimulation of the locus coeruleus (Segal & Sandberg 1977) or the nucleus raphe magnus (Oliveras et al. 1979) produces an inhibition of dorsal horn neurons. Following dorsal rhizotomy, however, stimulation of these areas produces excitation, rather than inhibition, in half the cells studied (Hodge et al. 1983).

Further evidence that loss of inhibitory control mechanisms contributes to pathological processing after nerve injury is indicated by the development of transynaptic degenerative changes in small and medium size neurons in lamina I-III of the spinal cord dorsal horn (Sugimoto et al. 1990). It has been suggested that the degeneration of neurons is triggered by intense stimulation associated with ectopic discharges in damaged nerves, and that the affected cells include inhibitory interneurons (Kajander & Bennett 1992; Sugimoto et al. 1990). Importantly, both the degenerative changes (Sugimoto et al. 1990) and behavioral indices of enhanced pain sensitivity in animals (Yamamoto & Yaksh 1993) are accentuated by intrathecal administration of strychnine, which, as a glycine antagonist, blocks inhibitory postsynaptic potentials.

Recently, several animal models of peripheral neuropathy have been developed which produce behavioral signs of hyperalgesia, allodynia, and spontaneous pain or dysesthesia. These behavioral signs bear a striking resemblance to symptoms of nerve injury-related pain in humans. These models involve placing either loosely constrictive ligatures around the entire rat sciatic nerve (Bennett & Xie 1988), tight ligatures around ⅓ to ½ the rat sciatic nerve (Seltzer et al. 1990), or tight ligatures around the L5 and L6 spinal dorsal nerves (Kim & Chung 1992), and observing the behavioral symptoms associated with the nerve pathology that develops over the next several days or weeks. Behavioral symptoms include hyperalgesia to radiant heat or mechanical stimulation, allodynia in response to warm or cold temperature stimulation, guarding of the affected limb (Attal et al. 1990; Bennett & Xie 1988; Kim & Chung 1992; Seltzer et al. 1990), and extraterritorial hyperalgesia and allodynia in adjacent uninjured tissue supplied by the saphenous nerve (Tal & Bennett 1994). In addition to these behavioral signs, recent evidence suggests that these nerve constriction injuries produces profound changes in spinal cord physiology, including transynaptic degeneration (Sugimoto et al. 1990), increases in c-*fos* expression (Kajander et al. 1990) and the growth associated protein GAP 43 (Cameron et al. 1991), as well as decreases in tachykinin immunoreactive staining (Bennett et al. 1989; Cameron et al. 1991) in the dorsal horn. They also produce increased spontaneous activity and increased excitability (lowered thresholds to mechanical stimulation, and afterdischarges to suprathreshold stimuli) of spinothalamic tract cells (Palecek et al. 1992), as well as spontaneous discharges (Kajander et al. 1992; Xie et al. 1995) and increases in immunoreactivity of tyrosine hydroxylase (Chung et al. 1993) and nitric oxide synthase (Steel et al. 1994) in the DRG cells. Furthermore, the constriction injury leads to a dramatic increase in spinal cord metabolic (2-DG) activity in both the ipsilateral and contralateral spinal cord (Mao et al. 1992a). Since metabolic activity is increased in the absence of additional peripheral stimulation, it has been argued that the behavioral symptoms are driven by sustained alterations in spinal cord function. This notion is supported by the finding that there is a reduction in the hyperalgesia that develops following constriction injury of the sciatic nerve if the nerve is locally anesthetized at the time of injury (Dougherty et al. 1992a). Finally, while hyperalgesia and spontaneous pain produced by nerve constriction are reduced by post-injury local anesthesia of the sciatic nerve (Mao et al. 1992b) or sympathectomy (Kim et al. 1993; Shir & Seltzer 1991) suggesting a peripheral contribution to the pain pathology, they are also reduced by systemic or intrathecal administration of NMDA antagonists (Davar et al. 1991; Mao et al. 1992b), suggesting a central contribution as well.

4.3. Focal points and the maintenance of central sensitization. Associated with neuropathic pain is the development of focal points, which when stimulated produce painful sensations that are referred to a remote area. Livingston (1943) reported that pain, allodynia, and hyperalgesia associated with nerve injury can sometimes be alleviated by injections of local anesthetic into a focal point. Recently, Gracely et al. (1992) has presented experimental data from four patients who demonstrated a focus of unusually great cutaneous sensitivity, as well as allodynia and hyperalgesia in a remote region, at a distance from the focus. Stimulation in the focus produced intense pain in the allodynic and hyperalgesic skin. Local anesthesia of the focus eliminated the ongoing pain as well as the allodynia and hyperalgesia. They proposed a model of neuropathic pain in which ongoing nociceptive input from the focus maintains altered central processing that accounts for various sensory and motor abnormalities. Similar to the model we have discussed thus far in relation to cutaneous hyperalgesia after tissue injury and referred pain, the Gracely model relies on an interaction between alterations in central processing produced by an initiating stimulus and maintaining inputs from an ongoing peripheral source. According to Gracely et al. (1992), the original nerve injury alone, or in combination with noxious inputs due to subsequent surgical procedures, produces central neural changes, that are dynamically maintained for prolonged periods by ongoing input from peripheral pathological sources. Importantly, they suggest that altered central processing is a normal process, and not a pathology, but is maintained by pathological peripheral inputs. Other models of the central consequences of peripheral injury have also been proposed. A 1991 consensus statement highlights the role of a variety of peripheral triggers in inducing and maintaining central sensitization (Devor et al. 1991). Furthermore, recent data from Koltzenburg et al. (1994) indicates that the degree of allodynia or brush-evoked pain in patients with neuralgia is closely correlated with the degree of ongoing pain present in the affected limb. They also found that brush-evoked pain was prevented when ongoing pain in the affected limb was relieved by a regional guanethidine block or by local anesthetic blocks of nerves supplying the symptomatic skin.

Ongoing inputs which maintain the altered central processing can arise from any number of peripheral sources of the kind we have discussed previously. Since a specific

peripheral source is not defined within Gracely's model, then similar neuropathic pains can be experienced by patients with very different peripheral pathologies. Thus, in SMP, activity in sympathetic efferents would contribute to the peripheral source, and in SIP, there would be another peripheral source (e.g., ephaptic connections, collateral sprouting, etc.). This may explain why patients with SMP and SIP often present with similar symptoms, and may also explain why patients on occasion have only some symptoms relieved by sympathetic blocks. According to Gracely et al., the key underlying mechanism is the altered central processing which can be maintained entirely, partially or not at all by abnormal peripheral sensitivity to sympathetic efferent activity.

Another similarity between Gracely et al.'s model and that presented in Figure 2 is the reliance of altered central processing on inputs associated with the test stimuli. Thus, Gracely described a case where repeated stimulation of the focus area resulted in reports of increased spontaneous pain and allodynia from the remote area. The altered central processing depends not only on an initial injury for its development and on ongoing peripheral inputs for its maintenance, but also on the magnitude of the peripheral inputs for its expression. They also argued that since pain, allodynia, hyperalgesia, and the sensitivity of the focus reappeared after local anesthesia of the focus area wore off, the altered central processing persists in a silent state until reactivated by the peripheral source. This raises the possibility that long-term blocking of the peripheral source may be necessary to reverse the altered central processing, and could produce a prolonged pain free period that outlasts the peripheral block. Evidence for this comes from clinical reports that a series of anesthetic or peripheral sympathetic blocks sometimes produces long-lasting pain relief in patients with neuropathic pain (Benedetti 1993; Bonica 1979; Livingston 1943).

Using our own terms, Gracely et al.'s hypothesis could be redefined as being reliant on both initial central sensitization and ongoing central sensitization. The two components contribute to net central sensitization, along with additional peripheral inputs from test stimuli or other sensory stimuli which contribute to the net activity of dorsal horn cells underlying neuropathic pain. Although there are many similarities between Gracely et al.'s model and our own, the main difference is their proposal that the altered central processing depends critically on maintaining inputs from peripheral sources (although they do propose that with prolonged peripheral input the altered central processing may become autonomous of the peripheral input). In contrast, we propose that if the initial central sensitization is great enough, then the net central sensitization will allow pain and hyperalgesia to persist in the absence of ongoing central sensitization from a peripheral source. Thus, secondary stroking hyperalgesia after faradic skin stimulation, or hyperalgesia to punctuate stimuli after capsaicin injection, will persist after local anesthesia of the injured skin.

4.4. Pain in phantom limbs and deafferented structures. Mitchell (1872) coined the term "phantom limb" to describe the well know phenomena in which amputees continue to report a sensory awareness of a limb that has been amputated. The term phantom limb pain (PLP) is now commonly used to describe pains which are referred to the phantom limb. The idea that pain can be referred to a phantom limb clearly implies that central neural mechanisms are involved. Unlike the peripheral theories of referred pain, there is absolutely no opportunity for PLP to result from a branching of nerves from an injured region to the area of reference. The simplest hypothesis is that inputs from damaged nerve trunks, stump neuromas, or DRG cells generate signals which are transmitted to the CNS where they are interpreted as coming from the amputated limb, so that pain is projected to that region. This explanation is similar to the convergence-projection hypothesis of referred pain, in that pain is projected to the amputated limb because of a misinterpretation of the origin of the input; however, with PLP there is no need to propose converging inputs, since the transected nerves in the stump continue to innervate the same spinal cord cells after amputation.

Some of the original descriptions of Mitchell (1872), as well as an accumulating body of more recent evidence, suggests that PLP depends not only on a mislocation of the origin of input, but also a sensitization of central neurons prior to, or during, amputation. A striking property of PLP is the persistence of a pain that existed in a limb prior to its amputation (Melzack 1971). Case studies of amputees (see Katz & Melzack 1990) have demonstrated pain "memories" of painful diabetic and decubitus ulcers, gangrene, corns, blisters, ingrown toenails, cuts, and deep tissue injury. In addition, the phantom limb may assume the same painful posture as that of the real limb prior to amputation (Katz & Melzack 1990). Mitchell described this phenomena in one of his patients: "Another class has the hand constantly in some painful position which it occupied before the operation, so that the last real sensation is so stamped upon the sensorium as to forbid its erasure by any future impression."

It has been reported that as many as 79% of amputees describe their phantom pains as similar to pains felt in the limb before amputation (Katz & Melzack 1990). Reports of pain memories in phantom limbs appear to be less common when there has been a discontinuity, or a pain-free interval, between the experience of pain and the amputation. This may explain why relief of preamputation pain by continuous epidural block for 3 days prior to amputation (Bach et al. 1988), as well as after amputation (Jahangiri et al. 1994), decreases the incidence of PLP 6 months later. Furthermore, there appears to be a higher probability that pain will persist in the phantom limb if pain is experienced at or near the time of amputation (Jensen et al. 1985; Katz & Melzack 1990), or if preamputation pain is very intense or of long duration (Jensen et al. 1985; Jensen & Rasmussen 1994).

There is also a literature on the persistence of painful and nonpainful sensations associated with removal or deafferentation of body structures other than the limbs, including breasts (Kroner et al. 1989), teeth (Hutchins & Reynolds 1947; Reynolds & Hutchins 1948), and internal and special sense organs. Ulcer pain has been reported to persist after vagotomy (Szasz 1949) or subtotal gastrectomy with removal of the ulcer (Gloyne 1954). Similarly, patients have reported labor pain and menstrual cramps following total hysterectomy (Dorpat 1971), rectal and hemorrhoid pain following removal of the rectum (Ovensen et al. 1991), the burning pain of cystitis after complete removal of the bladder (Brena & Sammons 1979), and the pain of a severely ulcerated cornea after enucleation of an eye (Minski 1943).

Pain also persists in patients with deafferentation that does not involve amputation. Patients with brachial plexus avulsions (Jensen & Rasmussen 1994; Reisner 1981) and spinal cord injuries often experience pain in the anesthetic, deafferented region (Berger & Gerstenbrand 1981; Conomy 1973). For example, Nathan (1962) described a patient who continued to feel the pain of an ingrown toenail after a complete spinal cord break. In addition, patients with injuries of the brachial plexus (Jensen & Rasmussen 1994; Reisner 1981) or spinal cord (Berger & Gerstenbrand 1981; Conomy 1973;) sometimes report that a limb is in the same uncomfortable, and often painful, posture it was in prior to the injury or block.

PLP or deafferentation pain is not entirely independent of peripheral inputs. In some instances there is a reactivation of pain experienced before amputation that is brought on by peripheral stimulation. Leriche (1947a; 1947b) described a patient who did not experience PLP until 6 years after amputation, when an injection into the stump instantly, and permanently, revived the pain of a former painful ulceration of the Achilles tendon. Nathan (1962; 1985) reported a similar phenomenon when applying noxious stimulation to the stump of an amputee who later reexperienced the pain of an ice-skating injury he had sustained 5 years earlier when the leg was intact. Noordenbos and Wall (1981) also described 7 patients with partial peripheral nerve injury and subsequent pain, who underwent complete nerve resection and graft or ligation. Following regeneration and a pain-free period, all redeveloped pain of the same quality and in the same location as the pain they had experienced prior to nerve resection, although in some patients the recurrence of pain was restricted to a smaller area within the originally painful region. These studies and case reports indicate that previous pains may be reactivated months or even years after the original injury, in some cases by a peripheral trigger which provides the input required to activate the sensitized central neurons. In the case of amputation phantoms, likely candidates for peripheral triggers include ectopic output from neuromas and DRGs.

4.5. Phantom-like pain in animals. Deafferentation by peripheral neurectomy or dorsal rhizotomy in rodents is followed by self-mutilation (autotomy) in which the animals bite and scratch the insensate paw to the point of amputation (Wall et al. 1979). There is evidence that autotomy behavior is produced by ongoing pain or dysesthesia, associated with increased neuronal activity, which is referred to the anesthetic region (Blumenkopf & Lipman 1991; Coderre et al. 1986; however, also see Rodin & Kruger 1984; Sweet 1981). Autotomy behavior is dramatically affected by alterations in the level of noxious input present at the time of, or prior to, nerve section. Thus, noxious chemical (Coderre et al. 1986; Dennis & Melzack 1979), thermal (Coderre & Melzack 1985; 1987; Katz et al. 1991), and electrical (Katz et al. 1991; Seltzer et al. 1991) stimulation prior to nerve sections significantly increases the severity of autotomy following neurectomy or rhizotomy. These findings suggest that the prior injury produces central changes which influence nociceptive behavior, after nerve sections, at a time when inputs from the injured region are no longer capable of transmitting their message centrally. In contrast to the effect of increasing noxious inputs at the time of nerve injury, reducing or eliminating the afferent barrage induced by nerve section produces a dramatic reduction in autotomy. When the afferent barrage induced by nerve cuts in rats is blocked by treating the sciatic and saphenous nerves with local anesthetics prior to sectioning them, there is a significant reduction in the incidence and severity of autotomy (González-Darder et al. 1986; Seltzer et al. 1991). It has also been shown that intrathecal treatment with morphine 1 hr before, but not 15 min after sciatic nerve section, resulted in a significant reduction in the severity of autotomy lasting for least 28 days (Puke & Weisenfeld-Hallin 1993).

Katz et al. (1991) recently developed an animal model which parallels the observation that human amputees report similar pains in a limb before and after amputation. In this animal model, rats selectively initiated autotomy in either the lateral or medial half of a hindpaw if that particular half had been given a thermal injury prior to sciatic and saphenous nerve sections. The selective attack on the previously injured region, despite the fact that the entire foot was deafferented, suggests that the rats were responding to pain referred to the injured area, which was associated with the prior injury and the central sensitization it produced. Rats injured after neurectomy did not show a similar preference indicating that the rats were not responding simply to peripheral cues associated with the injury.

4.6. Role of ongoing inputs in phantom limb pain (PLP). Both the human and animal data suggest that the initial central sensitization is critical to the development of PLP or pain in deafferented structures. However, as described previously, PLPs which resemble pre-amputation pain sometimes require a peripheral stimulus to trigger their onset. In other cases, PLP can be relieved by local anesthesia of focal points in the stump (Livingston 1943). Thus, while PLP may be initiated by central sensitization associated with an injury barrage, there is a role for peripheral inputs and potentially ongoing central sensitization for the maintenance of PLP. In many cases, PLPs resolve themselves within a few months of amputation. It is expected that in these cases PLP is originally driven primarily by inputs from ectopic activity in stump neuromas or DRG cells acting on dorsal horn neurons sensitized by the pre-amputation injuries and/or the injury barrage associated with the nerve sections at the time of amputation. The PLP subsides when the damaged nerves heal adequately, minimizing ongoing inputs from the stump neuromas or DRG cells. PLP may continue if the initial sensitization is of sufficient intensity, if peripheral tissues do not heal adequately, or if stump neuromas or DRG cells develop a sensitivity to sympathetic efferent activity so that ongoing peripheral inputs produce an ongoing sensitization. Furthermore, the fact that PLP can be triggered months or years after the amputation suggests that initial central sensitization produces a lasting influence on central processing, which can reinstate a painful condition if appropriate dorsal horn neurons and/or more rostral sensory structures are activated by peripheral triggering inputs.

5. Postoperative pain

Early this century, Crile (1913) first proposed that CNS changes produced by tissue damage and noxious inputs associated with surgery could contribute to postoperative

pain. However, it was only after the recent finding of Woolf and Wall (1986) provided a sound justification for preemptive treatment, that this idea began to receive the clinical attention it deserves. Woolf and Wall (1986) demonstrated in experimental animals that opioids are much more effective at reducing stimulus-induced increases in the excitability of the dorsal horn if they are administered prior to, rather than following, C-fiber electrical nerve stimulation. Recent clinical evidence supports the hypothesis that the administration of analgesic agents prior to surgery may prevent the central sensitizing effects of the surgical procedure. Thus, it may be possible to reduce postoperative pain intensity or lower post-operative analgesic requirements for periods much longer than the duration of action of the preoperatively administered agents.

5.1. Pre-emptive analgesia. A growing body of clinical data shows that preoperative local (Jebeles et al. 1991; Rademaker et al. 1991; Ringrose et al. 1984; Tuffin et al. 1989; Tverskoy et al. 1990) or spinal (Bugedo et al. 1990; Heard et al. 1992; Tverskoy et al. 1990) anesthesia, or the epidural preadministration of analgesic agents (Campbell et al. 1990; Kiss & Killan 1992; McQuay et al. 1988; Richmond et al. 1993), can significantly reduce postoperative pain or postoperative opioid requirements (see Woolf & Chong 1993 for review). The analgesic effects of such preoperative treatments are assumed to depend on the ability of the pretreatment to preempt the surgically induced sensitization of central nervous system (CNS) neurons; the term preemptive analgesia has been coined for such treatments (Wall 1988). Although there is considerable evidence to show that peripheral injury, as would occur with surgery, leads to a sensitization of CNS neurons (Hylden et al. 1989; Kenshalo et al. 1979; McMahon & Wall 1984; Perl 1976; Simone et al. 1991; Woolf & King 1990), the evidence for preemptive treatments to attenuate postoperative pain to a clinically significant degree is less convincing (Dahl 1994; Katz 1995; McQuay 1995). The conclusiveness of the evidence is hampered by the failure to adequately address whether the same treatment started *after* surgery could produce the same therapeutic benefit. Initial studies examined the effects of preemptive treatments on postoperative pain as compared with no treatment. Although studies of pretreatment versus no treatment were overwhelmingly suggestive of a beneficial effect of preemptive analgesia, its value became less obvious when compared with the same treatment initiated after surgery (i.e., pre- vs. postsurgery). Studies comparing the effectiveness of pre- versus postsurgical treatment using local anesthetic infiltrations, systemic analgesia or regional administration of opioids or local anesthetic agents have produced equivocal results, with some studies indicating a small beneficial effect (Ejlersen et al. 1992; Katz et al. 1992a; 1994; Richmond et al. 1993) and others no effect (Dahl et al. 1992b; Dierking et al. 1992; Pryle et al. 1993; Rice et al. 1990).

One explanation for the lack of clinically significant benefits of presurgical administration of opioids or local anesthetic agents has been that in some clinical trials pre- or intraoperative opioids are used routinely as part of the general anesthetic regimen in both pre- and postsurgical treatment groups (Katz et al. 1992b). Thus, it is possible that the pre/intraoperative opioid use may confound the results, since they may themselves produce a preemptive effect that reduces postoperative pain (Katz et al. 1996; Yashpal et al. 1996). Another explanation is that in some instances postoperative pain may depend more heavily on the peripheral inflammation that follows surgery than on central sensitization that occurs during surgery (Coderre et al. 1993; Woolf & Chong 1993), and consequently postsurgical treatments may be as effective as presurgical treatments.

There is now recent evidence that the preemptive effect is mediated by the NMDA receptor-ion channel complex (Roytblat et al. 1993; Tverskoy et al. 1994) since patients administered intraoperative ketamine (a clinically available anesthetic with NMDA channel blocking properties) but not a placebo show a reduction in postoperative mechanical hyperalgesia at the incision site two days after surgery (Tverskoy et al. 1994), and significantly reduced morphine requirements during the first 24 hrs after surgery (Roytblat et al. 1993). In both these studies, the preemptive effects of ketamine were observed at least 24 hrs after the duration of ketamine's pharmacological action.

5.2. Animal models of pre-emptive analgesia. In the animal literature, the formalin test has been used as a model of injury-induced central sensitization (Coderre et al. 1990), and as a model for studying the mechanisms underlying preemptive analgesia (Abram & Yaksh 1993; 1994; Goto et al. 1994). Subcutaneous injection of dilute formalin into a rat's paw produces a biphasic response including an early intense response in the first 5 min, and a later moderate response that is expressed from 20 to 60 min after injection (Dubuisson & Dennis 1977). The nociceptive response to subcutaneous formalin injection is matched by a corresponding biphasic increase in the activity of dorsal horn neurons after such injection (Dickenson & Sullivan 1987a). It has been demonstrated that intrathecal (i.t.) administration of either lidocaine (Abram & Yaksh 1994; Coderre et al. 1990) or opiates (Abram & Yaksh 1993; Dickenson & Sullivan 1987a) abolishes behavioral and dorsal horn neuron responses to subcutaneous formalin, if they are administered prior to, but not immediately after, the early phase of the formalin response. This suggests that neural activity generated during the early phase of the formalin response is capable of producing changes in CNS function which in turn influence nociceptive processing during the late phase. The ability of the preinjury treatment with i.t. lidocaine or opiates to suppress the late phase response to formalin has been described as an animal model of preemptive analgesia, since the pretreatments are able to preempt the central sensitization which contributes to persistent postinjury nociceptive behaviors.

Recently, we have demonstrated that the ability of i.t. lidocaine to preempt postinjury nociception in the formalin test was lost as the concentration of formalin was increased from 2.5 to 5% (Yasphal et al. 1996). A strong preemptive effect (i.e., a significant reduction in nociceptive scores) of lidocaine was obtained in rats given 2.5% formalin. This preemptive effect was reduced (resulting in significantly higher nociceptive scores) in a concentration-dependent manner in rats given 3.75 and 5% formalin. In the same study, we found that while a significant and concentration-related degree of inflammation (plasma extravasation) was produced by 3.75 and 5% formalin, the degree of inflammation produced by 2.5% formalin was not significantly different than that produced by the same volume (50 μl) of saline, and was only slightly, but not significantly higher

than no injection at all (Yasphal et al. 1996). Thus, the preemptive effects of i.t. lidocaine were greatest when there was little or no inflammation, and decreased directly with increases in peripheral inflammation. One implication of these findings is that it may be difficult to demonstrate a significant effect of preemptive analgesia after surgical procedures which produce extensive peripheral injury, accompanied by considerable local inflammatory changes. In this manner, the peripheral inflammatory changes and afferent input associated with postoperative inflammation may over-ride much of the beneficial effect of blocking afferent inputs at the time of surgery (but also see Katz et al. 1994). A second implication of this finding is that it is important to pay careful attention to the concentration of formalin that the investigators have used, when comparing the results of different studies using the formalin test. This distinction may explain the recent finding that there is no difference between pre- and posttreatment with i.t. lidocaine or excitatory amino acid antagonists (Chapman & Dickenson 1993) or opioids (Chapman et al. 1994) on the dorsal horn neuronal responses to a peripheral injection of 5.0% formalin to rats' toes, and why posttreatment with the NMDA antagonist AP5 produced a significant reduction in nociceptive responses to hindpaw injection of 10% formalin in mice (Murray et al. 1991). Furthermore, it may explain why late phase dorsal horn neuronal responses to 5.0% formalin are significantly reduced by local anesthesia of the injected area at the time of testing (Dickenson & Sullivan 1987b), but not by a prior local anesthesia of the injected area during the early phase (Haley et al. 1990).

5.3. Preemptive analgesia and models of persistent pain in animals. In addition to providing experimental evidence for a possible explanation as to why preemptive treatments may not always be effective for reducing of postoperative pain, our results may also explain why pretreatments are not always more effective than posttreatments for the alleviation of nociception after peripheral tissue injury in animals. We have previously (Coderre 1993) pointed out, for example, that the effectiveness of pre- versus posttreatment with the N-methyl-D-aspartate receptor antagonist MK-801 as an antinociceptive agent is very much dependent on the type of nociceptive test that is used. In the case of peripheral neuropathy following nerve constriction injury in rats (Davar et al. 1991; Mao et al. 1992a; 1992b; Yamamoto & Yaksh 1992a), it has been shown that MK-801 blocks or reduces hyperalgesia regardless of whether the drug administration is initiated prior to or following the injury. Furthermore, MK-801 effectively relieves hyperalgesia associated with carrageenan-induced inflammation when given as a posttreatment (Ren et al. 1992; Yamamoto et al. 1993). However, in the rat formalin pain model, NMDA antagonists have differential effects depending on whether they are administered pre- or postinjury. MK-801 significantly reduces nociceptive behaviors following formalin injury of a rat's hindpaw if administered prior to, but not following, the early phase response to formalin (Coderre & Melzack 1992; Yamamoto & Yaksh 1992b). Furthermore, the effects of pretreatment with MK-801 or other NMDA antagonists are much more pronounced on the late phase responses to formalin, than on early phase responses, either on behavioral measures of nociception (Coderre & Melzack 1992; Millan & Seguin 1993; Yamamoto & Yaksh 1992b), or on electrophysiological measures of dorsal horn neuronal activity (Haley et al. 1990). We suggest that either pre- or posttreatments are effective in nociceptive tests involving neuropathy or significant peripheral inflammation where there is ongoing afferent input, whereas pretreatments are more effective than posttreatments in the case of low concentration formalin-induced nociception, which depends to a large extent on an initial afferent barrage and central sensitization that occurs at the time of formalin injection. However, as discussed in the previous section this assumption may break down as the concentration of formalin is increased.

5.4. Contribution of ongoing inputs to postoperative pain. It is expected that with cutaneous hyperalgesia, referred pain, and neuropathic pain, postoperative pain is influenced both by central sensitization associated with an injury barrage during surgery as well as by ongoing peripheral inflammatory inputs that contribute to central sensitization after surgery. Thus, the effects of preemptive analgesia are dependent not only on the intensity of the initial barrage, but also on the degree of peripheral inflammation that develops after surgery. Although preemptive effects may be expected to be best demonstrated in cases of extensive surgical trauma in which there is a strong initial central sensitization, these may also be the same cases where a postsurgery treatment will also be effective by alleviating ongoing central sensitization due to ongoing inputs from inflamed peripheral tissue (Katz et al. 1993; Woolf & Chong 1993). It would be interesting to assess whether preemptive treatments would be most effective in minimizing postoperative pain for surgical procedures that result in damage to major nerves, thus producing a larger initial afferent barrage. Furthermore, it has recently (Katz et al. 1995) been demonstrated in a 2-year follow up of patients who had undergone lateral thoracotomy that the incidence of chronic post-thoracotomy chest wall pain was pronounced (>60%), whether patients received preemptive analgesic or postincisional treatments. These results suggest that although preemptive treatments may reduce post-operative pain and analgesic consumption in the immediate postoperative period, these short-term beneficial effects have little or no bearing on the development of chronic post-thoracotomy pain. It may be useful to extend preemptive treatment into the postoperative period to prolong the initial advantage conferred by the preoperative blockade in order to protect against long-term postoperative pain problems. Thus, the use of balanced analgesia (Dahl et al. 1990), with multiple agents and routes of administration to block nociceptive activity in the pre-, intra- and postoperative periods (e.g. Jahangiri et al. 1994), may be more useful from a long-term clinical perspective, than brief, preemptive treatments restricted to the pre- or intraoperative period.

6. Conclusions

We have examined clinical and experimental evidence in four key areas of pain research: cutaneous hyperalgesia, referred pain, neuropathic pain, and postoperative pain. In each there is evidence that persistent pain depends not only on central sensitization, but also on inputs from damaged peripheral tissue. Central sensitization may be comprised of both initial and ongoing components, each driven by variable levels of input from peripheral sources. Each of these factors – initial sensitization, ongoing central sensitiz-

ation, and input from peripheral sources – contributes to the net activity in dorsal horn neurons and thus influences the expression of persistent pain or hyperalgesia. Since each contributes to persistent pain and hyperalgesia in varying degrees across individual patients and diverse pain conditions, it is possible to find evidence for and against the role of central or peripheral neural mechanisms in the persistent pain problems we have discussed. Since each factor can contribute to both the initiation and maintenance of persistent pain, therapies should target both peripheral and central sources of pathology.

ACKNOWLEDGMENT

This work was supported by grants from the Medical Research Council of Canada to both authors.

BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 420–425
Printed in the United States of America

Central inhibitory dysfunctions: Mechanisms and clinical implications

Z. Wiesenfeld-Hallin
Karolinska Institute, Department of Medical Laboratory Sciences and Technology, Section of Clinical Neurophysiology, S-141 86 Huddinge, Sweden
Electronic mail: *zswh@bimd01.hs.sll.se*

H. Aldskogius
Karolinska Institute Biomedical Center, Department of Anatomy, 57173 Uppsala, Sweden

G. Grant
Karolinska Institute, Department of Neuroscience, 17177 Stockholm, Sweden

J.-X. Hao
Karolinska Institute, Department of Medical Laboratory Sciences and Technology, Section of Clinical Neurophysiology, S-141 86 Huddinge, Sweden

T. Hökfelt
Karolinska Institute, Department of Neuroscience, 17177 Stockholm, Sweden

X.-J. Xu
Karolinska Institute, Department of Medical Laboratory Sciences and Technology, Section of Clinical Neurophysiology, S-141 86 Huddinge, Sweden

Abstract: Injury to the central or peripheral nervous system is often associated with persistent pain. After ischemic injury to the spinal cord, rats develop severe mechanical allodynia-like symptoms, expressed as a pain-like response to innocuous stimuli. In its short-lasting phase the allodynia can be relieved with the γ-aminobutyric acid (GABA)-B receptor agonist baclofen, which also reverses the hyperexcitability of dorsal horn interneurons to mechanical stimuli. Furthermore, there is a reduction in GABA immunoreactivity in the dorsal horn of allodynic rats.

Clinical neuropathic pain of peripheral and central origin often cannot be relieved by opiates at doses that do not cause side effects. The loss of sensitivity to opiates may be associated with the up-regulation of endogenous antiopioid substances, such as the neuropeptide cholecystokinin (CCK). CCK and its receptor (CCK-R) protein is normally not detectable in rat dorsal root ganglion cells. After peripheral nerve section, both CCK and CCK-R are up-regulated in the dorsal root ganglia. Furthermore, CI 988, an antagonist of the CCK-B receptor, chronically coadministered with morphine, reduces autotomy, a behavior that may be a sign of neuropathic pain following peripheral nerve section. Thus, opiate insensitivity may be due to the release of CCK from injured primary afferents. Similarly, in the chronic phase of the spinal ischemic model of central pain, the allodynia-like symptom is not relieved by systemic morphine, but is significantly reversed by the CCK-B antagonist. Consequently, up-regulation of CCK and CCK-R in the CNS may also underlie opiate drug insensitivity following CNS injury. Thus, dysfunction of central inhibition involving GABA and endogenous opioids may be a factor underlying the development of sensory abnormalities and/or pain following injury to neural tissue.

Keywords: axotomy; cholecystokinin; dysinhibition; γ-aminobutyric acid (GABA); ischemia; morphine; neuropathy; opioids; pain; spinal cord

1. Introduction

Injury of the peripheral nervous system (PNS) or central nervous system (CNS) is often associated with pain. Dysfunction of central inhibition involving the inhibitory neurotransmitters γ-aminobutyric acid (GABA) and endogenous opioids, as well as the antiopioid neuropeptide cholecystokinin (CCK), may be a factor in the development of sensory abnormalities and/or pain following injury to neural tissue. In this review the possible common dysfunction of these inhibitory systems in the generation of neuropathic pain as a consequence of PNS and CNS tissue injury will be discussed.

0140-525X/97 $9.00+.10

2. Models of neuropathic pain following PNS and CNS injury

2.1. Experimental neuropathic pain following peripheral nerve injury. After peripheral nerve section (axotomy), a neuroma develops when the proximal nerve stump has no possibility of regenerating and reinnervating the periphery. Wall and Gutnick (1974) were the first to demonstrate that an ongoing abnormal discharge originating in the neuroma develops rapidly after peripheral nerve section. The neuroma is also sensitive to mechanical pressure and stimulation by adrenergic agonists (Wall & Gutnick 1974). Furthermore, the deafferented dorsal root ganglion (DRG) cells also become generators of ongoing discharges (Wall & Devor 1983). Abnormal discharges can also be recorded from spinal cord interneurons following deafferentation by dorsal rhizotomy (Lombard & Larabi 1983). Thus, there are functional changes at various levels of the nervous system following peripheral nerve injury that may underlie the development of neuropathic pain.

A model of experimental neuropathic pain following injury to primary afferents is autotomy behavior (self-mutilation of the deafferented body region) after section of major peripheral nerves. This behavior was proposed by Wall et al. (1979) to be a response to unpleasant and perhaps painful paresthesias referred to the denervated limb. The development of autotomy has also been observed after multiple dorsal rhizotomy (Albe-Fessard et al. 1979), after destruction of DRG (Wiesenfeld-Hallin et al. 1987), and after spinal nerve lesions (Lombard et al. 1979; Wiesenfeld & Lindblom 1980). An advantage of the axotomy model of neuropathic pain over other models of partial nerve injury that have more recently become available (e.g., Bennett & Xie 1988; Seltzer et al. 1990) is its robustness and reproducibility in a large number of laboratories. However, autotomy does not represent all forms of neuropathic pain of peripheral origin, but only deafferentation pain corresponding to phantom limb pain and anesthesia dolorosa. Pain due to partial nerve injury has other characterstics, including hyperalgesia to thermal and mechanical stimuli (see Bennett, 1994, for review).

2.2. A model of central pain following spinal cord ischemia. Central pain is caused by a lesion or dysfunction of the CNS (Merskey 1986). Application of excitatory agents on the dorsal horn or the trigeminal nucleus can induce hyperalgesic and allodynic responses in animals (Beyer et al. 1985; Black 1974; Dyken 1965; Yaksh 1989). These models do not, however, involve a lesion to the CNS and because of the short duration of the pain-like symptoms they do not represent a clinical model of central pain. We have studied the effects of a photochemically induced spinal ischemia, which appears to be a useful model of central pain with a number of features that are similar to clinical pain. (Hao et al. 1991a; 1991b; 1992b; 1992c; 1992d; 1992e; Xu et al. 1992a; see Wiesenfeld-Hallin et al., 1994, for review). The major sensory symptom is the development of mechanical allodynia, a pain-like response to innocuous stimuli, which has a short-lasting (days) and chronic (permanent) phase. The response to thermal stimuli is unaffected, indicating that the allodynia-like symptom is mediated by low threshold mechanoreceptors, rather than heat nociceptors. The neuronal dysfunction related to the development of allodynia probably involves excitotoxicity through the activation of the N-methyl-D-aspartate (NMDA) receptor by glutamate, since pretreatment with the NMDA-receptor antagonist MK-801 prevented the development of this sensory dysfunction (Hao et al. 1992a). There is also considerable evidence that the short-lasting allodynia is mediated by a loss of spinal GABA-ergic inhibition, whereas the chronic allodynia is due to a dysfunction of the endogenous opioid system (see below).

3. The role of GABA-ergic mechanisms in pain following injury to the PNS and CNS

3.1. Dysfunction of GABA-ergic inhibition after peripheral nerve injury. Immunohistochemical studies, using antisera to glutamate decarboxylase (GAD) and conjugates of GABA, have established that GABA-like immunoreactive (GABA-LI) neurons are present in the mammalian dorsal horn of the spinal cord (see Todd & Spike, 1993, for review). The GABA-LI cells are particularly concentrated in Rexed's laminae I–III. It is well documented that GABA-ergic interneurons in the dorsal spinal cord regulate the activity of many types of sensory afferents, predominantly through presynaptic inhibition on input from both A- and C-primary afferents, although postsynaptic actions of GABA agonists have also been reported (see Willis & Coggeshall, 1991, for review). Primary afferent depolarization (PAD) is an important mechanism that reduces transmitter release from the terminals of primary afferents into the spinal cord through presynaptic inhibition (Eccles et al. 1963). PAD is believed to involve spinal GABA inhibitory interneurons since nonspecific GABA antagonists reduce PAD (Banna & Jabbur 1969). Axotomy of peripheral nerve is associated with decreased capacity of the spinal cord to generate PAD (Wall & Devor 1981), which may be associated with the possible dysfunction of GABA-ergic interneurons following peripheral nerve injury (Castro-Lopes et al. 1993). Thus, after nerve section, loss of GABA-ergic inhibition may be one of the factors leading to the development of neuropathic pain. This was supported by the finding that the $GABA_B$ receptor agonist baclofen reduced the persistent expression of the immediate early gene c-fos in axotomized rats (Basbaum et al. 1991). Furthermore, degenerated ("dark") neurons have been found in spinal cord laminae I–II in rats with chronic constriction injury of the sciatic nerve. This effect was accentuated by the repeated administration of strychnine (Sugimoto et al. 1990). It is possible that these "dark neurons" are GABA-ergic and may underlie loss of endogenous inhibition following peripheral nerve injury since the $GABA_B$ receptor agonist baclofen has been found to cause antinociception in this model (Smith et al. 1994).

3.2. Dysfunction of the GABA-ergic system after spinal cord ischemia. The pharmacological basis of short-lasting behavioral allodynia has been investigated (Hao et al. 1991b; 1992e; Xu et al. 1992b). A large number of analgesics applied systemically, including morphine at nonsedative doses, were ineffective, but the $GABA_B$ receptor agonist baclofen dramatically reduced both tactile-evoked agitation and increased vocalization threshold to mechanical pressure in a dose-related manner. However, the $GABA_A$ agonist muscimol was ineffective.

The effect of baclofen on the response of single dorsal horn wide dynamic range (WDR) neurons was tested in normal and allodynic rats (Hao et al. 1992c; 1992d). In normal rats, subcutaneous electrical stimulation of the

receptive field or the sciatic nerve at an intensity that activated both myelinated (A) and unmyelinated (C) fibers evoked a biphasic response in all WDR neurons with a short-latency A-fiber response and a long-latency C-fiber response. In the majority of WDR neurons recorded in allodynic rats, electrical stimulation evoked a single burst with no separation between responses to A- and C-fiber input. Detailed analysis of the poststimulus histogram showed that the number of discharges evoked in WDR neurons in response to electrical stimulation in allodynic rats was significantly higher than in normal rats at all poststimulus intervals. The results also suggested that the myelinated A-fiber input was prolonged and enhanced during allodynia. The response of WDR neurons in normal rats to pressure by von Frey hairs increased linearly. Nearly all WDR neurons recorded in allodynic rats exhibited increased sensitivity to mechanical stimulation. The pressure–response curve was exponential and dramatically shifted to the left, with a significant lowering of the threshold for evoking neuronal responses. In contrast, the response of WDR neurons to heat stimulation was similar in allodynic and normal rats, indicating a lack of involvement of C-afferent input to the spinal cord in the geneneration of abnormal responses and confirming the results of behavioral tests.

Pretreatment with baclofen in normal rats did not alter the response pattern of WDR neurons to electrical stimulation. However, baclofen significantly depressed the responses of WDR neurons to intense, but not to innocuous, mechanical stimuli. The mechanical threshold of WDR neurons in normal animals pretreated with baclofen did not differ from that of rats without any drug treatment. In allodynic rats, pretreatment with baclofen normalized the response pattern of the neurons to electrical stimulation. The hypersensitivity of the WDR neurons to low-intensity mechanical stimulation in allodynic animals was totally reversed by baclofen and the depression of the response to intense stimuli was the same as in normal rats.

From the results of behavioral, physiological, and pharmacological studies, it is apparent that allodynia following spinal cord ischemia is mediated predominantly by abnormal input from myelinated afferents, as both behavioral and electrophysiological studies demonstrated the presence of mechanical hypersensitivity. In contrast, the behavioral response and the physiological response of dorsal horn WDR neurons to noxious heat stimulation was unchanged. Systemic low-dose baclofen relieved both behavioral allodynia and neuronal hypersensitivity, suggesting that the neuronal hyperexcitability underlying short-term behavioral allodynia is induced by loss of GABA-ergic inhibitory control (Game & Lodge 1975; Price et al. 1987), which may result from a high susceptibility of GABA-ergic neurons to EAA-mediated neurotoxicity (Sloper et al. 1986). These results thus indicate that inihibition of myelinated, low-threshold afferents is mediated through $GABA_B$ receptors and operates tonically under normal conditions (Fig. 1).

The role of the GABA-ergic system in short-lasting allodynia following spinal ischemia was recently examined with immunohistological methods (Zhang et al. 1994). The number of GABA-like immunoreactive cells in laminae I–III of the lumbar dorsal horn was significantly decreased bilaterally during the presence of allodynia compared to cervical levels and sham-operated controls. The number of

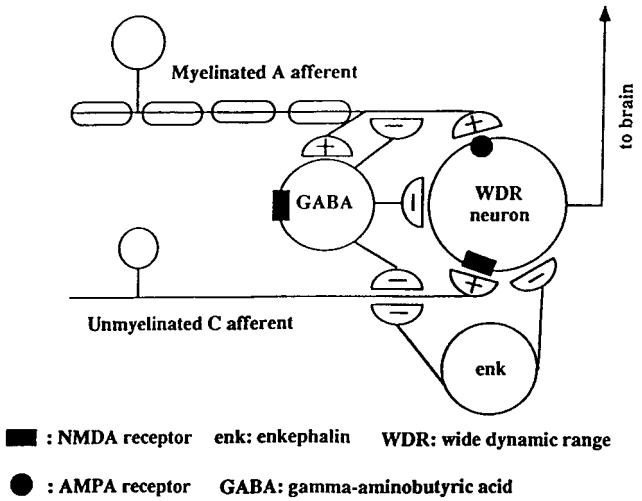


Figure 1. Illustration of the mechanisms underlying short-term allodynia. The normal organization of the inhibitory and excitatory mechanisms of the dorsal horn of the spinal cord that influence the response of wide dynamic range cells is shown. For the sake of simplicity, descending systems have been omitted. After reversible spinal cord ischemia, the GABA-ergic interneurons become dysfunctional because of the excitotoxic effect of glutamate acting on NMDA receptors. The presynaptic inhibitory function of GABA-ergic interneurons on low-threshold myelinated afferents is disrupted, leading to short-term mechanical allodynia. Hyperalgesia to thermal stimuli, mediated by unmyelinated afferents, does not occur because the endogenous opioid system is not disturbed by the ischemia that leads to short-term allodynia. Reproduced from Wiesenfeld-Hallin et al., 1994, with permission.

GABA-immunoreactive cells was restored after recovery from allodynia, indicating a strong correlation between the spinal level of GABA and the presence of pain-like response in corresponding dermatomes. The recovery of GABA-immunoreactivity indicated that the dysfunction of GABA-ergic interneurons was temporary and recovery of function is possible. The reduction of GABA immunoreactivity in the spinal cord following ischemia is paralleled by reduced spinal GABA levels following ischemia as measured with biochemical technique (Martiniak et al. 1991), indicating that the immunohistochemical results may actually reflect reduced amino acid levels.

4. The pharmacological basis of opiate insensitivity following injury to the PNS and CNS

4.1. Opiate insensitive forms of pain in the clinic. In the clinic it is desirable that a thorough examination and analysis of the patient's pain preceeds treatment. This may include quantitative sensory testing of peripheral nerve function (Lindblom 1985; Lindblom & Hansson 1991), the morphine test (Arnér 1991a), the phentolamine test (Arnér 1991b), and pain drawings (Schwartz & DeGood 1984). Consequently, nociceptive, neurogenic, idiopathic, and psychogenic forms of pain can often be identified and treatment can be adjusted to suit the patient (Arnér 1991a). For example, patients suffering from acute nociceptive and inflammatory pain after surgery are usually successfully treated with traditional drugs, including opiates. But pain

originating from damage to peripheral nerve or the CNS is often long-lasting or chronic and presents a severe problem in the pain clinic – and must be treated through alternative approaches. Thus, low efficacy of opiates in treating pain after infiltration of peripheral nerve or nerve plexus by cancerous tumors has been reported (Arnér & Arnér 1985). Other examples of opiate-insensitive pain are phantom limb pain following amputation (Sherman et al. 1980) and pain after spinal injury with deafferentation (Glynn et al. 1986). Although many different treatments have been suggested to reduce neurogenic pain, once established it is still a serious and difficult clinical problem (Arnér 1991a).

The question of opiate insensitive forms of pain has been intensively debated in the clinical literature; some authors maintain that opiate treatment is useful if very high doses are used (McQuay 1988; Portenoy et al. 1990). Arnér and Meyerson (1988; 1991) have proposed that opiates have less efficacy in neuropathic than nociceptive pain. Thus, the presence or absence of opiate insensitivity needs to be established for each patient. For patients unresponsive to analgesic doses of opiates that do not cause undesirable side effects, alternate methods of treatment must be considered. In milder forms of neurogenic pain transcutaneous nerve stimulation and antidepressant drugs are important therapeutic tools (Meyerson 1990). In the pain clinic, intraspinal (epidural or intrathecal) administration of α_2 adrenoceptor agonists like clonidine or local anaesthetics have been used as complements to morphine (Glynn et al. 1993).

4.2. Experimental evidence for opiate insensitivity. Although opiate insensitivity is an important clinical problem, there are relatively few experimental studies examining this issue. Furthermore, the results of these studies have been inconclusive. A high dose of morphine (240 μg/day) applied intrathecally (i.t.) continuously, starting at the time of axotomy and continuing for 14 days, was associated with decreased level of autotomy in rats (Wiesenfeld-Hallin 1984). These results speak against the presence of opiate insensitivity. However, in this study the positive effect of morphine on autotomy may have been due to the preemptive analgesic effect of the drug (Puke & Wiesenfeld-Hallin 1993), since morphine was injected in conjunction with the nerve injury. Indirect evidence for reduced sensitivity to morphine in neuropathic pain is the observation that after deafferentation a reduction of μ, δ, and κ opioid receptors have been reported in laminae I–III of the dorsal horn (Gouardères et al. 1991). The reduction of opioid receptors seems to be correlated to the degree of deafferentation: rhizotomy > sciatic nerve section > constriction nerve injury (Besse at al. 1992). There are data from rats with nerve constriction injury indicating that pain-related behavior in this model can be effectively relieved with opiates (Attal et al. 1991; Jazat & Guilbaud 1991). However, the down-regulation of opioid receptors after peripheral nerve section seems to be temporary (Besse at al. 1992), suggesting that reduced opiate sensitivity after nerve section may only partly depend on down-regulation of opioid receptors.

In support of the reduction of the analgesic effect of morphine following nerve section, the threshold dose of morphine required to depress the flexor reflex was 3 to 5 times higher in autotomizing rats following sciatic nerve section than in rats with intact nerves or axotomized rats that were not autotomizing (Xu & Wiesenfeld-Hallin 1991). However, higher doses of morphine were equally effective in depressing the flexor reflex in all groups. These results indicate that the sensitivity to, rather than the effectiveness, of morphine-induced antinociception was reduced following axotomy. Evidence will be presented that this phenomenon may be a result of an intrinsic antiopioid mechanism involving the neuropeptide CCK.

4.3. The possible role of cholecystokinin in opiate insensitivity following PNS injury. CCK is normally present in the brain as the sulphated C-terminal fragment CCK-8 and fulfills many of the criteria for a neurotransmitter with multiple functions (Vanderhaegen & Crawley 1985). CCK has been documented to have an opiate antagonistic property (Baber et al. 1989; Dourish et al. 1990; Faris et al. 1983; Itoh et al. 1982; Stanfa et al. 1994; Watkins et al. 1984; Wiesenfeld-Hallin & Duranti 1987; Wiesenfeld-Hallin et al. 1990). CCK receptors have been found to be heterogeneous; in rodents the CCK-A receptor has been found primarily in peripheral tissue, and the CCK-B receptor has been found in the CNS (Moran et al. 1986). Potent and highly selective antagonists of the CCK-B receptor have made it possible to investigate the role of CCK in the CNS. Thus, CCK antagonists have been found to enhance the analgesic effect of morphine and to reduce the development of morphine tolerance (Baber et al. 1989; Dourish et al. 1990; Xu et al. 1994a). We have provided evidence that plasticity of the CCK system may be involved in the reduced effect of morphine in the autotomy model of experimental neuropathic pain.

We examined the effect of sciatic nerve section on the presence of CCK mRNA in rat DRG cells. Furthermore, the effect of CI 988, an antagonist of the CCK-B receptor (Hughes et al. 1990), which was previously shown to potentiate the analgesic effect of morphine (Wiesenfeld-Hallin et al. 1990) on autotomy behavior was also examined (Xu et al. 1993). After sciatic nerve section, up-regulation of CCK mRNA in the ipsilateral L4 and L5 DRG was observed with *in situ* hybridization technique. Only a few CCK mRNA positive cells were seen in control ganglia and in ganglia contralateral to the nerve transection. Fourteen days after axotomy, up to 30% of all DRG cells on the nerve-sectioned side were CCK mRNA positive. Most of these cells were of the small type. CCK receptor protein also became expressed on DRG cells of all sizes (see Hökfelt et al., 1994, for review).

During an observation period of 15 days, the occurrence, development, and severity of autotomy were examined in five groups of rats injected with i.t. saline, s.c. (subcutaneous) saline, s.c CI 988, s.c. saline followed by i.t morphine, or s.c. CI 988 followed by i.t. morphine. The drugs were administered twice daily, starting 24 hours after axotomy. The rats injected with CI 988 and morphine autotomized significantly less than the other four groups during the 15-day observation period. The results of this behavioral study therefore suggested that coadministration of CI 988 and morphine produced significant suppression of autotomy behavior after peripheral nerve section.

As discussed above, the decreased sensitivity of the spinal cord to morphine is probably not just a result of a reduction of μ-receptors in the spinal cord. Therefore down-regulation of opioid receptors does not seem to be an important factor in decreased sensitivity to morphine after

peripheral nerve section. Development of morphine tolerance could also decrease the effectiveness of morphine analgesia. Tolerance to morphine could be expected if an initially decreased autotomy was followed by increased autotomy. No such effect was observed. Thus, there may be other mechanisms involved in the lack of effect of morphine after nerve section.

CCK normally exists in spinal cord interneurons and descending pathways, but rarely in primary afferents, in rat (Hökfelt et al. 1988; Ju et al. 1986; Williams et al. 1987; Zhang et al. 1993). Verge et al. (1993) have recently observed that mRNA for CCK is up-regulated in rat DRG after axotomy and even CCK-B receptor mRNA is increased in rat DRG after axotomy (Zhang et al. 1993). The results of a parallel behavioral study (Xu et al. 1993) indicated that the combination of morphine and CI 988 significantly suppressed autotomy behavior after sciatic nerve section. The tendency to reduced rate of autotomy in the group that received only CI 988 may be the result of increased efficacy of the intrinsic opioid systems during treatment with the CCK-B antagonist, indicating that a tonic CCK-ergic control of the endogenous antinociceptive system may be present following peripheral nerve injury. This is supported by recent physiological data indicating that, like in normal animals, CI 988 potentates the antinociceptive effect of morphine in axotomized rats (Xu et al. 1994b).

These results suggest that after peripheral nerve section there is a markedly increased production of CCK and CCK-receptor protein in the DRG. The altered expression of CCK is in line with other reports showing complex changes in the levels of neuropeptides and their receptors in primary sensory neurons and spinal cord after peripheral nerve section (Hökfelt et al. 1994). The increased expression of CCK in DRG after nerve section and the observation that CI 988 in combination with morphine can reduce autotomy behavior strongly suggest that the ineffectiveness of morphine in this experimental neuropathic pain model is related to the CCK system. Thus, after peripheral nerve injury it is possible that up-regulation of CCK synthesis and an increased release of CCK from primary afferents in the spinal cord antagonizes the effect of exogenously administered or endogenously released opioids. CCK may be also involved in the development of tolerance to morphine, because chronic coadministration of CI 988 with morphine prevented the development of tolerance to the analgesic effect of the opiate (Xu et al. 1992a) and CI 988 administered to tolerant rats reinstated the analgesic effect of morphine (Hoffmann & Wiesenfeld-Hallin 1994).

4.4. The possible role of the endogenous opioid system and CCK in central pain. In addition to short-lasting allodynia as described above, we have observed a chronic pain-related syndrome in rats after spinal cord ischemia (Xu et al. 1992b). This syndrome only developed in a subpopulation of rats after severe ischemia and with an extensive spinal cord lesion. The main symptom of this chronic pain-related syndrome is mechanical allodynia, which is more severe than during tonic allodynia. Similar symptoms have been described clinically in patients after spinal cord injury. We have also observed autotomy of the hindpaws in some allodynic animals, which may indicate the presence of phantom pain. In accordance with clinical experience, the chronic allodynia-like symptoms developed with a delay of several days to 1-1/2 months after the initial injury and persisted without signs of remission. Also in accordance with clinical experience, the chronic allodynia-like symptom was not responsive to most pharmacological treatments, including systemic morphine, clonidine, carbamazepine, pentobarbital, baclofen, muscimol, and diazepam. Since baclofen effectively relieved short-lasting but not chronic allodynia, the mechanisms underlying this pain-like state have presumably undergone a plastic change during transition from tonicity to chronicity.

Interestingly, in rats with spinal cord lesions but without chronic allodynia, systemic naloxone consistently provoked allodynia with characteristics very similar to the symptoms observed in allodynic rats (Xu et al. 1994b). Thus, the endogenous opioid system may be tonically active in these animals and may suppress the expression of allodynia. However, naloxone did not influence the tail flick latency in these rats, suggesting that the action of the activated endogenous opioid system is selective, that is, it did not raise the nociceptive threshold in general. Consequently, the endogenous opioid system may be involved in the suppression of abnormal pain-related sensations after spinal cord lesion, and disruption of such control may lead to the emergence of chronic allodynia. The disruption of this control may involve the CCK system since the CCK-B antagonist CI 988, but not the CCK-A antagonist CAM 1481, increased the vocalization threshold in chronically allodynic rats (Xu et al. 1994b). This effect of CI 988 was not due to general antinociception, as it had no effect on vocalization threshold to mechanical pressure in normal rats and its effect on the tail flick latency was slight and only observed at a high dose. CI 988 did not produce sedation at the doses tested and since diazepam failed to relieve allodynia, the anxiolytic property of CI 988 (Hughes et al. 1990) cannot be responsible for the observed effects. Therefore, the effect of CI988 on the vocalization threshold of spinally injured rats reflects an analgesic effect against chronic allodynia. The effect of CI 988 was reversed by naloxone, suggesting that it may be mediated through an opioidergic link. Thus, the analgesic effect of CI 988 on chronic allodynia may reflect an ongoing antagonism by the CCK system on the endogenous opioid system, which could suppress the exhibition of allodynia, as in nonallodynic spinally injured rats (Xu et al. 1994b).

Based on these results, we can propose a mechanism for the emergence of chronic allodynia in rats after spinal cord injury. Injury to the spinal cord may interrupt normal transmission and integration of sensory information. These abnormal sensory inputs may activate the endogenous opioid system, resulting in enhanced inhibitory control and suppression of the development of allodynia-like symptoms. Such enhanced opioidergic control could, however, eventually be interrupted in some rats by an up-regulated endogenous CCK system, leading to the development of chronic allodynia. Indeed, a recent study has demonstrated that, when the activity of endogenous opioid system was enhanced by blocking the degradation of endogenous opioids, a subsequent increase in activation of CCK-B receptors was observed (Ruiz-Gayo et al. 1992). The involvement of the endogenous CCK system in the development of chronic allodynia may explain why morphine had little effect upon central pain, because endogenous CCK may antagonize exogenously applied opiates as well. We are

currently investigating where in the CNS changes in the endogenous opioid and CCK systems take place that may underlie opioid insensitivity in chronic allodynia.

5. Conclusions

The expression of experimental neuropathic pain after peripheral nerve axotomy and spinal cord injury may involve dysfunction of the GABAergic system and an altered interaction between endogenous opioids and CCK. Increased understanding of the mechanisms underlying plasticity of neurotransmitter systems and their function following PNS and CNS injury may lead to improved treatment strategies for neuropathic pain.

ACKNOWLEDGMENTS

The present work was supported by the Swedish Medical Research Council (grant no. 07913), Astra Pain Control AB, the Bank of Sweden Tercentenary Foundation, and research funds of the Karolinska Institute.

BEHAVIORAL AND BRAIN SCIENCES (1997) **20,** 426–434
Printed in the United States of America

Sympathetic nervous system and pain: A clinical reappraisal

Helmut Blumberg
Department of Neurosurgery, University of Freiburg, 79106 Freiburg, Germany
Electronic mail: *boeger@nz11.ukl.uni-freiburg.de*

Ulrike Hoffmann
Department of Anaesthesiology, University of Freiburg, 79106 Freiburg, Germany

Mohsen Mohadjer
Department of Neurosurgery, University of Freiburg, 79106 Freiburg, Germany

Rudolf Scheremet
Department of Neurosurgery, University of Freiburg, 79106 Freiburg, Germany

Abstract: The target article discusses various aspects of the relationship between the sympathetic system and pain. To this end, the patients under study are divided into three groups. In the first group, called "reflex sympathetic dystrophy" (RSD), the syndrome can be characterized by a triad of autonomic, motor, and sensory symptoms, which occur in a distally generalized distribution. The pain is typically felt deeply and diffusely, has an orthostatic component, and is suppressed by the ischemia test. Under those circumstances, the pain is likely to respond to sympatholytic interventions. In a second group, called "sympathetically maintained pain" (SMP) syndrome, the principal symptoms are spontaneous pain, which is felt superficially and has no orthostatic component, and allodynia. These symptoms, typically confined to the zone of a lesioned nerve, may also be relieved by sympathetic blocks. Since the characteristics of the pain differ between RSD and SMP, the underlying kind of sympathetic–sensory coupling may also vary between these cases. A very small third group of patients exhibits symptoms of both RSD and SMP. The dependence or independence of pain on sympathetic function reported in most published studies seems to be questionable because the degree of technical success of the block remains uncertain. Therefore, pain should not be reported as sympathetic function independent until the criteria for a complete sympathetic block have been established and satisfied.

Keywords: causalgia; pain; quality control; reflex sympathetic dystrophy; skin temperature; sympathetic block; sympathetic nervous system; sympathetic reflexes; sympathetic–sensory coupling; sympathetically maintained pain

1. Introduction

In the past, different terms have been used for clinical conditions in which pain was likely to be maintained by ongoing activity of the sympathetic nervous system, for example, causalgia, algodystrophy, Sudeck's atrophy, and so forth (see Bonica 1990). Such disorders were later subsumed under the global term "reflex sympathetic dystrophy" (RSD), which was first introduced by Evans (1946). Later, the term "sympathetically maintained pain" (SMP) was introduced by Roberts (1986) and subsequently used by others (Campbell et al. 1992; Frost et al. 1988; Meyer et al. 1992). This, however, did not end the discussion of terminology. Recently, "complex regional pain syndrome" (CRPS) was suggested as a new term for RSD/SMP and causalgia (Merskey & Bogduk 1994).

Historically, the idea of blocking sympathetic activity to treat these patients was mainly based on the observation of cold skin. This observation was related to sympathetic hyperactivity, which was thought to excite nociceptive fibres within the symptomatic area (Leriche 1923; 1949). At the same time, this sympathetic–sensory coupling was believed to be part of a vicious circle, where the abnormal afferent activity – via sensitized central nervous (spinal cord) structures – could maintain sympathetic hyperactivity and consequently pain (Livingston 1943/1976). This model has been used to this date to explain the relationship between the sympathetic system and pain in RSD/SMP. For each main aspect of this hypothesis there is supportive evidence, obtained by experiments on animals and by experimental investigations on humans (Bennett & Xie 1988; Bond et al. 1991; Devor et al. 1991; 1994; Jänig 1990; Jänig & Koltzenburg 1991; Jänig & Schmidt 1992; Levine et al. 1985; Wall & Devor 1983; Willis 1992; Woolf et al. 1992).

With this coupling as a cause of pain, however, there are certain observations that remain a matter of discussion. First, the clinical picture of patients who do respond to sympathetic blocks may differ among each other quite considerably and form different groups of patients (Blumberg & Jänig 1993). Second, a number of patients do not

0140-525X/97 $9.00+.10

show signs of sympathetic hyperactivity. To the contrary, they may even exhibit warm and dry skin, but nevertheless get pain relief upon sympathetic blocks (Blumberg et al. 1994; de Takats 1943). Third, in a number of patients with cold extremities the pain may not be relieved by sympathetic blocks (Drummond et al. 1991). Finally and most importantly, there is no study that defines the criteria for a complete interruption of the sympathetic innervation of a blocked extremity. Consequently, there is a problem regarding the quality control of sympathetic blocks, which also makes it difficult to judge the outcome of sympatholytic strategies in RSD/SMP as far as pain mechanisms are concerned.

2. Pain and the sympathetic system – a clinical reappraisal

As the above observations suggest, there are good reasons for a clinical reappraisal of the relationship between the sympathetic nervous system and pain. In order to do this and to describe the patients in whom a sympathetic component of pain may be assumed, we divide these patients into three groups.

The patients of the first group are characterized by the occurrence of a rather complex clinical symptomatology, in whom the distal part of the extremity is affected predominantly, where the symptoms – independently of the kind and location of a preceding lesion – appear in a glove- or sock-like distribution manner. In accordance with the literature, these patients are diagnosed as having RSD (CRPS type I, Merskey & Bogduk 1994). The relatively rare patients of the second group are diagnosed as having SMP (Blumberg 1992; Blumberg & Jänig 1993). This group is characterized by the occurrence of two major symptoms, spontaneous pain and allodynia, which typically are found within the zone of a lesioned nerve. To illustrate the clinical picture of these two patient types, representative case reports will be given first, to be followed by discussions of the related clinical aspects of pains and the sympathetic system in these conditions. In addition, there is a very small group of patients who show signs of both RSD and SMP. This condition, labelled causalgia, will be discussed briefly. Finally, we discuss the possible kinds of sympathetic–sensory coupling in these conditions.

3. Reflex sympathetic dystrophy

3.1. Case report 1

On the September 7, 1993, a 72-year-old working man underwent surgical intervention because of Dupytren's contracture, involving the long flexor tendon of the left fourth finger. The day after surgery he complained about strong pain, which he felt localized under the palmar wound suture. As the pain persisted and signs of a deep haematoma developed at that site, the suture was reopened three days after the initial intervention to treat the pain and the haematoma. At that time, there was no edema outside the treated area, no signs of inflammation, and the fingers, except the treated one, could be moved freely.

In the night following this secondary intervention, the patient suddenly experienced a drastic change of symptoms. The entire hand became swollen and warmer compared with the other hand, together with deep and severe pain felt diffusely inside the hand and also affecting most finger joints. The fingers and the wrist could no longer be moved actively, and he developed tremor of the hand and fingers. He got some pain relief by elevating the affected arm and by applying ice to the hand.

There were no signs of local or systemic inflammation, and oral analgetics and physical therapy were applied for treatment. These, however, did not significantly ameliorate the symptoms. Later, intermittent exaggeration of the pain developed and brachial plexus anaesthesia was introduced. This stopped the pain for the duration of the anaesthesia and the swelling diminished. However, after the anaesthesia wore off, the symptoms and swelling returned. Finally, RSD was suspected and the patient was sent to us, 4 weeks after the first surgical intervention.

Upon neurological investigation, the following findings were observed.

3.1.1. Autonomic symptoms. The entire left hand was severely swollen (Fig. 1A/1B), and its volume was about 50% increased compared with the healthy hand (Fig. 2). The entire left hand was found to be warmer than the right one (see also sect. 3.1.4), and the patient reported that this hand did not cool as much as the right one when exposed to cold (see sect. 3.1.5). The palmar side showed dry skin on touch (hypohydrosis). Signs of trophic disturbances (e.g., increased hair or disturbed nail growth) were absent.

3.1.2. Motor symptoms. Active movements of the fingers and hand joints were severely impaired, including flexion (Fig. 1A) and extension (Fig. 1B) functions. Muscular strength was absent (plegia) and there was strong postural tremor of the left hand and fingers.

3.1.3. Sensory symptoms. The patient complained about severe and diffuse pain, which he felt deep inside the entire hand, and which increased upon hanging down the hand and was somewhat alleviated when elevating the arm above the heart level (orthostatic component of the pain). Touch sensation of the left hand was impaired (hypoaesthesia) in a glove-like distribution, as was sensation of pin prick (hypoalgesia). Allodynia or hyperpathia (as defined by Merskey & Bogduk 1994) were not found.

3.1.4. Measurement of finger tip temperatures. Left/right hand: 35.4/34.7; 34.8/33.8; 35.4/33.8; 35.7/34.0; 35.7/34.0°C. Each finger of the affected left side showed higher skin temperature (SkT) values than the corresponding spot on the healthy side, which gives rise to a so-called systematic side difference of SkT (Blumberg 1991).

3.1.5. Further diagnostic procedures. During ischemia of the left hand that was maintained for a couple of minutes (see sect. 3.3.1.), pain was absent. Conventional three-phase bone scanning in all phases showed diffuse increase of radionuclide uptake in the affected hand (not illustrated).

Since the patient did show abnormal behaviour of SkT under environmental thermal load, reflexive behaviour of SkT was investigated experimentally under whole body warming and cooling. Consistent with the case history, diminished skin vasoconstriction upon cooling was found in the sick hand compared with the healthy hand (Fig. 3A). To test a possible sympathetic contribution to the pain, an

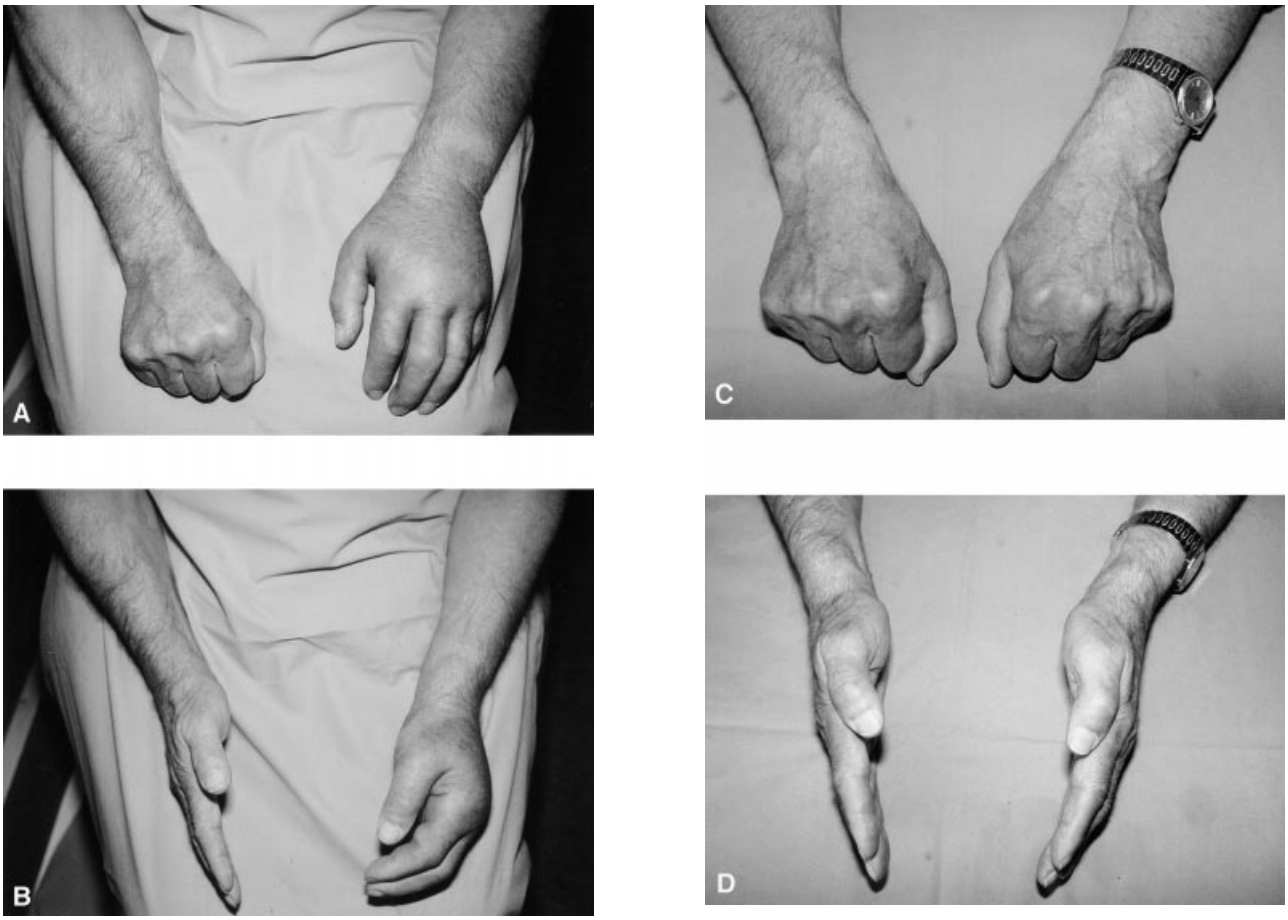


Figure 1 (A–D). Patient with RSD of the left hand following surgical intervention of Dupytren's contracture (see case report 1) before (A/B) and after therapy (C/D). Note the severe edema of the left hand, accompanied by the inability to close the fist (A) and to extend the fingers (B), as well as relief of these symptoms 3 months after therapy (C/D).

intravenous regional guanethidine blockade (Hannington-Kiff 1993) with 2.5 mg of the drug was applied, using some modifications of this technique (Hoffmann & Blumberg 1994). This yielded pain relief for more than one day, which was accompanied by reduction of the edema (Fig. 2).

3.1.6. Treatment and follow up. Since pain and swelling returned, similar blocks were used several times, as indicated in Figure 2, after which the patient finally became free of pain and swelling along with an improvement of all the other symptoms. At the end of treatment, about four weeks later, he went back to work. At 3-month (Fig. 1C/1D) and 9-month follow up, swelling was absent, motor functions were normal, and there were no symptoms of RSD; SkT was regular under resting conditions (lack of systematic side difference of SkT) and under thermal load (Fig. 3B).

3.2. On the description and incidence of RSD symptoms

Currently, it is uncertain which clinical approach should be applied in order to get a systematic description of the relevant symptoms of RSD. According to our approach, which is in agreement with the guidelines of the new classification of chronic pain (Merskey & Bogduk 1994) regarding RSD, the changes associated with RSD occur distally in the affected extremity, showing a triad of autonomic, motor, and sensory symptoms (Blumberg 1991; Blumberg & Jänig 1993; Kurvers et al. 1995). On the basis of our observations (Blumberg & Jänig 1993), the type and incidence of major symptoms in 121 RSD cases with an affected upper extremity are indicated in Table 1, column A.

Other studies list the symptoms of RSD without putting them into certain categories (for reference see Schwartzmann & McLellan 1987), or relate them to categories like inflammatory, neurological, and sympathetic (as done by Veldman et al. 1993). By applying the scheme of the RSD triad to the latter study of 829 cases with RSD (Table 1, column B), and comparing the incidence of symptoms with that of our study, a similar incidence is found for most symptoms, especially for the motor symptoms (Table 1). Interestingly, allodynia, which has a low incidence in our study, was not even mentioned as a symptom of RSD in the study by Veldman et al. (1993), but had a high incidence in other studies (Bonelli et al. 1983; Drummond et al. 1991; Kurvers et al. 1995; Price et al. 1992). The reasons for these differences remain uncertain.

The following discussion of RSD will concentrate on the various aspects of pain and the sympathetic system in this condition.

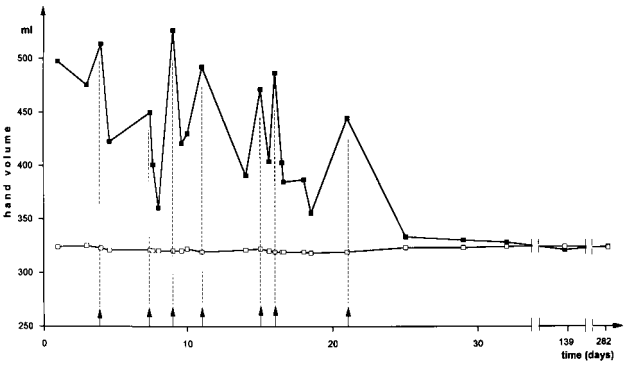


Figure 2. Effect of sympatholytic therapy on the edema in RSD (case report 1). The hand volume (edema) was calculated by measuring the water volume (ml) that was replaced from a filled vessel after each hand was submerged to a certain anatomic level (just proximal to the hand joint). Days of sympatholytic interventions (guanethidine blocks, using 2.5 mg of the drug; see Hoffmann & Blumberg 1994) are indicated by black arrows, connected with dotted lines. Note that each sympatholytic intervention was followed by a decrease of the edema, which returned several times, however, but finally subsided. The edema did not return during the entire period of follow up (9 months). Compare with Figure 1A/1B before and Figure 1C/1D after therapy, same patient. Filled squares: volumes of the left (sick) hand. Empty squares: volumes of the healthy hand.

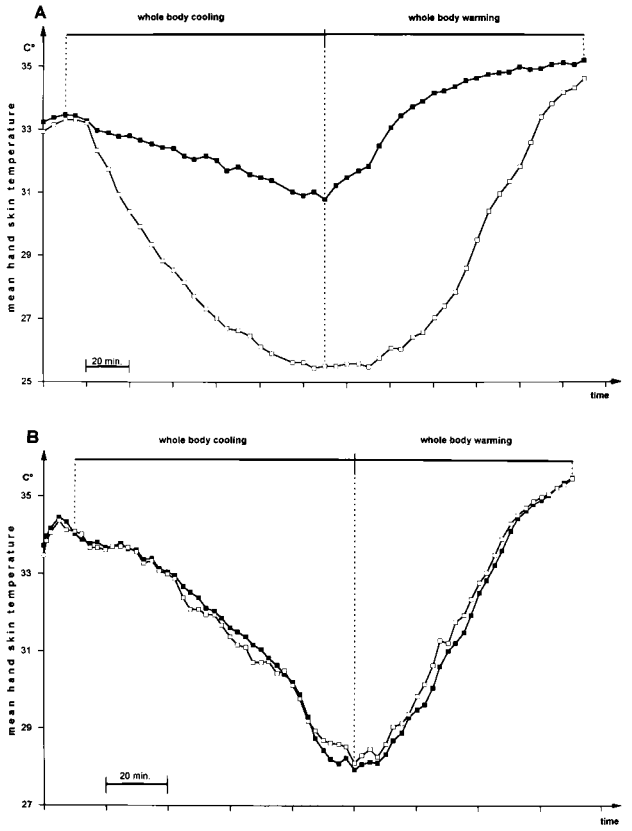


Figure 3 (A–B). Behaviour of SkT under thermal load in RSD before (A) and after therapy (B), same patient as in Figures 1 and 2. Thermal load was applied by means of whole body warming and cooling. To this end, the subject was lying in a suit lined with tubes, in which running water of different temperatures was used to cool or warm the body (water temperature: 12°C and 50°C, respectively). During the experiment, SkTs of the dorsal side of both hands were monitored by thermography at 2.5 minutes intervals. From the stored Thermograms, the thermography system for each body side calculated a mean hand SkT from an area as defined by the investigators (including all fingers and the distal part of the middle hand). First temperature values of each curve shown represent the initial SkT as obtained under room temperature conditions. For each side, consecutive mean temperature values as obtained under stimulation were connected with lines. Filled squares: values of the left (sick) hand. Empty squares: values of the healthy hand. *A:* Before therapy, under cooling conditions the SkT decreased more slowly in the sick hand than in the healthy hand. Under warming, SkT increased on both sides, which finally showed about the same level of temperatures. Note that the measurements on the sick hand were done at an area (dorsal side of the hand/fingers) that was not affected by the preceding lesion. *B:* After therapy, under the entire period of thermal load, the SkT showed a regular behaviour, exhibiting no relevant side differences of SkT (see Blumberg 1988).

3.3. Characteristics of the pain and quality control of sympathetic blocks in RSD

The pain in RSD may be of throbbing, shooting, burning or other quality (see Bonica 1990). Typically, the pain is reduced by elevating the affected extremity and worsens when the extremity is allowed to hang freely. This phenomenon, for which no diagnostic importance has been attached in the literature, was called recently the orthostatic component of the pain in RSD (Blumberg 1988; Blumberg & Jänig 1993). In the following subsections, the response of the pain in RSD upon regionally applied ischemia and upon sympatholytic strategies will be discussed. Such strategies also need to be discussed with respect to their quality control.

3.3.1. Pain behaviour under regionally applied ischemia. It has been observed occasionally that interrupting the circulation of the distal part of the affected extremity in cases with RSD leads to complete relief of pain for the duration of the interruption (de Takats 1943; Gracely et al. 1990; Loh et al. 1981). Later on in RSD this phenomenon was studied systematically by wrapping an Esmarch bandage (or an equivalent) around the hand or foot (to reduce its volume), followed by the application of a suprasystolic cuff directly proximal to this bandage (to stop arterial inflow). This so-called ischemia test, which does not block nerve conduction, typically leads to prompt – within one minute – and complete pain relief for the duration of the cuff (Blumberg & Hoffmann 1992). The mechanisms of the pain-suppressing (positive) effect of this test are not clear, but changes of the microcirculation might be involved since they are induced by the test procedure. An interesting finding was that the outcome of the ischemia test was positive in 38 out of 40 RSD patients who also had an orthostatic component in their pain (Blumberg & Hoffmann 1994).

3.3.2. Effect of sympatholytic strategies. Beneficial results are often reported when sympatholytic strategies are applied for RSD. Accordingly, textbooks recommend this kind of therapy (Bonica 1990; Cousins & Bridenbaugh 1988; Wall & Melzack 1993). In all of these studies, however, there were also nonresponders, sometimes as many as 25% to 50% of patients (Arnér 1991; Bonelli et al. 1983; Driessen et al. 1983; Wang et al. 1985), with no obvious differences associated with the kind of sympathetic block chosen

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

Table 1. *Type and incidence of symptoms in RSD*

	A	B
I. Autonomic symptoms		
distally generalized swelling	94%	69%
warm/cold affected extremity	68%/9%	58%/39%
hypohydrosis/hyperhydrosis	21%/32%	0/47%
II. Motor symptoms		
active movement reduced	87%	88%
muscular strength diminished	93%	95%
tremor (postural or action)	55%	49%
III. Sensory symptoms		
deep spontaneous pain	74%	93%
hypoalgesia/hyperalgesia	36%/55%	0/75%
hypoaesthesia/hyperaesthesia	58%/12%	69%/76%
mechanical allondynia	6%	0

Source: Adapted from Blumberg & Jänig (1993) (A); Veldman et al. (1993) (B).

(e.g., stellate ganglion or guanethidine block). In these studies, it was not mentioned whether the clinical picture of these nonresponders differed from that of the responders. Concerning this aspect, recent findings in 40 RSD patients seem to indicate that a positive ischemia test predicts a positive response to sympathetic block in RSD, both procedures leading to a similar and statistically significant reduction in pain (Blumberg & Hoffmann 1994).

In any case, the degree of technical success of sympathetic blocks should be known in order to accurately interpret the test results. Assessment of the outcome of sympathetic blocks is discussed in the next section.

3.3.3. Quality control of sympathetic blocks. To fulfill the aim of the procedure, sympatholytic interventions should lead to complete interruption of sympathetic activity – in case of stellate ganglion or lumbar chain blockades with local anaesthetics – or to complete depletion of noradrenaline from sympathetic postganglionic vasoconstrictor endings – in case of intravenous regional guanethidine blocks. However, these interventions do not lead inevitably to a complete block. Thus, one has to prove the technical efficacy of sympathetic blocks in every single case.

To achieve this purpose, the behaviour of the sympathetic effector organs needs to be evaluated within the painful area. For cases with RSD in the upper extremity the textbooks note that a Horner sign (ptosis and miosis at the ipsilateral eye) can be used to assess the quality of the stellate ganglion block (Bonica 1990; Cousins & Bridenbaugh 1988; Wall & Melzack 1993). Yet there is no study that has shown that the occurrence of that sign is necessarily combined with a sympathetic block of the ipsilateral extremity. For the lower extremity, a similar sign does not exist.

As a more reasonable way to evaluate a sympathetic block, measurements of skin blood flow/ SkT have been suggested (Bonica 1990; Wall & Melzack 1993). Among the methods available (e.g., laser doppler flowmetry, thermometry, pulse plethysmography), thermometry is the only one that delivers absolute values – in terms of SkT. This method is thus favorable for the standardization of the quality control of sympathetic blocks.

The SkT should increase under each kind of sympathetic block due to the interruption of the sympathetic vasoconstrictor action on the vessels. However, the level of SkT to be reached under the block is uncertain. Textbooks, for example, state that an increase of a few degrees is a sufficient sign of a sympathetic block (Bonica 1990; Cousins & Bridenbaugh 1988; Wall & Melzack 1993), which has been taken up by related studies (Bonelli et al. 1983; Drummond et al. 1991; Erickson & Hogan 1993; Hoffman et al. 1993; Olcott et al. 1991; Price et al. 1992; Torebjörk et al. 1995; Uematsu et al. 1981; Wang et al. 1985).

Such notions are too vague: if the block is applied under room temperature conditions, the possible range of SkT changes is defined by the initial SkT, which may be rather low in the case of RSD with cold skin (e.g., around 20°C; see Fig. 4A), and by the body core temperature (around 36.5°C). According to physiological principles, under a complete block core temperature should nearly be reached at the skin (see Fig. 4C). However, previous studies show inconclusive results with variable values of SkT under a sympathetic block, including spinal anaesthesia, with about 25°C at the lower level and seldom exceeding 35°C (Bengsston 1984; Irazuzta et al. 1992; Treede et al. 1992). This important aspect of the quality control of sympathetic blocks thus remains to be clarified. Furthermore, beside complete or failed interruption of sympathetic activity as a consequence of sympatholytic strategies, one also has to consider the possibility of certain partial effects (see Fig. 4B).

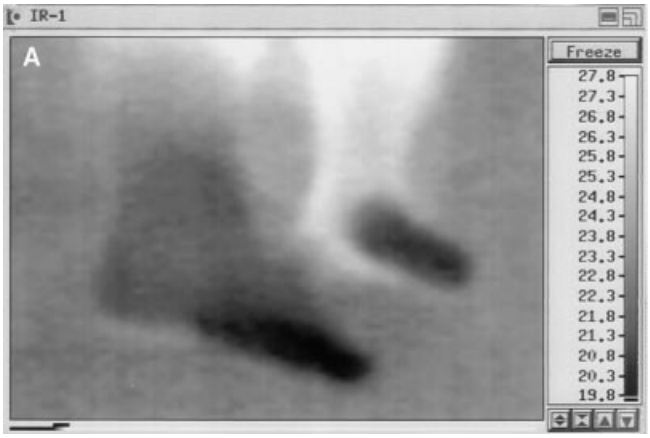
In most studies of sympatholytic strategies in RSD, measurements of sympathetic reflexes have not been conducted to prove completeness of the blocks (Arnér 1991; Blanchard et al.1990; Blumberg & Hoffmann 1994; Bonelli et al. 1983; Cooper et al. 1989; Davidoff et al. 1988; Drummond et al. 1991; Erickson & Hogan 1993; Evans 1946; Hoffmann & Blumberg 1994; Hoffman et al. 1993; Irazuzta et al. 1992; Olcott et al. 1991; Price et al. 1992; Uematsu et al. 1981; Wang et al. 1985; Wilder et al. 1992). Therefore, it remains inconclusive whether patients who did not benefit from a sympathetic block failed to benefit because the block was incomplete.

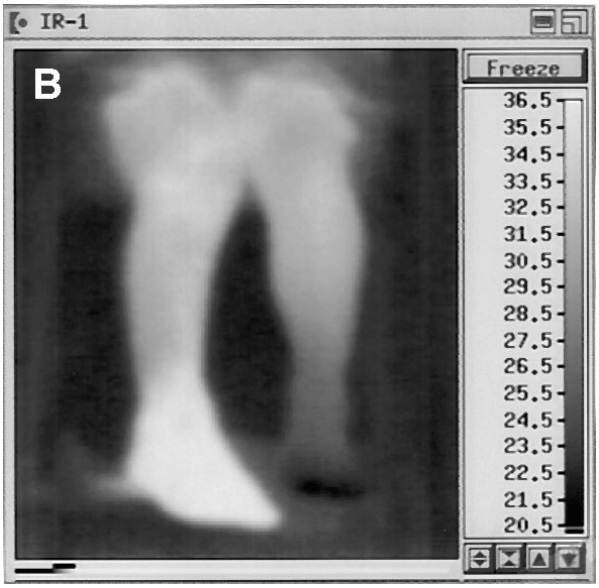
To summarize, more work is needed to define the necessary standard of the quality control of sympathetic blocks in RSD. The nature of the pain in RSD will be discussed below (see sect. 6.).

3.4. Indications for disturbed sympathetic functions in RSD

It is generally assumed that sympathetic functions are disturbed in RSD, based on the clinical observation of side differences of SkT and of sweating in these patients. Tests of resting and reflex behaviour of sympathetic effector organ activity have been done in patients with RSD and the results of such studies, as well as of studies related to the nature of the edema in RSD, are discussed in the following sections.

3.4.1. Behaviour of resting SkT in RSD. In many clinical studies reporting on changes of SkT in terms of warm or cold skin in RSD, the criteria underlying these reports have not been explicitly stated (Arnér 1991; Blanchard et al. 1990; Davidoff et al. 1988; Evans 1946; Pak et al. 1970; Patman et al. 1973; Price et al. 1992; Veldman et al. 1993; Wilder et al. 1992). In only a few studies, SkT measure-





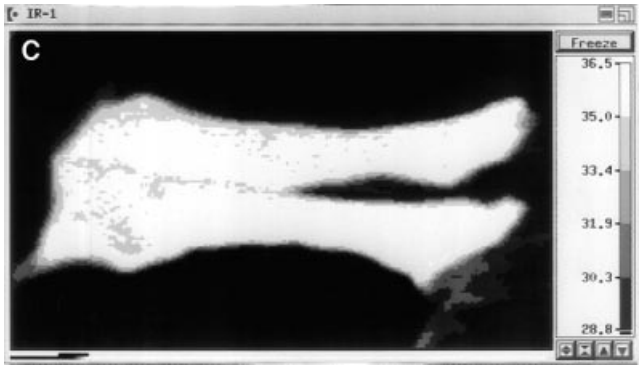


Figure 4 (A–C). Behaviour of SkT in RSD prior to and during sympatholytic interventions with various degrees of technical success. Patient: a 22-year-old female with cold RSD at the right lower extremity following lesion at the right knee (duration of RSD: 9 years). In the Thermograms, values on the scale at the right side indicate degree Celsius, which are related to the colours of the Thermograms. For Thermograms A and B, the highest and the lowest temperature values read by the Thermogram are indicated by the value on the top and the bottom of the scale, respectively. Note the different colour scale used for Thermogram C: In order to improve the colour contrast between the SkT of the legs and the surrounding area, all temperatures below 28.8°C were transferred to black colour by the thermography system (Agema Thermovision 900). *A:* Sock-like distribution of cold skin in RSD before any sympathetic block. *B:* Partial effect of right lumbar sympathetic block at L3 level. Note that the sympathetic block only affected the more distal part of the right leg, with a maximal SkT of 36.5°C. Focal measurements of SkT, for example, at the dorsum at the right foot or big toe, in this case would suggest a sufficient sympathetic blockade of the entire right leg. *C:* Effect of sympatholytic spinal anaesthesia. Somatic level of anaesthesia: Th 10. Note that nearly the entire legs showed SkT values above 35°C, with a maximal SkT of 36.5°C, as also obtained at the foot during the lumbar sympathetic block (compare with Thermogram B).

ments were applied in RSD, most of which were done focally (Bonelli et al. 1983; Irazuzta et al. 1992; Kurvers et al. 1995; Wang et al. 1985). Only occasionally were studies done in a more systematic manner, for example, by measuring the SkT of all fingers or toes of both sides (Blumberg 1988; 1991; Blumberg & Jänig 1993), or more extensive areas were monitored by applying thermography (Feldman 1991; Pochaczevsky 1987; Sherman et al. 1994).

According to the studies with a systematic measurement of SkT, the skin in most of the patients on the affected extremity was warmer compared with the healthy side (Blumberg & Jänig 1993). This was especially true for early RSD with duration not exceeding 10 days, where patients only rarely exhibit cold skin (Blumberg 1991). Naturally, the observation of warm skin speaks against the assumption of generally abnormal high sympathetic tone in RSD. This is also true for the finding of lowered plasma catecholamine levels – which, interestingly, often was combined with cold skin – and of increased subcutaneous blood flow in RSD extremities compared with the healthy side (Christensen & Henriksen 1983; Drummond et al. 1991).

3.4.2. Reflex behaviour of SkT in RSD. Patients with RSD may recognize an abnormal behaviour of SkT, especially when exposed to low environmental temperature. This was verified under experimental conditions, in which whole body cooling by means of a thermal suit was used reflexively to stimulate the skin sympathetic vasoconstrictor activity. Under this condition, RSD patients developed either more or less vasoconstriction on the affected side compared with the healthy side, giving rise to side differences of SkT, which in 38 cases with RSD statistically proved to be significantly higher compared with a control group of 18 cases (Blumberg 1988; see also Fig. 2). Other studies, using some other kind of thermal stimulation, report similar results (Bej & Schwartzman 1991; Cooke et al. 1993; Leriche 1923; Morsier 1947; Trostdorf 1956).

These findings, which were made in skin with intact innervation, support the idea that RSD may be associated with an abnormal reflex pattern of skin sympathetic vasoconstrictor neurones. Direct evidence for this phenomenon, as could possibly be obtained by bilateral recordings of these neurones under such conditions, is missing in RSD. On the other hand, this interpretation of these results is consistent with experimental findings obtained in animals, which show that the reflex pattern in skin vasoconstrictor neurones may change after a nerve lesion (Blumberg & Jänig 1985; Jänig & Koltzenburg 1991). However, there is no indication that changes of SkT, as possibly caused by abnormal sympathetic activity, are related to the pain in

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

RSD, since, for example, it is hard to see how diminished skin vasoconstriction upon cooling (see Fig. 3) may contribute to pain.

3.4.3. On the nature of the edema in RSD. The edema seems to be a major symptom in RSD (Blumberg 1992; Blumberg & Jänig 1993; Davidoff et al. 1988; Evans 1946; Kurvers et al. 1995; Pak et al. 1970; Patman et al. 1973; Veldman et al. 1993), but it is uncertain whether the edema should be described as an autonomic (sympathetic) symptom (Blumberg 1988; Blumberg & Jänig 1993; Kurvers et al. 1995) or as an inflammatory sign, for example (Veldman et al. 1993).

The edema was measured in only a few studies to see whether it could be diminished by sympatholytic strategies. In one study, the edema was reduced at the end of treatment, without showing acute effects (Davidoff et al. 1988). In another study, which is part of an ongoing investigation at our institution, acute effects of various kinds of sympatholytic interventions on the edema in RSD could be demonstrated (Blumberg et al. 1994; see also Fig. 2). On the other hand, treatment with corticoids also results in diminished edema in RSD, but no acute effects are reported (Christensen et al. 1982; Kozin et al. 1976).

The mechanisms of the edema in RSD are unknown. Regarding the effect of sympatholytic interventions upon this symptom of RSD, a contribution of the sympathetic system may be assumed. This assumption is part of a hypothesis that tries to explain this effect and, at the same time, the nature of the pain in RSD, which will be discussed in section 6.

4. Sympathetically maintained pain syndrome

4.1. Case report 2

A 59-year-old man suffered from steadily increasing paraesthesias for more than a year in his left arm, radiating from the side of the elbow to the fifth finger. There was no pain and no motor deficit in that hand. As even light movements of the elbow caused these paraesthesias, a neurological investigation, including electrophysiological measurements, was performed. Evidence for slight partial lesion of the ulnar nerve was found, probably caused by a tumor that was situated close to the ulnar sulcus. Due to this finding, a neurosurgical intervention was performed. Most of the tumor was extirpated, and a neurinoma was diagnosed histologically.

The patient noticed spontaneous pain for the first time at the ulnar part of the left hand about 8 days following that intervention. Inside the same area pain developed upon brief light touch, but not upon light constant touch. As these symptoms did not disappear during the following 2 months, and could not be treated by analgesic oral drug treatment, the patient finally was referred to us with the suspicion of having ulnar nerve neuralgia.

Upon neurological investigation, the following findings were obtained in the left arm in comparison with the right arm: tendon jerks at the upper extremities were regular and there was no motor deficit, including the muscles that were supplied by the left ulnar nerve. There was no edema and, upon touch, no abnormality of sweating or of SkT; the patient reported no change of SkT under environmental thermal load. He reported spontaneous burning pain, which was felt superficially within the zone of the left ulnar nerve, also affecting the ulnar part of the fourth finger. The pain was constantly present during the entire day with about the same intensity, and was not influenced by elevating or hanging down the affected extremity (no orthostatic component). Within the area of the spontaneous pain, mechanical allodynia was present, accompanied by diminished sensation on touch (hypoaesthesia) and on pin prick (hypoalgesia). There was no hyperpathia. Measurements of SkT in all finger tips showed no systematic side differences. There was no pain relief under the ischemia test.

4.1.1. Treatment and follow up. Due to the combination of spontaneous pain and allodynia, strictly confined to the zone of a partially lesioned nerve (see Fig. 5), SMP syndrome was assumed and a diagnostic ipsilateral stellate ganglion block was performed. Following the block, pain and allodynia were absent for a couple of days. Upon their return, a second block was done, which again relieved both symptoms. Thereafter, the patient was free of any pain sensation for the entire period of follow up (6 months), with the sensory signs of the partial lesion of the ulnar nerve still being present.

4.2. Clinical picture of SMP syndrome – comparison with RSD

In contrast to RSD, patients with SMP show a less complex clinical picture, its principal symptoms being spontaneous pain and (mechanical/cold) allodynia (Blumberg & Jänig 1993; Frost et al. 1988; Leriche 1949; Loh & Nathan 1978; Meyer et al. 1992; Nathan 1947; Treede et al. 1992; Wahren et al. 1995). For the most part these symptoms are found in the zone of a partially lesioned nerve or nerve root; this spatial relationship seems to be the key feature of the SMP syndrome.

The spontaneous pain in SMP is mostly felt superficially, does not show an orthostatic component, and, according to preliminary results (Blumberg & Hoffmann, in preparation), is not suppressed by the ischemia test, all these

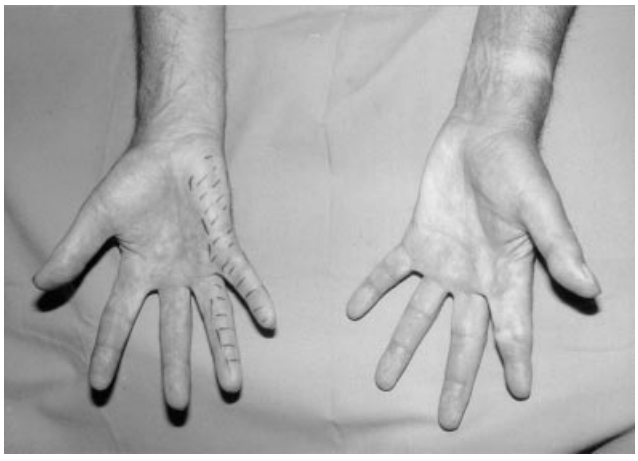


Figure 5. Patient with SMP with partial right ulnar nerve lesion (see case report 2). The area of spontaneous pain and mechanical allodynia is marked. Note that there is no impairment of finger spreading, indicating regular function of the related muscles supplied by the ulnar nerve.

characteristics contrasting to those of the pain in RSD. Motor symptoms, if present at all, can be explained by lesions of motor axons. Consistent signs of sympathetic hyperactivity are not found in SMP and autonomic changes, which may be present inside the lesioned area (Frost et al. 1988), can be related to denervation and reinnervation of sympathetic effector organs. In Table 2, the main clinical features of SMP are compared with those of RSD. The distinct differentiation made between these conditions may not always be found clinically as there may be transient states between them, in the extreme case leading to causalgia (see sect. 5).

4.3. Effect of sympatholytic strategies in SMP

Sympatholytic strategies should, by definition, relieve pain in SMP. Consequently, sympathetic blocks should yield positive results if applied for diagnostic reasons, as has been reported for the test with local anaesthetics (applied to paravertebral ganglia), the intravenous phentolamine test (Treede et al. 1992; Raja et al. 1991), and the guanethidine block (Torebjörk et al. 1995; Wahren et al. 1995).

The phentolamine test is reported not to produce signs of a sympathetic block in patients suspected of having SMP (nor in RSD; see Arnér 1991), but nevertheless induces pain relief in some patients (Meyer et al. 1992; Raja et al. 1991). Thus, in terms of quality control the efficacy of the test to block alpha adrenergic receptors remains uncertain, at least for the site of neurovascular transmission. For the nonresponders of such studies, this offers the possibility that the doses of the drug applied may have been insufficient to interrupt the possible sympathetic– sensory coupling in these cases. Thus, it seems hard to conclude that these nonresponders did have sympathetically independent pain (SIP), as has been stated (Meyer et al. 1992; Treede et al. 1992).

Individuals with suspected SMP who fail to respond to conventional sympathetic blocks should not be diagnosed as having SIP, if the blocks are incomplete, as judged by SkT (Treede et al. 1992). Regarding the diagnosis of SIP as a consequence of a negative response upon guanethidine block (Torebjörk et al. 1995; Wahren et al. 1995), one also has to consider the possibility that such a (distal) regional block does not reach more proximal areas, as recently has been pointed out by Wall (1995). According to findings obtained in animal experiments, proximal areas like the dorsal root ganglion, for example, may be the site of sympathetic– sensory coupling following nerve lesion (Devor et al. 1994; Wall & Devor 1983).

5. Causalgia – a combination of RSD and SMP?

Regarding the clinical picture of patients that are described as suffering from causalgia and respond to sympathetic blocks (see Bonica 1990), there are two main characteristics: they have a partial nerve lesion, which is acutely followed by pain within the zone of the lesioned nerve. The pain syndrome of these cases consists of spontaneous pain and of pain upon light touch, now called allodynia. In addition, and also occurring acutely or somewhat later in the course of the disorder, they often develop symptoms outside the territory of the lesioned nerve. These symptoms consist, for example, of diffuse pain and distally generalized swelling, changes of skin blood flow and sweating, and motor impairments.

In the light of the differentiation of SMP and RSD as described above, the initial symptoms of causalgia are rather similar to those of SMP, whereas the later symptoms are similar to the clinical picture of RSD. Thus it seems that causalgia is a combination of RSD and SMP. Such very rare cases can also be observed today (Blumberg & Jänig 1993; Blumberg & Hoffmann, in preparation).

6. On the nature of the pain in RSD/SMP

In most cases, the deep and diffuse distal pain in RSD is associated with evidence of disturbed microcirculation, as indicated by the presence of edema. The relationship between the pain and the state of the microcirculation is indicated by the orthostatic component of the pain and the positive effect of the ischemia test. In order to explain this relationship and the effect of sympatholytic strategies on the pain – and on the edema – in RSD, a hypothesis is presented below (Blumberg 1988; Blumberg & Jänig 1993; Blumberg et al. 1994). It is a specification of a general model regarding the pathophysiology of RSD, which also considers a key role of the sympathetic system in RSD (Devor et al. 1991; Jänig 1985; 1990).

This hypothesis states that nociceptive afferent input, which is generated by the lesion preceding the onset of RSD, sensitizes spinal circuits, thereby leading to an ab-

Table 2. *Comparison of the main features of reflex sympathetic dystrophy (RSD), sympathetically maintained pain (SMP) syndrome, and causalgia*

	RSD	SMP	Causalgia
Incidence	–common	–rare	–very rare
Etiology	–any kind of lesion	–partial nerve lesion	
Localization	–distal part of extremity	–confined to lesioned nerve	Combination of the symptomatology
Spontaneous pain	–common	–obligatory	
	–deep, diffuse, with orthostatic component	–superficial, without orthostatic component	of SMP and RSD, the condition
Allodynia	–rare (<10%)	–obligatory	typically starting with SMP, followed
Autonomic, motor, and sensory symptoms	>90% distally generalized (glove-/sock-like distribution)	related to nerve lesion	by the RSD type of symptoms

https://doi.org/10.1017/S0140525X97261480 Published online by Cambridge University Press

normal discharge pattern in sympathetic vasoconstrictor neurones. This generates edema by an increased vasoconstriction on the postcapillary side compared with the precapillary side, inducing increased filtration pressure. The related increased interstitial pressure may activate afferent fibres (e.g., deep nociceptors). Thus, an indirect sympathetic–sensory coupling takes place via the vascular system. As the characteristics of the sympathetic outflow with its diffuse distribution are not related to nerve zones or dermatomes, the edema – and thereby this coupling – at the affected extremity will occur in a diffuse distribution. Consequently, unlesioned regions are affected. The coupling maintains the central sensitization process, thereby producing a vicious circle.

Sympathetic blocks will lead to a decrease of the edema, mainly by interrupting vasoconstrictor activity and thereby opening the venules. Due to the accompanied decrease of interstitial pressure, activity of afferent (nociceptive) fibres is reduced. As a result, central sensitization is abolished and the vicious circle is interrupted. Consequently, sympathetic outflow normalizes (for further details of the hypothesis see Blumberg et al. 1994).

It is uncertain whether similar pain mechanisms are also active in SMP, in which the pain also can be abolished by sympathetic blocks. However, two main arguments favour a more direct sympathetic–sensory coupling in SMP compared with RSD: first, the pain in SMP does not seem to be related to the state of microcirculation, as typically there is no edema and no orthostatic component. Second, typically the pain in SMP is distributed within the zone of a lesioned nerve. Thus it seems more likely that the coupling occurs at the lesioned nerve, which may have increased expression of noradrenergic (alpha 1) receptors (Campbell et al. 1992; Devor et al. 1991; Torebjörk et al. 1995). At the same time, spinal neurones (e.g., wide dynamic range neurones) may have become sensitized, as indicated by the presence of allodynia (Gracely et al. 1992; Roberts 1986), and this sensitization may be maintained by the nociceptive input due to sympathetic–sensory coupling at the lesioned nerve. This coupling is interrupted by sympathetic blocks, finally leading to relief of pain in SMP (for further discussion see Blumberg & Jänig 1993; Jänig & Koltzenburg 1992). In causalgia, both kinds of proposed pain mechanisms may be active at the same time.

Other pain mechanisms may also develop in the course of RSD/SMP (Price et al. 1989; Wahren et al. 1995). In RSD this will happen especially when the condition is associated with other disorders like carpal tunnel syndrome or psoriasis arthritis (see Conca et al. 1995), or if the syndrome is artificially maintained as in cases with Munchausen's syndrome (Rodriguez–Moreno et al. 1990). In addition, one has to consider the possibility of so-called psychogenic pain mechanisms. In all such cases, which may be called atypical RSD, even complete sympathetic blocks may give unsatisfactory results.

7. Conclusion

This paper describes the clinical picture of patients who may develop pain related to the sympathetic system. In one group, labeled RSD, patients exhibit a very complex syndrome, which can be characterized by a triad of autonomic, motor, and sensory symptoms that appear in a distally generalized distribution. The accompanying pain is typically felt deeply and diffusely inside the symptomatic area. It usually has an orthostatic component and is suppressed by the ischemia test. In other, rare cases, which are called SMP, the symptoms are more localized and consist for the most part of spontaneous pain, felt superficially without orthostatic component, plus allodynia. A very small number of patients exhibits symptoms of both RSD and SMP. They fit the description of causalgia.

In RSD, indirect evidence for disturbed sympathetic function is present on the basis of clinical and experimental findings, but there is no indication that these disturbances are related to the pain. This pain seems to respond to sympatholytic strategies, but there are no clinical findings that allow the prediction of the outcome of a sympathetic block in a given patient with RSD. To date, it is uncertain whether a positive ischemia test, if combined with an orthostatic component of the pain, may be such a predictor. New observations indicate that sympathetic blocks may be followed by acute and lasting relief of the edema in RSD. For the small group of patients with SMP, evidence for disturbed sympathetic functions has not been reported; as in RSD, safe clinical criteria that would allow the prediction of the outcome of sympatholytic interventions in these patients do not exist.

Criteria necessary to define sympathetic blocks as being complete do not seem to be sufficiently elaborated, and the degree of technical success remains uncertain for most reported blocks in RSD/SMP. Hence it also remains uncertain whether the effect of such blocks on the pain in these conditions depends on the quality of the blocks, both for the responders and the nonresponders. Finally, one must consider the existence (or coexistence) of other (or additional) pain mechanisms in RSD/SMP, which may lead to the unsatisfactory results of sympathetic blocks. All these circumstances may be relevant to the discussion of sympathetic–sensory coupling as a possible cause of the pain in RSD/SMP.

ACKNOWLEDGMENTS

This work was supported by the Bundesministerium für Forschung und Technology. We wish to thank Mrs. A. Boeger-Koch for her technical assistence.