Supplementary Information

A. Predicted seismic structures for non-deforming oceanic upper mantle

For an oceanic upper mantle that is not deforming, we assume that the grain size is constant. We consider different grain sizes (1 mm – 10 cm) and calculate their respective seismic structures (blue lines, Figure S1). As expected, Q values (Figure S1a) within the upper mantle are larger for larger grain size, resulting in faster seismic velocities (Figure S1b) than for smaller grain size. Notably, a LVZ can be produced (Figure S1b) but not the low Q zone (Figure S1a).

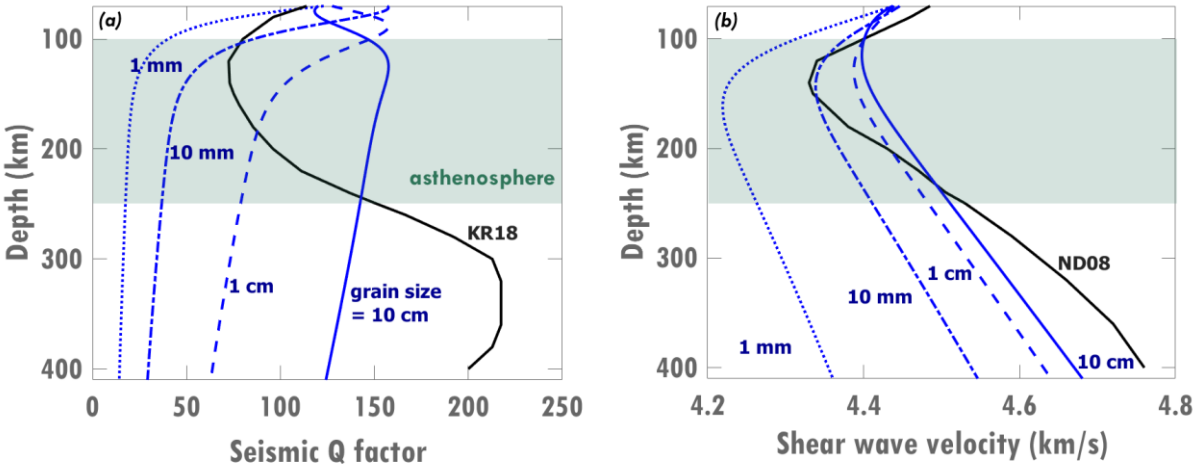

Figure S1. Seismic structures for oceanic upper mantle that is not deforming. (a) The seismic Q structures are calculated using Faul and Jackson's (2010) formulation for 100 s period, where Q is sensitive to a chosen grain size (values given), which is assumed constant in the absence of deformation. The global KR18 model is from Karaoglu & Romanowicz (2018). (b) The associated forward shear wave velocities are estimated using Karato's (1993) formulation, again for constant chosen grain size. The global ND08 model is from Nettles & Dziewonski (2008).

B. Analytical solution for 1-D rheology-dependent mantle flow in N layers

To implement composite rheology in the upper mantle, we must combine both Newtonian Poiseuille flow (*PFn1*) and plug flow (*PFn3*) models. For an assigned Newtonian rheology for the mantle transition zone, we only use the *PFn1* model. We

apply Equations (3) and (6.2) and the boundary conditions shown in Figure S2 and summarized below:

$$v_{x,1}(z_0) = U_p \tag{S1}$$

$$v_{x,N}(z_N) = 0 \tag{S2}$$

$$\text{at } z_i: \quad \tau_i(bottom\ boundary\ of\ ith\ layer) = \tau_{i+1}(top\ boundary\ of\ ith+1\ layer) \tag{S3}$$

$$\text{at } z_i: \quad v_{x,i}(bottom\ boundary\ of\ ith\ layer) = v_{x,i+1}(top\ boundary\ of\ ith+1\ layer) \tag{S4}$$

This yields a set of equations:

$$\frac{\partial p}{\partial x} z_i \ + \ C_i = \frac{\partial p}{\partial x} z_i \ + \ C_{i+1} \rightarrow C_i = C_{i+1} \tag{S5}$$

$$A_{PFn1,i}\left[\frac{1}{2}\frac{\partial p}{\partial x}{z_i}^2 + \ C_i z_i\right] + \ A_{PFn3,i}\begin{bmatrix}\frac{1}{4}\left(\frac{\partial p}{\partial x}\right)^3 {z_i}^4 + C_i\left(\frac{\partial p}{\partial x}\right)^2 {z_i}^3 \\ +\frac{3}{2}{C_i}^2\frac{\partial p}{\partial x}{z_i}^2 + {C_i}^3 z_i\end{bmatrix} + k_i =$$

$$A_{PFn1,i+1}\left[\frac{1}{2}\frac{\partial p}{\partial x}{z_i}^2 + \ C_{i+1} z_i\right] + \ A_{PFn3,i+1}\begin{bmatrix}\frac{1}{4}\left(\frac{\partial p}{\partial x}\right)^3 {z_i}^4 + C_{i+1}\left(\frac{\partial p}{\partial x}\right)^2 {z_i}^3 \\ +\frac{3}{2}{C_{i+1}}^2\frac{\partial p}{\partial x}{z_i}^2 + {C_{i+1}}^3 z_i\end{bmatrix} + k_{i+1} \tag{S6}$$

We linearize the equations by grouping the terms in Equations (S5) and (S6) such that the terms with first degree C's and k's (constants of integration) are on the left side of the equation and the remaining terms are on the right side. Then, we can express the boundary conditions for the layered system as $M\mathbf{R}=\mathbf{A}$ where vector **R** contains the constants of integration (C's and k's) and vector **A** has the higher degree C's:

28
$$\begin{pmatrix} a_{PFn1,1}+a_{PFn3,1} & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ a_{PFn1,1}+a_{PFn3,1} & 1 & -(a_{PFn1,2}+a_{PFn3,2}) & -1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{PFn1,2}+a_{PFn3,2} & 1 & -(a_{PFn1,3}+a_{PFn3,3}) & -1 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & a_{PFn1,N-1}+a_{PFn3,N-1} & 1 & -(a_{PFn1,N}+a_{PFn3,N}) & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & a_{PFn1,N}+a_{PFn3,N} & 1 \end{pmatrix} \begin{pmatrix} C_1 \\ k_1 \\ C_2 \\ k_2 \\ C_3 \\ k_3 \\ \vdots \\ C_{N-1} \\ k_{N-1} \\ C_N \\ k_N \end{pmatrix}$$

29
$$= \begin{pmatrix} -(b_{PFn1,1}+b_{PFn3,1})+U_p \\ -(b_{PFn1,1}+b_{PFn3,1})+(b_{PFn1,2}+b_{PFn3,2}) \\ 0 \\ -(b_{Pn1F,2}+b_{PFn3,2})+(b_{PFn1,3}+b_{PFn3,3}) \\ 0 \\ \vdots \\ -(b_{PFn1,N-1}+b_{PFn3,N-1})+(b_{PFn1,N}+b_{PFn3,N}) \\ 0 \\ -(b_{PFn1,N}+b_{PFn3,N}) \end{pmatrix} \quad \text{(S7)}$$

30 where,

31
$$a_{PFn1,i} = A_{PFn1,i}\, z_i \quad \text{(S8)}$$

32
$$a_{PFn3,i} = A_{PFn3,i} \left(\frac{\partial p}{\partial x}\right)^2 {z_i}^3 \quad \text{(S9)}$$

33
$$b_{PFn1,i} = \frac{1}{2} A_{PFn1,i} \frac{\partial p}{\partial x} {z_i}^2 \quad \text{(S10)}$$

34

35
$$b_{PFn3,i} = A_{PFn3,i} \begin{bmatrix} \frac{1}{4}\left(\frac{\partial p}{\partial x}\right)^3 {z_i}^4 + \\ +\frac{3}{2} {C_i}^2 \frac{\partial p}{\partial x} {z_i}^2 + {C_i}^3 z_i \end{bmatrix} \quad \text{(S11)}$$

36 The terms $A_{PFn1,i}$ and $A_{PFn3,i}$ for the upper mantle are defined in Equations (5.3) and
37 (5.4), respectively. For the mantle transition zone (MTZ), $A_{PFn1,i} = 2/\eta_{MTZ}$ and
38 $A_{PFn3,i} = 0$. The higher degree C_i terms in Equation (S11) or in vector **A** are considered
39 constant and we initially guess them to be the same for every layer i (as in Equation
40 S5) to determine the C_i and k_i in vector **R** by inversion (**R**=M^{-1}**A**). Then, the C_i in vector

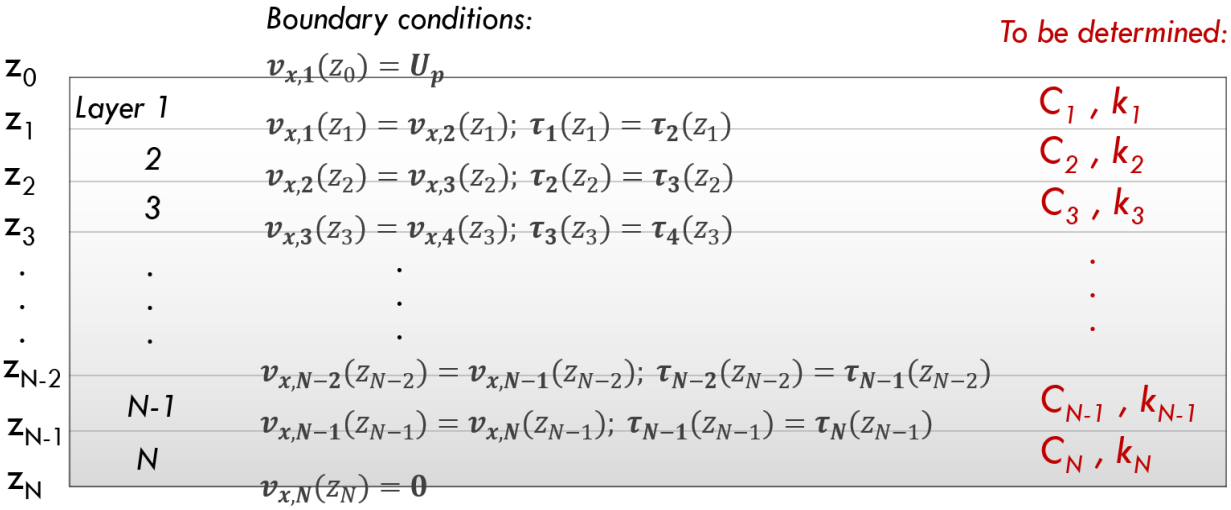

Figure S2. The boundary conditions for 1D model with N layers in terms of stress τ_i and flow horizontal velocity $v_{x,i}$ where i is the layer number**.** The C_i and k_i integration constants in Equation (3) for stress and Equation (6.2) for flow velocity are determined, which allows us to solve stresses and flow velocities within the model.

A is updated in every iteration with the calculated C_i in vector **R** until their absolute difference is $\leq 10^{-6}$. Then, stresses (Equation 3) and velocities (Equation 6.2) with depth can be calculated using the derived C_i and k_i from vector **R**.

C. Iteration scheme to compute steady state grain size and stress evolution

The $A_{PFn1,i}$ and $A_{PFn3,i}$ parameters used in calculating stress τ and horizontal velocity v_x (Section A) are dependent on grain-size, which evolves with time (Equation 7). Both τ and v_x reach a steady state, which is determined by employing the scheme below:

$$\boldsymbol{t_0}: \quad assume\ constant\ d_0 \rightarrow calculate\ v_{x,t_0}\ and\ \tau_{t_0}$$

$$\boldsymbol{t_1 = t_0 + \Delta t}: \quad calculate\ \Delta d_1\ and\ d_1 \rightarrow calculate\ v_{x,t_1}\ and\ \tau_{t_1}$$

$$\boldsymbol{t_2 = t_1 + \Delta t}: \quad calculate\ \Delta d_2\ and\ d_2 \rightarrow calculate\ v_{x,t_2}\ and\ \tau_{t_2}$$

$$\vdots$$

$$\boldsymbol{t_k = t_{k-1} + \Delta t}: \quad calculate\ \Delta d_k\ and\ d_k \rightarrow calculate\ v_{x,t_k}\ and\ \tau_{t_k}$$

where

$$t_k = k\Delta t = grain\ size\ evolution\ time$$

$$\Delta t = change\ in\ time\ or\ time\ interval$$

$$d_k = new\ grain\ size\ structure\ after\ t_k\ (Equation\ (S12))$$

$$\Delta d_k = change\ in\ grain\ size\ after\ t_k\ (Equation\ (S13))$$

$$v_{x,t_k} = horizontal\ velocity\ profile\ of\ the\ flow\ at\ t_k$$

$$\tau_{t_k} = stress\ profile\ induced\ by\ the\ flow\ at\ t_k$$

After time t_k (which is $t_{k-1} + \Delta t$), we determine the new grain size structure d_k:

$$d_k = d_{k-1} + \Delta d_k \quad \text{(S12)}$$

where Δd_k is estimated by multiplying the grain-size change rate $\dot{d}_{k-1}$ at t_{k-1} by Δt:

$$\Delta d_k = \Delta t\left[\dot{d}_{k-1}\right] = \Delta t\left[\dot{d}_{gg,k-1} - \dot{d}_{dr,k-1}\right] \quad \text{(S13)}$$

Here $\dot{d}_{k-1}$ is estimated using Equation (7) where $\dot{d}_{gg,k-1}$ is the grain growth term and $\dot{d}_{dr,k-1}$ is the dynamic recrystallization term (the first and second terms on the right hand side of Equation (7), respectively). The constants used in the calculation of $\dot{d}$ (as described by Equations (7) and (S13)) are summarized in Table S1. Using the new d_k, we recalculate the horizontal velocity, shear stress, and viscosity structures. We iterate this process until a steady state grain size is reached at steady-state time t_{ss} (typically << 1 Myr, criterion is discussed in Section D).

Table S1. Grain size evolution parameters are taken from Behn et al. (2009) since they are calibrated to laboratory data, and the flow law parameters are from Hirth and Kohlstedt (2003).

Symbol	Description	Value		Units
$\dot{d}_{gg}$	Grain growth rate			m/s
$\dot{d}_{dr}$	Dynamic recrystallization rate			m/s
τ	Shear stress			Pa
σ	Differential stress (2τ)			Pa
p_g	Grain growth exponent	3		
G_o(dry)	Grain growth constant for 50 ppm H/Si	1.5×10^{-5}		$m^{p_g} s^{-1}$
G_o(wet)	Grain growth constant for 1000 ppm H/Si	4.5×10^{-4}		$m^{p_g} s^{-1}$
E_g	Activation energy for grain growth	350		kJ/mol
V_g	Activation volume for grain growth	4×10^{-6}		m^3/mol
λ	Reciprocal of strain required for new grain size	10		
χ	Fraction of work done by dislocation to ground boundary area	0.1		
c	Geometrical constant	3		
γ	Average specific grain boundary energy	1		J/m^2
$\dot{\varepsilon}_{disl}$	Dislocation creep strain rate	*For olivine*		s^{-1}
		DRY	WET	
A_{disl}	Dislocation creep prefactor	1.1×10^5	30	$MPa^{-3.5} s^{-1}$
n_{disl}	Dislocation creep stress exponent	3.5	3.5	
p_{disl}	Dislocation creep grain size exponent	0	0	
r_{disl}	Dislocation creep water exponent	0	1.2	
α_{disl}	Constant for melt factor	45	45	

E_{disl}	Dislocation creep activation energy	530	480	kJ/mol
V_{disl}	Dislocation creep activation volume	15×10^{-6}	11×10^{-6}	m^3/mol
$\dot{\varepsilon}_{diff}$	Diffusion creep strain rate	*For olivine*		s^{-1}
		DRY	WET	
A_{diff}	Diffusion creep prefactor	1.5×10^9	1×10^6	$MPa^{-3.5}s^{-1}$
n_{diff}	Diffusion creep stress exponent	1	1	
p_{diff}	Diffusion creep grain size exponent	3	3	
r_{diff}	Diffusion creep water exponent	0	1	
α_{diff}	Constant for melt factor	30	30	
E_{diff}	Diffusion creep activation energy	375	335	kJ/mol
V_{diff}	Diffusion creep activation volume	6×10^{-6}	4×10^{-6}	m^3/mol

D. Convergence criterion for grain size evolution

To determine the steady-state time t_{ss}, we employ a convergence criterion of:

$$\frac{\Delta d_{norm}}{d_{norm}} \leq \vartheta \tag{S14}$$

where ϑ is the limit for convergence, Δd_{norm} is the depth-averaged norm of grain size change, and d_{norm} is the depth-averaged norm of grain size. As a convergence criterion, we use Equation (S15) for a chosen timestep Δt. At time t_k, the parameters in Equation (S14) are calculated as:

$$\vartheta = 5 \times 10^{-4} \left(\frac{\Delta t}{1000\ yr}\right) \tag{S15}$$

$$\Delta d_{norm} = \frac{\sqrt{\sum_{i=1}^{N+1} (d_k - d_{k-1})^2 \Delta z}}{\sum_{i=1}^{N+1} \Delta z} \tag{S16}$$

$$d_{norm} = \frac{\sqrt{\sum_{i=1}^{N+1} {d_k}^2 \Delta z}}{\sum_{i=1}^{N+1} \Delta z} \quad \text{(S17)}$$

When the criterion in Equation (S14) is met, $t_k \sim t_{ss}$.

E. Additional analyses at steady-state

E.1 Effect of initial grain-size

We compare two steady-state calculations that are the same except for different initial olivine grain-sizes (1 mm or 10 mm), which produces flow via Couette flow (*CF*) above a 10^{21} Pa·s mantle transition zone as shown in Figure S3a. Such a flow configuration

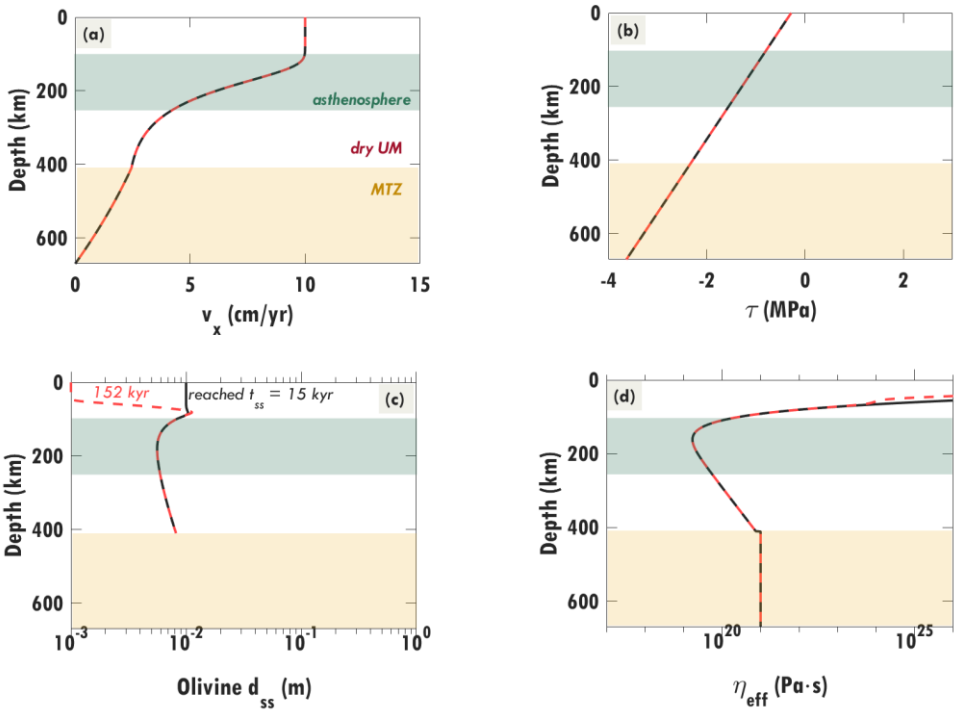

Figure S3. Effect of initial grain size (1 mm and 10 mm for the red dashed and black lines, respectively) on the steady-state (a) upper mantle flow, (b) induced shear stresses, (c) grain-size structure, and (d) effective viscosity. We assume dry conditions, and that the 60 Myr old oceanic upper mantle and mantle transition zone are deformed by plate motion of 10 cm/yr and a pressure gradient of -5kPa/km. Until the flow reaches steady state, grain size changes according to the grain size evolution AE07 model (Austin and Evans, 2007). Flow additionally alters the grain-size structure, which in turn changes the flow and rheology with time. The flow eventually reaches steady state after a time t_{ss}=152 kyr for an initial grain size of 1 mm and t_{ss}=15 kyr for an initial grain size of 10 mm (see Supplementary Information D). The timesteps Δt used for 1-mm and 10-mm flow models are 10 yr and 100 yr, respectively.

dominates because of large viscosities in the upper mantle and mantle transition zone (Figure S3d). Initially smaller (1 mm) and larger grain-sizes (10 mm) evolve to the same steady-state grain-size structure (except for the stiff undeforming lithosphere, Figure S3c) and the same steady-state upper mantle flow (Figure S3a) with the same stress profile (Figure S3b). Clearly, the choice of initial grain-size does not affect the system's eventual steady-state but it does affect the time it takes the grain size to reach steady state. A larger initial grain size (i.e., 10 mm) stabilizes faster (15 kyr) compared to a smaller grain size (1 mm, 152 kyr), because large grain-sizes subdivide rapidly (Equation 7).

E.2 Effect of grain-size evolution model

Hall & Parmentier (2003) provide another grain-size evolution model (HP03 model):

$$\textit{HP03 model}: \quad \dot{d} = {p_g}^{-1} d^{1-p_g} G_o \exp\left(-\frac{E_g + PV_g}{RT}\right) - \lambda \dot{\varepsilon}_{disl} d \tag{S18}$$

The grain-size structure stabilizes faster when using AE07 model ($t_{ss} = 478\ kyr$) compared to using the HP03 model (598 kyr) because of AE07's strong dependence on grain-size (Figure S4c). Although the HP03 model (red dashed line, Figure S4c) predicts larger grain sizes than does the AE07 model (black line), their flow configurations (Poiseuille flow or *PF*, Figure S4a), stress profiles (Figure S4b), and viscosities (Figure S4d) are nearly the same.

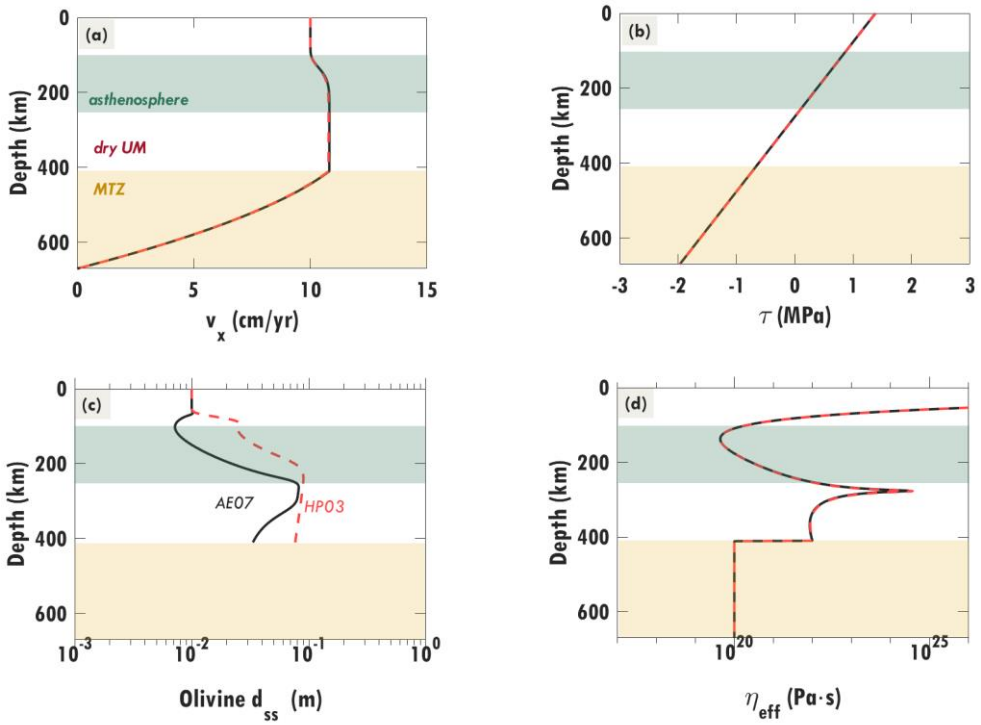

Figure S4. Effect of grain size evolution model (HP03 (Hall & Parmentier, 2003) for the red dashed line and AE07 (Austin and Evans, 2007) for the black line) on the steady-state (a) upper mantle flow, (b) induced shear stresses, (c) grain-size structure (initially 10 mm grain size), and (d) viscosity. The flow conditions considered are the same as in Figure S3. The timesteps Δt used for HP03 and AE07 grain-size evolution models are 1000 yr and 100 yr, respectively.

E.3 Effect of contrasting rheologies between upper mantle and MTZ

In Section 4, the comparable effective viscosities of upper mantle and mantle transition zone result in a *CF*-dominated dry upper mantle and a *PF*-dominated wet upper mantle. However, with contrasting rheologies, the dry upper mantle can accommodate a *PF* configuration (case ii, Figure S5a) and a *CF* configuration in wet upper mantle (case i, Figure S5d). The less viscous mantle transition zone below dry upper mantle (Figure S5c) allows a pressure-driven flow within the upper mantle (case ii, Figure S5a). In contrast, the more viscous mantle transition zone below the wet upper mantle (Figure S5f) can shut down pressure-driven flow unless the pressure gradient is large enough (cases ii and iii, Figure S5d) to drive *PF* that exceeds the plate-driven flow.

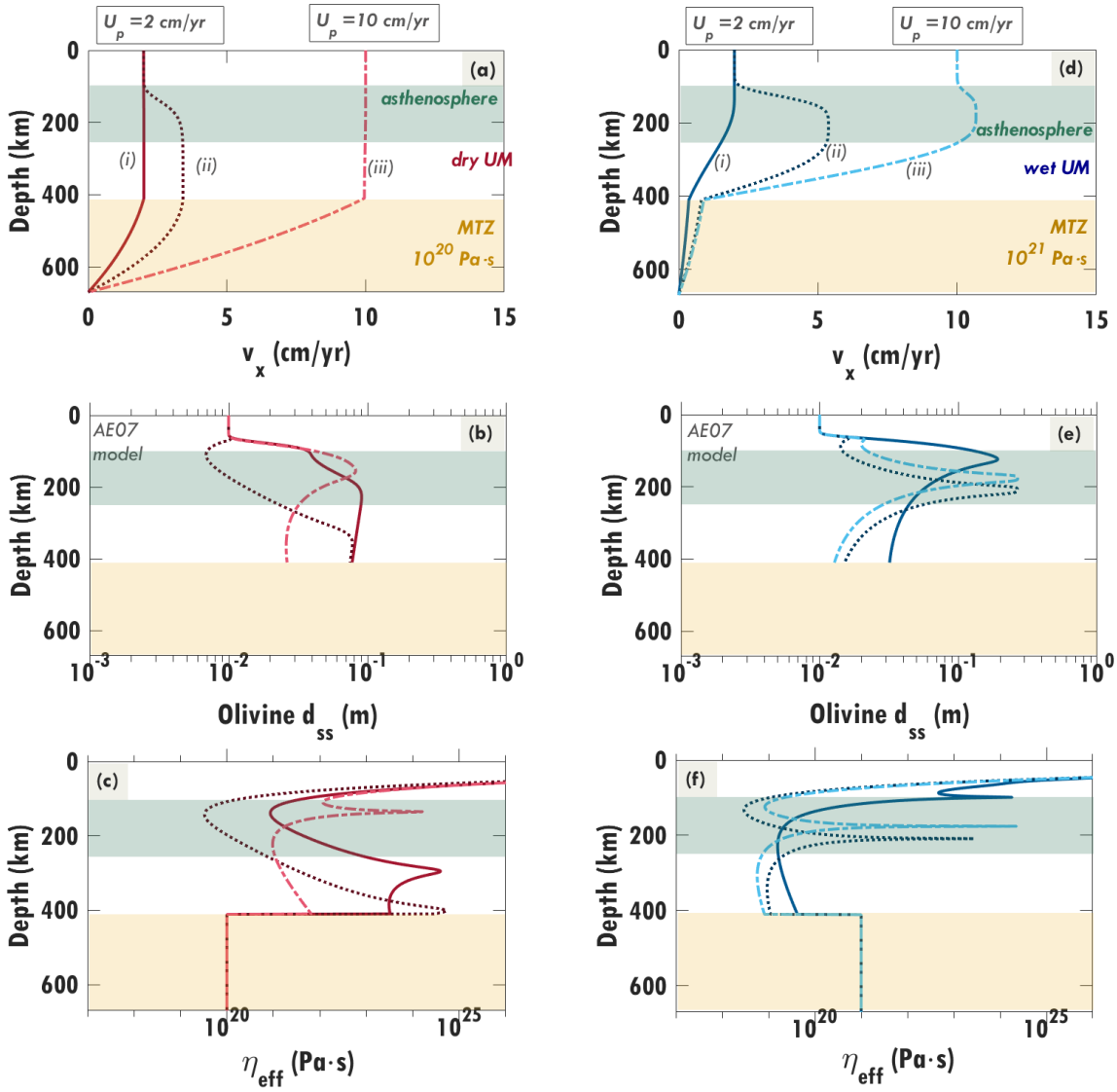

Figure S5. Effect of contrasting rheologies between upper mantle (UM) and MTZ on the steady state (a, d) flow, (b, e) grain size structure, and (c, f) viscosity for (a-c) dry and (d-f) wet (d-f) conditions. Different combinations of plate velocity and horizontal pressure gradient (labeled as i, ii and iii) are considered, and are the same as in Figure 4. For dry upper mantle, the assigned mantle transition zone (MTZ) viscosity is 10^{20} Pa·s, and 10^{21} Pa·s MTZ for wet upper mantle. The less viscous MTZ viscosity allows for a PF configuration to dominate in the more viscous upper mantle. Otherwise, CF may dominate unless the pressure gradient is large enough to drive PF that exceeds plate velocity. The initial grain-size for each calculation is 10 mm. A timestep Δt of 1000 yr is used for case (i), and 100 yr for cases (ii) and (iii).

E.4 Effect of water and small melt fraction on upper mantle flow and rheology

Varying the water distribution (blue line in Figure S6 vs. case (i) wet upper mantle in Figure 4) clearly alters the rheology and grain-size structures, and thus also the flow pattern. In contrast, adding a small amount of melt (< 0.1%, Figure 6a.2) in the

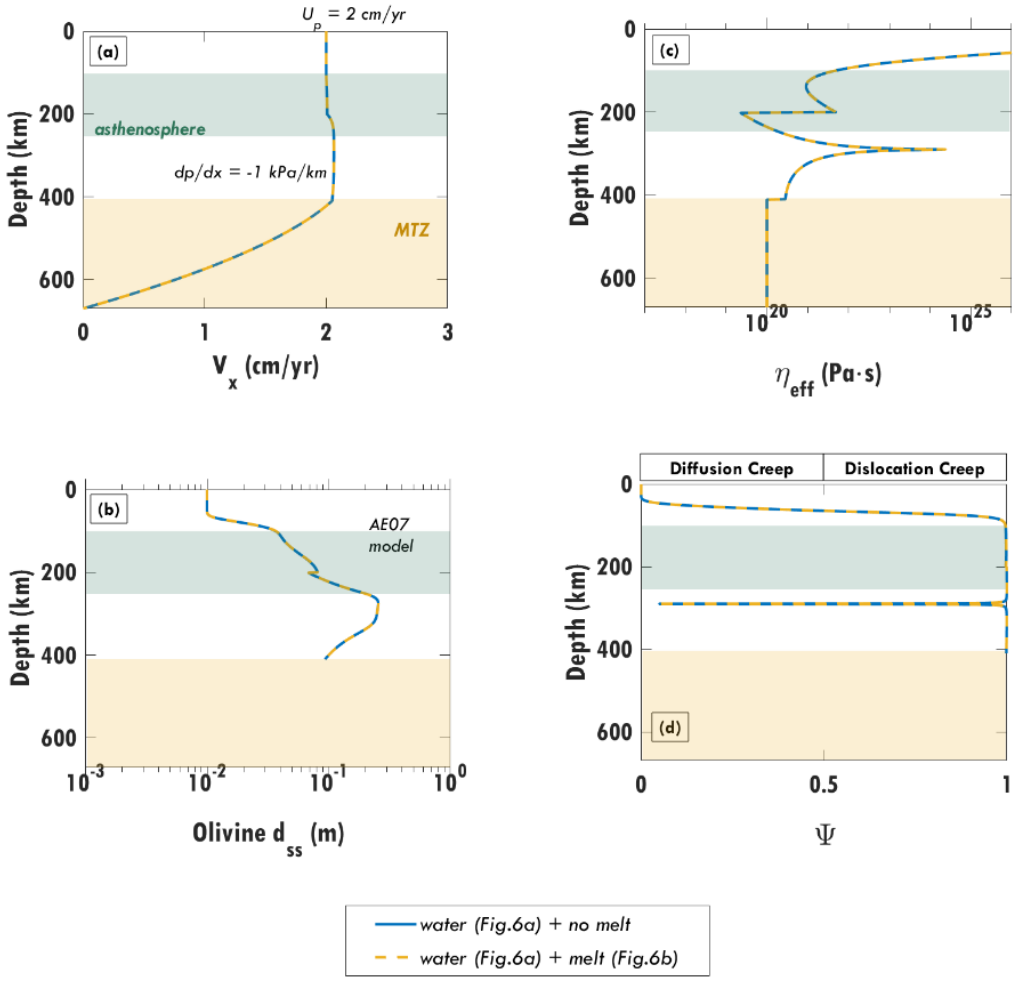

Figure S6. Effect of excluding melt (blue line) vs. including melt (dashed orange line) on steady-state upper mantle (a) flow, (b) olivine grain sizes, (c) viscosity, and (d) dominant deformation mechanism. A 60 Myr old oceanic upper mantle is considered with 2 cm/yr plate velocity and a -1 kPa/km pressure gradient. The grain-size evolution model used is AE07 (Austin and Evans, 2007) with a timestep Δt of 1000 yr.

asthenosphere reduces the viscosity by at least a factor of ~0.8, which yields negligible changes to the flow pattern and grain-size structure (compare blue and dashed orange lines, Figure S6).

E.5 Different flow configurations in dry and wet upper mantle

The flow configurations and rheological structures for the different plate velocity and pressure gradient combinations, and MTZ viscosities, shown in Figures S7 and S8 for dry and wet conditions, respectively, as discussed in Section 6.3.

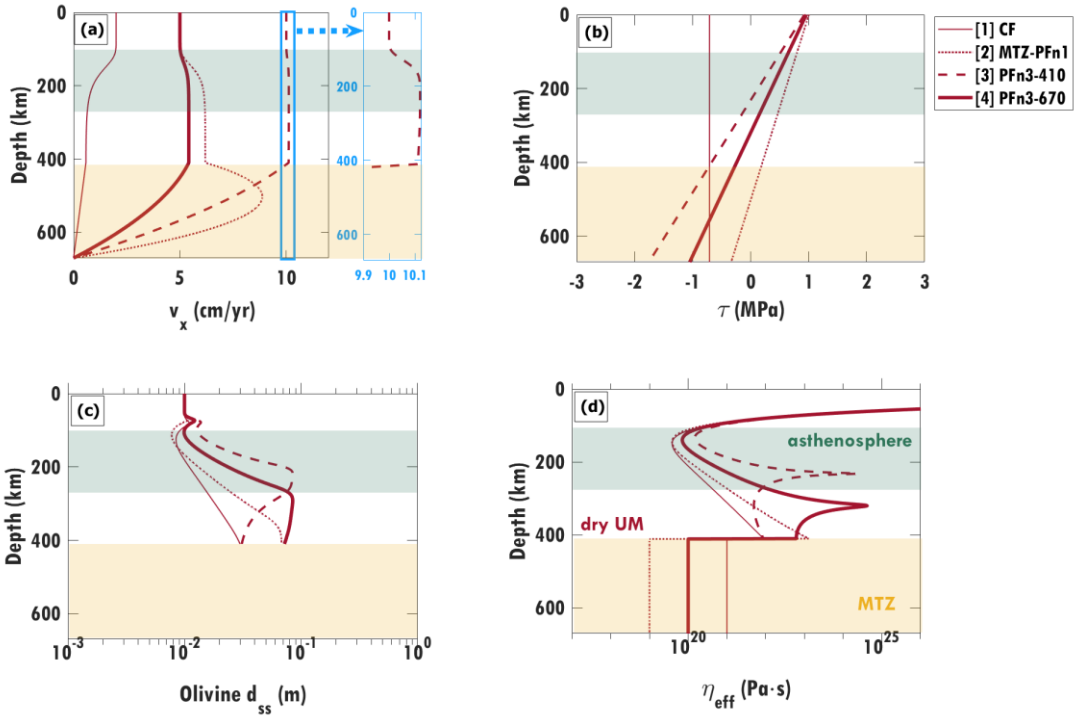

Figure S7. The (a) four flow configurations for dry (50 ppm H/Si) upper mantle and their associated (b) stresses, (c) grain-sizes, and (d) viscosity structures. The corresponding plate speed and pressure gradient combinations used to produce such flows are shown in Figure 7b.1 (in red rectangles).

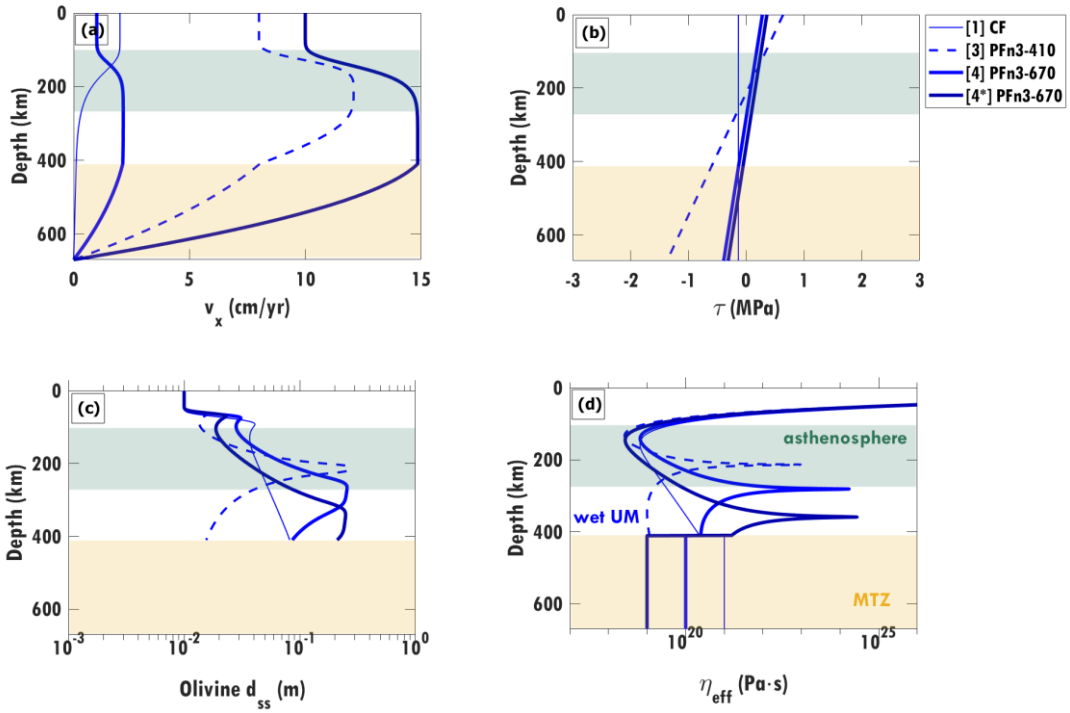

Figure S8. The (a) four flow configurations plausible for wet (1000 ppm H/Si) upper mantle and their associated (b) stresses, (c) grain-sizes, and (d) viscosity structures. The corresponding plate speed and pressure gradient combinations used to produce such flows are shown in Figure 7b.2 (in blue rectangles).