Zoned mantle convection

By Francis Albarède\(^1\) and Rob D. van der Hilst\(^2\)

\(^1\)Ecole Normale Supérieure de Lyon, 69007 Lyon, France
\(^2\)Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Published online 27 September 2002

We review the present state of our understanding of mantle convection with respect to geochemical and geophysical evidence and we suggest a model for mantle convection and its evolution over the Earth’s history that can reconcile this evidence. Whole-mantle convection, even with material segregated within the \(D^0\) region just above the core–mantle boundary, is incompatible with the budget of argon and helium and with the inventory of heat sources required by the thermal evolution of the Earth. We show that the deep-mantle composition in lithophilic incompatible elements is inconsistent with the storage of old plates of ordinary oceanic lithosphere, i.e. with the concept of a plate graveyard. Isotopic inventories indicate that the deep-mantle composition is not correctly accounted for by continental debris, primitive material or subducted slabs containing normal oceanic crust. Seismological observations have begun to hint at compositional heterogeneity in the bottom 1000 km or so of the mantle, but there is no compelling evidence in support of an interface between deep and shallow mantle at mid-depth.

We suggest that in a system of thermochemical convection, lithospheric plates subduct to a depth that depends—in a complicated fashion—on their composition and thermal structure. The thermal structure of the sinking plates is primarily determined by the direction and rate of convergence, the age of the lithosphere at the trench, the sinking rate and the variation of these parameters over time (i.e. plate-tectonic history) and is not the same for all subduction systems. The sinking rate in the mantle is determined by a combination of thermal (negative) and compositional buoyancy and as regards the latter we consider in particular the effect of the loading of plates with basaltic plateaux produced by plume heads. Barren oceanic plates are relatively buoyant and may be recycled preferentially in the shallow mantle. Oceanic plateau-laden plates have a more pronounced negative buoyancy and can more easily founder to the very base of the mantle. Plateau segregation remains statistical and no sharp compositional interface is expected from the multiple fate of the plates.

We show that the variable depth subduction of heavily laden plates can prevent full vertical mixing and preserve a vertical concentration gradient in the mantle. In addition, it can account for the preservation of scattered remnants of primitive material in the deep mantle and therefore for the Ar and \(^{3}\)He observations in ocean-island basalts.

**Keywords:** mantle convection; deep-mantle composition; deep-mantle discontinuities; oceanic plateaux; lithospheric-plate buoyancy

One contribution of 14 to a Discussion Meeting ‘Chemical reservoirs and convection in the Earth’s mantle’.

1. Introduction

Views on modern mantle convection and the thermochemical evolution of Earth’s mantle over geological time have strongly divided Earth scientists. On the basis of strong mass-balance arguments, geochemists have in general been strongly supportive of separated convective regimes with little or no mass flux between them. Together with the less radiogenic character of helium in most ocean-island basalts (OIBs) with respect to mid-ocean-ridge basalts (MORBs), the apparent imbalance between the terrestrial budgets of $^{40}\text{K}$ and $^{40}\text{Ar}$ was long considered the strongest evidence that the lower mantle is not degassed. The terrestrial imbalance between heat production and surface heat flow on the one hand, and the isotopic budget of lithophilic radiogenic elements in chondritic Earth on the other, also seems to require the presence of substantial mantle volumes that are not sampled by oceanic basalts.

The sharp, global seismic discontinuity at a depth of 660 km was initially chosen as a major hindrance to whole-mantle convection on the basis that very few earthquakes seemed to originate from below this depth and, therefore, that lithospheric plates do not penetrate it. An additional justification of layering at this depth was that subtracting the inventory of incompatible elements of the crust and the upper mantle from the inventory of the bulk silicate Earth (BSE) leaves a nearly chondritic distribution of the lithophilic refractory elements in the lower mantle. Even the terrestrial Nd and Sr isotopic budgets seemed to be consistent with a lower mantle with primitive Sm/Nd and Rb/Sr ratios (Allègre et al. 1980; De Paolo & Wasserburg 1976; Jacobsen & Wasserburg 1979; O’Nions et al. 1979). In contrast, physical arguments raised by convectionists (Davies 1988a, b; Hager 1984; Tackley 2000; Tackley et al. 1994) and seismic imaging (e.g. Grand 1994; van der Hilst et al. 1991, 1997) indicate that this discontinuity is far more permeable to the penetration of lithospheric plates.

In this paper we review the evidence supporting each point of view. We argue that the different geochemical components represent a rather continuous spectrum resulting from the interplay of a range of geological processes in a zoned, but not strictly layered, mantle. We introduce a new thermochemical, whole-mantle convection model in which the eventual fate of the slabs of subducted lithosphere depends on their thermal structure and composition, and hence on the plate-tectonic setting at the Earth’s surface prior to subduction and the evolution over geological time of the convergent margin involved. It is shown that, in particular, recycling of normal oceanic lithosphere in the shallow mantle and selective accumulation of ancient plume heads at the bottom of the mantle can help to resolve the apparent contradictions between the conventional models of layered and whole-mantle convection.

2. Isotopic heterogeneities in the Earth’s mantle

The lack of any mantle sample whose isotopic properties may reflect closed-system evolution since the accretion of the Earth indicates that mantle convection and plate tectonics obliterated the traces of the initial mantle differentiation. With the possible exception of $^{129}\text{Xe}$ and $^{132}\text{Xe}$ anomalies (Allègre et al. 1983), the Earth seems also to have lost any other isotopic heterogeneity that would result from the decay of short-half-life nuclides. The pattern of isotopic heterogeneities observed
for long-lived radioactive systems therefore requires that the parent and daughter nuclides, such as $^{87}\text{Rb}^{87}\text{Sr}$, $^{238}\text{U}^{206}\text{Pb}$, $^{147}\text{Sm}^{143}\text{Nd}$ and $^{176}\text{Lu}^{176}\text{Hf}$, went for aeons through igneous and surface processes that redistributed them unevenly amongst ‘reservoirs’ (Armstrong 1968; Hofmann & White 1980). Examples of such processes are MORB extraction from the upper mantle, subduction-zone magmatism, hydrothermal activity and erosion-sedimentation, which produce geodynamic units as diverse as the depleted upper mantle, the continental crust and the altered oceanic crust.

Geochemists observe the outcome of these processes, that is the isotopic heterogeneities, and use either empirical or statistical methods (Allègre et al. 1987; Hart et al. 1992; White 1985; Zindler & Hart 1986) to define the minimum number of ‘mantle components’ that can account for the modern isotopic variability of oceanic basalts. Establishing and understanding the correspondence between each mantle component and a specific geological process that shaped its geochemical characteristic has been a major task for igneous geochemistry over the 1980s and the 1990s. This correspondence has recently been reviewed in detail (Carlson 1994; Hofmann 1997; White 1985; Zindler & Hart 1986) but has remained speculative enough to appear to some as pointless taxonomy.

New stable and radiogenic evidence on Hawaiian basalts has improved the geochemical interpretation of mantle components substantially, however. Since significant shifts in isotopic compositions require the intervention of low-temperature processes, values of $\delta^{18}\text{O}$ deviating significantly from the upper-mantle value indicate that part of the mantle source of basalts contains rock precursors that were exposed to surface conditions. The positive correlation observed between $\delta^{18}\text{O}$ and radiogenic isotopes linked, for the first time, radiogenic ingrowth to the temperature of the processes that formed the mantle components (Eiler et al. 1996). The most remarkable feature is probably the existence of a strong positive correlation between the $^{187}\text{Os}/^{188}\text{Os}$ versus $\delta^{18}\text{O}$ array that Lassiter & Hauri (1998) found in Hawaiian basalts. The fact that the $^{187}\text{Os}/^{188}\text{Os}$ and $\delta^{18}\text{O}$ do not pass through the modern mantle values excludes contamination of the plume basalts by the modern lithosphere. High $\delta^{18}\text{O}$ values indicative of low-temperature exchange of the source protolith with sea water have been associated with rocks that evolved for aeons with high Re/Os ratios, i.e. ancient basalts. Over the last few years, all the essential petrographic ingredients of recycled oceanic crust have thus been identified in the source of Hawaiian basalts: altered basalt (Lassiter & Hauri 1998), gabbro (Sobolev et al. 1998) and even pelagic sediments (Blichert-Toft et al. 1999). Evidence of recycled material has likewise been argued for the source of Icelandic basalts (Chauvel & Hemond 2000).

Geochemical evidence indicates that geochemical reservoirs coexist on a range of scales. Zindler et al. (1984) showed that most of the mantle’s isotopic variability can be observed on the scale of a single seamount. Although the recycled components are best identified in hotspot basalts, there is little doubt that they also leak through mid-ocean ridges and, in particular, the Mid-Atlantic Ridge. OIB-type material appears distinctly in some MORBs and its abundance correlates with ridge bathymetry (Schilling et al. 1983). As it is unlikely that the Mid-Atlantic Ridge rests over a line of hotspots, some material has to leak to the surface through the ridge crest from a reservoir with OIB characteristics but without plume dynamics.

3. Evidence against whole-mantle convection

The terrestrial abundances of the refractory lithophilic elements—notably the continental crust, the depleted mantle and the core—cannot be recombinated into a bulk Earth with properties that resemble those known from the least differentiated bodies of the Solar System, i.e. chondrites. The underlying assumption of all these models is that the terrestrial abundances of the refractory lithophilic elements are chondritic (e.g. Drake & Righter 2002). The relative proportions of refractory elements do not vary much amongst the different classes of chondrites (Wasson & Kallemeyn 1988) and this is usually considered strong evidence that the terrestrial ratios of refractory elements must also be nearly chondritic.

(a) The undegassed character of the lower mantle

The decay of the long-lived isotope \( ^{40}K \) produces \( ^{40}Ar \). The planetary inventory of \( ^{40}Ar \) indicates that the mantle is not completely degassed (Allègre et al. 1996; Turekian 1959; Turner 1989). Using a U content of the BSE of 21 ppm and a K/U ratio of 10 000 to 125 000, Wasserburg et al. (1964) and Jochum et al. (1983) concluded that ca. 50% of the terrestrial \( ^{40}Ar \) must reside outside of the atmosphere, the continental crust and the terrestrial mantle. A standard view is that the missing Ar is sequestered in the lower mantle, which, as a consequence, must be undegassed and cannot participate in the formation of the oceanic lithosphere. Attempts to characterize the \( ^{40}Ar \) content of OIBs failed because of clear indications of subsurface outgassing and atmospheric contamination (Dixon et al. 1997; Farley & Craig 1994). Albarède (1998a) and Davies (1999) argued, however, that a terrestrial K/U ratio lower by 30–40% than commonly accepted values would suffice to account for the missing argon.

The decay of the long-lived isotopes \( ^{238}U, ^{235}U \) and \( ^{232}Th \) produces \( ^{4}He \). The \( ^{4}He/^{3}He \) ratio increases proportionally to the ratio \( ^{238}U/^{3}He \) (and to a lesser extent the Th/U ratio), i.e. in much the same way as the rate of \( ^{143}Nd/^{144}Nd \) radiogenic ingrowth is proportional to the \( ^{147}Sm/^{144}Nd \) ratio. Helium is much more radiogenic in MORBs \( ^{4}He/^{3}He \approx 90 000 \) or \( ^{3}He/^{4}He \approx 8R_A \) than in OIBs; extreme values of \( ^{4}He/^{3}He \approx 20 000 \) \( ^{3}He/^{4}He \approx 30R_A \) have been reported from several hotspots, notably in basalts from the Loihi seamount in Hawaii (Kurz et al. 1983). This difference requires that the \( ^{238}U/^{3}He \) ratio was fractionated in the source of MORBs with respect to that of OIBs. The standard model (e.g. Allègre et al. 1996; Kurz et al. 1983) ascribes this difference to the MORB source having lost a substantial fraction of its initial \( ^{3}He \) over the geological ages through repeated cycles of magma extraction and outgassing, which increases the \( ^{238}U/^{3}He \) ratio of the residue and, therefore, its modern \( ^{4}He/^{3}He \) ratio. Graham et al. (1990) pointed out that, as long as a gas phase is not present, He is probably more compatible than U during melting and therefore low \( ^{4}He/^{3}He \) (high \( ^{3}He/^{4}He \)) ratios may signal a residual rather than outgassed mantle. Magmas ponding at the base of the oceanic lithosphere are unlikely to evolve bubbles simply because CO\(_2\) saturation is not reached at this pressure for any reasonable carbon content in basalts (Dixon & Stolper 1995). Under such circumstances, rare gases do not behave differently from any other incompatible element and their abundances in melt and residuum are only controlled by mineral/melt partitioning. Preliminary experiments (e.g. Brooker et al. 1998) actually support the weakly incompatible character of all the rare gases, but reliable He-partitioning
data are not, to date, available. Evidence that high $^3$He/$^4$He hotspots come from undegassed mantle is therefore strong but not conclusive.

Strong indication of inefficient mantle outgassing also comes from a comparison between heat flow and the rate of $^4$He and $^{40}$Ar outgassing (Allègre et al. 1996; O’Nions & Oxburgh 1983). Heat production in the Earth is quantitatively accounted for by the very same nuclides as those producing $^4$He and $^{40}$Ar. Using standard tables of heat production (van Schmus 1995), the fluxes $J_{\text{He}}$ and $J_{\text{Ar}}$ of $^4$He and $^{40}$Ar, respectively, can be related to heat production $A$ through

$$A \text{ (Wm}^{-2}\text{)} = 6.0610^{11} J_{\text{He}} \text{ (mol m}^{-2}\text{ s}^{-1}) + 7.1410^{10} J_{\text{Ar}} \text{ (mol m}^{-2}\text{ s}^{-1}).$$

If $J_{\text{He}}$ and $J_{\text{Ar}}$ are taken as those of the modern surface flux of $^4$He (O’Nions & Oxburgh 1983) and $^{40}$Ar (Allègre et al. 1996), only 15% of the modern surface heat flow is accounted for (see below). The difference in Xe-isotopic composition between the atmosphere and the mantle (Allègre et al. 1983) and the presence of Ne with possible solar isotopic compositions (Honda et al. 1991; Moreira et al. 2001) do not necessarily indicate that the lower mantle is fully undegassed, only that it is not well mixed. This problem will be addressed below. How $^3$He became tied up in a protolith that went through surface processes is a major conundrum of recent isotopic geochemistry (Anderson 2000; Kurz et al. 1996).

(b) Planetary heat balance

A strong case for hidden radioactive sources is made by the Urey ratio, $Ur$, which is the ratio of the heat production by the radioactive elements U, Th and K to the modern surface heat flow of 42 TW (Sclater et al. 1980). Parametrized cooling models of the Earth indicate that, by now, most of the heat inherited from the terrestrial accretion and short-lived radioactive nuclides must have decayed away. With the value of the present day heat flow ($J_q = 82$ mW m$^{-2}$), and the assumption that the heat flow at the core–mantle boundary (CMB) represents ca. 10% of the surface heat flow ($J_c = 0.1J_q$), the heat-balance equation yields an approximate mantle cooling rate of

$$\frac{dT}{dt} = -(0.9 - Ur) \times 150 \text{ K Gyr}^{-1},$$

where $\bar{T}$ is the mean temperature of the mantle (e.g. Schubert et al. 2001). Realistic cooling rates require that the Urey number is close to unity ($Ur = 0.7$–0.8) and therefore that the Earth is close to thermal equilibrium with its radioactive heat sources. This value of the Urey number is much higher than expected from the source of MORB, which implies that heat sources must be sequestered in volumes of the mantle not tapped by MORB (Kellogg et al. 1999).

It is unlikely that sources other than radioactive decay contribute significantly to the heating of the mantle. The two modes of transformation of mechanical and potential energy into thermal energy (dissipation) are shear heating (Hewitt et al. 1975; Jarvis & McKenzie 1980; Turcotte et al. 1974) and gravitational stratification (Vacquier 1991). These heating modes are not primary sources of heat, since they originate from energy (heat) introduced at the CMB or from radioactive heating, but these transformations are irreversible and constitute a local source of entropy. Overall, however, they do not appear to contribute significantly to mantle heating, except perhaps during catastrophic mantle overturns.

Phil. Trans. R. Soc. Lond. A (2002)
At least 30–50% of the heat-producing elements are therefore still concealed in reservoirs essentially untapped by basaltic magmas. Whether the missing heat resides at the bottom of the mantle or in the core is still unclear: more potassium dissolved in the core than is currently assumed could certainly help to power the terrestrial dynamo. A potassium-rich core, however, does not help to fill the gap of missing heat sources and would worsen the situation for the terrestrial $^{40}$Ar inventory. Moreover, increasing the heat flow above 10% at the CMB would require that a substantial fraction of upwelling material does not reach the surface at hotspots. So far, the strongest argument supporting this idea is the existence of plume-type geochemical anomalies correlated with bathymetric highs along the Mid-Atlantic Ridge (Schilling et al. 1983) but the amount of hotspot material involved seems too small to account for such a substantial part of the terrestrial heat flow. The notion that hotspots may underestimate plume flux is also supported by recent evidence from seismological imaging, which suggests that lower-mantle low-wave-speed anomalies may branch off and connect to slow anomalies beneath mid-ocean ridges (Romanowicz & Gung 2002).

(c) Geochemical ‘paradoxes’

The missing-heat argument is compounded by a variety of geochemical ‘paradoxes’, a catchword for the observation that, for some elements or isotopic systems, the known inventories do not add up to a chondritic Earth. In each case, these paradoxes call for unaccounted domains in the deep mantle that would make the Earth’s relative abundance of lithophilic elements chondritic and point to the geochemically enriched character of this reservoir, i.e. its relative enrichment of the most incompatible elements with respect to the less incompatible elements.

The oldest of these paradoxes is the observation made by Allègre (1969) that, in a $^{207}$Pb/$^{204}$Pb versus $^{206}$Pb/$^{204}$Pb plot, the lead-isotope compositions of nearly all terrestrial rocks plot to the right of the meteorite primary isochron (geochron). But other systems can also be used. Blichert-Toft & Albarède (1997) determined the modern isotopic composition of hafnium in various classes of chondrites and concluded that the $^{176}$Hf/$^{177}$Hf ratio of the BSE is independent of the nature of the chondrite class chosen to best represent modern terrestrial material. They pointed out that both parent–daughter nuclides of the Lu–Hf and Sm–Nd systems are lithophilic refractory elements, which is not the case for U–Th–Pb, Re–Os, Rb–Sr and K–Ar systems. Such a property makes Hf- and Nd-isotope compositions ideal for discussing planetary inventories. Blichert-Toft & Albarède (1997) determined that the $^{176}$Hf/$^{177}$Hf versus $^{143}$Nd/$^{144}$Nd mixing relationships do not make it possible to recombine the continental crust with the depleted upper mantle to obtain a chondritic composition of Nd and Hf, regardless of which chondrite class is considered (figure 1). Finally, Turcotte et al. (2001) pointed out that, when tested against the terrestrial inventory of radioactive heat sources indicated by plausible Urey ratios, the Th/U systematics in the Earth also require that the deep mantle has a different value from the Th/U of the upper mantle, which they interpret as requiring layered convection.

4. Geophysical arguments against a layered mantle

Geophysical evidence for substantial structural complexity in the upper-mantle transition zone, on the one hand, and for the penetration of some slabs into the deeper
Figure 1. Nd–Hf isotope systematics require the presence in the deep mantle of a hidden reservoir with a non-chondritic condition. BSE denotes bulk silicate Earth (modern chondrites); CC, continental crust; DM, depleted mantle (MORB source). Mantle array: the small open circles represent ca. 1500 measurements for oceanic basalts (reproduced courtesy of J. Blichert-Toft). The curvature of the mixing hyperbola between CC and DM is well constrained by the nearly chondritic Hf/Sm ratio of the various silicate end members and the $^{147}\text{Sm}/^{144}\text{Nd}$ of the mantle and the crust. Recombining CC and DM into BSE is therefore not possible unless a third component, hidden in the lower mantle, is included in the isotopic budget. This component is not isotopically equivalent to the primitive component, subducted oceanic crust, continental debris or any combination thereof.

mantle, on the other, weighs against the end-member model of strict layering at a depth of 660 km and that of unhindered whole-mantle flow (Schubert et al. 2001; van der Hilst et al. 1991).

(a) Complex slab morphology in the upper-mantle transition zone

Detailed analyses of subduction zone seismicity (e.g. Giardini & Woodhouse 1984; Lundgren & Giardini 1994; Okino et al. 1989) suggest that beneath some convergent margins slabs can be severely deformed in the upper-mantle transition zone. Furthermore, the cessation of deep-focus earthquakes at a depth of 700 km has often been used as an argument against the penetration of subducting slabs beyond the 660 km seismic discontinuity and, thus, in favour of the layering of mantle convection at that depth, even though the pioneering work on deep-slab seismicity by Isacks & Molnar (1971) allowed alternative views (see Helffrich & Brodholt 1991; Wortel & Vlaar 1988).

With its ability to also image aseismic parts of the slab, seismic tomography paints a fundamentally different picture from that above: some slabs of subducted lithosphere indeed appear trapped in the upper mantle, whereas others seem to penetrate...
Figure 2. Series of mantle cross-sections through the recent P-wave model of Kárason & van der Hilst (2000) to illustrate the structural complexity in the upper-mantle transition zone and the regional variation in the fate of the slabs. Dashed lines are drawn at depths of 410, 660 and 1700 km, respectively. The model is based on short-period, routinely processed P, pP and PKP travel-time residuals (Engdahl et al. 1998) and a large number of PP-P and PKP-Pdiff differential times measured by waveform cross-correlation from long-period seismograms. The global model was parametrized with an irregular grid of constant-wave-speed cells, which allows high resolution in regions of dense data coverage, and three-dimensional finite frequency sensitivity kernels were used to account for different periods at which the measurements were made. With this technique, the low-frequency data can constrain long-wavelength mantle structure without preventing the short-period data from resolving small-scale heterogeneity.

into the lower mantle (figure 2). This complexity was first borne out by regional studies (e.g. Fukao et al. 1992; van der Hilst et al. 1991; Zhou & Clayton 1990) but has since been confirmed by high-resolution global inversions (Bijwaard et al. 1998; Fukao et al. 2001; Kárason & van der Hilst 2000; van der Hilst et al. 1997). Slab deflection (and accumulation) probably occurs beneath Izu Bonin (figure 2), the Banda and New Hebrides arcs and several parts of the Mediterranean; deep slabs, sometimes severely deformed in the transition zone, have been detected beneath the Mariana, Central Japan, Tonga-Kermadec, Sunda and northern Kurile arcs, the Philippines, parts of the Aegean and Central and South America (figure 2). Apparently the fate of subducted slabs is more complex than expected from the end-member models of either unobstructed whole-mantle flow or convective layering at a depth of 660 km (Gu et al. 2001; Lay 1994; van der Hilst et al. 1991).

Inspired by Kincaid & Olson (1987), van der Hilst & Seno (1993), van der Hilst (1995) and Griffiths et al. (1995) argued that the observed complexity does not imply layering at depths of 660 km but can be caused by interplay between relative plate motion (i.e. lateral trench migration) and slab deformation upon encountering resis-
tance (e.g. higher viscosity or a depressed phase boundary). Numerical simulations support this view (Christensen 1996; Davies 1995; Gurnis & Hager 1988; Zhong & Gurnis 1995), but selective weakening of the descending plate by grain-size reduction upon phase transformation may also contribute (Riedel & Karato 1997). (Note that the complexity caused by interaction of downwellings with the upper-mantle transition zone does not appear to be restricted to above the 660 km discontinuity but persists to depths of ca. 1000 km, i.e. the base of layer ‘C’ in Bullen’s classical subdivision of Earth’s interior (Fukao et al. 2001; van der Hilst & Kárason 1999).)

(b) Geophysical evidence against convective layering at a depth of 660 km

Several lines of geophysical evidence support slab flux into the lower mantle. The magnitude of lateral variations in depth to the 660 km discontinuity (e.g. Flanagan & Shearer 1998; Richards & Wicks 1990) is consistent with an endothermic phase change in olivine that is by itself too weak to stratify flow (Christensen & Yuen 1984). Furthermore, long-wavelength gravity anomalies and dynamic topography can be explained by models of deep circulation with a high-viscosity lower mantle (e.g. Davies 1988b; Davies & Richards 1992; Hager 1984).

Travel times of high-frequency waves projected to the so-called ‘residual sphere’ have been used to argue for the presence of subducted slab in the lower mantle beneath convergent margins in Central America and the western Pacific (Creager & Jordan 1984, 1986; Fischer et al. 1991; Jordan & Lynn 1974). Most of these early claims for slab penetration across the 660 km discontinuity were confirmed by later tomographic studies.

Van der Hilst et al. (1997) and Grand et al. (1997) showed the similarity between P- and S-wave models of the mantle (except at very large depths) and pointed out that many slabs may have sunk beyond depths of 660 km—and even 1000 km—into the lower mantle beneath the Americas, southern Asia and Indo-China and Tonga (figure 2). Even though it has not yet been quantified from the tomographic models, it seems that slab flux into the lower mantle far exceeds the 1% or 2% level inferred from some of the more restrictive geochemical mass balances (e.g. O’Nions & Tolstikhin 1996).

(c) Seismological evidence against global mid-mantle interfaces

Van der Hilst & Kárason (1999) noted that, in the current snapshot of convection, not all the slabs that sink across the 660 km discontinuity may also reach the CMB. This observation could simply reflect subduction history or effects of spherical symmetry on planar downwellings (van der Hilst et al. 1997). But, along with emerging seismological evidence for compositional heterogeneity (see below) and the geochemical and heat-balance arguments for incomplete mantle mixing (see above), Kellogg et al. (1999) used it to propose a model, hereafter denoted by KHH99, in which the mantle beneath most of the deepest slabs has remained compositionally distinct. The KHH99 model implies a mid-mantle interface or thermal boundary. In search of evidence for such a structure, we examined short-period data from deep earthquakes recorded at dense receiver arrays in the western US. Specifically, we searched for evidence of scattering or phase conversion at interfaces in the mid-mantle using time-slowness and azimuth-slowness stacks (Castle & van der Hilst 2002). In accord
with earlier findings (Kaneshima & Helffrich 1999), we detected scattered energy from depths of \( \text{ca.} 1600 \) km beneath the western Pacific, but examination of data from earthquakes beneath Tonga-Kermadec and South America did not yield robust evidence for scattering in the mid-mantle. Subtle gradients can be overlooked and the need for dense receiver arrays and deep earthquakes restricts such investigations to a few relatively small regions in the deep mantle, but we consider it unlikely that a global interface exists between depths of 1000 and 2000 km in the mantle (Castle & van der Hilst 2002; Vidale et al. 2001).

5. Composition of the deep mantle

(a) Seismological evidence for lateral variations in bulk composition

In recent years, much seismological research has focused on determining the relative variations in density and compressional, shear and bulk sound speed (see Masters et al. (2000) for a review). The ratio of relative variations in shear versus compressional wave speed, often expressed as \( R = \frac{d \ln V_S}{d \ln V_P} \), increases with depth, in particular in the lowermost mantle. It is debated what values of \( R \) necessarily imply compositional variations, but values less than \( \text{ca.} 2.5 \), which occurs over most radial profiles, may still be compatible with effects of pressure and temperature (S. Karato 2001, personal communication).

A focus on the radial dependence of \( R \) can be misleading, however, and aspherical variations may be substantial and important. From the analysis of EHB (Engdahl et al. 1998) travel-time data, Saltzer et al. (2001) demonstrated that \( R \) increases more rapidly with depth (to values in excess of 2.5) away from regions of Post-Mesozoic subduction than it does beneath the major convergent margins, where \( R \) remains less than 2.0 (figure 3). This result concurs with the results of Masters et al. (2000) and suggests that the \( R \) values that are most diagnostic of compositional heterogeneity are primarily associated with the lowermost mantle regions characterized by lower than average \( V_P \) and, in particular, \( V_S \). Many of these regions are also marked by a negative correlation between shear- and bulk-sound speed (Ishii & Tromp 1999; Kennett et al. 1998; Masters et al. 2000; Saltzer et al. 2001; Su & Dziewonski 1997), which is hard to explain by thermal effects alone. Some of the low shear speeds in the deep mantle may thus be due to higher intrinsic density, for instance, due to iron enrichment (van der Hilst & Kárason 1999).

The morphology of anomalous domains and their height above the CMB has remained enigmatic. While the most anomalous values probably occur near to the base of the mantle, the \( R \) profiles for regions near to and away from major downwellings begin to diverge near a depth of 1500 km (Masters et al. 2000; Saltzer et al. 2001), which suggests that compositional differences may extend to far above the CMB. In a regional study, Saltzer et al. (2002) determined variations in elastic parameters from the Earth’s surface to the CMB along a transect from Japan, across Alaska, to North America. They selected P and S (and PnP and ScS) waveforms from the Data Management Center of the Incorporated Research Institutions for Seismology (IRIS) and determined differential travel times by waveform cross-correlation. The high quality of data allowed the determination of the Poisson’s ratio, \( \sigma \), and relative variations in \( V_P \), \( V_S \) and bulk sound speed as a function of depth. Using published equations of state and wave-speed-to-temperature derivatives, Saltzer et al. (2002) showed that a small increase in iron

Figure 3. Lateral variation in the depth dependence of the relative variation of shear and compressional wave speed, expressed as $R = \frac{d \ln V_S}{d \ln V_P}$. Light-grey curve, $R = R(z)$ in the mantle below regions of Post-Mesozoic convergent margins (depicted by grey in the map inset on the right); dark grey curve, $R = R(z)$ away from recent downwellings. For comparison we also show the approximate value of $R$ inferred from mineral-physics experiments, assuming variations in temperature only. The dashed line labelled ‘K93’ is from Karato (1993). This comparison shows that ‘normal’ $R$ values prevail to a depth of ca. 1500 km, and that in the deep mantle the slab values remain normal, but away from the slabs $R$ gradually increases to values in excess of 2.5, which seems hard to explain in terms of temperature alone. The bottom 300 km or so of the mantle is poorly sampled by the P and S data considered here. (Modified from Saltzer et al. (2001).)

(1–3%) in the bottom 1000 km or so of the mantle can explain the observations, whereas neglect of compositional variations would require unrealistically large thermal effects.

We remark that the geographical region involved in their study is relatively small and establishing the spatial extent and volumetric significance of anomalous regions remains a challenge for modern seismological imaging. Furthermore, there could be changes in deep-mantle mineralogy or phase chemistry (e.g. Shim et al. 2001) that have not been accounted for.

(b) Geochemical evidence for a compositionally distinct lower mantle

Mass balancing between the different mantle reservoirs, as attempted by a number of authors, most recently by Helffrich & Wood (2001), is fraught with large uncertainty. Melting mechanisms are not understood well enough for the main composition of the upper mantle to be evaluated from MORB chemistry with sufficient precision. Therefore, here we use a different procedure to substantiate our claim that the deep
mantle is enriched and that the enrichment level of the more incompatible elements is higher than that of the more compatible elements.

We first evaluate the mean composition of the mantle, hereafter referred to as the bulk mantle (BM), by removing for each element the part of the BSE inventory that is presently located in the continental crust. This fairly robust estimate is based on the continental-crust composition of Taylor & McLennan (1995) and Rudnick & Fountain (1995) and on the volume of crust determined in the CRUST 5.1 model of Mooney et al. (1998). We then normalize the mean N-MORB composition of Hofmann (1988) to the BM (figure 4) and observe that the normalized mean MORB is significantly less enriched in very incompatible elements, such as Th, Ba and La, than in the more compatible elements, such as Sm and Yb. For instance, the La/Sm ratio of MORB (1.0) is substantially smaller than that of the BM (1.4), regardless of the La/Sm of continental crust adopted (here 4.6). These estimates are robust because different continental-crust models show a strong correlation between their incompatible-element contents. The La/Sm ratio of the MORB source is predictably smaller than that of MORB themselves. A deep-mantle reservoir with an La/Sm ratio larger than that of the BM, which largely escapes observation, therefore exists.

Straightforward implications are that the MORB source must be even more depleted in very incompatible elements than the BM with respect to the more compatible elements and therefore that an enriched complement (clearly the OIB source) exists that hosts the missing Th, Nb, Ba, La, etc. Since the latter enriched reservoir cannot be located in the upper mantle, where it would feed the mid-ocean ridges, it has to be located in the lower mantle. The corollary of this argument is that a MORB-like lower mantle would leave a complementary enriched upper mantle unfit for generating MORB.

The effect of interaction between the oceanic crust and sea water is unlikely to affect the present conclusion, which is based on refractory lithophilic elements largely unaffected by alteration. Extraction of felsic melts from the crust at subduction zones strengthens the depleted character of the subducted oceanic crust even further. For three reasons, the enrichment of the lower mantle does not seem to be due to sprinkled terrigenous material. First, the geochemical expression of such a material is well known as the EM2 component of oceanic basalts (White 1985; White & Duncan 1996; White & Hofmann 1982). This component, best represented in the Society Islands, is far too uncommon to account for the enrichment of large mantle volumes. Second, subduction of large amounts of continental crust would leave large amounts of unsupported atmospheric $^{40}$Ar. Coltice et al. (2000) calculated that, as a consequence, a mass equivalent to at most 30% of the modern continental crust was recycled into the mantle by subduction of terrigenous sediments. Third, Blichert-Toft & Albarède (1997) pointed out that the modern $^{176}$Hf/$^{177}$Hf and $^{143}$Nd/$^{144}$Nd systematics of oceanic basalts requires a hidden component, most obviously segregated in the deep mantle, that cannot be accounted for by any combination of continental detritus and recycled oceanic crust.

We conclude that the mantle is necessarily zoned and that its lower parts concentrate the most incompatible elements, such as Th (and most likely U), Ba, Nb and La, with respect to the more compatible elements. The hidden reservoir is located in the deep mantle and is concentrated in heat-producing elements, but it cannot be the graveyard of ordinary oceanic lithospheric plates. The original model of Hofmann & White (1982) (see also Christensen & Hofmann 1994; Coltice & Ricard 1999), with
Figure 4. Bottom graph shows the composition of the bulk mantle (BM) (BSE stripped from the core and model continental crust) in lithophilic refractory incompatible elements normalized to the BSE composition of McDonough & Sun (1995). Two different models of continental-crust composition have been considered (RF, Rudnick & Fountain 1995; TML, Taylor & McLennan 1995). The top graph shows MORB compositions normalized to the BM. Both the MORB composition of Hofmann (1988) (AWH) and the mode of the samples compiled in the Petrological Database (PETDB) have been used. Elements are organized from left to right by decreasing incompatibility. Elements mobile during interaction of the oceanic crust with sea water, such as U, Ba and Rb, have not been considered. MORBs are depleted in the most incompatible elements (e.g. low La/Sm) relative to the BM. As a result of fractionation during melting, the MORB source must be even more depleted, which calls for a complement, most likely to be buried in the deep mantle, enriched in the most incompatible elements with respect to the BM.

delaminated oceanic crust lining the bottom of the mantle, is therefore difficult to reconcile with MORB geochemistry.

6. Thermochemical mantle convection and variable-depth subduction

Although our studies support some of the main aspects of the KHH99 model, it is unlikely that a thermal-boundary layer or sharp compositional interface exists between a depth of 1000 km and the top of the $D''$ region. Here we present an alternative mechanism for producing and preserving compositional heterogeneity in the deep mantle without the need for a thermal boundary layer (figure 5). Central to our model is the fact that slabs of subducted lithosphere are thermochemical features; there is thus a balance between negative thermal buoyancy and positive compositional buoyancy, and neutral buoyancy can be reached at a depth that is not the same for all slabs. Although other compositional effects can perhaps be
considered, the variation in composition due to loading of the plate with ocean plateaux or plume heads is of particular interest.

Oceanic plateaux (Ben-Avraham et al. 1981; Coffin & Eldholm 1994), usually representing enormous eruptions of plume-head basalts, are prominent features of the ocean floor. With basalts converted to (heavy) eclogite upon subduction, slabs that carry oceanic plateaux may preferentially sink to greater depth than slabs that are on average more ‘barren’. This system of thermochemical convection could produce and maintain poorly mixed and compositionally distinct domains with the required geochemical signatures in the deep mantle.

Selective subduction is unlikely to lead to a layered mantle. Geochemical differences across a typical mantle column are likely to be progressive and irregular from one place to another. For a given thermal structure, plates with a heavier basaltic load, i.e. plates loaded with abundant plume-head material, should, on average, sink to deeper levels than plates overlaid by ordinary oceanic crust and vice versa: for a given composition, the older and colder plates are likely to sink more deeply than young, relative warm slabs. In this scenario, the plate-tectonic setting and its evolution over time thus play a more important role than in models of purely thermal convection.

The model of thermochemical convection with variable fate of slabs of subducted lithosphere has characteristics that fit the conditions listed above.

(i) Deeper layers in the mantle remain geochemically more enriched than shallow layers. They receive more incompatible elements, and in particular more U, Th and K, because the oceanic plateau lithosphere has more basaltic components. Deep domains also have lower Sm/Nd and Lu/Hf ratios and higher Rb/Sr and Re/Os because plume heads come from mantle sources that have been geochemically enriched for long periods of time.

(ii) The mantle does not become homogeneous through time as a result of mantle convection and mixing. Heterogeneity is continually maintained by selective subduction at different levels and by the modest proportion of plates that carry enough plume material to sink all the way down to the CMB. Since much, perhaps most, of the recycling takes place in the low-viscosity shallow mantle, this model overcomes one of the consequences most difficult to reconcile with geochemical evidence: unimpeded convection would rapidly fill-up the deep mantle with young plates and push older material upwards into the shallow mantle (Christensen & Hofmann 1994; Coltice & Ricard 1999; van Keken & Ballentine 1999).

(iii) The chemistry of the lowermost layers should be different from that of the overlying mantle. Iron, titanium and incompatible elements, excess Nb and Ta, and all the other elements normally enriched in OIB, notably U, Th and K, should be more abundant near the base of the mantle. A precise evaluation of the deep-mantle composition and of the composition gradient is, however, not possible because the dilution of fertile material by interstitial residual peridotite is essentially unconstrained.

(iv) An interface between the deep mantle and the shallow mantle is not required, nor is a thermal boundary layer.
Figure 5. The model of selective subduction and zoned mantle convection. Barren oceanic lithospheric plates (right) quickly reach neutral buoyancy and are dispersed in the shallow mantle. Plates with a heavy load of plume heads (centre) are negatively buoyant and sink to the CMB. Selective subduction is the geodynamic expression of thermochemical convection: it maintains the enrichment of the deep mantle in incompatible and fertile elements and counteracts the vertical mixing driven by the overall mantle circulation. Because of secular cooling, plume instabilities must have been more prevalent in the Archean. Thermochemical convection must therefore have been an essential geodynamic aspect of the early mantle evolution.

7. Discussion

In many respects, the chemical state of the mantle resembles that of the ocean, in the sense that they both lack interfaces while thermochemical convection maintains vertical chemical gradients across the system. Geochemical evidence that the mantle is poorly mixed vertically is overwhelming. The modern extraction rate of lithophilic incompatible elements into the oceanic lithosphere is reasonably well constrained. Most of the lithospheric inventory of these elements is concentrated in the basaltic part of the crust. The residence time of an element in the mantle with respect to ridge processes, which is simply the reciprocal of its probability of extraction into the ridge system per unit time, can be estimated as the ratio of its total amount in the BM divided by its flux into the oceanic crust. Depending on the element, the residence time varies from 6 to 14 Gyr (figure 6); such a low probability of extraction of incompatible elements into MORB illustrates that mid-ocean ridges tap a part of the mantle that is depleted with respect to the bulk mantle. Vertical mixing in the mantle is therefore inefficient. How far back in time did such a situation prevail? Heat production declines with radioactive decay, and the Earth has been cooling over time. Parametrized convection models with internal heating suggest that plate velocity changes with $Ra^{1/2}$, where $Ra$ is the Rayleigh number, while the surface heat flow (Nusselt number) increases as $Ra^{1/4}$ (McKenzie et al. 1974). Plate velocity, and therefore subduction flux, increases with the square of the surface heat flow (see also Phipps Morgan 1998). Radiogenic heat production was larger in the Early Archean.

Phil. Trans. R. Soc. Lond. A (2002)
Figure 6. Residence time of refractory lithophilic elements in the BM. It has been assumed that 3.5 km$^2$ of oceanic crust 6 km thick is extracted each year at ridge crests. The long residence times indicate that the modern mantle is not homogenized by mantle convection at its modern pace. Because of radioactive decay, plate velocities, and therefore residence times, were about 10 times faster than today, but were not fast enough to have extracted all the primordial nuclides into the ridges. Although thermal evolution of the planet requires that these residence times were about an order of magnitude shorter in the Early Archean, the persistence of primitive isotopic signals in the source of oceanic basalts is therefore not excluded by the dynamics of the system.

than today by about a factor of three and it is therefore likely that the residence times of incompatible elements were a factor of 10 shorter. The values averaged over the entire Earth’s history were certainly intermediate. Although it may have been different in the past, the modern mantle as a whole can be neither homogeneous nor in a chemical steady state.

We believe that the deep part of the mantle is only rarely sampled by OIBs, which is why the denomination of ‘hidden reservoir’ remains appropriate. Hotspot basalts, such as those from Hawaii and Iceland, contain recycled components that resemble the normal oceanic lithosphere, with higher-than-chondritic integrated Sm/Nd and Lu/Hf ratios. We have suggested elsewhere (Gasperini et al. 2000) that hotspot basalts containing a substantial fractional of EM1 component (notably Pitcairn and Sardinia basalts) sample a source rich in oceanic plateaux. This interpretation is based on

(1) Sr, Ba and Eu evidence that their protolith contains plutonic precursors that fractionated in the stability field of plagioclase, i.e. at a very shallow level;

(2) their U/Pb, Sm/Nd and Lu/Hf integrated ratios are slightly lower than chondritic values, and therefore slightly enriched; and

(3) the high Th/U ratio of their mantle source (approximately 4.0) incompatible with recycled MORB.

We therefore consider that the deep mantle is home to EM1. It could also be home to the Focus Zone (Hart et al. 1992) or C (Hanan & Graham 1996) components but these have so far more the character of a mean value of mantle isotopic properties rather than that of a well-defined mantle source with sharply identified properties.

Oceanic plateaux make up $29 \times 10^6$ km$^2$, or ca. 10% of the sea-floor surface area (Nur & Ben-Avraham 1982; Schubert & Sandwell 1989). The relatively small proportion of the ocean-floor surface covered by plume heads (oceanic plateaux) is somewhat compensated by thicker crust: typical oceanic plateaux are underlaid by basaltic crust 25 km or more thick (Abbott et al. 1997), in contrast to a 6 km-thick normal oceanic crust, and they may extend over surfaces of $10^6$ km$^2$. When a significant part of a plateau crosses into the eclogite facies, it exerts a strong pulling force on the host lithospheric plate. The gravitational pull of the normal oceanic lithosphere topped with a thin layer of oceanic crust is expected to be smaller, and the light, barren plates are more likely to be dislocated and dispersed in the shallow mantle. We can try to evaluate how fast eclogitised plateaux could sink into the lower mantle. As a simple illustration intended to assess whether the magnitude of such a mechanism is not unreasonable, let us consider that the plateau material forms spherical blobs and that the sinking rate of these blobs is given by Stokes’s law. We used conservative estimates with a plateau thickness as an estimate of the radius $R = 30$ km, a density contrast $\Delta \rho = 50$ kg m$^{-3}$ (1% density contrast (Kesson et al. 1998)) between the eclogitised plateau and the barren oceanic lithosphere, and a lower-mantle viscosity $\mu = 6 \times 10^{21}$ Pa s as determined by post-glacial rebound (Mitrovica 1996; Nakada & Lambeck 1989). For these values, a blob of plateau material would sink at a velocity $v$ that is 0.1 cm yr$^{-1}$ higher than the normal lithosphere, i.e. at a differential rate that is significant (1–10%) with respect to subduction rates. In comparison, small blobs with a diameter similar to the thickness of the ordinary oceanic crust would sink at a rate close to normal. Large blobs of basaltic material forming major oceanic plateaux are therefore more likely to segregate in the lower mantle than the thin veneer of oceanic crust normally associated with the oceanic lithospheric plates.

It has been argued that continents grow by the lateral accretion of oceanic plateaux to pre-existing continental nuclei (e.g. Albarède 1998a; Boher et al. 1992; Kimura et al. 1993; Stein & Hofmann 1994) and whether plateaux recycle or accrete to continents then becomes an issue. The Ontong-Java plateau, for instance, seems to resist subduction, but the Hawaiian track trails into the Aleutian trench with no visible accretion of a plume head. Most plateaux must be entrained with the subducting plates, however. Since the mean age of the modern sea floor is only ca. 60 Myr (Parsons 1982) it can safely be assumed that the modern oceans are not cluttered with old buoyant plateaux. Let us call the surface area free of oceanic plateaux $S_f$, the fraction of sea floor covered by plateau $\varphi$, the rate (3 km$^2$ yr$^{-1}$) at which new sea floor is created $k$, and the rate at which it is covered by hotspot eruptions $P$. The conservation of plateau-free surface area (accretion minus subduction and plateau invasion) requires

$$\frac{dS_f}{dt} = k - (1 - \varphi)(k + P).$$

Steady state (i.e. $dS_f/dt = 0$) gives $P = k\varphi/(1 - \varphi)$. If oceanic plateaux were irreversibly buoyant, the sea floor would have been entirely filled up with oceanic plateaux in ca. 650 Myr, and plate tectonics could no longer operate. Although
oceanic plateaux are occasionally accreted to the continents to form juvenile terrains, most of them must therefore be subducted. The mechanisms that condition the dual fate of plateau material are not well understood but may be related to the rare occurrence of ultramafic rocks in accreted plateaux (Burke 1988): the plateaux accreted to the continents may preferentially be those for which the lithospheric root is detached from the buoyant crust.

The irreversible foundering or subduction of geochemically enriched—and perhaps Fe-rich—ancient crust was previously considered by Chase & Patchett (1988) as the main cause for Nd in the Early Archean mantle being anomalously radiogenic with respect to the modern mantle. The present model has also some points in common with that of Phipps-Morgan (1998), since plume-head material tends to bob up and down and loop between the surface and the deep mantle, largely decoupled from the normal oceanic lithosphere (MORB source), which loops between the surface and the upper mantle. It can perhaps be reconciled with the blob model of Becker et al. (1999) if the blobs are rich in subducted plume heads. Moreover, to prevent contamination of the MORB source by blob material, the blobs must be kept from dispersing into the upper mantle upon accumulation of subducted material at the base of the mantle. This problem is compounded by the increased buoyancy of large blobs (higher than 800 km) caused by radioactive heating.

In such a context how may some tracers still indicate that primordial material is preserved in the lower mantle? Even if interpretations exist that do not require a primordial source exist for He (see above), this is not the case for the survival of solar neon (Honda et al. 1991; Moreira et al. 2001). The rather long residence times of incompatible lithophilic elements (6–14 Gyr) suggest that a substantial part of the deep mantle never went through the ridge system. As a first approximation, extraction of an element into the ridge system is a Poisson process and the residence times of each individual atom of any species \( i \) should be exponentially distributed. The fraction \( F_i(\theta) \) of atoms of \( i \) that have resided for a time in excess of \( \theta \) in this reservoir is given by

\[
F_i(\theta) = \exp\left(-\frac{\theta}{\tau_i}\right),
\]

where \( \tau_i \) is the usual (mean) residence time averaged over the entire population of \( i \).

For a mean residence time in the mantle of 5 Gyr, 45% of the atoms in the mantle will escape extraction into oceanic crust for 4.5 Gyr. This proportion falls to 0.01% for a residence time of 500 Myr more relevant for the Early Archean mantle and the mean value for the entire history is probably somewhere in between these values. Even if the zone over which melting extracts incompatible elements is broad and if convective mixing stretches material into thin streaks, primordial isotopic signatures should survive at small scales.

8. Conclusion

The composition of the mantle varies with depth. We favour a deep mantle made of recycled plume heads but perhaps preserving small proportions of relatively primitive mantle. In contrast, the shallow mantle collects the lighter barren oceanic lithospheric plates. The fate of subducted plates therefore depends on their buoyancy, which in turn depends on their thermal structure and composition, for example the presence
of plume heads, both of which relate to plate-tectonic setting and its changes over geological time. The present model does not necessitate the presence of an interface between the deep mantle and the shallow mantle.

Kerstin Lehnert (Petrological Database, PETDB) and Bärbel Sarbas (GEOROC) helped us with the retrieval of essential data from Lamont and Mainz databases. Janne Blichert-Toft shared her unpublished data and provided continuous scientific feedback. Discussions with Nicolas Coltice, Al Hofmann and Yanick Ricard helped shape the arguments. Support to F.A. (as Crosby Lecturer) by MIT made it possible to initiate this work; R.D.v.d.H. acknowledges support by the David and Lucile Packard Foundation.

References


Phil. Trans. R. Soc. Lond. A (2002)


Saltzer, R. L., Stuzmann, E. & van der Hilst, R. D. 2002 Poisson’s ration beneath Alaska from the surface to the core-mantle boundary. (Submitted.)


Phil. Trans. R. Soc. Lond. A (2002)


