The role of the thalamus in the flow of information to the cortex

S. Murray Sherman and R. W. Guillery

1Department of Neurobiology, State University of New York, Stony Brook, NY 11794-5230, USA
2Department of Anatomy, University of Wisconsin School of Medicine, 1300 University Avenue, Madison, WI 53706, USA

The lateral geniculate nucleus is the best understood thalamic relay and serves as a model for all thalamic relays. Only 5–10% of the input to geniculate relay cells derives from the retina, which is the driving input. The rest is modulatory and derives from local inhibitory inputs, descending inputs from layer 6 of the visual cortex, and ascending inputs from the brainstem. These modulatory inputs control many features of retinogeniculate transmission. One such feature is the response mode, burst or tonic, of relay cells, which relates to the attentional demands at the moment. This response mode depends on membrane potential, which is controlled effectively by the modulator inputs. The lateral geniculate nucleus is a first-order relay, because it relays subcortical (i.e. retinal) information to the cortex for the first time. By contrast, the main thalamic relay of visual information, the pulvinar region, is largely a higher-order relay, since much of it relays information from layer 5 of one cortical area to another. All thalamic relays receive a layer-6 modulatory input from cortex, but higher-order relays in addition receive a layer-5 driver input. Corticocortical processing may involve these corticothalamicocortical ‘re-entry’ routes to a far greater extent than previously appreciated. If so, the thalamus sits at an indispensable position for the modulation of messages involved in corticocortical processing.

Keywords: pulvinar; bursts; corticothalamic; drivers; modulators

1. INTRODUCTION

Virtually all information that reaches the cerebral cortex must first pass through the thalamus, and yet the thalamus is often seen as a simple machine-like relay. This suggests that nothing would be lost if information were passed directly from peripheral receptors, such as the retina, to the neocortex. However, the known complexity of thalamic circuitry points strongly to a significant role for thalamic processing, and details about that role have emerged during the past decade or so, showing that the thalamus can dynamically alter the information relayed in a manner that reflects various behavioural states, such as attention and drowsiness. Some of the better understood of these functions will be reviewed here, and we argue that the thalamus plays a crucial role in controlling the flow of information to the cortex. For thalamic relays that we classify as 'first-order relays', this information comes from the sensory periphery (visual, auditory, tactile, etc.) and from other parts of the brain such as the cerebellum or the mammillary bodies; whereas for other thalamic nuclei that we have classified as 'higher-order relays', this information comes from the cerebral cortex itself (Guillery 1995; Sherman & Guillery 2001; Guillery & Sherman 2002). These higher-order relays form the largest part of the thalamus in primates. An important feature of the argument presented here is that this part of the thalamus plays a significant role in corticocortical communication, again, acting as a dynamic control of information that is being passed through the thalamus from one cortical area to another.

Among thalamic relays, we know most about the thalamic relay of visual information, which involves the lateral geniculate nucleus and pulvinar region.1 The lateral geniculate nucleus provides the relay of retinal information and innervates striate cortex as well as some extra-striate areas. The pulvinar region innervates all, or nearly all, known extra-striate visual areas, and, as we shall argue below, the main information it relays derives from the cortex itself, as it is a higher-order relay and relays messages from one visual area to another. The lateral geniculate nucleus, a first-order relay, will serve as the primary exemplar of thalamic relay functions in the first section of this paper. In the second section, the pulvinar region will serve as an exemplar of higher-order thalamocortical circuits. Both parts will focus on the thalamic relays in the cat.


The lateral geniculate nucleus of the cat, like that of all mammals, is laminated, each layer being innervated by one or the other eye (for details of the innervation of cats and other mammals, see Sherman 1985; Casagrande & Norton 1991). Several functionally distinct, parallel pathways run from the retina through the geniculate relay to the visual cortex. In the cat, these are known as the W-, X- and Y-pathways, and in primates, as koniocellular,
parvocellular and magnocellular pathways (reviewed in Sherman 1985; Casagrande & Norton 1991; Hendry & Reid 2000). They have their origin in the retina from structurally and functionally distinct retinal ganglion cell types. Their termination, in several distinct layers of the lateral geniculate nucleus, varies with species. Whereas the X- and the Y-pathways are mingled in the major geniculate layers of the cat (layers A and A1; see figure 1), in primates, the magnocellular and parvocellular pathways terminate in distinct layers, called parvocellular and magnocellular. In cats, the W-pathways have a separate set of small-celled geniculate layers (layers C1 and C2 in figure 1), but in most primates the pathways relating to the smallest cells, the koniocellular pathways, terminate close to, and mingle with, both the parvocellular and magnocellular layers. In the bush baby (Galago), the magnocellular, parvocellular and koniocellular pathways all have a distinct layer for each eye.

One important point about these parallel pathways is that, at present, there is no evidence that they significantly interact in their primary thalamic relay to cortex, irrespective of their laminar distribution. The layering of the lateral geniculate nucleus is vital for understanding the separation of left eye from right eye inputs, since each eye has its own set of layers. Further, it shows that there is a tendency for functionally distinct cell types to be segregated even though the functional significance of this segregation is still not understood. Figure 1 shows that the several representations of the visual field within the layers are in register with each other so that one can draw lines of projection through the nucleus that run roughly perpendicular to the layers and that represent single directions (from the eye) in visual space (Sanderson 1971). The functionally distinct retinal afferents are arranged in the layers in a sequence that is characteristic for any one species. This arrangement of functionally distinct cell types along single lines of projection has provided a useful basis for many studies of the functional organization of the lateral geniculate nucleus, and we encounter the concept again in the section on the pulvinar. Here, we focus primarily on the major layers of the cat’s lateral geniculate nucleus, the A-layers, where the X- and the Y-pathways terminate. This is because the A-layers of the lateral geniculate nucleus of the cat are the best studied of any thalamic relay, and thus these will serve as the model for exploring the details of the functional organization of thalamic relays.

(a) Cell types in the A-layers

Figure 2 shows the three cell types in the A-layers: X and Y relay cells and interneurons. The relay cells differ in their morphology. Y-cells have larger cell bodies and thicker dendrites. The dendrites tend to be smooth and contained in a roughly spherical arbor. X-cells usually have clustered appendages on proximal dendrites, often near primary branch points, and these are interesting because they mark the postsynaptic sites of retinal inputs and triads (see § 2b(vii)). The arbors of X-cells tend to be bipolar in shape, oriented perpendicular to the layering. Interneuronal dendrites are thin and their arbors, like those of X-cells, are oriented perpendicular to the layering. They have numerous axon-like terminal swellings that are actually presynaptic terminal boutons that give rise to the F2 terminals described below. Thus, much of the synaptic output of these cells is from their dendrites, and these terminal boutons are both presynaptic and postsynaptic. Interneurons also have a traditional, single axonal output that ramifies close to the cell body.

Cable modelling suggests an important functional difference between relay cells on the one hand and interneurons on the other, and this is related to the unusual structure of the interneuronal dendrites (Bloomfield & Sherman 1989). Both X and Y relay cells are electrotionically compact, implying that synapses on even the most peripheral dendritic sites can produce sizeable PSPs at the cell body. By contrast, interneurons are electrotionically extended, implying that peripheral inputs, and particularly those onto the dendritic terminal boutons, will have negligible effect on the cell body. This suggests an interesting hypothesis for interneuronal functioning: the axonal output is controlled by inputs to proximal dendrites in a conventional manner. However, the more peripheral synaptic inputs have an insignificant effect on the axon, but instead act on the dendritic outputs in an independent manner (Bloomfield et al. 1987; Cox & Sherman 2000).

(b) Circuitry

(i) Inputs

Figure 3 schematically summarizes circuitry involving the A-layers of the lateral geniculate nucleus of the cat (Sherman & Guillery 2001). The major inputs to relay cells, in addition to glutamatergic retinal afferents, are GABAergic inputs from local neurons (reticular cells and interneurons), glutamatergic inputs from layer 6 of the cortex, and cholinergic inputs from the parabrachial region of the midbrain. Curiously, the cholinergic parabrachial inputs also employ nitric oxide as a transmitter. Minor inputs (not shown in figure 3) include noradrenergic inputs from the parabrachial region (cholinergic and noradrenergic cells afferent to relay cells...
Thalamic relay functions  S. M. Sherman and R. W. Guillery 1697

Figure 2. Reconstruction of an X-cell, a Y-cell and an interneuron from A-layers of the cat’s lateral geniculate nucleus. The insets for the X-cell show the grape-like clusters appended to the dendrites, mostly near the primary branch points, and the inset for the interneuron shows the presynaptic bouton terminals. Scale bar for main figure, 50 μm, scale bar for the insets for the X-cell and interneuron, 10 μm.

Figure 3. Neuronal circuitry related to A-layers of the lateral geniculate nucleus of the cat. The various inputs are displayed, including the neurotransmitters associated with them and the type of receptor, ionotropic or metabotropic, each activates. Also, driver versus modulator inputs are shown. Abbreviations: LGN, lateral geniculate nucleus; PBR, parabrachial region; TRN, thalamic reticular nucleus. (See text for details.)

are intermixed in the parabrachial region, serotonergic inputs from the dorsal raphe nucleus, and histaminergic inputs from the tuberomamillary nucleus of the hypothalamus.

(ii) Mapping

Figure 1 shows that there is a precise retinotopic map in the lateral geniculate nucleus. The afferents from the cortex, from the thalamic reticular nucleus and from interneurons are connected in accord with this retinotopy, but the other afferents are more diffusely organized. This implies that cortical and local inputs can act with local sign on a limited part of the visual field, whereas the other inputs have a more global effect on relay cell responses for the whole of the visual field.

(iii) Postsynaptic receptors

Both ionotropic and metabotropic receptors are postsynaptic to the above mentioned inputs in relay cells (figure 3). They are both complex proteins located in the postsynaptic membranes. While many differences between these receptor types exist, only a few concern us here (for details see Nicoll et al. 1990; Mott & Lewis 1994; Recasens & Vignes 1995; Pin & Duvoisin 1995; Conn & Pin 1997; Brown et al. 1997).

Ionotropic receptors include AMPA receptors for glutamate, GABA, receptors, and nicotinic receptors for acetylcholine and these are directly linked to specific ion channels. Transmitter binding leads to a rapid conformational change that opens an ionic channel and produces a PSP that is fast, with a short latency (less than 1 ms) and a brief duration (a few tens of ms). Metabotropic receptors include various metabotropic glutamate receptors,
GABA_\textsubscript{B} receptors, and various muscarinic cholinergic receptors. These are not directly linked to ion channels. Instead, transmitter binding produces a series of biochemical reactions that ultimately lead to the opening or closing of an ion channel, which, for thalamic cells, is usually a K\textsuperscript+ channel; when opened, this produces an IPSP as K\textsuperscript+ flows out of the cell and, when closed, produces an EPSP as K\textsuperscript+ leakage is reduced. These postsynaptic responses are slow, with a long latency (at least 10 ms) and a prolonged duration (hundreds of ms or more). Metabotropic receptors also usually require higher firing rates to be activated.

Retinal input activates only ionotropic glutamate receptors. However, all non-retinal inputs can activate both ionotropic and metabotropic receptors (figure 3), but it is not clear whether any individual non-retinal axon can activate both. Nonetheless, the activation of metabotropic receptors means that these inputs can create sustained changes in baseline membrane potential, which, among other things, means that these inputs can have sustained effects on the overall responsiveness of relay cells. Other consequences of these sustained postsynaptic responses are considered below.

(iv) Synaptic structure

The various afferents have distinct ultrastructural features (Guillery 1969; Sherman & Guillery 2001). Retinal terminals are the largest, have spherical vesicles, form multiple, asymmetric synaptic contacts and are never postsynaptic to other processes. They are known as ‘RL’ (for round vesicle and large profile) terminals. The terminals from cortical layer 6 and the parabrachial region resemble each other: they are relatively small, have spherical vesicles, form asymmetric contacts (usually one per terminal) and are never postsynaptic. They are known as ‘RS’ (for round vesicle and small profile) terminals. The local GABAergic inputs form two distinct types of synaptic terminal, although both types are relatively small, exhibit flattened or pleomorphic vesicles and form symmetric contacts (again, usually one per terminal). Thus, these are called ‘P’ (for flattened vesicle) terminals. One type, called F1, derives from axons of interneurons and reticular cells. Like the other terminals so far described, these are strictly presynaptic. The other type, called F2, derives from dendrites of interneurons (see description of interneurons in §2a), and is both presynaptic and postsynaptic; these are the only processes in the thalamus that are both presynaptic and postsynaptic. F2 terminals are also set apart by their involvement in triadic synaptic arrangements, which are described in §2b(vii).

(v) Drivers and modulators

Although geniculate relay cells are obviously in the business of relaying retinal input to the cortex, it is important for understanding thalamic relays (and possibly other relays) to appreciate that these retinal inputs, seen morphologically as RL terminals, provide only ca. 5–10% of all synaptic inputs to relay cells (Van Horn et al. 2000). Local GABAergic inputs, cortical inputs and brainstem inputs each comprise ca. 30%, and the remainder (noradrenergic, serotonergic and histaminergic inputs) represent less than 5% of the total. While such anatomical information about pathways and synaptic connections is a requisite key to understanding functional circuitry, information about the numbers of synaptic inputs from various sources onto relay cells underscores the folly of assuming that in any pathway the numerically largest component must be functionally the most important. Such logic would lead to the conclusion that the lateral geniculate nucleus relayed parabrachial inputs to the cortex and that the retinal input played a minor role. Of course, with functional data, we are better informed, and this, in turn, means that the functional importance of an input cannot be ranked on the basis of size.

The evident importance of the relatively small retinal input to the lateral geniculate nucleus has previously led us to propose the notion of dividing inputs into drivers and modulators (Sherman & Guillery 1998, 2001). In the lateral geniculate nucleus and in other thalamic relays, the driver input, which has the characteristic RL terminals, brings the main information to a cell or cell group, and in each of the major thalamic sensory relays dominates the receptive field properties of the target cells. The modulator input, which is the remainder, including the RS terminals, modulates how that input is handled. Because of the many varieties of modulation, and the frequent need for fine grain in this process, it is logical that more synapses are devoted to this process than to coding the basic information, so there are more modulator inputs, but the driver inputs must have a powerful postsynaptic effect in order to transmit the information. This seems to be the case, since retinal EPSPs produced by the retinal inputs are relatively large.

Interestingly, for the cells in layer 4 of the striate cortex that are postsynaptic to the geniculate inputs, a similar pattern emerges. The driver inputs to these are the geniculate axons (Reid & Alonso 1995, 1996; Ferster et al. 1996; Chung & Ferster 1998), and yet these represent only 5–10% of the synaptic inputs to these cells in cats and monkeys (Ahmed et al. 1994; Latawiec et al. 2000). The remarkably similar number for retinal drivers to geniculate relay cells and geniculate drivers to layer-4 cortical cells may be a coincidence, but it is plausible that the same general rules that apply to the thalamus as regards drivers and modulators also apply to the cortex. We shall resume this theme later.

(vi) Location of inputs onto relay cell dendrites

As shown schematically in figure 4, the inputs to relay cells are not distributed evenly on their dendrites (Guillery 1969; Wilson et al. 1984; Erisir et al. 1997). Retinal and parabrachial inputs are limited to proximal dendrites (ca. 100–150 \mu m from the cell body), while cortical inputs are located more distally. Inputs from interneurons (both from axons and dendrites) are more concentrated in the proximal zone, amongst retinal and parabrachial inputs, while reticular inputs are concentrated in the distal zone, amongst cortical inputs. Some functional consequences of this differential distribution of the various afferents are considered in §2b(vii).

(vii) Differences between X- and Y-cells: triads

While the above distributions of inputs onto dendrites applies to both X and Y relay cells, there is a difference in the nature of retinal and parabrachial inputs to these two relay cell types. A particular kind of synaptic arrange-
Figure 4. Schematic view of distribution of synapses on X and Y relay cells of the lateral geniculate nucleus of the cat. The T-channels are also shown; the line thickness indicates relative density, a thicker line for dendrites where T-channels are denser, and a thinner line for the soma, where T-channels are sparser. They are found throughout the cell membranes but are denser on the dendrites than on the cell body. Note that retinal, interneuronal and parabrachial inputs contact proximal dendrites, while cortical and reticular inputs are concentrated on peripheral dendrites. Abbreviation: TRN, thalamic reticular nucleus.

Figure 5. Schematic view of triadic circuitry in a glomerulus of the lateral geniculate nucleus in the cat. The arrows indicate presynaptic to postsynaptic directions. The question marks postsynaptic to the dendritic terminals of interneurons indicate that it is not clear whether GABA_b (i.e. metabotropic) receptors exist there. Abbreviation: PBR, parabrachial region.

ment known as the triad is common to all thalamic relays in cats and monkeys. In the lateral geniculate nucleus, triads are effectively limited to X- but not Y-cells. Triads are not found in most thalamic nuclei of rats and mice, which lack interneurons (Arcelli et al. 1997). There are two types of triad, shown schematically in figure 5, and both contain the dendritic terminal of interneurons as a central element. One, known as the retinal triad, involves a retinal terminal that contacts both an interneuronal dendritic terminal and a relay X-cell dendrite, with the dendritic terminal contacting the same relay cell dendrite. The other, known as a parabrachial triad, involves two terminals from one parabrachial axon, with one of the parabrachial terminals contacting an interneuronal dendritic terminal and the other contacting a relay X-cell dendrite, with the dendritic terminal contacting the same relay cell dendrite.

Figure 5 offers some insights into the functioning of these triads derived from investigations into the postsynaptic receptors involved (Cox & Sherman 2000). For a retinal triad, the retinal input to the relay cell operates through ionotropic receptors, while the feed-forward, disynaptic inhibitory input to the relay cell operates from the retinal input through metabotropic (and possibly) ionotropic receptors. This means that the monosynaptic EPSP will grow more or less smoothly with firing rate in the retinal afferent, as will the disynaptic IPSP, starting with retinal input to the interneuronal terminal if ionotropic receptors are also involved there. However, the presence of metabotropic receptors there suggests that, as the retinal afferent firing rate reaches a certain threshold level to activate these receptors (see above), there will be a significant increase in the IPSP, which will continue to grow more strongly with increased afferent firing than will the EPSP. The retinal firing rate increases with the contrast of the visual stimulus (Kaplan et al. 1987; Thejomayen & Matsubara 1993), which, in turn, suggests that the increase in the EPSP slows down markedly with higher-contrast stimuli. This would have the effect of extending the range of contrasts to which the relay cell can respond in a graded fashion before saturation in the response kicks in. This can be considered a form of ‘contrast gain control’, which is usually attributed to cortical rather than thalamic circuitry (Ohzawa et al. 1982; Truchard et al. 2000; Przybyszewski et al. 2000).

For the effects of the parabrachial inputs, a similar scenario can be imagined. The direct inputs to relay cells begin activating an EPSP at even the lowest firing rate in the afferent because of the presence of nicotinic (ionotropic) receptors on the relay cell. Only after firing of the parabrachial afferent increases to some threshold level will the disynaptic IPSP be affected, because such a high firing rate would be needed to activate the muscarinic receptors on the interneuronal terminal, but now the effect would be to reduce any background GABA release from that terminal. The overall effect would be a marked increase in the EPSP with higher firing rates in the parabrachial afferent.

Because triadic circuitry is effectively limited to geniculate relay X-cells, Y-cells would not be expected to show these effects.

(c) Response modes of relay cells: tonic and burst firing

(i) Voltage dependency of response mode

All thalamic relay cells are able to respond to excitatory inputs, including driver inputs, in one of two very different response modes, known as tonic and burst. This ability is due to the presence in these cells of voltage-dependent low threshold Ca^{2+} spikes that are based on T (for
transient) type Ca$^{2+}$ channels (reviewed in Sherman 2001; Sherman & Guillery 2001). The voltage dependency of
these channels is as follows (see figure 6a,b). At relatively
depolarized membrane potentials, the channels are inacti-
nated, and they play no role in the neuronal firing proper-
ties. Under these conditions, the cell fires a stream of
unitary action potentials that lasts as long as the supra-
threshold activation, and this is the tonic firing mode
(figure 6a). T-channel inactivation is removed (or, the
channels are de-inactivated) by hyperpolarization, and

\[ \text{Figure 6. Properties of burst and tonic firing. (a,b) Voltage dependency of the low threshold spike for a geniculate relay cell recorded intracellularly in vitro. Responses are shown to the same depolarizing current pulse administered intracellularly but from two different initial holding potentials. With relative depolarization (a), } I_T \text{ is inactivated, and the response is a barrage of unitary action potentials lasting for the duration of the suprathreshold stimulus. This is the tonic mode of firing. With relative hyperpolarization (b), } I_T \text{ is de-inactivated, and the response is a low threshold spike with four action potentials riding its crest. This is the burst mode of firing. (c) Input–output relationship for another geniculate relay cell recorded intracellularly in vitro. The input variable is the amplitude of the depolarizing current pulse, and the output is the evoked firing frequency determined by the first six action potentials of the response, since this cell usually exhibited six action potentials per burst in this experiment. The initial holding potentials are shown: −47 mV and −59 mV reflects tonic mode, whereas −77 mV and −83 mV reflects burst mode. (d) Tonic and (e) burst responses to visual stimulation of a relay cell recorded intracellularly in vivo from the cat’s lateral geniculate nucleus. Shown in each condition are average response histograms to four cycles of a drifting sinusoidal grating (ii) and during spontaneous activity (i). The contrast changes resulting from the drifting grating are shown below the histograms. Current injected through the recording electrode was used to bias membrane potential to more depolarized (−65 mV), producing tonic firing (a), or more hyperpolarized (−75 mV), producing burst firing (b).} \]
Once de-inactivated, they are primed to be activated by the next suprathreshold depolarization. Activation of these channels results in an inward $\text{Ca}^{2+}$ current, $I_T$, which in turn leads to an all-or-none voltage spike, known as the ‘low threshold spike’. This spike, typically, is large enough to fire a high frequency cluster of action potentials, and this is the burst-firing mode (figure 6b).

Switching between modes is accomplished by a suitable shift in baseline membrane potential. However, there is also a time dependency to these shifts, because the inactivation state of the $T$-channels is a complex function of voltage and time (Jahnsen & Llinás 1984; Smith et al. 2000). Roughly speaking, to shift from burst to tonic mode requires a depolarization that is sufficient in amplitude and duration, and the reverse shift from tonic to burst mode requires a similar, sustained hyperpolarization; in both cases the polarization shift must be sustained for more than ca. 100 ms. Both firing modes can be present in the awake animal, although the more alert the animal, the more the tonic firing prevails (Ramcharan et al. 2000; Swadlow & Gusev 2001), but nonetheless the presence of both modes has an impact on the nature of the information relayed.

(ii) Linearity

Figure 6a,b demonstrates that the very same excitatory stimulus produces two very different signals relayed to the cortex, and the difference depends on the initial membrane potential of the relay cell, because this determines the inactivation state of $I_T$. The stimulus in this example is a current pulse, but the same would apply to a sufficiently large EPSP. It is also important to note that the low threshold spike is activated in an all-or-nothing manner (Zhan et al. 1999). One implication of this for the difference in input-output relationships between burst and tonic firing is shown in figure 6c. This relationship is fairly linear for tonic firing, because there is a direct link between the input depolarization and activation of action potentials, leading to the monotonic relationship, as shown. However, the relationship is indirect for burst firing, since it is the low threshold spike that controls firing, and its all-or-nothing nature means that larger activating inputs do not produce larger low threshold spikes and thus do not produce larger responses.

From the cellular properties described above and in figure 6c, it is clear that tonic firing represents a more linear relay mode. This is also seen in the responses of geniculate relay cells to visual stimuli. A clear example is shown in figure 6d,e, which shows the responses to a drifting sinusoidal grating of a relay cell recorded in vivo in an anaesthetized cat. When the cell is in tonic mode, the response to the grating has a sinusoidal profile (figure 6d(ii)). This means that the response level closely matches the changes in contrast, indicating a very linear relay of this input to the cortex. However, when the same stimulus is applied to the same cell, but now in burst mode, the response no longer looks sinusoidal (figure 6e(ii)), indicating considerable nonlinear distortion in the relay. Thus, tonic mode is better at preserving linearity in the relay of information to the cortex (Sherman 1996, 2001).

(iii) Detectability

Figure 6d(i),e(i) shows that spontaneous activity is lower during burst than tonic firing. Higher spontaneous activity helps to preserve response linearity, because it minimizes rectification of the response to inhibitory phases of visual stimulation, and rectification is nonlinear. Perhaps more interesting is the notion that spontaneous activity represents firing without a visual stimulus and can thus be considered a noisy background against which the signal—the response to the visual stimulus—must be detected. Therefore, the signal-to-noise ratio is higher during burst firing, and a higher signal-to-noise ratio implies greater stimulus detectability. This has been confirmed through the use of a method from signal-detection theory involving the calculation of receiver-operating characteristic curves (Green & Swets 1966; Macmillan & Creelman 1991) showing that stimulus detectability is improved during burst firing compared with tonic firing (Sherman 1996, 2001).

(iv) Bursting as a ‘wake-up call’

The above differences in firing modes in respect of linearity and detectability suggest the following proposition (Sherman 1996, 2001). Tonic firing is better for faithful and accurate relay of the retinal input, because it avoids the nonlinear distortions created during burst firing that compromise the accuracy of the messages relayed through the thalamus. Burst firing, however, is better for initial stimulus detectability. As an example, during drowsiness, it might be useful to have geniculate relay cells in burst mode to maximize detection of a new visual stimulus, and after detection, the relay can be switched to tonic firing for more faithful stimulus analysis (for details of this hypothesis, see Sherman 1996, 2001). Consistent with this is evidence from studies of the somatosensory thalamus of awake, behaving rabbits, that relay cells in burst mode are much more likely to activate their cortical target cells than when these relay cells fire in tonic mode (Swadlow & Gusev 2001). Also, Swadlow et al. (2002) have shown in this preparation that burst firing produces much more postsynaptic activity in cortical columns that is much longer lasting than does tonic firing. Nonetheless, it must be emphasized that this notion of bursting as a ‘wake-up call’ remains hypothetical and requires further testing.

(v) Control of response mode

For this theory to be plausible, thalamic circuitry must be capable of controlling firing mode. In fact, the circuitry shown in figure 3 provides this requirement. As noted above, to switch the inactivation state of $I_T$ requires a change in membrane voltage that must be sustained for more than 100 ms: sustained depolarization to inactivate $I_T$ and sustained hyperpolarization to de-inactivate $I_T$. Activation of ionotropic receptors with their fast PSPs is poorly suited to this task, because without extensive temporal summation, the resultant changes in membrane polarization would be too transient to significantly affect the inactivation state of $I_T$. Activation of metabotropic receptors with their slow PSPs is poorly suited to this task, because without extensive temporal summation, the resultant changes in membrane polarization would be too transient to significantly affect the inactivation state of $I_T$. Activation of metabotropic receptors from cortex or muscarinic receptors from the parabrachial region produces a sufficiently long EPSP to
inactivate $I_T$ and switch the firing mode from burst to tonic. Comparably, activation of GABA$_B$ receptors, from activation of reticular and/or interneuronal inputs, produces a sufficiently long IPSP to de-inactivate $I_T$ and switch the firing mode from tonic to burst (for details, see Sherman & Guillery 1996, 2001).

From figure 3, it is clear that the cortical and parabrachial inputs ultimately control firing mode via their direct inputs to relay cells, which promote tonic firing, and their indirect inputs, via reticular and/or interneuronal inputs, which promote burst firing. Both inputs seem to have the same cellular effects. However, the corticogeniculate pathway, as well as its reticular and interneuronal relay, is topographic and purely visual, so that this pathway presumably controls firing mode for discrete geniculate cell populations based on such properties as different locations or different class (i.e. X or Y). The parabrachial input is diffusely organized, suggesting that it has more diffuse effects, such as would be relevant for overall levels of attention (e.g. more bursting exists during states of drowsiness; see Ramcharan et al. 2000; Swadlow & Gusev 2001).

(vi) Relationship of inputs to $T$-channels

An additional insight into the control of firing mode may be gleaned from reconsideration of figure 4. The $T$-channels that underlie $I_T$ are found throughout the neuronal membranes but are more numerous and denser on dendrites, including peripheral dendrites (Zhou et al. 1997; Destexhe et al. 1998). This means that brainstem and interneuronal inputs, which are located proximally near retinal inputs, can affect the spike-generating region of the cell and also the voltage of local membranes containing $T$-channels but are also near enough to retinal inputs to provide a relatively direct influence on the development of retinal EPSPs. By contrast, cortical and reticular inputs are so distally located that they are unlikely to have much direct influence on the spike-generating region or retinal inputs. Instead, they may mainly affect the postsynaptic cell through controlling membrane voltage where voltage-sensitive ion channels, such as $T$-channels, are concentrated.

3. A HIGHER-ORDER RELAY: THE PULVINAR REGION

(a) Afferents to and efferents from the pulvinar region

In general we know less about the organization of the afferents to the pulvinar region than we know about those going to the lateral geniculate nucleus, and here we will focus specifically on the afferents that come from the cerebral cortex and the superior colliculus, looking particularly at those that are likely to be acting as drivers. From the limited data available, it appears that the modulatory inputs to the pulvinar region are generally arranged like those to the lateral geniculate nucleus (i.e. the non-retinal inputs in figure 3; see Feig & Harting 1998). So far as the corticothalamic component is concerned, the critical difference between the lateral geniculate nucleus and the pulvinar region is that the lateral geniculate nucleus receives afferents from cortical layer 6 of the visual cortex (areas 17, 18 and 19) but not from any other cortical layers (Gilbert & Kelly 1975), whereas the pulvinar region receives afferents from layers 5 and 6 of several different cortical areas (Abramson & Chalupa 1985). Correspondingly, whereas corticothermal afferents to the lateral geniculate nucleus all have the appearance and synaptic relationships of RS terminals (see § 2b(iv)), corticothalamic axons to the pulvinar region from two distinct populations: one that resembles the RS axons of the lateral geniculate nucleus and another that resembles the retinogeniculate axons; that is they are RL terminals (Mathers 1972; Ogren & Hendrickson 1979; Rockland 1998).

For all cortical areas where the relationships have been studied, the corticothalamic RL axons come from cortical layer 5, not layer 6, and the RS axons come from layer 6 (Deschênes et al. 1994; Bourassa & Deschênes 1995; Bourassa et al. 1995; Rouiller & Welker 2000; Ojima 1994). We have argued previously that the RL axons in thalamic nuclei are generally drivers, whereas the RS axons are modulators (Sherman & Guillery 1998, 2001). This argument is based, in part, on the morphological resemblance of the axon terminals and their synaptic connections in the thalamus, the RL terminals that come from cortical layer 5 being similar to the retinal afferents in appearance and connections to more proximal sectors of dendrites, often forming triads and arranged in glomeruli (Feig & Harting 1998; Patel & Bickford 1997; Carden & Bickford 2002). In these several respects they also resemble other known driver afferents to the thalamus from the medial lemniscus, the inferior colliculus, the mammillary bodies or the cerebellum (Jones & Powell 1969; Ralston 1969; Jones & Rockel 1971; Harding 1973; Somogyi et al. 1978; Linsky & Kiltsas-Linsky 1990).

The argument that layer-6 afferents are modulators and layer-5 afferents are drivers is also based on the observation that, where a thalamic nucleus is innervated from layer 6 alone, silencing the relevant cortical area does not abolish the receptive field response (see above for the lateral geniculate nucleus, and see Diamond et al. (1992) for the ventral posterior nucleus), whereas the receptive field responses of thalamic cells in nuclei that receive layer-5 afferents can be abolished by silencing the cortex (Bender 1983; Chalupa 1991; Diamond et al. 1992). In what follows, we shall treat the RS axons that innervate the pulvinar region from layer 6 as modulators and the RL axons from layer 5 as drivers. It is important to recognize that, contrary to the claims made by Jones (2002), the terminations of the layer 5 axons in higher-order nuclei are well localized and cannot be regarded as providing a 'diffuse' projection to the thalamus (Deschênes et al. 1994; Bourassa & Deschênes 1995; Bourassa et al. 1995; Rockland 1996; Guillery et al. 2001).

On the basis of this evidence, the pulvinar region can be seen to contain higher-order circuits, that is, circuits that receive their driving inputs from layer 5 of one or more cortical areas and pass this information to other (higher) areas of the cortex. Evidence for comparable higher-order circuits can be seen in several other thalamic nuclei that receive cortical afferents that resemble RL afferents or originate in cortical layer 5 (summarized by Sherman & Guillery 2001; Guillery & Sherman 2002). These transthalamic corticocortical pathways provide potentially important, but often unrecognized, pathways for corticocortical communication. Their relationship to
the more widely studied direct corticocortical routes (Van Essen et al. 1992; Kandel et al. 2000) is almost entirely unexplored. Since we know that, essentially, all areas of the neocortex receive a thalamic input, and that for cortical areas that have been studied in most detail (e.g. visual cortical area 17 and auditory cortex) the thalamic afferents provide the key driver afferents, it is reasonable to expect that the thalamocortical pathways from higher-order thalamic relays such as the pulvinar region will make a significant contribution to the functional organization of the higher cortical areas that they innervate. For this reason, it becomes important to define the nature of the driving afferents that a higher-order relay like the pulvinar region receives. We need to identify the driving afferents that come from the cortex to determine whether there are any subcortical driving afferents, and then to relate the distribution of these inputs to the pattern of thalamocortical outputs. We have seen that layer-5 inputs to the pulvinar region arise from several different cortical areas (Abramson & Chalupa 1985). These include visual areas 17, 18 and 19, and several more lateral areas related to the suprasylvian sulcus in the cat. In addition to these cortical driving afferents, a second source of driving afferents to the pulvinar region may come from the tectum and pretectum.

For the tectal and pretectal afferents, a critical issue that remains unresolved is whether some, all, or none of these afferents are, indeed, driving afferents. If any of them are drivers, then the pulvinar region would have to be considered as a mixture of first-order and higher-order circuits. Some fine structural studies describe tectopulvinar terminals as R_L, and therefore as putative drivers (see § 2b(5)), whereas others show RS type terminals, putative modulators (Mathers 1971; Partlow et al. 1977; Robson & Hall 1977). Experimental evidence about the functional role of the tectopulvinar afferents suggests that, if there are drivers present, they may not be making a major contribution. That is, when cortical areas that innervate the pulvinar region are silenced, this is generally reported as a successful method of abolishing the characteristic receptive field responses of the cells in the pulvinar region or of other cortical areas to which the pulvinar region projects, and this is in contrast to silencing tectum, which has less dramatic effects on the receptive field properties of cells in the pulvinar region (Chalupa et al. 1972; Bender 1983; Chalupa 1991). The evidence that some tectopulvinar axons may be RL (that is driver type) terminals comes from fine structural studies of degenerative changes produced by tectal lesions. Since the corticopulvinar axons that come from layer 5 and have RL terminals also send large branches to the tectum (see Guillery et al. 2001), the possibility that some retrograde changes produced by the tectal lesions may have affected the appearance of the RL terminals in the pulvinar region indirectly has to be considered as a possible explanation of the apparently contradictory fine structural evidence about the tectopulvinar pathway. Apart from the tectal and pretectal inputs, there are no other known sources of pulvinar afferents that might serve to drive cells in the pulvinar region and provide their characteristic receptive field properties. The very small direct retinal input to the pulvinar demonstrated by Iyaya & Van Hoesen (1983) and Nakagawa & Tanaka (1984) are not relevant candidates, not only because they provide a very small and scattered input that would only rarely be encountered by a recording electrode in the pulvinar, but also because, as direct retinal afferents, they would not provide the receptive field properties that resemble cortical rather than retinal cells.

The cortical afferents from layer 5 must be regarded as at least one key driving input to the pulvinar region and will be the focus of the following discussion. Recordings from cells in the pulvinar region are readily compared with recordings from cells in layer 5 of the visual cortex. The receptive field properties resemble those of complex cells in the visual cortex, are binocular, and are orientation and direction selective (Mason 1981; Casanova et al. 1989; Chalupa & Abramson 1989; Merabet et al. 1998), confirming that an important driving input is coming from the visual cortex. There is some indication that the response properties vary somewhat with position in the pulvinar region (Mason 1981; Chalupa & Abramson 1989), but in view of the great variety of cortical sources of layer-5 afferents to the pulvinar region, the relative uniformity of responses is surprising. It is possible that details of some of the relevant variables of the response properties characterizing the several distinct cortical inputs still remain to be defined.

Efferents from the pulvinar region have a widespread distribution to several cortical areas. In the cat, axons have been traced to areas 17, 18, 19, 21a, and to both banks of the middle and the lateral suprasylvian sulci (Symonds et al. 1981; Abramson & Chalupa 1985; Miceli et al. 1991). Whereas the projections to areas 17 and 18 pass to layer 1 of the cortex, the other cortical areas receive these afferents in layer 4. Similarly, in the monkey there is a widespread distribution of axons from the pulvinar region to the cortex, with cortical areas that receive pulvinar afferents including V1, V2, V4, and MT and area 7a of the posterior parietal cortex (Rockland et al. 1999; Darian-Smith et al. 1999; Adams et al. 2000). The axons distribute primarily to layer 3, but also go to adjacent layers and some go to layer 1. In the cat and the monkey, there is evidence that single cells from the pulvinar region can have branching axons that go to more than one cortical area (Kaufman et al. 1984; Miceli et al. 1991).

It is reasonable to conclude that in the areas receiving afferents from the pulvinar region, important aspects of cortical function will depend on these thalamic inputs. In order to understand the information that the pulvinar region is transmitting to the cortex it will be necessary to define the nature of the messages that layer-5 cells in several functionally distinct cortical areas are sending to this region, and then to define how any one set of inputs relates to each of the several outputs. Although we are far from understanding these relationships, defining them will, in the long run, depend on understanding the ground rules that govern the internal organization of the pulvinar region.

(b) The functional organization of the pulvinar region

Our current view of the functional organization of the relay in the pulvinar region is limited because we do not have sufficiently detailed information about the afferent and efferent connections of the different subdivisions of
the pulvinar region. In order to understand the nature of the messages that are relayed through the pulvinar region from one cortical area to another, we need to understand how these messages relate to the functional properties of the cortical layer-5 cells that innervate the relay cells of the pulvinar region from several different cortical areas, and how those properties, in turn, relate to the functional properties of the cortical areas that receive an innervation from the pulvinar region. To a significant extent, lack of agreement as to how the region should be subdivided has prevented a systematic study of connectivity patterns.

Subdivisions of the pulvinar region have been based on a number of different experimental approaches. Cytoarchitectonic and immunohistochemical studies have shown that different regions have distinctive staining properties (Berson & Graybiel 1983; Gutierrez et al. 2000; Adams et al. 2000); connectional studies have related cortical, tectal and pretectal afferents to identifiable subdivisions (Updyke 1983; Shipp 2001), and have demonstrated distinctive, topographically organized maps within some of these pathways; maps based on receptive field position in the visual fields have been plotted, showing that there are several more or less complete representations of the contralateral visual hemifield in the pulvinar region, each representing a functionally distinguishable entity, that can often, but not invariably, be related to the connectional studies. The relationships have been considered in detail for the cat by Updyke (1983), who based his analysis on cytoarchitectonic studies, visual field maps, and on detailed plots of the topographically mapped inputs from several cortical areas to the proposed subdivisions. He identified 'isolocation columns' in the pulvinar region as columns of cells that connected to a shared cortical locus in terms of the cortical afferents. Examples of these columns (called 'isocortical columns' by Guillery et al. 2001) are shown schematically in figure 7 for two subdivisions of the cat's pulvinar region, the pulvinar nucleus and the lateral part of the lateral posterior nucleus, and are there compared with the projection columns of the lateral geniculate nucleus. Comparable relationships between visual field maps and the subdivisions of the pulvinar region have been more difficult to define for the monkey (Adams et al. 2000), although Shipp (2001) has recently proposed a somewhat comparable method of analysis.

It has to be stressed that the identification of isocortical columns (Guillery et al. 2001), which are roughly equivalent to Updyke's isolocation columns (Updyke 1983) or to Shipp's lines of isorepresentation (Shipp 2001), provide a useful framework for a connectional analysis, but that there are a number of difficulties about using them as a basis for a rigorous, detailed analysis of connectional patterns in the pulvinar region. This is because the identification of these columns is based on several different experimental approaches, and the different names that have been used reflect this. Guillery et al. (2001) used an analysis of layer-5 driver afferents that come from the cortex, and these show a rough scatter around the lines defined by Updyke (1983) who had correlated visual field maps with maps of corticothalamic axon terminations based on autoradiographic tracing methods. That is, Updyke's tracing experiments included both drivers (from layer 5) and modulators (from layer 6). The layer-6 inputs to the pulvinar region are more scattered around the isocortical columns than are the layer-5 inputs (Guillery et al. 2001; and figure 7), although for some corticothalamic pathways both show roughly comparable distributions. Further, where both types of corticothalamic afferent arise from the same small cortical area, they terminate in largely non-overlapping but adjacent zones (Guillery et al. 2001). Shipp's (2001) analysis, using horseradish peroxidase as a combined anterograde and retrograde tracer, introduces a further complexity, since his analysis is based on a mixture of layer-5 and layer-6 corticothalamic afferents and also on the retrogradely labelled thalamic cells. His study makes no distinctions between the projection patterns of the two types of corticothalamic axon and the thalamocortical pathways. The assumption that all are in register may be justified in very general broad outlines, but will break down when details of connectional patterns are needed.

The problem of defining exactly how the input pathways to the pulvinar region relate to the cortical distribution of the thalamocortical outputs to several different cortical areas is entirely undefined at present, and must be regarded as one of the key issues that needs to be addressed if we are to understand how cortical areas communicate with each other through the thalamus. It is not enough to show how cells in the pulvinar distribute their axons to different cortical areas and to different cortical laminae. For any one population of individual cells in the pulvinar region that is identified on the basis of a projection to a particular cortical area, or where the thalamocortical axons branch to a group of areas, it will be essential to show the nature of the cortical afferents (driver primarily but also modulator) to the same cells.

Figure 7 illustrates an isocortical column passing obliquely through the lateral part of the lateral posterior nucleus of the cat and shows the distribution of driver and modulator afferents from areas 17, 18 and 19. Inputs from other cortical areas are known, but are not illustrated in figure 7. The most rostral and dorsal parts of the column receive no driver afferents from areas 17, 18 or 19, only modulators from area 19. There is evidence for projections from other cortical areas, including parietal cortex (areas 5 and 7; see Heath & Jones 1971; Kawamura et al. 1974; Robertson & Cunningham 1981) to this part of the nucleus, but at present it is not known whether these are from layer 5 or layer 6. The most caudal and ventral parts of the column receive inputs from visual cortical areas of the lateral suprasylvian cortex (Updyke 1983) and these represent layer-5 as well as layer-6 afferents (Abramson & Chalupa 1985); some of those coming from area PMLS have the fine structural characteristics of the layer-5 driver afferents (S. Feig, B. K. August and R. W. Guillery, unpublished observations). An important feature of the connections illustrated in figure 7 is that the layer-5 inputs from cortical areas 17, 18 and 19 are mingled within the ventral and posterior two thirds of the illustrated isocortical column (purple), but that cells in the column as a whole can be compared with cells that are grouped around a single line of projection in the lateral geniculate nucleus. Although there is no lamination in the lateral posterior nucleus, the nature of the input changes, as for the lateral geniculate nucleus, from one end of the column to the other. This is represented in the lateral posterior nucleus by the changing cortical origin of the afferents and in the

Phil. Trans. R. Soc. Lond. B (2002)
lateral geniculate nucleus by the change in the functional type of the retinal afferents (X, Y, W). It is also represented by the structure of the terminals. In each nucleus, the more dorsal terminals have a relatively open structure, with terminal swellings widely spaced, but have a more tightly packed structure ventrally near the optic tract. For the lateral posterior nucleus, this applies to all three of the inputs illustrated in figure 7. That is, in terms of functionally distinct afferents to the lateral posterior nucleus we have to consider two types of functional distinction. One is the difference between cortical areas. We expect the layer-5 output cells of each cortical area to have a more or less distinctive set of functional properties, although there may well be significant overlap from one area to the next. The other distinction that may also play an important role is in the different functional characteristics that may characterize the layer-5 output cells of any single cortical area. There is no reason for thinking that all layer-5 cells that project to the lateral posterior nucleus from any one cortical area will have the same functional properties. The fact that the columns in the lateral posterior nucleus are not uniform all along their length, either in terms of cortical afferents or in the structure of the terminals, suggests that any one cortical area may resemble the retina in having a mosaic of functionally distinct layer-5 cells, which, like the retinal ganglion cells, project to well defined subdivisions of any one column. This is currently only speculation, but it could be useful if it stimulates a search for functional distinctions that characterize either the cells of layer 5 or the properties of the cells in the lateral posterior nucleus along the axis of any one column.

4. CONCLUSIONS

The complex cell and circuit properties of the lateral geniculate nucleus leave little doubt that the relay of retinal information to the cortex is an active, mutable process. Insofar as thalamic relays share a common pattern of axonal types and synaptic connections (Sherman & Guillery 2001), generalizations about dynamic relay properties gleaned mostly from the lateral geniculate nucleus will apply to most thalamic relays, first-order as well as higher-order. We have offered some specific suggestions about how circuit properties control a voltage-dependent conductance, $I_T$, in relay cells to control responsiveness, and...
how this could affect the nature of information relayed to the cortex. However, it should be appreciated that this property, related to tonic- and burst-response modes, may be one of many mechanisms by which thalamic relays can control the flow of information to the cortex and that \( I_T \) is but one of many voltage-dependent conductances that is under the control of modulatory afferents (for a more complete description of these conductances, see Sherman & Guillery 2001).

The lateral geniculate nucleus is a first-order relay, meaning that it is responsible for transferring subcortical (i.e. retinal) information to the cortex. By contrast, much, or all, of the pulvinar region is a higher-order relay, transferring information between cortical areas. Thus, much of corticocortical communication involves a route through the thalamus. This provides corticocortical communication with the same advantages that the thalamus provides for the relay of retinal information to the cortex. The alternative route for corticocortical communication—direct connections among areas—needs to be reconsidered since possibly many, and perhaps all, are modulatory in nature, and the transthalamic pathway can provide the main information transfer. Thus, the full impact of the thalamus may be much more than simply controlling flow of information from the periphery and from other parts of the brain to the cortex: it may remain an active partner in all cortical computations. In order to understand the functional organization of the transthalamic pathway it will be necessary to define how the connections and the functional properties of the cortical inputs to higher-order relays relate to the thalamocortical afferents. For the pulvinar region we show that afferents from several distinct cortical areas are intermingled within connectionally defined columns. We compare the organizational principles that emerge from a study of cortical afferents to these columns in the pulvinar region with those that characterize the retinal input to the lateral geniculate nucleus.

ENDNOTES

1The thalamic region involved with the extra-geniculate relay of visual information has confusingly different names across species. For instance, it includes the lateral posterior nucleus and the pulvinar in the cat, which may correspond to much or all of the pulvinar in a primate. For simplicity, we shall refer to this thalamic region as the ‘pulvinar region’.

2We have been asked about indirect evidence for non-cortical driver afferents to the pulvinar based on observations on the cortical area known as MT (for middle temporal). Here, Rodman et al. (1989, 1990) found that lesions of the primary visual cortex or of the superior colliculus alone produce only moderate changes in the receptive fields of the cortical cells, whereas combined lesions of both render MT cells silent. The connectional basis of these changes is not known. One possibility is that the small direct afferent component from the retina to the pulvinar (Itaya & Van Hoesen 1983; Nakagawa & Tanaka 1984) provides the extra-geniculostriate input to area MT and that the tectum provides a crucial modulatory role. Alternative explanations are also possible, but as the tectal lesions alone do not affect the cortical responses in MT, and as we do not know the source of the extra-striate afferents to MT, we cannot regard the evidence about receptive fields in MT as critical for understanding the nature of the tectal inputs to the pulvinar.

REFERENCES


Darian-Smith, C., Tan, A. & Edwards, S. 1999 Comparing
thalamocortical and corticothalamic microstructure and spatial reciprocity in the macaque ventral posterolateral nucleus (VPLc) and medial pulvinar. J. Comp. Neurol. 377, 535–549.


Glossary

EPSP: excitatory postsynaptic potential

IPSP: inhibitory postsynaptic potential

PSP: postsynaptic potential