The Earth's mantle

George R. Helffrich*† & Bernard J. Wood†

** Earth and Planetary Science, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-Ku, Tokyo 152-8551, Japan*
† Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK

Seismological images of the Earth's mantle reveal three distinct changes in velocity structure, at depths of 410, 660 and 2,700 km. The first two are best explained by mineral phase transformations, whereas the third—the D″ layer—probably reflects a change in chemical composition and thermal structure. Tomographic images of cold slabs in the lower mantle, the displacements of the 410-km and 660-km discontinuities around subduction zones, and the occurrence of small-scale heterogeneities in the lower mantle all indicate that subducted material penetrates the deep mantle, implying whole-mantle convection. In contrast, geochemical analyses of the basaltic products of mantle melting are frequently used to infer that mantle convection is layered, with the deeper mantle largely isolated from the upper mantle. We show that geochemical, seismological and heat-flow data are all consistent with whole-mantle convection provided that the observed heterogeneities are remnants of recycled oceanic and continental crust that make up about 16 and 0.3 per cent, respectively, of mantle volume.

The Earth's mantle comprises 82% of its volume and 65% of its mass. It constitutes virtually all of the silicate part of the Earth, extending from the base of the crust (0.6% of Earth's silicate mass) to the top of the metallic core (Fig. 1). When the core segregated from the silicate and gas of the proto-Earth, it incorporated high concentrations of the **siderophile** elements, leaving **lithophile** elements in the silicate mantle. (Words in bold are explained in the glossary; see Box 1.) Thus, the current composition of the mantle has core formation imprinted on it—as pronounced depletions in, for example, Fe, Ni, S, W, Pt, Au and Pb relative to the **chondritic meteorites**[1,2], which are used to constrain the composition of the whole Earth. Owing to the fractionation of lithophile radioactive parent isotopes such as ^{238}U and ^{182}Hf from their siderophile daughters ^{206}Pb and ^{182}W, core formation can be dated as the time at which the evolution of the isotopic compositions of Pb and W diverged from the meteorite trend. The result (about 4.5 Gyr ago) corresponds to 50–100 Myr after the formation of the oldest meteorite bodies in the Solar System[3,4].

Mantle rocks that occur occasionally at the surface, either as tectonic fragments (kilometre scale) or as inclusions in explosive eruptives (centimetre scale) are predominantly **peridotites**. One of the main questions facing the Earth sciences is whether a peridotite composition, representative of the upper 150 km of the mantle, can also be assumed to represent the remainder of the mantle down to 2,900 km depth. Many geochemical data suggest not. For example, the current heat flux at the Earth's surface is about 44 TW (44×10^{12} W), most of which can be reasonably attributed to radioactive decay of K, U and Th in the mantle[5]. The upper-mantle source region of mid-ocean ridge **basalt** (MORB) is depleted in these elements, however, and only produces 2–6 TW. The simplest explanation of the shortfall is that there is a lower layer enriched in the heat-producing elements that is only sporadically involved in the production of surface rocks[6–8]. Similarly, the flux of the

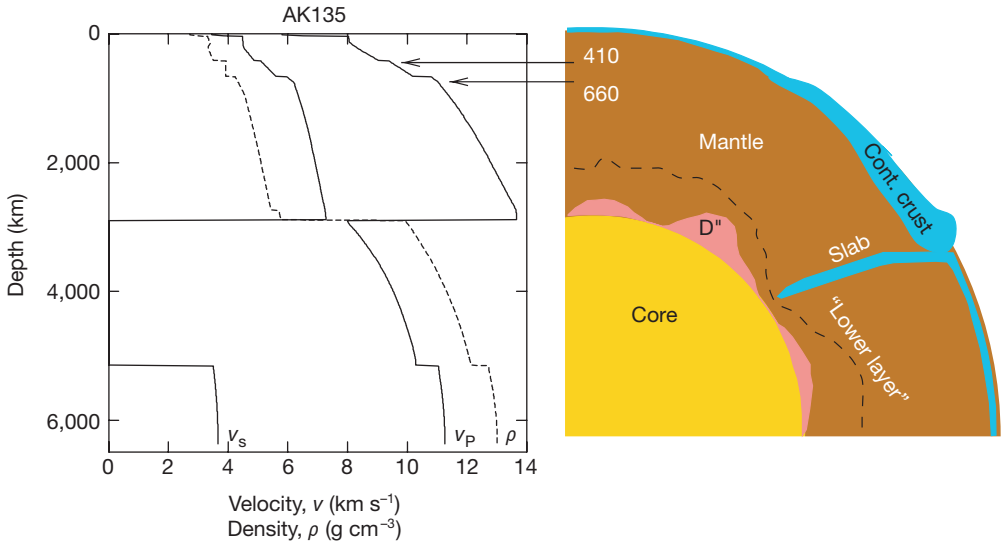

Figure 1 Speeds of seismic waves in the Earth, showing the major discontinuities at 410 and 660 km depth, and the D″ layer at the base of the mantle. Also shown (right) is a cut-away view of the Earth, showing a region of continental crust and a subducted lithospheric slab extending into the lower mantle. The D″ layer is irregular in thickness. A hypothetical ''lower layer''[5] is shown with a dashed line.

radioactive products ^{40}Ar (from K) and ^{4}He (from U, Th) into the atmosphere should be predictable from the surface heat-flow and the approximate abundances of K, U and Th in the bulk Earth. In fact, about 50% of the ^{40}Ar produced over the age of the Earth is missing from the atmosphere[9], while the flux of ^{4}He from the oceans is only 5% of that predicted from oceanic heat flow[10]. These observations again lead naturally to the concept of a region in the deep Earth, rich in the products of radioactive decay, which exchanges heat but little mass with the convecting upper mantle (Fig. 1).

The object of this Review is initially to summarize the data relating to the structure and composition of the mantle with an emphasis on those observations in favour of, or against, chemical layering. We will then present some possible ways in which conflicting evidence might be reconciled, and suggest some future directions that will help resolve the remaining questions.

Mantle structure

The principal data relating to the deep interior are those of seismology. Global one-dimensional seismic models[11] (Fig. 1) clearly show the main features that constrain the mineralogical and chemical constitution of the deep interior. More detailed regional studies provide important additional information. Figure 1 shows that there are two main seismic discontinuities in the mantle, at 410 and 660 km depth. In the lower mantle, below 660 km, seismic velocities increase monotonically with depth until the D″ region is reached at about 200 km above the core–mantle boundary. The latter is a region of low gradient in seismic wave speeds, and contains large-scale low-velocity structures and regional discontinuities[12]. Although the mantle discontinuities (410, 660 and D″) are all, potentially, due to chemical layering of the mantle, pressure-induced phase transformations in peridotite are more plausible explanations for the two shallower discontinuities.

Phase transformations. Bernal[13] was the first to propose that rapid increases in seismic velocity in the mantle might be due to phase transformations rather than a change in composition. Experiments in the mid-1960s[14] showed that the olivine component of peridotite undergoes successive pressure-dependent transformations to the spinel structure (ringwoodite), and ultimately breaks down to form $(Mg,Fe)SiO_3$ perovskite plus (Mg,Fe)O (refs 15,16):

$(Mg,Fe)_2SiO_4$ = $(Mg,Fe)_2SiO_4$		Pressure, 13–14 GPa; depth, 410 km
Olivine	Wadsleyite	
$(Mg,Fe)_2SiO_4$ = $(Mg,Fe)_2SiO_4$		Pressure, 18 GPa; depth, 520 km
Wadsleyite	Ringwoodite	
$(Mg,Fe)_2SiO_4$ = $(Mg,Fe)SiO_3$ + $(Mg,Fe)O$		Pressure, 23 GPa, depth, 660 km
Ringwoodite	Perovskite + Magnesiowüstite	

The transformations of olivine to wadsleyite, and ringwoodite to perovskite (+ oxide), clearly correlate with the two major global seismic discontinuities, and must generate at least some part of the seismic 'signal'. But do phase transformations at constant peridotite bulk composition completely explain the seismic phenomena, or are chemical discontinuities also required? Consider the responses of the three transformations to changes in temperature. Figure 2 (ref. 17) shows a pressure–temperature (P–T) diagram for pure Mg_2SiO_4 that illustrates the main points. The transformation of olivine to wadsleyite has a positive slope, with dP/dT of approximately +3 MPa K^{-1}. In contrast, the breakdown of ringwoodite to perovskite plus magnesiowüstite has a negative slope of –2 MPa K^{-1}. Thus, if the 410- and 660-km discontinuities are entirely due to these phase transformations, regions of abnormally low temperature such as subduction zones should correspond to elevation of the '410' to lower depths and depression of the '660' to greater depths. We address the testing of this prediction in the next section.

A more direct approach to correlating mineralogy and seismic properties is through the elastic properties of the whole mineral assemblage. This requires consideration of transformations in the non-olivine component (30%) through the transition zone (410–660 km) and into the lower mantle.

Phase changes in the pyroxene and garnet components of mantle peridotite are gradual, and lead to changes in slope of the curves of seismic velocity versus depth rather than to discrete discontinuities. The principal transformations involve, first, dissolution of pyroxene into the garnet structure at pressures corresponding to 350–500 km depth[18,19]. Then, at about 580 km depth, a few per cent of $CaSiO_3$ perovskite begins to exsolve from the garnet, which at that point constitutes about 30% of a bimineralic (garnet + ringwoodite) mantle:

Box 1
Glossary

Basalt A rock made by 5–20% melting of peridotite. It is produced under mid- ocean ridges (mid-ocean ridge basalt, MORB) by upwelling of hot mantle. Ocean island basalts (OIBs) are products of similar processes under ocean islands.
Chondritic meteorites A class of meteorites that record the early history of the Solar System. They are thought to represent primitive planetary materials.
EM-1, EM-2 basalts OIBs having isotopic similarities to recycled continental materials.
HIMU basalts OIBs having isotopic similarities to recycled oceanic crust.
Incompatible/compatible When rock starts to melt, some elements (incompatible) are concentrated in the liquid silicate while others (compatible) remain in the solid minerals.
Lithophile/siderophile Elements that dissolve readily in metallic iron under moderately reducing conditions are siderophile, and are concentrated in the core. Those that more readily bond with oxygen are lithophile, and are concentrated in the silicate Earth.
Peridotites Rocks made up predominantly of olivine $(Mg,Fe)_2SiO_4$, with lesser amounts of orthopyroxene $(Mg,Fe)SiO_3$ and clinopyroxene $Ca(Mg,Fe)Si_2O_6$. Although rarely found at the surface, they appear to be the dominant rock type of the mantle.
Refractory Refractory elements condense from gas to solid at high temperatures.

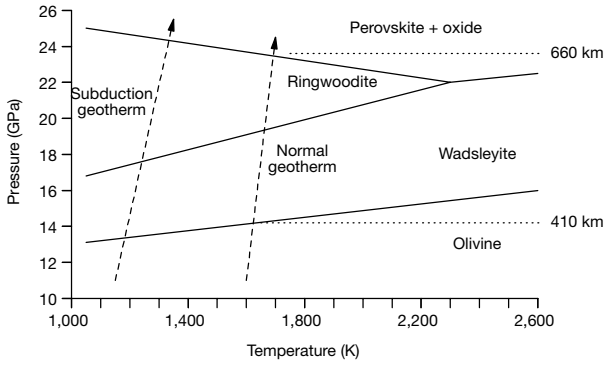

Figure 2 Schematic phase diagram of Mg_2SiO_4 olivine in mantle peridotite. Solid lines show the phase stability regions for olivine, wadsleyite, ringwoodite, and perovskite + oxide. Dashed lines show approximate temperature increase with depth in the mantle and inside a subduction zone. Cooler temperatures in subduction zones move the phase transitions either to shallower depths (olivine → wadsleyite) or deeper depths (ringwoodite → perovskite + oxide).

$$(Mg, Fe, Ca)_3(Si, Al)_2Si_3O_{12} = (Mg, Fe, Ca)_3(Si, Al)_2Si_3O_{12} + CaSiO_3$$

Si-rich garnet Reduced-Ca garnet Perovskite

Finally, between 23 and 26 GPa (660 and 750 km depth), garnet dissolves into the perovskite phase that had initially been produced by breakdown of ringwoodite[20]:

$$(Mg, Fe)SiO_3 + (Mg, Fe, Ca)_3(Si, Al)_2Si_3O_{12} = (Mg, Fe, Al)(Al, Si)O_3 + CaSiO_3$$

Low-Al perovskite Garnet High-Al perovskite Ca-perovskite

Predicted mineralogy should enable calculation of seismic wave speeds, which, if the chemical composition is correct, should match the average one-dimensional wave speed structure. Unfortunately, however, individual elastic constants of the high-pressure minerals have not yet been measured at pressures (13–23 GPa) and temperatures (1,400–1,800 °C) relevant to the Earth's interior. Values of bulk modulus K and density ρ are available, even under lower-mantle conditions[21,22], so $\Phi = K_S/\rho = v_P^2 - (4/3)v_S^2$, rather than the individual seismic velocities v_P and v_S, provides a well-constrained comparison between seismology and mineralogical properties. (Here v_P and v_S are the velocities of P and S waves, respectively.) Results of such calculations are disappointing, however. The data are not accurate enough to discriminate between a deep mantle of peridotitic composition and one which is enriched in Fe or Si, as has been proposed by numerous authors[23–25].

Seismological constraints. Seismological results stimulated most of the ideas on mantle structure discussed above. Historically, radial variations in seismic velocities were derived from analysis of long-period seismic-wave travel times[26–28]. More recently, seismologists have used both short- and long-period data for tomographic models of wave-speed variations[29–31]. Additionally, regional short-period seismic arrays comprising hundreds of instruments have revolutionized the study of small-scale mantle structures—such as the widths of the 410- and 660-km discontinuities[32]. These studies complement the long-period global results by providing detail at small, 10-km, length scales.

The mineralogical reactions responsible for the 410- and 660-km discontinuities have characteristic pressure–temperature slopes ($dP/dT = \Delta S/\Delta V$ where ΔS and ΔV refer to entropy and volume changes of reaction, respectively.) (Fig. 2) that can, in conjunction with short-period seismic studies, be used to constrain mantle properties[33,34]. In the vicinity of a predictable change in mantle temperature structure a phase transformation must change depth, leading to predictable and testable discontinuity topography. For example, the temperature difference (~700 K) between the interior of a subducted lithospheric slab and the surrounding mantle[35] should lead to displacements of approximately −60 km and +30 km of the 410- and 660-km discontinuities, respectively, if they are solely due to the isochemical transformations of olivine to wadsleyite and ringwoodite to perovskite + magnesiowüstite[36]. A study of the depth of the 410-km discontinuity[37] in the Izu-Bonin subduction zone, based on data from short-period arrays, indicated maximum uplift to about 350 km. Detailed study of the 660-km discontinuity in the same region revealed regional depression of about +30 km, and no further mantle features below this depressed transition for a further 300 km (ref. 38). Near subduction zones, the width of the depressed '660 km' transition is a few hundred kilometres[37–39], approximating the dimensions of the 'thermal halo' around a cool slab. The depth and width of the depression is thus in agreement with the simple phase-transformation model of the 660-km discontinuity, and contrary to the predicted dynamical behaviour of subducted material intruding into (or deflected by) a compositionally different layer[40,41]. In the latter case, depression of the discontinuity is predicted to be of the order of several hundred kilometres. The extent of discontinuity topography near subduction zones is thus consistent with isochemical phase changes, and supports the hypothesis that there are no compositional changes at 410 or 660 km depth. The results are also consistent with penetration of slabs into the lower mantle and with deep recycling of former lithospheric material.

Global seismic tomographic images, based on both short- and long-period data, also strongly suggest slab penetration into the lower mantle. A variety of modern studies show narrow features of high seismic velocity extending from sites of present-day subduction deep into the lower mantle[29,30,42]. The implication of these seismological observations, in combination with dynamical modelling of mantle mixing[43,44], is that there is no major change of chemical composition between the upper and lower mantles. But in contrast to the well-resolved high-velocity structures extending into the lower mantle, the diffuse form of warm low-velocity structures in tomographic images do not as yet convincingly delineate return flow from the lower to the upper mantle.

An important feature of the lower mantle that long-period tomographic studies do not resolve is small-scale elastic heterogeneity, which has been discovered using scattered waves[45–48]. These studies observed bodies distributed throughout the lower mantle that were smaller than 10 km in size, and whose wave speeds varied by 1% or more from that of the ambient mantle. In the few cases where individual scatterers were identifiable, they had seismic velocities at least 4% slower than their surroundings[48] and arrayed themselves in linear structures resembling heterogeneities stretched by convective stirring in the mantle[49–51]. These 'scatterers' must be chemical rather than thermal heterogeneities because they are too small (assuming thermal diffusivity of 1×10^{-6} m^2 s^{-1}; ref. 52) to maintain a temperature difference from the mantle for longer than about 200,000 years. Given the tomographic evidence for deep subduction, the most likely explanation is that they are remnants of lithospheric slabs.

Substantial changes in seismic wave speeds occur locally in the lowermost part of the mantle, D″. This approximately 200-km-thick region varies in thickness, and contains discrete discontinuities[53], scattering features[46,54–56], and possibly, thin low-velocity layers against the core–mantle boundary[57,58]. D″ is also anisotropic, and the variations in P-wave and S-wave speeds are uncorrelated[59,60]. In combination, these features suggest that D″ is a region with pronounced lateral chemical heterogeneity that differs chemically from the overlying mantle.

These seismological properties of mantle structure may be summarized in a map (Fig. 3) of the angular correlation function of seismic heterogeneity provided by analysing recent mantle tomographic studies[61]. The angular correlation function reveals the lateral extent of correlated variations in wave speed at a given depth in the mantle. Lateral heterogeneity in the upper mantle is significant, and is organized into broad regions (continents and oceans)[60], shown by the wide correlation lengths. In the lower mantle, the correlation length diminishes to an approximately constant level until the base of the mantle is approached, when it increases again owing to the structure of D″. There are no changes in correlation length in the lower mantle attributable to boundaries between elastically different material. Thus, whatever heterogeneity is present in the lower mantle, the observational seismic data only require it to be homogeneously distributed, without undergoing any significant scale reorganization with depth.

In summary, modern seismological studies show that the mantle is chemically heterogeneous at the 10-km length scale, but give no clear evidence of radial stratification other than at the boundary with the D″ region. Both tomographic and seismic array studies are consistent with deep subduction and hence with whole-mantle convection.

Heat flow and geochemistry

Current global heat loss is 44.2×10^{12} W (44.2 TW)[62]. This is made up of contributions from radioactive decay of K, U and Th and from secular cooling of the Earth. Measurements of the compositions of the most primitive, unfractionated, mantle peridotites indicate that

refractory lithophile elements such as Ca, Sc, Ti, the rare-earth elements (REE), U and Th are in chondritic ratio one to another in the silicate part of the Earth[1,63]. This leads to concentrations of about 20 p.p.b. (parts per billion) U and 80 p.p.b. Th in bulk silicate Earth (that is, mantle + crust). Chemical analysis of igneous rocks indicates that their K/U ratios are generally about 12,500, irrespective of rock type or tectonic settting[64]. This implies a K content of bulk silicate Earth of about 12,500 times the U content, or 250 p.p.m.

This inventory of radioactive elements is at present producing 20 TW, or about 45% of the total heat flow[65]. But the mantle has a long thermal time constant, and advective removal of heat is linked to the vigour of plate-tectonic motions that vary over time, so that the heat loss measured now was generated by decay in the distant past. Although the age of the current heat-loss is, of course, model-dependent, we can make defensible estimates based on the apparent ages of isotopic heterogeneities in the mantle, such as those of modern **HIMU basalts**[8]. These are being generated now from melting of oceanic crust that was recycled from the surface into the deep interior 1–2 Gyr ago. Using the latter figures as a measure of the age of the heat reaching the surface yields a radioactive heat contribution of 24.9–32.9 TW (ref. 65). Parametrized convection models of the cooling of the Earth's interior yield secular cooling rates of 30–100 K Gyr^{-1} (refs 66, 67). This loss of the Earth's heat energy produces heat fluxes between 5.9 and 20 TW, which are capable of making up the balance from the geochemically inferred radioactive heat production.

By mass-balancing the different parts of the silicate Earth to the overall bulk silicate composition[1], we can determine how much 'hidden' layering is necessary to make the whole. We will concentrate on the heat-producing elements, and begin by considering two major reservoirs whose concentrations of K, U and Th are reasonably well-known—the continental crust[68] and the mantle source of MORB[69]. From the major-element compositions of MORB, it appears that these voluminous volcanic rocks are produced by about 10% melting of peridotite mantle[70,71]. Assuming that K, U and Th act as **incompatible** elements, this means that the mantle contains about 0.1 times the concentrations found in MORB. In that case the mantle source can only have about one-third the concentrations of K, U and Th (and other incompatible elements) postulated for the bulk silicate Earth[1]. It is therefore 'depleted' mantle. As the continental crust is highly enriched in these incompatible elements, the simplest assumption is that enriched crust and depleted mantle are complementary, and formed by differentiation of bulk silicate Earth[69]. The result (Table 1) indicates that there is insufficient enriched crust for the depleted peridotite mantle to comprise the whole mantle—it turns out to constitute only 50% of the whole mantle. The calculation can be repeated with isotope systems such as Sm–Nd with very similar results[72]. Mantle layering (Fig. 1) is the logical conclusion.

The simple layered model is, however, compromised by seismological evidence in favour of whole-mantle convection, or at least of deep subduction of slabs into the lower mantle and of heterogeneities introduced by subduction. That these heterogeneities are real and long-lasting is attested to by seismological evidence of lower-mantle 'scatterers', and by enriched ocean island basalts (OIBs) that contain isotopic evidence of contributions from old subducted oceanic crust (HIMU) and from continental crust and lithosphere (**EM-1**, **EM-2**)[8]. The current rate of subduction[73] of 3 km^2 per year means that approximately 18 km^3 of oceanic crust—and, assuming 10% melting by mass, 140 km^3 of highly depleted, 'sterile' peridotite—is being returned to the mantle each year. Rather more difficult to estimate is the amount of continental sediment that is returned to the mantle with the subducted oceanic plates. Recent estimates vary between 0.7 and 1.5 km^3 per year[74]. Therefore, assuming a constant rate of subduction for 4 Gyr, we find that the mantle contains 5% recycled oceanic crust, 45% recycled 'sterile' mantle and about 0.3% recycled continental material. If the higher rate of heat production in the past (3.6 times current values at 4 Gyr ago) was compensated by higher heat loss through formation of oceanic lithosphere, then it seems likely that almost the entire mantle has been through the cycle of partial melting, oceanic crust formation and sediment recycling. In that case a better view of the mantle today would be as a heterogeneous mixture of 'sterile' highly depleted mantle containing 'blobs' of oceanic and continental crust of different ages and sizes. Such 'blobs' are real, and recorded as chemical heterogeneities in studies of scattered seismic waves.

In Fig. 4 we show how a heterogeneous model of the mantle would fit with the chemical composition of the silicate Earth. Assuming that there is no primitive undepleted mantle, the silicate Earth consists of continental crust (0.6% by mass), 'sterile' mantle (> 45%), and recycled continental and oceanic material. As the depleted mantle produces basalt by about 10% partial melting[70,71], we can consider it a mixture of 10% recycled oceanic crust and 90% 'sterile' peridotite. There are two end-member approaches for mass-balancing K, U and Th in the silicate Earth that differ in their complexity. Either yields agreement with K, U and Th abundances, so we consider the differences they predict in other trace-element abundances to discriminate between them. One end-member assumes that average rates of oceanic crust production over the past 4 Gyr were twice the present rate, processing the whole mantle through the MORB production process, with 0.4% recycled continental crust material. The other end-member assumes a lower rate

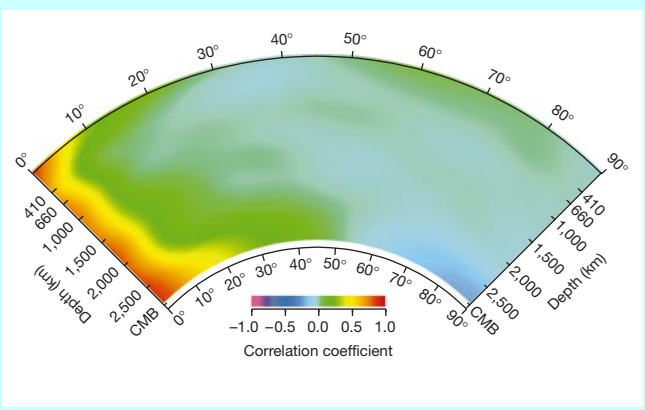

Figure 3 Angular correlation function of a shear-wave tomographic model[30] as a function of depth. This characterizes the similarity of wave speed of two points at the same depth in the mantle separated by some angular distance. Every point in the mantle is perfectly correlated with itself, resulting in the maximum correlation for points separated by zero distance (left side of figure). A drop of e^{-1}, or about 0.4, indicates a significant correlation decrease (the yellow band). Blue colours representing zero or negative correlations arise at large separations, and a degree-two pattern at the mantle's base is made evident by the maximum negative correlation at ~90°. The correlation length is greatest in the upper mantle and in D″. In the upper mantle, this probably corresponds to continent/ocean differences[60], while the structure in D″ possibly reflects the pattern of heat flow out of the core into the mantle. In the lower mantle above D″ the correlation length is approximately constant, indicating no reorganization of heterogeneity scale length in the mid–lower mantle. CMB, core–mantle boundary.

Table 1 Mass-balance calculation of the fraction of mantle depleted by continental crust formation

Element	Bulk silicate Earth (p.p.m.)	Continental crust (p.p.m.)	Depleted mantle (p.p.m.)	Fraction of mantle that is depleted
K	250	15,800	85	0.50
U	0.02	1.4	0.0065	0.54
Th	0.08	5.6	0.0164	0.46

of oceanic crust production, no recycled continental crust, and that the 16% remaining, yet-to-be-processed mantle mass fraction is excess—or 'old'—material left over from an episode of early differentiation and oceanic crust extraction, as Hf isotope evidence suggests[75]. In either case, modern MORB comes from a 'marble-cake'[50] mixture of 'sterile' mantle and recycled oceanic crust. OIB has a similar origin, but the melting process must involve more of the recycled oceanic and continental crust in order to generate its higher incompatible trace-element content[8,50]. We can test the two end-members by mass-balancing the bulk silicate Earth contents of elements such as Ti, Zr and La, which are geochemically similar to K, U and Th. The end-member with no 'old' recycled oceanic crust requires about 1% additional oceanic crust to mass-balance (Fig. 4). The end-member with 'old' excess crust overestimates the trace element abundances by an amount equivalent to 10% additional oceanic crust. Best overall agreement is achieved with a 75%/25% mixture of the end-member models (0.3% recycled continental material and 5% 'excess' recycled oceanic crust). In either, the recycled material in depleted mantle (>90% of the mantle) is distributed as the small, kilometre-scale blobs discussed earlier. Recycled continental and 'excess' oceanic crust is largely isolated from the MORB reservoir, suggesting either a thin lower-mantle layer or that the lowermost mantle contains a high proportion of very long-lived blobs.

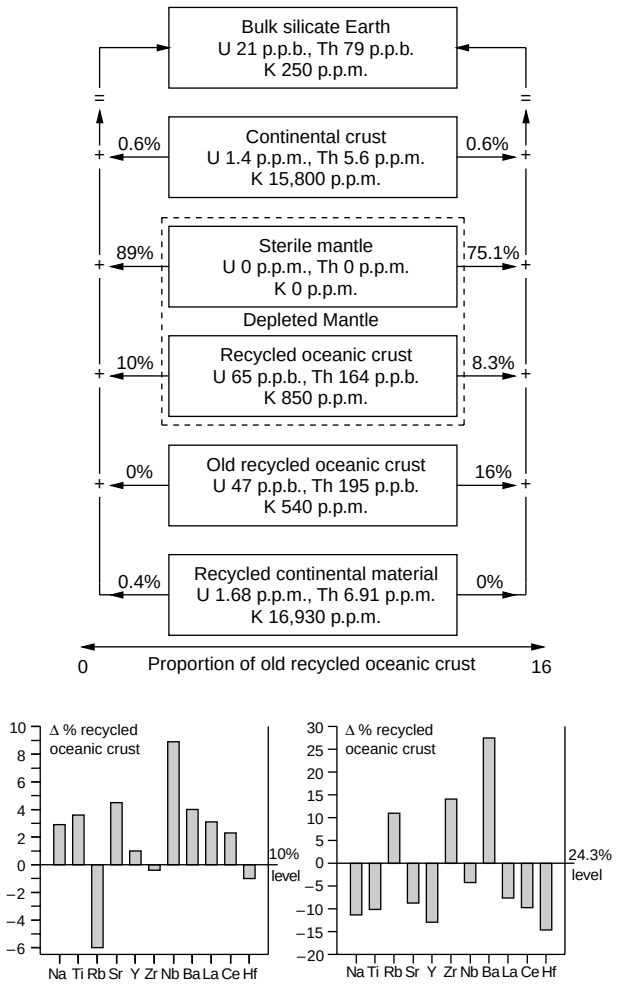

Figure 4 Approximate mass balance for a bulk silicate Earth made up of continental crust, 'sterile' mantle, recycled oceanic crust, old recycled oceanic crust, and recycled continental material. The proportions satisfy seismological, heat-flow and most trace-element geochemical constraints. Diagrams at bottom show misfit to trace-element abundances if the total amount of recycled oceanic crust is varied. The large Nb value is symptomatic of the "missing niobium problem"[8,87]. Ba and Rb values are very sensitive to the amount of recycled continental crust and to unknown effects of subduction zone fluids.

This simple model is a starting point for considering the mantle as a laterally heterogeneous, rather than a vertically stratified, system. In its simplicity, there are some geochemical aspects that it ignores. It ignores the role of delaminated continental crust, a possible component of EM-1 type OIB[8], and the selective extraction and recycling of K and U by fluids in subduction zones[76]. Nor do we consider any isotopic constraints on the necessity for primitive 'untapped' reservoirs somewhere in the Earth. Although both Sm–Nd and Rb–Sr isotope systematics can be fitted into the recycling scheme presented here, noble-gas isotopes should provide stronger constraints. The high ratio of primordial ^{3}He to the radiogenic (from U, Th decay) ^{4}He in some OIBs and the decoupling of heat from ^{4}He flux are considered evidence of an undegassed deep reservoir[7,10,84,85]. The bulk silicate Earth[1] has a K content which should have, over 4.5 Gyr, produced about twice as much ^{40}Ar as is observed in the atmosphere, crust and depleted mantle[9], also implying that there is an isolated reservoir somewhere. These conclusions are, however, based on the assumption that the noble gases are highly incompatible and if this is not the case, they could be invalidated. For example, the ^{3}He/^{4}He ratio of all parts of the Earth is decreasing with time owing to production of ^{4}He from U and Th. If, as seems possible, He is more compatible in mantle crystals than U and Th[86], then very old subducted mantle, which had undergone partial melting in the distant past, would have a relatively high ^{3}He/U ratio. This region would retain a high ^{3}He/^{4}He ratio through geological time because of the low rate of ^{4}He production. Some OIB might tap such a source rather than a primitive layer. Further complexities will clearly be required as the behaviour of the noble gases in OIB and MORB generation becomes better understood.

An integrated view of the mantle

Seismological and chemical evidence shows that heterogeneity is ubiquitous in the mantle, and that subduction is the key process producing it[8,77]. This heterogeneity extends from top to bottom of the mantle under different guises. In the upper mantle, it takes the form predominantly of cold subducted lithospheric slabs. Their excess density carries them into the lower mantle where they either stagnate through thermal equilibration or sink to the core–mantle boundary. Simultaneously, the material mixes back into the mantle more effectively with time as thermal equilibration reduces rheological and density differences. Heterogeneity diminution by stretching at flow stagnation points yields a spectrum of sizes whose average diminishes with time but which still retains large fragments[49,51]. D″ might also play a homogenizing role as a storage and entrainment site for oceanic crust[44,78,79]. Rheological differences between the admixed material and the peridotitic mantle tend to inhibit and delay homogenization[80]. In consequence, lower-mantle heterogeneity almost certainly exists at many scales, representing all stages in the mixing process. This is seismically expressed in the lower mantle as discrete lithospheric slabs, as 8-km-scale scatterers and in the unusual array of D″ properties.

How is this lower-mantle heterogeneity distributed and what proportion of it is likely to be seismically observable? To answer this we use the constraints of seismic scattering[81]. The volumetric fractional scattered power of a heterogeneity, $\langle\Phi\rangle^2/VA^2$, depends on its volume a^3, the average velocity perturbation v in the body, the scattering angle θ and wavenumber k, and decreases with distance r from the scatterer:

$$\left[\frac{\langle\Phi\rangle^2}{VA^2}\right] = a^3\left[\frac{k^4 v}{6\pi r^2}\right]\left[\frac{3\cos\theta + 1 + 2\cos^2\theta}{1 + 4k^2a^2\sin^2(\theta/2)}\right]$$

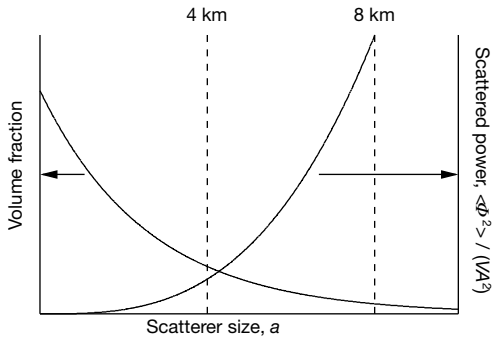

Figure 5 Visibility of mantle heterogeneity by seismic scattering. The abundance of mantle heterogeneities decreases with size exponentially (left axis), while the scattering power increases with the cube of the scatterer size (right axis). Scatterers smaller than ~4 km are not likely to be visible on account of the attenuation of high-frequency body waves in the upper mantle (wavenumber k ∝ frequency) that would be sensitive to them. Consequently, only the largest mantle heterogeneities yield observable scattering. With values of $\kappa = 2.25 \times 10^{-3}\,\mathrm{km}^{-1}$ and $f_0 = 2.04 \times 10^{-3}$ that satisfy the observational scattering constraints and the abundance of non-'sterile' mantle components, 93 vol.% of heterogeneity should be < 4 km in size, and hence invisible to seismic scattering.

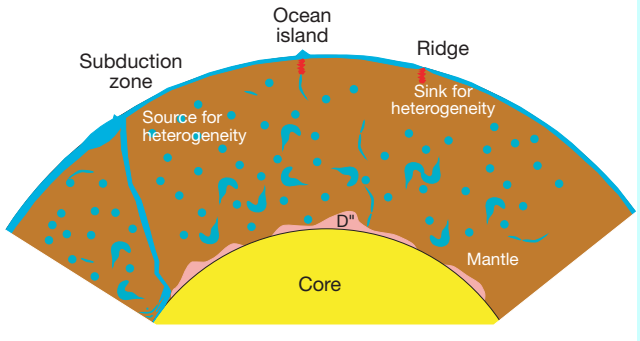

Figure 6 Sketch of proposed model of a chemically unstratified mantle. Subduction of oceanic lithosphere introduces heterogeneity into the mantle. Mixing by convective stirring of the mantle disaggregates the subducted lithosphere and minor continental material, producing isolated heterogeneities that scatter seismic energy but are too small to be observed tomographically. Melting at mid-ocean ridges and at ocean islands produces basalts and homogenizes the two types of mantle material, one enriched in incompatible elements and the other 'sterile'.

Models of the convective stretching and dispersal of heterogeneities indicate that average size becomes smaller with time, with an exponential dependence of fractional volume f on size a (ref. 49): $f(a) = f_0 \exp(-\kappa a)$. The unknowns f_0 and κ can be constrained because the integral of f over heterogeneity sizes 0–8 km must yield the volume fraction of the mantle occupied by heterogeneities (around 11–16% from geochemical arguments). Also, the fraction at a = 8 km must agree with the number of seismically observed heterogeneities of that size. The latter is unknown, because much of the mantle has yet to be systematically explored for scatterers, but it should be of the order of the volume of subducted oceanic crust currently *en route* to the base of the mantle. This volume is about 0.2% of the mantle, assuming the present global rate of mass transfer into subduction zones and a vertical descent rate in the mantle of 30 mm yr^{-1} (refs 30, 48). Both constraints are satisfied by $\kappa = 2.25 \times 10^{-3}\,\mathrm{km}^{-1}$ and $f_0 = 2.04 \times 10^{-3}$, leading to the prediction that 93% (by volume) of the heterogeneity is smaller than 4 km, which would be invisible to short-period wave scattering (Fig. 5).

From a seismological viewpoint, the mantle therefore appears to exhibit an overlapping collection of vertically homogeneous regions that coexist with heterogeneity on a range of length scales (Fig. 6). However, most of the heterogeneity volume is represented by very small bodies (< 4 km). This homogeneous distribution of heterogeneity extends from the base of the mantle into the MORB melting region, 40–80 km below the surface, where it is destroyed by melting and melt extraction. The latter melting region acts as a near-surface 'heterogeneity sink'—the larger heterogeneities can only be inferred from their chemical signatures. Despite having to traverse the region of homogenization, however, large peridotite bodies tectonically emplaced in the continental crust still exhibit a wide range of metre- and centimetre-scale heterogeneities[82,83].

Perspective and outstanding problems

Early interpretations of the seismic structure of the Earth, incorporating what was known about the compositions of meteorites, suggested a gradual change from silicate to iron-rich compositions with depth. This view was abandoned in favour of a layered mantle, once the nature of the core–mantle boundary was recognized, and finite strain theories became available to model changes in material elasticity with depth. The seismic discontinuities at 410 and 660 km depth initially appeared to be likely boundaries for the compositional layering. Recent studies of discontinuity topography and tomographic images of slab penetration into the lower mantle have, however, forced revision of this view. The latter results are most consistent with the discontinuities corresponding to phase transformations under isochemical conditions in a peridotite mantle. Stratification seems, therefore, to be an increasingly difficult position to defend. Lateral heterogeneity is, however, well established in the upper mantle, and has recently been shown to be present in the lower mantle. The principal issue is becoming how to interpret the heterogeneity seen at all depths in the mantle and to integrate it with the Earth's chemical composition. We have presented here a simple geochemical model based on four components—'sterile' peridotite, recycled oceanic crust, and continental crust part of which is recycled—that is one way to explain most of the geophysical and geochemical observations.

With present techniques, seismic tomography, which averages over large volumes in the lower mantle, cannot resolve heterogeneities smaller than 400 km in size. Scattering, however, probes subwavelength scales in the mantle (~10 km), and the challenge will be to integrate this into whole-Earth studies. The aim now, guided by convective mixing models, will be to decipher the style and rate of mantle mixing, placing constraints on the residence times of the chemical components of the mantle in their respective reservoirs. The challenge to experimental and isotope geochemists is clearly to understand better the sources and sinks of rare gases in the Earth, and the mechanisms by which they are extracted from the mantle and deposited in the atmosphere. □

1. McDonough, W. F & Sun, S.-S. The composition of the Earth. *Chem. Geol.* **120,** 223–253 (1995).
2. Newsom, H. E. in *Global Earth Physics: Handbook of Physical Constants* Vol. 1 (ed. Ahrens, T. J.) 159–189 (Reference Shelf Series, American Geophysical Union, Washington DC, 1995).
3. Allègre, C. J., Manhès, G. & Gopel, C. The age of the Earth. *Geochim. Cosmochim. Acta* **59,** 1445–1456 (1995).
4. Halliday, A., Rehkämper, M., Lee, D.-C. & Yi, W. Early evolution of the Earth and Moon: New constraints from Hf-W isotope geochemistry. *Earth Planet. Sci. Lett.* **142,** 75–89 (1996).
5. Kellogg, L. H., Hager, B. H. & van der Hilst, R. D. Compositional stratification in the deep mantle. *Science* **283,** 1881–1884 (1999).
6. Schilling, J.-G. Iceland mantle plume: geochemical evidence along Reykjanes Ridge. *Nature* **242,** 565–571 (1973).
7. Allègre, C. J., Staudacher, T. & Sarda, P. Rare-gas systematics—formation of the atmosphere, evolution and structure of the Earth's mantle. *Earth Planet. Sci. Lett.* **81,** 127–150 (1987).
8. Hofmann, A. W. Mantle geochemistry: the message from oceanic volcanism. *Nature* **385,** 219–229 (1997).
9. Allègre, C. J., Hofmann, A. W. & O'Nions, R. K. The argon constraints on mantle structure. *Geophys. Res. Lett.* **23,** 3555–3557 (1996).
10. O'Nions, R. K. & Oxburgh, R. Heat and helium in the Earth. *Nature* **306,** 429–431 (1983).

11. Kennett, B. L. N., Engdahl, E. R. & Buland, R. Constraints on seismic velocities in the Earth from traveltimes. *Geophys. J. Int.* **122,** 108–124 (1995).
12. Lay, T., Williams, Q. & Garnero, E. J. The core–mantle boundary layer and deep Earth dynamics. *Nature* **392,** 461–468 (1998).
13. Bernal, J. D. Commentary. *Observatory* **59,** 265–269 (1936).
14. Ringwood, A. E. & Major, A. High-pressure transformations in pyroxenes. *Earth Planet. Sci. Lett.* **1,** 241–245 (1966).
15. Liu, L. The post-spinel phase of forsterite. *Nature* **262,** 770–772 (1976).
16. Ito, E. & Takahashi, E. Postspinel transformations in the system Mg_2SiO_4-Fe_2SiO_4 and some geophysical implications. *J. Geophys. Res.* **94,** 10637–10646 (1989).
17. Fei, Y., Saxena, S. & Navrotsky, A. Internally consistent thermodynamic data and equilibrium phase relations for compounds in the system MgO-SiO_2 at high pressure and temperature. *J. Geophys. Res.* **95,** 6915–6928 (1990).
18. Akaogi, M. & Akimoto, S. Pyroxene-garnet solid-solution equilibria in the systems $Mg_4Si_4O_{12}$ - $Mg_3Al_2Si_3O_{12}$ and $Fe_4Si_4O_{12}$ - $Fe_3Al_2Si_3O_{12}$ at high pressures and temperatures. *Phys. Earth Planet. Inter.* **15,** 90–106 (1977).
19. Ringwood, A. E. Phase transformations and their bearing on the constitution and dynamics of the mantle. *Geochim. Cosmochim. Acta* **55,** 2083–2110 (1991).
20. Wood, B. J. Phase transformations and partitioning relations in peridotite under lower mantle conditions. *Earth Planet. Sci. Lett.* **174,** 341–354 (2000).
21. Funamori, N., Yagi, T. & Utsumi, W. Thermoelastic properties of $MgSiO_3$ perovskite determined by in situ X-ray observations up to 30 GPa and 2000 K. *J. Geophys. Res.* **101,** 8257–8269 (1996).
22. Fiquet, G., Dewaele, A., Andrault, D., Kunz, M. & LeBihan, T. Thermoelastic properties and crystal structure of $MgSiO_3$ perovskite at lower mantle pressure and temperature conditions. *Geophys. Res. Lett.* **27,** 21–24 (2000).
23. Anderson, D. L & Bass, J. D. Transition region of the Earth's upper mantle. *Nature* **320,** 321–328 (1986).
24. Jeanloz, R. & Knittle, E. Density and composition of the lower mantle. *Phil. Trans. R. Soc. Lond. A* **328,** 377–389 (1989).
25. Stixrude, L., Hemley, R. J., Fei, Y. & Mao, H. K. Thermoelasticity of silicate perovskite and magnesiowüstite and stratification of the Earth's mantle. *Science* **257,** 1099–1101 (1992).
26. Jeffreys, H. & Bullen, K. E. *Seismological Tables* (British Association Seismological Investigators Committee, London, 1940).
27. Shearer, P. M. Global mapping of upper-mantle reflectors from long-period SS precursors. *Geophys. J. Int.* **115,** 878–904 (1993).
28. Flanagan, M. P. & Shearer, P. M. Global mapping of topography on transition zone velocity discontinuities by stacking SS precursors. *J. Geophys. Res.* **103,** 2673–2692 (1998).
29. van der Hilst, R., Engdahl, E., Spakman, W. & Nolet, G. Tomographic imaging of subducted lithosphere below northwest Pacific island arcs. *Nature* **353,** 37–43 (1991).
30. Grand, S. P. Mantle shear structure beneath the americas and surrounding oceans. *J. Geophys. Res.* **99,** 11591–11621 (1994).
31. Masters, G., Johnson, S., Laske, G. & Bolton, H. A shear velocity model of the mantle. *Phil. Trans. R. Soc. Lond. A* **354,** 1385–1410 (1996).
32. Vidale, J. E & Benz, H. M. Upper mantle seismic discontinuities and the thermal structure of subduction zones. *Nature* **356,** 678–683 (1992).
33. Helffrich, G. R. & Bina, C. R. Frequency dependence of the visibility and depths of mantle seismic discontinuities. *Geophys. Res. Lett.* **21,** 2613–2616 (1994).
34. Stixrude, L. Structure and sharpness of phase transitions and mantle discontinuities. *J. Geophys. Res.* **102,** 14835–14852 (1997).
35. McKenzie, D. P. Speculations on the consequences and causes of plate motions. *Geophys. J. R. Astron. Soc.* **18,** 1–32 (1969).
36. Bina, C. R. & Helffrich, G. Phase-transition Clapeyron slopes and transition zone seismic discontinuity topography. *J. Geophys. Res.* **99,** 15853–15860 (1994).
37. Collier, J. & Helffrich, G. Topography of the "410" and "660" km seismic discontinuities in the Izu-Bonin subduction zone. *Geophys. Res. Lett.* **24,** 1535–1538 (1997).
38. Castle, J. C & Creager, K. C. Seismic evidence against a mantle chemical discontinuity near 660 km depth beneath Izu-Bonin. *Geophys. Res. Lett.* **24,** 241–244 (1997).
39. Wicks, C. W. & Richards, M. A. A detailed map of the 660-kilometer discontinuity beneath the Izu-Bonin subduction zone. *Science* **261,** 1424–1427 (1993).
40. Christensen, U. & Yuen, D. The interaction of a subducting lithospheric slab with a chemical or phase boundary. *J. Geophys. Res.* **89,** 4389–4402 (1984).
41. Kincaid, C. & Olson, P. An experimental study of subduction and slab migration. *J. Geophys. Res.* **92,** 13822–13840 (1987).
42. Bijwaard, H., Spakman, W. & Engdahl, E. R. Closing the gap between regional and global travel time tomography. *J. Geophys. Res.* **103,** 30055–30078 (1998).
43. Davies, G. & Richards, M. Mantle convection. *J. Geol.* **100,** 151–206 (1992).
44. Christensen, U. R. & Hofmann, A. W. Segregation of subducted oceanic crust in the convecting mantle. *J. Geophys. Res.* **99,** 19867–19884 (1994).
45. Hedlin, M. A. H., Shearer, P. M. & Earle, P. S. Seismic evidence for small-scale heterogeneity throughout the Earth's mantle. *Nature* **387,** 145–150 (1997).
46. Vidale, J. E & Hedlin, M. A. H. Evidence for partial melt at the core-mantle boundary north of Tonga from the strong scattering of seismic waves. *Nature* **391,** 682–685 (1998).
47. Castle, J. C. & Creager, K. C. Topography of the 660-km seismic discontinuity beneath Izu-Bonin: Implications for tectonic history and slab deformation. *J. Geophys. Res.* **103,** 12511–12527 (1998).
48. Kaneshima, S & Helffrich, G. Dipping low-velocity layer in the mid-lower mantle: Evidence for geochemical heterogeneity. *Science* **283,** 1888–1891 (1999).
49. Olson, P., Yuen, D. A. & Balsiger, D. Convective mixing and the fine structure of mantle heterogeneity. *Phys. Earth Planet. Inter.* **36,** 291–304 (1984).
50. Allègre, C. J & Turcotte, D. L. Implications of a two-component marble-cake mantle. *Nature* **323,** 123–127 (1986).
51. Gurnis, M. & Davies, G. F. Mixing in numerical models of mantle convection incorporating plate kinematics. *J. Geophys. Res.* **91,** 6375–6395 (1986).
52. Chai, M., Brown, J. M. & Slutsky, L. J. Thermal diffusivity of mantle minerals. *Phys. Chem. Minerals* **23,** 470–475 (1996).
53. Weber, M. & Davis, J. P. Evidence of a laterally variable lower mantle structure from P- and S-waves. *Geophys. J. Int.* **102,** 231–255 (1990).
54. Earle, P. S. & Shearer, P. M. Observations of PKKP precursors used to estimate small-scale topography on the core-mantle boundary. *Science* **277,** 667–670 (1997).
55. Yamada, A. & Nakanishi, I. Short-wavelength lateral variation of a D" P-wave reflector beneath the southwestern Pacific. *Geophys. Res. Lett.* **25,** 4545–4548 (1998).
56. Thomas, C., Weber, M., Wicks, C. W. & Scherbaum, F. Small scatterers in the lower mantle observed at German broadband arrays. *J. Geophys. Res.* **104,** 15073–15088 (1999).
57. Revenaugh, J. & Meyer, R. Seismic evidence of partial melt within a possibly ubiquitous low-velocity layer at the base of the mantle. *Science* **277,** 670–673 (1997).
58. Garnero, E. J. & Helmberger, D. V. Further structural constraints and uncertainties of a thin laterally varying ultralow-velocity layer at the base of the mantle. *J. Geophys. Res.* **103,** 12495–12509 (1998).
59. Kendall, J. M. & Silver, P. G. Constraints from seismic anisotropy on the nature of the lowermost mantle. *Nature* **381,** 409–412 (1996).
60. Su, W.-J. & Dziewonski, A. M. Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle. *Phys. Earth Planet. Inter.* **100,** 135–156 (1997).
61. Grand, S. P., van der Hilst, R. D. & Widiyantoro, S. Global seismic tomography: A snapshot of convection in the Earth. *GSA Today* **7,** 1–7 (1997).
62. Pollack, H. N., Hurter, S. J. & Johnston, J. R. Heat flow from the Earth's interior—analysis of the global data set. *Rev. Geophys.* **31,** 267–280 (1993).
63. Hart, S. R. & Zindler, A. In search of a bulk-Earth composition. *Chem. Geol.* **57,** 247–267 (1986).
64. Jochum, K. P., Hofmann, A. W., Ito, E., Seufert, H. M. & White, W. M. K, U and Th in mid-ocean ridge basalt glasses and heat production, K/U and K/Rb in the mantle. *Nature* **306,** 431–436 (1986).
65. Van Schmus, W. R. in *Global Earth Physics: Handbook of Physical Constants* Vol. 1 (ed. Ahrens, T. J.) 283–291 (Reference Shelf Series, American Geophysical Union, Washington DC, 1995).
66. Jackson, M. J. & Pollack, H. N. On the sensitivity of parameterized convection to the rate of decay of internal heat sources. *J. Geophys. Res.* **89,** 10103–10108 (1984).
67. Christensen, U. Thermal evolution models for the Earth. *J. Geophys. Res.* **90,** 2995–3007 (1985).
68. Rudnick, R. L. Making continental crust. *Nature* **378,** 571–578 (1995).
69. Hofmann, A. W. Chemical differentiation of the Earth: the relationship between mantle, continental crust and oceanic crust. *Earth Planet. Sci. Lett.* **90,** 297–314 (1988).
70. Kinzler, R. J. & Grove, T. L. Primary magmas of midocean ridge basalts. 2. Applications. *J. Geophys. Res.* **97,** 6907–6926 (1992).
71. Klein, E. M. & Langmuir, C. H. Global correlations of ocean ridge basalt chemistry with axial depth and crustal thickness. *J. Geophys. Res.* **92,** 8089–8115 (1987).
72. Allègre, C. J. Chemical geodynamics. *Tectonophysics* **81,** 109–132 (1982).
73. Chase, C. G. The n-plate problem of plate tectonics. *Geophys. J. R. Astron. Soc.* **29,** 117–122 (1972).
74. Plank, T. & Langmuir, C. H. The geochemical composition of subducting sediment and its consequences for the crust and mantle. *Chem. Geol.* **145,** 325–394 (1998).
75. Blichert-Toft, J. & Albarède, F. The Lu-Hf isotope geochemistry of chondrites and the evolution of the mantle-crust system. *Earth Planet. Sci. Lett.* **149,** 243–258 (1997).
76. Elliott, T., Zindler, A. & Bourdon, B. Exploring the kappa conundrum: the role of recycling in the lead isotope evolution of the mantle. *Earth Planet. Sci. Lett.* **169,** 129–145 (1999).
77. Davies, G. F. Earth's neodymium budget and structure and evolution of the mantle. *Nature* **290,** 208–213 (1981).
78. Hofmann, A. W. & White, W. M. Mantle plumes from ancient oceanic crust. *Earth Planet. Sci. Lett.* **57,** 421–436 (1982).
79. Sleep, N. H. Gradual entrainment of a chemical layer at the base of the mantle by overlying convection. *Geophys. J. Int.* **95,** 437–448 (1988).
80. Metcalfe, G., Bina, C. R. & Ottino, J. M. Kinematic considerations for mantle mixing. *Geophys. Res. Lett.* **22,** 743–746 (1995).
81. Chernov, L. A. *Wave Propagation in a Random Medium* (Dover, New York, 1960).
82. Kornprobst, J. Le massif ultrabasique des Beni Bouchera (Rif Interne, Maroc): étude des péridotites de haute temperature et de haute pression, et des pyroxénites, à grenat ou sans grenat, qui leur sont associées. *Contrib. Mineral. Petrol.* **23,** 283–322 (1969).
83. Nicolas, A., Bouchez, J. L. & Boudier, F. Interprétation cinématique des déformations plastiques dans le massif de lherzolite de Lanzo (Alpes Piémontaises)—comparaison avec d'autres massifs. *Tectonophysics* **14,** 143–171 (1972).
84. Porcelli, D. & Wasserburg, G. J. Mass-transfer of helium, neon, argon, and xenon through a steady-state upper-mantle. *Geochim. Cosmochim. Acta* **59,** 4921–4937 (1995).
85. O'Nions, R. K & Tolstikhin, I. N. Limits on the mass flux between lower and upper mantle and stability of layering. *Earth Planet. Sci. Lett.* **139,** 213–222 (1996).
86. Broadhurst, C. L., Drake, M. J., Hagee, B. E. & Bernatowicz, T. J. Solubility and partitioning of Ar in anorthite, diopside, forsterite, spinel, and synthetic basaltic liquids. *Geochim. Cosmochim. Acta* **54,** 299–309 (1990).
87. Rudnick, R. L., Barth, M., Horn, I. & McDonough, W. F. Rutile-bearing refractory eclogites: Missing link between continents and depleted mantle. *Science* **287,** 278–281 (2000).

Correspondence and requests for materials should be addressed to G.R.H. (e-mail: george@geology.bristol.ac.uk).