INTRODUCTION

Auditory stimulation at different frequencies elicits an electroencephalographic (EEG) steady-state auditory response (SSAR) that peaks at the stimulation frequency. In humans it has the greatest amplitude around 40 Hz (Galambos et al. 1981). This oscillatory response appears to be generated in the supratemporal gyrus of the auditory cortex and modulated by thalamocortical systems according to both magnetoencephalography (MEG) studies and brain electrical source analysis (BESA) (Herdman et al. 2002; Makela and Hari 1987; Steriade et al. 1991). The origin and significance of the 40-Hz SSAR is unclear, although some authors discuss it in relation to sensory coordination of neuronal responses. Although fMRI has a low temporal resolution, the underlying model is based on hidden brain structures that play the role of relay stations.

We have previously shown an increase in cortical synaptic activity with 40-Hz stimulation in the auditory cortex, posterior superior temporal gyrus (STG), and superior temporal sulcus (STS) and bilateral activation of the cerebellar hemispheres using regional cerebral blood flow (rCBF) and positron emission tomography (PET) (Pastor et al. 2002). The activated areas were in the posterolateral portion of both cerebellar hemispheres, lateral to the paravermian region, Crus II (Schmahmann et al. 1999). Connectivity studies in the cat have shown that the bulk of afferents to this area originate in the temporal lobe with a relay in the pontine nuclei, before reaching the cerebellar cortex. In the primate, there are no direct projections from the primary auditory area; instead, cortico–pontine auditory fibers originate in the secondary auditory area A2 and adjacent association areas, although the most important cortico–ponto–cerebellar afferents are from multimodal areas in the upper bank of the STS (Schmahmann and Pandya 1991). These neurons project to the dorsolateral and lateral nuclei of the pons, which, in turn, project to the cerebellar area activated in our study (Brodal 1979). The cerebellum then returns connections to the cortex via the thalamus. These complete an auditory cortico–cerebellar–thalamic loop.

METHODS

Subjects

Ten right-handed healthy volunteers (mean age, 32.2 yr; SD, 5.6), with no neurological or hearing deficits, were studied, with approval of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
from the Institute of Neurology and National Hospital for Neurology and Neurosurgery Joint Ethics Committee (University College of London Hospital Trust). Subjects gave informed written consent after explanation of the experimental procedure.

**Stimuli**

Stimuli were trains of clicks delivered binaurally through headphones. The auditory stimuli were generated by Cogent 2000 software (Wellcome Trust Centre for Neuroimaging, London, UK). Stimulation comprised three types of click trains (three different frequencies: 40, 26, and 12 Hz) at 95-dB intensity. None of the frequencies included harmonics of the scanner noise (Fig. 2). Null events comprised only the scanner noise. To control attentional set, subjects were asked to make a motor response to white noise bursts. The trains (2-s duration) had the same sonority and the duration of every click was inversely correlated to frequency; clicks in the 40-Hz train were 1 ms; 26 Hz = 1.53 ms; and 12 Hz = 3.3 ms.

**fMRI study**

**EXPERIMENTAL FMRI DESIGN.** Each subject underwent three fMRI scanning sessions. We used an event-related design with continuous scanning. The subjects were naïve to the aim of the experiment and were instructed before the scanning, without a training period. During each session, each subject received 18-s epochs of 2-s click trains of the same frequency, intermingled with null events and sparse bursts of noise (Fig. 3). Each session comprised a series of epochs, in which the three frequencies (12, 26, and 40 Hz) were intermingled in a pseudo-random sequence.

A central fixation cross was displayed in front of the subject. Subjects were required to maintain central fixation throughout each session, and this was monitored with an eye-tracking device.

**DATA ACQUISITION.** Imaging was performed using a 3-Tesla head scanner (Siemens Allegra, Erlangen, Germany) equipped with a head volume coil. T1-weighted structural images were acquired using the following combination = 1 mm isotropic; field of view = 240 × 256 mm²; matrix = 240 × 256; 176 slices; repetition time (TR)/echo time (TE)/inversion time (TI) = 7.92/2.4/910 ms; flip angle = 15°; bandwidth = 195 Hz/pixel. The functional images, sensitive to blood oxygenation level–dependent contrast, were acquired by T2*-weighted echoplanar imaging. Each volume consisted of 38 transverse slices (3-mm thickness; 0.75-mm gap; matrix = 64 × 64, 3 × 3 mm pixels; TE = 65 ms) covering the whole brain. A total of 279 sequential volumes were acquired per session, with an effective scanning time of 34 min. To avoid systematic interactions between slice acquisition and stimulus presentation, the stimulus onset was randomly jittered with respect to the beginning of each volume acquisition.

**FMRI ANALYSIS.** We used the statistical parametric mapping (SPM2) software for image processing and analysis (http://www.fil.ion.ucl.ac.uk/spm/spm2.html). For each subject and session, the first four volumes were discarded to allow for T1 equilibration. The remaining 825 (3 × 275) volumes were realigned to the first image, sinc-interpolated over time to correct for phase-advance during volume acquisition, coregistered to the anatomical scan, and normalized to the Montreal Neurological Institute reference brain. The data were smoothed spatially with a Gaussian Kernel (8-mm full-width at half-maximum). We used a conventional (fixed-effect) SPM analysis and a DCM analysis for each subject. In the first (fixed-effects) analysis, the five event types (i.e., three frequencies, null events, and the noise targets) were modeled using appropriate stimulus functions convolved with a canonical hemodynamic response function (HRF). For each subject, we tested for the overall effect of listening to trains of clicks versus the background noise of the scanner using a multidimensional F-contrast. This contrast tests for the main effect of frequency and any interactions. This contrast highlights brain areas involved in repetitive auditory stimulation processing. The ensuing SPM was used to define the location of subject-specific regions that entered the DCM analysis (Supplemental Table S1)¹ (Friston et al. 2003).

To confirm the selective 40-Hz activation of the cerebellum, we tested for activations during the 40-Hz epochs, relative to the remaining two frequencies. The SPM results of this analysis are shown as maximum intensity projections and on a slice through the cerebellum (hemispheric clusters x, y, z: −40, −78, −30 and 38, −82, −32) (see Fig. 4I). The threshold used for display was P = 0.01 (uncorrected). The peak response in the cerebellar regions reached an uncorrected P < 0.0001 and survived a small volume correction at P < 0.05, using a spherical search volume of 16 mm centered on the location of the orthogonal main effect of all frequencies in the cerebellum (from the preceding SPM analysis).

¹ The online version of this article contains supplemental data.
We tested the hypothesis that 40-Hz-selective responses in the cerebellar Crus II auditory region could be explained by a selective enabling of connections in the auditory cortico–thalamic loop. To characterize cortico–cerebellar interactions we evaluated three DCMs that allowed for distinct frequency-specific modulation of different connections. As described in Friston and colleagues (Friston et al. 2003; Penny et al. 2004), each DCM is characterized by three sets of parameters: the $A$ parameters, which specify which regions are connected, are called endogenous connections; the set of $C$ parameters specify the influence of exogenous inputs (frequency-specific auditory stimulation) on each region; and, finally, the $B$ or bilinear parameters, which specify how endogenous connections are changed by exogenous auditory inputs. These encode as bilinear or modulatory effects. Together, these parameters characterize the effective connectivity or architecture of a DCM.

In our study, the endogenous connections were specified as follows: first, bidirectional connections between auditory cortex and the medial geniculate complex of the thalamus. These were based on the fact that almost all of the projections of the thalamic nuclei to the cortex are reciprocated by cortico–thalamic fibers (Ramón y Cajal 1911); second, a unidirectional connection between auditory cortex and Crus II (Huang et al. 1991); and third, between Crus II and the thalamus (Schmahmann and Pandya 1997). We selected the most significant voxels in each subject’s SPM$_{F}$, in auditory STG/STS, Crus II in the cerebellum, and medial geniculate nuclei to obtain the center for regional volumes of interest. The volume of interest for the auditory cortex included the planum temporale, the STG, and adjacent multisensory integration areas (STS). In four individuals, STG showed the maximum activity during stimulation and, in two, the voxel of maximum activity was in STS (Fig. 4II). These three regions, with the individual voxel maxima displayed as a point in Talairach coordinates for each of the ten subjects, are shown in Fig. 4II, A–C. The activities in these regions were summarized with the principal eigenvariate of responses in voxels within 8 mm of the respective centers.

The specification of the modulatory or frequency-specific ($B$) connectivity varied according to three different DCMs. In the first DCM, the different frequencies were allowed to change the connections from auditory STG/STS to Crus II (Fig. 1A, Model 1). This model reflects the evidence for selective auditory cerebellar activation at 40-Hz induced oscillatory activity, which supports the inclusion of the auditory cerebellum in the network of cortical oscillatory-induced responses (Pastor et al. 2002). Activations with similar location in the cerebellum are found in PET studies of temporal auditory processing (Griffiths et al. 2000; Lockwood et al. 1999; Penhune et al. 1998; Ramnani et al. 2000). The auditory cortico–ponto–cerebellar projection, in contrast to other cortical inputs, projects to the ipsilateral Crus II (Brodal 1983). In the second (Fig. 1B, Model 2), we also allowed modulation of the connection from the medial geniculate body to the auditory cortex. The model allows for the participation of cortico–thalamo–cortical pathways (Van Horn and Sherman 2007). Finally (Fig. 1C, Model 3), we allowed for further frequency-specific changes in the connection from Crus II to the medial geniculate body, based on the evidence that disruption of auditory cerebellar input to the medial geniculate body reduces induced cortical oscillatory activity (Pastor et al. 2006). In all models, each frequency also served as an

---

**FIG. 2.** *Left:* depicts the frequency spectrum of echo planar imaging gradient waveforms. *Right:* a detailed look at the low-frequency spectrum, showing the relation with the auditory frequencies used in the experiment.

**FIG. 3.** Schematic of the experimental design; showing the temporal sequence of click trains during the scanner noise, intermingled with white noise bursts and ensuing motor responses.
exogenous or direct input into the medial geniculate body of the thalamus (C parameters).

In the DCM analyses, the exogenous inputs were the three frequency-specific stimulus functions used to form the regressors for the preceding SPM analyses.

**Model comparison**

For each subject we performed Bayesian model comparison using the evidence of each model (probability of the data given the model). A ratio of the evidence for two models (or difference in log-evidence) provides evidence for superiority of one of the three models over another, in terms of both accuracy and complexity (see Penny et al. 2004). In addition, using the best DCM, we also performed a classical (between-subject) analysis, using the coupling parameters as summary statistics in nonparametric significance tests. We used these tests to show our conclusions can be generalized to the population from which our subjects came.

**RESULTS**

**Model comparison**

We first report the results of the model comparison and then turn to a quantitative analysis of the effects under the best model. For all subjects, model 1—with modulation of the connection from auditory STG/STS to Crus II—was superior to models 2 and 3, according to the difference in log-evidences or log-Bayes factors (see Table 1). This suggests a remarkable consistency over subjects. The Bayes factor \( B_{ij} \) indicates that model \( i \) is significantly more likely than model \( j \), when its value is \( >20 \) (or \( >3 \) when using the log-Bayes-factor); i.e., the data are 20 times more likely under one model relative to the other. To compute the log-Bayes factors for each subject, we simply took the sum over the three sessions (this corresponds to multiplying the marginal likelihoods and is motivated easily by the fact the data came from independent sessions). The results of this model comparison are shown in Table 1. Note that the simplest model (model 1) has the greatest log evidence, in all subjects. We confirmed this using Akaike’s and Bayesian information criteria. This means that the additional complexity of adding modulatory effects to the DCM could not be compensated for by an increase in accuracy or fit to the data.

According to Table 1, model 1 is superior to models 2 and 3, which suggests that frequency modulates only the auditory STG/STS to Crus II connection. Therefore subsequent analyses were based on this model, using a between-subject analysis of

![Fig. 4. I: axial, coronal, and sagittal sections of the canonical space showing activations during the auditory 40-Hz epochs, relative to the remaining 2 frequencies. The statistical parametric mapping (SPM) results are shown as a slice through the cerebellum Crus II (Hemispheric clusters x, y, z: −40, −78, −30 and 38, −82, −32). SPM threshold \( p_{FWE} \) corrected = 0.05. II: axial, coronal, and sagittal sections of the canonical T1 MR template depicted as boundaries of the group of 10 individual activations in auditory STG/STS (A), the medial geniculate body (B), and cortical cerebellar hemisphere (C), used for DCM volumes of interest.](http://jn.physiology.org/)

**TABLE 1. Log-Bayes factors for each subject**

<table>
<thead>
<tr>
<th>Subject</th>
<th>( \ln B_{12} )</th>
<th>( \ln B_{13} )</th>
<th>( \ln B_{23} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S01</td>
<td>50.499584700729571</td>
<td>101.0504919158240</td>
<td>50.550907215094412</td>
</tr>
<tr>
<td>S02</td>
<td>50.285374210039790</td>
<td>100.8363549212949</td>
<td>50.550991251125424</td>
</tr>
<tr>
<td>S03</td>
<td>50.453069426240461</td>
<td>101.0039600631356</td>
<td>50.550890653915597</td>
</tr>
<tr>
<td>S04</td>
<td>50.499758532289411</td>
<td>101.0507195737283</td>
<td>50.550964722099309</td>
</tr>
<tr>
<td>S05</td>
<td>49.917815360998532</td>
<td>100.4685086597841</td>
<td>50.550691498785612</td>
</tr>
<tr>
<td>S06</td>
<td>49.981878187716745</td>
<td>100.5254082724381</td>
<td>50.54653084721326</td>
</tr>
<tr>
<td>S07</td>
<td>50.481854324116910</td>
<td>101.0327686489922</td>
<td>50.550914324875237</td>
</tr>
<tr>
<td>S08</td>
<td>50.492187938936695</td>
<td>101.0431228173830</td>
<td>50.55093487446299</td>
</tr>
<tr>
<td>S09</td>
<td>50.480668115902148</td>
<td>101.0316095570481</td>
<td>50.550941441145909</td>
</tr>
<tr>
<td>S10</td>
<td>50.334221433835182</td>
<td>100.8851178843870</td>
<td>50.550800354603553</td>
</tr>
</tbody>
</table>
The same test was performed over B values obtained for 12, 26, and 40 Hz; i.e., the distributions for each frequency were compared with zero. The only statistically significant result was at 40-Hz stimulation frequency, with a $P$ value of 0.0273 (Table 3B). Therefore we conclude that 40-Hz-selective responses in cerebellar Crus II auditory region are most likely to be due to a selective enabling of connections from Auditory STG/STS to Crus II.

To summarize, our results show a clear superiority of model 1 over models 2 and 3 (this was completely consistent over subjects). This indicates a special role for the cerebellum, in relation to the auditory task of our experiments; second, under this model there is a significant, frequency-selective modulation of the Auditory STG/STS to Crus II connection during 40-Hz processing.

**DISCUSSION**

The present study reveals first, that in humans, 40-Hz auditory input engages the coupling between auditory area-STG/STS and Crus II in a selective fashion. There is increasing evidence that the auditory steady-state response at 40 Hz represents an induced oscillatory brain activity rather than an evoked response (Ross et al. 2005). The anatomical substrate comprises the cortico–ponto–cerebellar projections, part of a closed loop with the cerebral cortex, in which the cerebellum returns projections to the cerebral cortex via the thalamus. The systems interacting with the cerebellum are not just somato-sensory and motor but include visual and auditory loops (Middleton and Strick 1994; Schmahmann and Pandya 1997).

It is interesting to note that the participation of cortico–thalamic loops in stimulus-induced oscillatory brain activity has been described previously (Ribary et al. 1991) in a more generic context. Furthermore, in PET studies of steady-state auditory responses we found, in addition to auditory cortex and cerebellar activation, that thalamic rCBF activation correlated with the SSAR amplitude (Pastor et al., unpublished results). Using MEG, the generator of SSAR has been located in auditory cortex (Engelien et al. 2000; Herdman et al. 2003; Pantev et al. 1996; Ross et al. 2002). The participation of the cerebellum in auditory-induced activity has not been studied by EEG and MEG techniques because they are relatively insensitive to cerebellar sources.

MEG does not see deep sources easily and the contribution of cerebellar sources remains unclear. Ross et al. (2005) characterized the recovery of the 40-Hz SSAR after a concurrent brief burst of noise, 200–300 ms poststimulus, with a 4- to 6-ms latency shift. A resonance state driving source at 6-ms latency admits several synapses, and thus it could have thalamic or cerebellar origin. Previous studies of SSAR reset (Ross and Pantev 2004), after a gap duration as small as 3 ms, suggest that the resetting mechanism is sensitive to changes in the auditory modality or even in other sensory modalities (Makeig and Galambos 1989; Rohrbaugh et al. 1990). Many authors conclude that the SSAR rests on a number of separate neural oscillations (Ross et al. 2005). Our finding of 40-Hz-selective responses in the cerebellar Crus II auditory region, explained by a selective enabling of connections in the auditory cortico–thalamic loop, is consistent with this notion. Using whole-scalp MEG during SSAR synchronization in the auditory cortex, medial parietal cortex and thalamus waveforms...
was found; however, the waveform morphology for activity attributed to cerebellum was distinct (Bish et al. 2004). This interesting finding needs further clarification to assess whether activity in the sensorimotor cerebellum could have concealed or overlapped the auditory cerebellar activity, during click-train stimulation.

We have previously shown an increase in synaptic activity in primary auditory cortex, STG/STS, and bilateral activation of the cerebellar hemispheres at 40-Hz simulation using measurements of rCBF. The magnitude of the rCBF increment was 8% in auditory cortex, comparing 40 with 12 Hz, and 5% comparing 40 with 30 Hz (Pastor et al. 2002).

Furthermore, we found that interference with cerebellar output, by repetitive transcranial magnetic stimulation, modifies functional responses associated with cortical auditory processing. The finding of greatest effects on the 40-Hz SSAR supports the notion that the cerebellar cortex has a role in the regulation of auditory cortical oscillatory activity (Pastor et al. 2006). In the cerebellum, the characteristic 30- to 40-Hz frequency of Purkinje cell firing, mediated by the mossy fiber system, can be modulated by the other afferent system, the climbing fibers (Lou and Bloedel 1992; Sato et al. 1992). Oscillations, including the gamma band (from 10 to 50 Hz) have been recorded over the cerebellar surface of humans, locked to electrical somatosensory stimulation independent of motor components (Teschke and Karhu 2000).

Our event-related fMRI experiment suggests the input from auditory cortex to the cerebellar hemisphere through cerebro–pontine pathways is conveyed, preferentially, at gamma-band frequencies. In other words, the cerebellum gates cortical output in the cortical–cerebellar–thalamic loop to preferentially boost 40-Hz responses. It is unlikely that any enhanced oscillatory activity reflects just the propensiveness of some brain regions to resonate at this frequency (Kapoor et al. 1991). The paradigm we used required subjects to attend to the stimuli: the correspondence between SSAR oscillatory activity and transient 40-Hz responses (Titiinen et al. 1993) points to a possible role in selective gating cortical output in the cortical–cerebellar–thalamic pathways in general. The cerebellum receives monaural auditory input, processes it bilaterally in the lateral hemispheres, and integrates visual signals in neighboring hemispheric areas (Pastor et al. 2003). The cerebellum, driven by the auditory STG/STS, fulfills the role of an epicenter within the attentional network that may modulate ongoing cortico–thalamic oscillatory activity, in this case the generation of the SSAR.

GRANTS

This work was supported by the Wellcome Trust. M. A. Pastor was supported by Ministerio de Educación, Cultura y Deporte/Fulbright Grant EX2001 46214114.

REFERENCES


