Compositional Heterogeneity in the Bottom 1000 Kilometers of Earth’s Mantle: Toward a Hybrid Convection Model

Rob D. van der Hilst* and Hrafnkell Kárason

Tomographic imaging indicates that slabs of subducted lithosphere can sink deep into Earth’s lower mantle. The view that convective flow is stratified at 660-kilometer depth and preserves a relatively pristine lower mantle is therefore not tenable. However, a range of geophysical evidence indicates that compositionally distinct, hence convectively isolated, mantle domains may exist in the bottom 1000 kilometers of the mantle. Survival of these domains, which are perhaps related to local iron enrichment and silicate-to-oxide transformations, implies that mantle convection is more complex than envisaged by conventional end-member flow models.

Despite recent progress in tomographic imaging (1–3) and numerical flow modeling (4), important aspects of convective stirring of Earth’s mantle have remained enigmatic. Analyses of trace elements and noble gas isotopes suggest that distinct mantle reservoirs have survived for a large part of Earth’s 4.5 × 10^9 year history (5–7). In some convection models a reservoir boundary at 660-km depth separates a depleted upper from an enriched lower mantle (5, 6), but computer simulations demonstrate that such layering of flow requires radial variations in intrinsic density and cannot be maintained on geological time scales by viscosity stratification and effects of isochemical phase transitions alone (8). There is, however, no compelling evidence for a change in bulk chemistry at 660-km depth (9, 10), and seismological evidence implies that slabs of subducted oceanic lithosphere can sink deep into the lower mantle (2, 3). Resolving the dilemma requires new views on isotopic and major element relationships (11) or on lower-mantle processes. We explore the latter and propose that segregation of mantle domains does occur but at much greater depth than the base of the upper-mantle transition zone.

A recent tomographic study (2) revealed two depth intervals, from ~400- to 1000-km depth and from ~1700- to 2300-km depth, where inferred mantle structure seems too complex for undisturbed whole-mantle convection. This complexity is borne out by the radial correlation of mantle structure (12) (Fig. 1). The shallow interval of reduced radial correlation (Fig. 1) comprises the upper-mantle transition region in a broad sense (13, 14). Analog (15) and numerical (16) flow modeling demonstrate that the complex flow trajectories revealed by seismic imaging (17) may merely represent local and transient layering (18) of a convective system that is otherwise characterized by deep, but not necessarily mantle-wide, circulation (2, 3). The persistence of structural complexity to near-1000-km depth (Fig. 1) corroborates previous inferences (19–21) and is perhaps related to the stability of Al-rich phases and silicate ilmenite into the topmost lower mantle (22). The base of this interval coincides with mid-mantle discontinuities as inferred from converted seismic waves and mantle impedance profiles (23), but there is no consensus on either the global significance or the cause of these discontinuities, and it is not obvious if and how they are related to the complexity of mantle structure.

Seismologically observed complexity in the lower mantle probably reaches a maximum in its bottom 300 km, the D′ region (24, 25). However, several lines of geophysical evidence suggest that structural and chemical heterogeneity is not confined to this region but extends up to at least 1000 km above the core-mantle boundary (CMB). High-resolution tomography has revealed a change in the spatial pattern of heterogeneity in the middle of the lower mantle (2, 3). To investigate the deep mantle in more detail, we enhanced data coverage by including high-quality differential travel-time residuals from core-refracted (PKP) waves (26). The new images confirm that at 1700 ± 200 km depth, the linear features of higher-than-average P wavespeed that are so prominent in the mid-mantle (Fig. 2A) begin to disintegrate (Fig. 2B), with only some fragments of them connecting to the D′ region (2). Beneath the depth interval from 1700 to 2300 km, which is devoid of spatially coherent structures (Fig. 3), a pattern emerges of long-wavelength structures that becomes more pronounced toward the base of the mantle (Fig. 2C). This shift to long-wavelength structures is consistent with results of other studies (27).

The changes in the pattern or spectrum of heterogeneity do not, by themselves, require spatial variations in bulk chemistry. Sheeplike downwellings may break up when they sink to larger depth in a spherical medium (28), and the bifurcated downwellings spread out when they approach Earth’s dense core, which explains the shift to long-wavelength structure. Radial variations in viscosity (4, 29) can force this to happen at shallower depths. Furthermore, seismic images are a “snapshot” of time-dependent processes (3), and the inferred change in the planform of mantle structure may be a transient feature related to past plate reorganization. Some of the slabs that are detected in the mid-mantle may not yet have reached 2000-km depth, whereas others may no longer be detectable (2). The global extent of the disruption (Fig. 3) can be explained if, with the exception of the most thermally inert downwellings—for example, beneath eastern Asia and central America (2, 3)—slabs lose their excess negative (thermal) buoyancy and assimilate in the middle of the lower mantle (9).

Fig. 1. Depth variation of two global diagnostics deduced from recent tomographic models (2, 32). The radial correlation (solid line) (12) is a measure of the continuity of structure in radial direction. For our purposes the absolute values are less important than the relative variations, which reveal a reduced correlation of structure between 400 and 1000 km and between 1600 and 2500 km depth. The dashed line depicts the radial variation of the root mean square (RMS) amplitude of relative variations of bulk sound and shear wavespeed, ζ = dlnv/dlnλ, after (32); see (30) for explanation of symbols. The open (gray) symbols mark the depth range where scattering of short-period PKP waves may occur (35).

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: hilst@mit.edu
Other observations are, however, more difficult to explain within the framework of purely thermal convection. First, joint interpretations of $P$ and $S$ data (30) have revealed radial changes in $\xi = d \ln \phi / d \ln \alpha$ (the ratio between relative variations in bulk sound, $\phi = \sqrt{\rho / \mu}$, and shear wavespeed, $\beta = \sqrt{\mu / \rho}$) (31, 32). The depths where the root mean square (rms) amplitude of $\xi$ changes most coincide with steep gradients in the radial correlation profile (Fig. 1). This observation suggests that the mantle comprises three dynamic regimes, and not two as suggested by conventional models of stratified flow. In the bottom third $\xi$ varies laterally, with positive and negative values contributing to the enhanced rms values (32). A recent study (33) revealed that the wavespeed ratio $\nu = d \ln \phi / d \ln \alpha$, with $\alpha$ the compressional wavespeed, reaches anomalous values in certain regions only—for example, in the central Pacific mantle deeper than 1800 km. The lateral, and hence isobaric, changes in sign and magnitude of $\xi$ or $\nu$ are hard to explain solely by thermal variations at high ambient pressures and suggest variations in composition (34). There is a caveat: Interference with core-refracted SKS waves can degrade the quality of the direct $S$ data for epicentral distances between, roughly, 80° and 90°, that is, $S$ wave turning points deeper than about 2000 km. Further research is thus required to substantiate the inferred changes in $\xi$ and $\nu$ in the deep mantle.

Second, analysis of high-frequency (1 Hz) precursors to the core phase $PKP$ has provided compelling evidence for the presence of small (approximately tens of kilometers) scatterers of seismic energy in the bottom 1000 km or so of the mantle (35). Both the sharpness of their interface and their size are inconsistent with a thermal origin. Short-period $P$-$S$ phase conversions at 1600-km depth have also been explained by compositional heterogeneity (36). The geographical distribution of either type of scatterer is, however, not yet established, so that it is not currently known if and how these observations relate to large-scale patterns of mantle flow.

Third, the seismologically slow regions that contribute to the long-wavelength pattern in the deep mantle—for example, beneath the Pacific and Africa—often extend far above the CMB (2, 3, 27, 33, 37). Image resolution is an issue, but the dimension and morphology of such “mega-plumes” differ from those expected for thermal events and, even with a range of boundary conditions and viscosity profiles, they are not easily reproduced by computer simulations of thermal convection (4, 38). Convective instabilities limit the temperature contrast across a purely thermal boundary layer, whose collapse can thus not produce sufficient temperature contrasts to explain the low wavespeeds (39, 40).

Finally, slight superadiabatic temperatures in the bottom 1000 km of the mantle have been proposed on the basis of (i) small radial variations of the inhomogeneity parameter (41, 42), (ii) an increase with depth of viscosity and, perhaps, density that is smaller than expected from adiabatic compression (43, 44), and (iii) numerical flow modeling with maximum viscosity near 2000-km depth (45). Fundamental uncertainties render these inferences inconclusive (43, 46), but if true, superadiabatic temperatures suggest ineffective removal of heat from the deep mantle. The implied sluggish convection can temporarily be achieved with pressure- and temperature-dependent rheology (45) but is difficult to maintain on geological time scales without gravitational stabilization if radiogenic heat...
production is significant. High-viscosity blobs in the deep mantle have been proposed to explain the survival of distinct, enriched-isotope reservoirs (47), but the enhanced heat production would rapidly destroy the viscosity contrast unless it is somehow controlled by compositional variations.

Each of these lines of evidence needs to be substantiated, but together they suggest changes in bulk composition in the bottom third of the mantle. An exciting possibility is that the anomalous domains contain relatively undpleted material that has remained isolated from the rest of the mantle since the early stages of Earth's evolution. The presence of enriched material implies enhanced heat production, and, analogously to the stabilization of D' material proposed by others (48, 49), the positive buoyancy must be compensated by increased intrinsic density if the reservoirs are to remain dynamically stable (7). The existence of dense but seismologically slow (and presumably hot) regions in the deep mantle has been confirmed by studies of Earth's free oscillations (49). The net density increase may be small, which allows substantial topography of the thermochemical boundary layer (7, 38). Some downwellings may depress the interface and sink to near the CMB—for example, beneath eastern Asia and central America. Elsewhere, the hot, dense material may rise up to large distances above the CMB, which can explain the “mega-plumes” beneath Africa and the western Pacific (38) and perhaps also the scatterers in the mid-mantle (36). It may not be easy to detect the interface by conventional seismic imaging on the basis of waveform stacks.

Understanding the above inferences in terms of high-pressure major element mineralogy and phase chemistry remains a challenge. Of potential importance is the breakdown of mantle silicates into dense oxides (14), along with, and perhaps as a mechanism for, iron enrichment. The thermodynamic stability of (Mg$_{1-x}$Fe$_x$)$_2$SiO$_4$ perovskite is fiercely debated (15, 14), along with, and perhaps also the scatterers in the mid-mantle.

References and Notes

10. The location and mobility of trace elements and rare gases may be unrelated to major-element chemistry, in which case explanation of the geochemical record does not require the identified reservoirs (5). This may reconcile the isotope record with geochemical constraints, but would not resolve the issue of heat production discussed in (7).
11. The radial two-point correlation function, calculated for our current P-wave model (26) after P. Puster and T. H. Jordan [J. Geophys. Res. 102, 7625 (1997)], is a measure of the correlation between wave speed anomalies at different depths; sharp reductions indicate that the heterogeneity planform changes from one depth range to another. The distance over which this correlation is determined is controlled by the thickness of the layers that parametrize our problem (about 150 km). At large depths the correlation function is not well constrained and depends on the seismological model used. We calculated radial correlation functions for different inversions and found that the relative reduction in the upper and lower mantle is robust, but that the absolute values are sensitive to the type and amount of regularization used. We stress that this is a global diagnostic; radial continuity is possible on a local scale even in regions of reduced global correlation. Likewise, in case an interface has significant topography, the minimum in radial correlation will not be sharp but will be spread out over a certain depth range.
12. Layer “C” as defined by K. E. Bullen [Cambridge Univ. Press, Cambridge, 1947].
19. In addition to the P, Pf, and pwp travel-time residuals from the catalog by E. R. Engdahl, R. D. van der Hilst, and R. P. Buland [Bull. Seismol. Soc. Am. 88, 722 (1998)], we use PKP$_{C60}$-PKP$_{C60}$ and PKP$_{AB}$-PKP$_{AB}$ differential times from two sources: the data set of T. McMenemy [thesis, University of Washington, Seattle (1995)]; who measured differential times from waveform cross-correlation, and the catalog by Eng- dahl et al. For the latter, we use the cross correlation that at least two PKP phases be identified by station operator, which appears to be a significant quality criterion. In the tomographic inversion, the core phases are given a larger weight than the P and P wave travel-time residual times. Differential travel times were previously used in studies of the core-mantle boundary region because they are primarily sensitive to structure in the lowermost mantle. However, the signal pertinent to the shallow mantle cannot be neglected and is exploited in our tomographic inver- sion for three-dimensional structure. The solutions were obtained after about 200 iterations, but most of the conver- gence was achieved within the first 20 iterations.
21. D. Becricogi, G. Schubert, G. Glatzmaner, Science 244,
Dipping Low-Velocity Layer in the Mid–Lower Mantle: Evidence for Geochemical Heterogeneity

Satoshi Kaneshima1* and George Helffrich2

Data from western United States short-period seismic networks reveal a conversion from an S to a P wave within a low seismic velocity layer (greater than or equal to the 4 percent velocity difference compared to the surrounding mantle) in the mid–lower mantle (1400 to 1600 kilometers deep) east of the Mariana and Izù-Bonin subduction zones. The low-velocity layer (about 8 kilometers thick) dips 30° to 40° southward and is at least 500 kilometers by 300 kilometers. Its steep dip, large velocity contrast, and sharpness imply a chemical signature that the temperature may seem “superadiabatic” although radial variations in density may be similar to those expected from adiabatic compression.

1Department of Earth and Planetary Sciences, Faculty of Science, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152, Japan. 2Department of Geology, University of Bristol, Queens Road, Bristol, UK.

*To whom correspondence should be addressed.

Planetary differentiation and convection create heterogeneities in the mantle, which are ultimately related to the cooling of Earth over the age of the solar system. Seismically, these heterogeneities express themselves in velocity heterogeneity due to variations in temperature, bulk composition, and phase changes in the mantle. In the upper mantle, all three mechanisms act, accounting for the greater velocity variations there (+5%) in comparison to the lower mantle (+0.5%) (I). In the lower mantle, it is not clear which mechanisms act, because the lower mantle seems comparatively homogeneous except for the...