Origins of cratonic mantle discontinuities: A view from petrology, geochemistry and thermodynamic models

Sonja Aulbach, Malcolm Massuyeau, Fabrice Gaillard

To cite this version:
Origins of cratonic mantle discontinuities: A view from petrology, geochemistry and thermodynamic models

Sonja Aulbach\textsuperscript{1,}\textsuperscript{*}, Malcolm Massuyeau\textsuperscript{2}, Fabrice Gaillard\textsuperscript{2}

\textsuperscript{1} Inst. Geosciences, Goethe-Universität, Frankfurt am Main, Germany
\textsuperscript{2} Institut des Sciences de la Terre d'Orléans, Orléans, France

*Corresponding author. \textit{E-mail address:} s.aulbach@em.uni-frankfurt.de

Word count: 10,586

Figures: 10
Tables: 1

Invited Review Article Lithos

Resubmitted: 2 November 2016
Table of Contents

1. Introduction

2. Craton definition, craton types and nature of the cratonic mantle sample
   2.1 Intact vs disturbed lithospheric roots and associated magmatism
   2.2 Accidental mantle samples: Xenoliths and xenocrysts

3. Modes of lithosphere interrogation
   3.1 Petrology, fabrics, geochemistry and thermodynamic modelling
   3.2 Geophysical observations

4. Discussion
   4.1 Lithosphere-asthenosphere boundary (LAB)
      4.1.1 Strong vs. weak or absent cratonic seismic LAB signals: The role of reworking and rejuvenation
      4.1.2 Nature of melt near cratonic LABs: Evidence from mantle samples
      4.1.3 Conditions for melt production below intact craton roots
      4.1.4 Melt retention and resupply at the LAB
   4.2 Mid-lithospheric discontinuities (MLDs): Geophysical expression of metasomatism, Palaeo-LABs or future LABs?
      4.2.1 MLDs related to fabric
         Radial anisotropy contrasts at the pervasive-to-focused melt flow transition
         Frozen rock fabrics: MLDs as palaeo-LABs
      4.2.2 Present-day ponded melts or fluids
      4.2.3 Deposition of metasomatic minerals during past metasomatism
         Metasomatic minerals accumulated during subduction
         Mantle metasomatism
         Observed and inferred accumulations of metasomatic minerals in cratonic mantle
         Phase equilibria and thermodynamics: Can metasomatism explain MLDs?

5. Summary and conclusions
1. Introduction

Geophysically detectible mid-lithospheric discontinuities (MLD) and lithosphere-asthenosphere boundaries (LAB) in Earth’s ancient continental masses (cratons; Fig. 1) have received much attention over recent years. A plethora of papers has been published on the topic of both, with a focus on geophysical observables (e.g. Selway et al., 2015; Rader et al., 2015; Hopper and Fischer, 2015), and on petrologic-geochemical data (e.g. O’Reilly and Griffin, 2010). The overarching goal of the various lines of investigating cratonic lithosphere, including current efforts to find and explain cratonic mantle discontinuities, is to elucidate the origin and evolution of continents, which bears on early terrestrial dynamics, such as the nature of crust and mantle differentiation or the viability of plate tectonics (Gerya, 2014).

Xenoliths, xenocrysts and magmas forming in or passing through cratonic mantle provide temporally and spatially punctuated information on the composition and fabric of cratonic roots at the time of magmatism, and on the nature and timing of events that formed them. Geophysical observations can provide regional or craton-scale insights, but require validation from actual sample properties, and must rely on background information afforded by mantle samples for correct interpretation. Although these different modes of lithosphere interrogation do not operate on the same spatial or temporal scales, they offer the opportunity to investigate lithospheric discontinuities through the joint investigation of all available physical, dynamic, chemical and petrologic constraints.

Perhaps not surprisingly, there is little agreement on the origins of cratonic mantle discontinuities, and the increasing integration of multiple geophysical observables with the mineralogy and composition of mantle samples has led to the recognition that universally applicable mechanisms are unlikely to explain all observations (e.g. Fischer et al., 2010;
Selway et al., 2015). It is clear, however, that the peculiarities of cratonic lithosphere composition and construction - cool geothermal gradients, compositional and rheological stratification, multiple metasomatic overprints in the course of their long existence - offer myriad possibilities for the generation, modification, translation and even obliteration of such discontinuities in ancient lithospheric keels.

In this review, we first provide a brief introduction into the cratonic lithosphere, its uniqueness compared to younger lithosphere, and the nature and shortcomings of the cratonic mantle sample (section 2). We then present the different lines of investigation that can be combined to make inferences on the nature and origin of cratonic lithosphere discontinuities (section 3): Petrographic-geochemical observations (microstructures, mineral modes, elemental and isotopic composition), geothermobarometry (to locate mantle samples in the lithosphere column), thermodynamic modelling and geophysical observations (seismic velocity and anisotropy, resistivity, reflectivity). We subsequently discuss possible origins of the two major mantle lithospheric discontinuities - LABs and MLDs (section 4). We conclude with a summary of the most important and, in some cases, novel insights gained from this exercise. In particular, we would like to highlight that the combined perspective from petrology, geochemistry and thermodynamic modelling indicates that most MLDs may be best explained by accumulations of seismically slow minerals deposited during past magmatic-metasomatic activity. Finally, we suggest that the sparseness of seismic LAB observation beneath stable cratons relate to the difficulty to form and retain melt beneath these typically thick (200-250 km) lithospheres.
2. Craton definition, craton types and nature of the cratonic mantle sample

Cratons are often defined as the cores of tectonomagmatically quiescent continents where the last major tectono-thermal event occurred in the Archaean or Palaeoproterozoic and which were amalgamated by Neoarchaean time (≥ 2.5 Ga); larger continental regions that arose by later collisions, as evidenced by Proterozoic mobile belts, are also sometimes referred to as cratons (Fig. 1). These mobile belts shielded the cratonic cores from later tectonomagmatic activity, ensuring their long-term stability (Lenardic et al., 2000), although in some cases later reactivation with minor to severe loss of cratonic lithosphere has been documented (e.g. O’Reilly and Griffin, 2010). Compared to younger continental lithosphere, cratonic mantle is characterised by exceptionally high degrees of melt extraction, which resulted in strong Fe- and H$_2$O-depletion (hence buoyancy and viscosity) that allowed the formation of exceptionally stable and thick lithosphere (sometimes referred to as “roots”; O’Reilly et al., 2001; Griffin et al., 2009; for a compilation of cratonic crustal and cratonisation ages as well as thermal thicknesses, see Table 1 in Artemieva, 2006) and that can be linked to warmer Archaean mantle potential temperatures T$_P$ (Davies, 2009; Herzberg et al., 2010). These properties play a pivotal role in the longevity of cratons (Jordan, 1988; Abbott et al., 1997; Lenardic et al., 2003; Afonso and Schutt, 2012; Wang et al., 2014).

In the context of this review, how cratons formed and evolved is important in that it affects the composition and fabric of the lithospheric mantle that may today be visible as discontinuities. Processes that have been put forward for the formation of diamondiferous – therefore thick – cratonic roots range from subduction-accretion or collisional settings to plume tectonics (Lee, 2006; Pearson and Wittig, 2008, 2014; Lee et al., 2011; Aulbach, 2012) (Fig. 2). The depth of the LAB may vary as a function of T$_P$ and composition (Fig. 3A), and
may increase due to thickening during the aforementioned accretionary and plume subcretion processes, or decrease due to a variety of destructive processes described in the next section.

2.1 Intact vs disturbed lithospheric roots and associated magmatism

Intact cratonic roots extend to ~200-250 km and are characterised by high seismic velocities, low electrical conductivities and low surface heat flow (Artemieva and Mooney, 2001; Priestley and McKenzie, 2006; Eaton et al., 2009; Gardès et al., 2014), consistent with a cold and thick lithospheric lid. Because they are thicker than the depths where dry peridotite crosses its solidus at ambient mantle potential temperatures, only small-volume, volatile-rich melts, such as kimberlites, lamproites and alkali basalts, are typically able to form beneath such lids in hot and/or oxidising, carbonated-(hydrous) mantle (Rohrbach and Schmidt, 2011; Dasgupta, 2013) (Fig. 3A). Occasionally, anomalously hot and/or fertile mantle can lead to the emplacement of higher-volume melts, despite the presence of a thick lithosphere (e.g. the 2.05 Ga Bushveld event; Wilson, 2012). Beneath some cratonic areas, investigation of mantle xenoliths and geophysical observations has revealed much thinner or strongly modified lithospheric roots. A number of processes can conspire to erode or destroy cratonic lithosphere. Heating, for example by impinging plumes, addition of water, for example by dehydration of seawater-altered slabs beneath cratons or in the transition zone, injection of wet melts at the LAB, or addition of Fe-rich melts during metasomatism (Fig. 4), lower the viscosity and density contrast between cratonic root and convecting mantle, imperilling its stability (Foley, 2008; Li et al. 2008; Peslier et al. 2010; Windley et al., 2010; Selway and Karato 2014; Wang et al., 2014, 2015; Pearson et al., 2014; Zheng et al., 2015). The consequences are (1) foundering of gravitationally unstable, deep, little depleted mantle residues of polybaric melting (Arndt et al., 2009), (2) foundering due to an inverted temperature gradient developing near the base of the lithosphere (Michaut et al., 2009), (3) the establishment of melt-percolation fronts atop anomalously hot mantle generating MLDs
Wölbern et al., 2012) and (4) densification and rheological weakening due to melt infiltration (Wang et al., 2015; Zheng et al., 2015) (Fig. 4). Such processes may also act at modern LABs, leading to thinning and, in cases, even complete past (or future) loss of deep cratonic roots.

Thinning cratons are characterised by a record of progressively more voluminous magmas, enabling increasing melting intervals to be achieved in upwelling convective mantle, starting with volatile-rich melts, such as carbonatites and kimberlites, and ending with silicate melts, such as melilitites and alkali basalts (e.g. North Atlantic craton: Tappe et al., 2007; western edge of the Tanzanian craton: Rosenthal et al., 2009) (Fig. 5A). Interaction of the lowermost 40 km of the Kaapvaal root with, likely plume-derived, asthenospheric melts linked to group I kimberlite magmatism led to severe refertilisation (O'Reilly and Griffin, 2010). Rift-related thinning, accompanied by melt infiltration and severe Fe enrichment, is evident for the margins of the Tanzanian craton (Ritsema et al., 1998; Lee and Rudnick, 1999) and suggested for Sao Francisco craton (Pinto et al., 2010). The Moroccan Atlas, part of the Saharan metacraton (Abdelsalam et al., 2002), may once have been underlain by a cratonic root prior to rifting and orogenesis, enabled by craton-edge driven convection currents (Kaislaniemi and van Hunen, 2014). The most striking example of cratonic lithosphere destruction may be the North China craton, where a diamondiferous root that was still present in the Palaeozoic, as sampled by kimberlites, was replaced by hotter and thinner mantle, as sampled by Cenozoic basaltic magmatism (Tang et al., 2013). Other examples include the Mojavia and Wyoming cratons, where interaction with the Farallon plate is evidenced by the presence of unusually fertile, thin and wet remaining root (Lee et al. 2001; Lee 2005; Li et al. 2008). Thus, emplacement of higher-volume, alkaline basaltic magmas betrays the presence of disrupted
cratonic roots, which allows longer melting columns to develop in decompressing convecting mantle.

2.2 Accidental mantle samples: Xenoliths and xenocrysts

Asthenosphere-derived magmas penetrating intact or disturbed cratonic roots often entrain fragments (xenoliths) of the wall-rock they traverse, which may disintegrate to single grains (xenocrysts, including diamond) that still bear the geochemical fingerprint of their source rocks (Griffin et al., 1999; Grütter et al., 2004). Combined with geothermobarometry, these samples provide invaluable information on the depth-dependent composition and fabric of the cratonic mantle keel, from which they were entrained, although constraints on structural relationships afforded by outcrops are precluded. Xenoliths and xenocrysts inevitably provide snapshots of the lithospheric mantle state at the time of magmatism that cannot be unquestioningly compared to present-day geophysical observation: Heating associated with metasomatism may dissipate, melts may freeze and fabric may be destroyed due to recrystallisation. In addition, the volume of cratonic mantle sampled by magmas is unlikely to be representative of the whole craton because magmas preferentially utilise multiply reactivated lithospheric weak zones as pathways, and the conduit walls are accordingly highly overprinted (Poudjom-Djomani et al. 2005; Afonso et al., 2008; Giuliani et al., 2015a).

Finally, the size of the mantle sample is limited by dike width and loading capacity of the host magma (Sparks et al., 2006), which restricts the accuracy, with which bulk chemical and modal compositions can be determined. This accidental and, in particular, unrepresentative nature of the cratonic mantle sample also implies that direct matching between sample-derived 1D lithosphere profiles and regional geophysical surveys is fraught with uncertainty.

3. Modes of lithosphere interrogation

3.1 Petrology, fabric, geochemistry and thermodynamic modelling
Over its long existence and subsequent to the partial melting event leading to its depletion, the cratonic mantle has been multiply refertilised and metasomatised (e.g. O’Reilly and Griffin, 2013; Pearson and Wittig, 2014). It is known that fertile peridotite changes its modal composition with increasing degree of melt extraction from lherzolite (olivine, orthopyroxene, clinopyroxene plus feldspar or Al-spinel or garnet, depending on pressure) to harzburgite (≤ 5 vol.% clinopyroxene) and dunite (≤ 10 vol.% pyroxenes) (Streckeisen, 1976), while refertilisation by melts reverses this process (Griffin et al., 2009) (although note that some dunites and harzburgites result from melt-peridotite interaction; Kelemen et al., 1997).

In particular the presence of atypical phases in primitive mantle, such as amphibole, phlogopite, carbonates, Ti minerals and others (“modal metasomatism”; Harte, 1983) attests to the former presence of fluids or melts that significantly affected the mantle lithosphere and possibly changed its thermophysical properties, such as density and seismic velocity (Afonso and Schutt, 2012; Selway et al., 2015; Rader et al., 2015). In addition, peridotite fabric generates anisotropy and therefore affects the seismic response of the mantle (Savage, 1999).

The most common process producing both shape- and crystal-preferred orientations associated with fabric is ductile deformation by dislocation creep (Tommasi et al., 2004). For example, foliation and lineation defined by elongate crystals and the sometimes associated crystallographic-preferred orientation (CPO), generated during mineral dissolution-reprecipitation in response to reactive melt migration, are detectible microscopically or using modern analytical techniques (e.g. Electron Back-Scattered Diffraction; Tommasi et al., 2004; Bascou et al., 2011; Baptiste et al., 2012). However, a straightforward relationship between fabric and CPO is not always evident (e.g. Tommasi et al., 2016). While crystal growth during static recrystallisation tends to destroy the shape-preferred orientation, the CPO can be preserved (Baptiste et al., 2012; Tommasi and Vauchez, 2015). In contrast, CPO is weakened by dynamic recrystallisation or melt-fluid-rock reactions leading to neo-crystallisation of
olivine (Tommasi et al., 2004; 2008), of pyroxenes (Le Roux et al., 2008; Soustelle and Tommasi, 2010) or of both (Hidas et al., 2016).

Partial melt extraction and modal metasomatism are usually accompanied by changes in both major- and trace-element composition, while cryptic metasomatism only affects the trace-element budget (Dawson, 1984). Metasomatic agents range from carbonated melts imparting strongly sinusoidal REE patterns in garnet (commonly in clinopyroxene-free harzburgites and in inclusions in diamond), to silico-carbonatite melts resulting in mildly sinusoidal REE patterns in garnet. Interaction with silicate melts leads to “normal” LREE-depleted patterns, with smoothly increasing abundances towards the HREE, which are indistinguishable from those expected in garnet from fertile (undepleted) lherzolites (Fig. 6A). The trace element composition of the lithosphere does not directly influence its thermophysical properties, with the possible exception of the effect of heat-producing elements on the geothermal gradient and of H⁺, which significantly affects rheology and other mantle properties (Hirth and Kohlstedt, 1996). The distinctive trace-element signatures imparted by the passage of different melts and fluids can alert us to changes of dynamic or tectonic regimes, which may in turn have affected lithospheric fabric and mineralogy (e.g. Simon et al., 2007; Agashev et al., 2013; Aulbach et al., 2013), providing important context for the interpretation of geophysical observations. Even when precise ages prove elusive, radiogenic isotope compositions constrain time-integrated enrichment and depletion in the cratonic lithosphere that can be correlated with tectonothermal events evident in the crustal record (e.g. Carlson, 1995; Pearson et al., 2005).

The chemical composition of mantle xenoliths, predominantly peridotite and subordinately eclogites and pyroxenites (Schulze, 1989), can be combined with experimentally calibrated
geothermobarometer combinations to locate them in the mantle column (Smith, 1999; Grütter, 2009; Nimis and Grütter, 2010). Geothermobarometry allows one-dimensional mantle stratigraphies to be constructed for each xenolith locality, which may be combined with other localities to map the lithosphere laterally (Griffin et al., 2004). Some lithologies implicated in the generation of MLDs (garnet- or orthopyroxene-free pyroxenites, phlogopitites; Wölbern et al., 2012), but also spinel peridotites and eclogites, lack mineral assemblages required for pressure calculation, although appropriate thermometers may exist. Their position in the lithosphere column is then inferred by assuming that they equilibrated to a known pressure-temperature array, typically obtained from garnet-bearing peridotites and pyroxenites from the same kimberlite, and by projecting calculated temperatures onto this geotherm (e.g. Kopylova et al., 2016). Uncertainties on temperatures and pressures obtained from thermobarometry of ~50º C and ~0.4 GPa, respectively, have been cited (Smith, 1999; Nimis and Grütter, 2010). Thus, uncertainties on the corresponding depths lie within the narrow depth interval over which a chemical, thermal or microstructural change has to occur to be geophysically detectible as a discontinuity (30-40 km; Selway et al., 2015).

Pressure-temperature arrays defined by appropriate samples also provide constraints on the thermal state of the mantle at the time of xenolith or xenocryst entrainment. Thus, stable cratons with thick keels typically have lower geothermal gradients than disrupted cratons with thinner keels, and older kimberlites sample a warmer cratonic lithosphere than younger ones (Artemieva, 2009; Grütter, 2009). Kinked geotherms (Lesotho, Kaapvaal craton: Boyd and Nixon, 1973), stepped geotherms (Slave craton: Griffin et al., 1999) or distributed temperatures over a short pressure interval are often associated with xenoliths showing strong fabric; they document the advection of heat by penetration of hot, asthenosphere-derived melts, probably during a magmatic event precursory to kimberlite magmatism (O’Reilly et al.,
2001; Grütter, 2009; O’Reilly and Griffin, 2010) (Fig. 5B). In contrast, differences in the initial thermal state, buoyancy and viscosity resulting from the formation of cratons in hot plumes vs. cool slab environments have dissipated by today (Eaton and Perry, 2013).

While thermodynamic modelling has long been employed to investigate the composition and quantity of melts produced as a function of temperature and volatile (H$_2$O) content in chemical equilibration with the convective mantle (Ghiorso and Sack, 1995; Ghiorso et al., 2002), the presence of carbonates in peridotite and their effects on the mantle solidus has until recently only been parameterised in empirical models (Dasgupta et al., 2013). A new thermodynamic model (Massuyeau et al., 2015) improves on this, based on the silica activity in the melt in presence of both CO$_2$ and H$_2$O. This approach is here for the first time applied to investigation of the stability of melts at the LAB (section 4.1) and of phlogopite at MLD depths (section 4.2).

3.2 Geophysical observations

A variety of geophysical tools are available to study the physical properties of the lithosphere, including discontinuities, and these have been reviewed in detail in several recent papers (Selway, 2014; Fischer, 2015; Hopper and Fischer, 2015; Selway et al., 2015; Rader et al., 2015). Discontinuities in seismic velocity and anisotropy in the crust and upper mantle are imaged using receiver functions (RF), based on different arrival times for primary and converted P- and S-waves (termed PRF and SRF, respectively) and back-azimuthal differences in amplitudes (Kind et al., 2012; Wirth and Long, 2014; Selway et al., 2015). These high-frequency body waves are sensitive to small structures, with a depth resolution of 10-15 km for velocity contrasts at mantle depth. In contrast, body wave tomography has a lower depth resolution, but the best lateral resolution (Leveque and Masson, 1999; Selway et al., 2015). Because the depth interval of the transition between seismically slow and fast
regions is poorly resolved, trade-offs between the depth range and the sharpness of the transition imply that seismic velocity reductions can be modelled both as a discontinuity (large contrast) or a gradient (Lebedev et al., 2009).

A common tool to study anisotropy is shear-wave splitting analysis of various teleseismic phases, typically originating as S-waves that are converted to P-waves in the core and reconverted to S-waves (splitting of core phases) (Savage, 1999). This type of analysis has good lateral but poor depth resolution. Surface waves can also be used to image anisotropy in wave speeds with better depth, but poorer lateral resolution, whereby the phase velocities of surface waves can vary in the direction of propagation (azimuthal anisotropy) and in vertically polarised Rayleigh and horizontally polarised Love waves (radial anisotropy) (Montagner, 1998; Becker et al., 2012). The juxtaposition of mantle layers with isotropic and anisotropic seismic wave speeds or with contrasting geometries of azimuthal or radial seismic anisotropies can produce conversions of S- and P-waves that appear as MLDs. Assuming radial anisotropy stratification, radial anisotropy could cause velocity reductions (negative phases) in SRF, whereas azimuthal anisotropy should cause both positive, negative or no phases, depending on the direction of wave propagation relative to the orientation of the anisotropy and the location of the seismic receiver (Selway et al., 2015).

Finally, magnetotelluric data provide a means to interrogate the conductivity in lithospheric mantle, where high resistivity (low conductivity) is expected for dry and depleted mantle and low resistivity may be caused by a melt, an interconnected film or layer of conductive minerals, or of water stored in nominally anhydrous minerals (NAMs) (Selway, 2014).

4. Discussion

4.1 Lithosphere-asthenosphere boundary (LAB)
Various definitions exist for the continental LAB (chemical, thermal, mechanical, seismic, electrical; Artemieva and Mooney, 2001; Eaton et al., 2009), implying that the thickness of cratonic roots can be measured by a variety of geochemical, petrologic and geophysical means. Compositionally, the LAB is typified by a transition from depleted to fertile (or refertilised) peridotite that has been mapped using xenoliths and xenocrysts (Gaul et al., 2000; Kopylova and Russell, 2000; Griffin et al., 2002, 2004). This approach may, however, underestimate the true base of the cratonic mantle root (O’Reilly and Griffin, 2010).

Thermally, ancient thick roots are characterised by low surface heat flow (Jaupart et al., 1998; Artemieva and Mooney, 2001). Physically, reduced shear (S)- and compressional (P)-wave speeds at ~170-250 km depth depict the transition from the cold and viscous lithosphere to the asthenosphere (e.g. Romanowicz, 2009; Eaton et al., 2009; Fischer et al., 2010; Ford et al., 2010). This may be accompanied by a reduction in electrical resistivity (Jones et al., 2009; Eaton et al., 2009) and an increase in seismic velocity anisotropy (Foster et al., 2014).

These changes in the physical and chemical properties at the base of the cratonic lithosphere do not always result in a geophysically detectible discontinuity. In particular the cratonic LAB is known to produce a much weaker seismic signature than the oceanic LAB, and is often characterised by a small and gradual velocity change with a weak and intermittent seismic signal (Eaton et al. 2009; Abt et al. 2010; Fischer et al., 2010; Hopper et al., 2014; Hopper and Fischer, 2015). Here, we discuss the origin of the variable signal from cratonic LABs with respect to the stability of small volumes of melt, which can satisfy both seismic and electrical anomalies, and in particular explore the conditions required for generating melt in the sublithosphere and maintaining it near LAB depth.

4.1.1 Strong vs. weak or absent cratonic seismic LAB signals: The role of reworking and rejuvenation
When present, seismically imaged discontinuities (localised in a <30 km depth interval) beneath cratons (e.g. Baltic shield, Canadian shield and parts of the North American craton; Rychert et al., 2005; Ford et al., 2010; Foster et al., 2014) cannot be produced by thermal effects alone and instead require contrasts in composition, fabric, water content or the presence of partial melt or volatiles (Rychert and Shearer, 2009; Fischer et al., 2010; Ford et al., 2010). For example, ponding of plume-derived hot material beneath the lithosphere, thermal insulation of the mantle beneath large continental masses or a contrast in anisotropy or grain size, enhanced by mantle flow at lithospheric edges, have been invoked (Lebedev and van der Hilst, 2008; Foster et al., 2014; Hopper and Fischer, 2015). Where discrete LABs are not detected at depths commensurate with the extent of seismically fast roots, the boundary is characterised by a velocity gradient and hence more accurately described as a transition zone that is spread out over a large depth interval (~50 km; Fischer et al., 2010; Ford et al., 2010; Yuan et al., 2011).

Oftentimes, highly disparate depths to the LAB are determined. For example, results for the Tanzanian lithosphere thickness vary from 150 to 200 km (Ritsema et al., 1998; Weeraratne et al., 2003; Fishwick, 2010). The shallower range of these estimates may represent the tops of live melt percolation fronts that may manifest themselves as MLDs (section 4.2). Topography (that is, a non-flat LAB with variable depths) due to thermochemical erosion (Fig. 4) will further prevent coherent seismic signals from the LAB as melt may accumulate in depressions in the LAB, or may be produced due to additional decompression in ascending mantle at these depressions. Accordingly, weak seismic velocity gradients near the LAB have been attributed to thermal and melt-related disruptions (Hopper et al., 2014; Hopper and Fischer, 2015). These disruptions eventually lead to lithosphere rejuvenation accompanied by Fe-enrichment, which may explain low resistivity in the deep lithosphere along with observed
positive geoid and Bouguer gravity anomalies, as applies to the São Francisco craton (Pinto et al., 2010).

The Kaapvaal craton is a particularly well-studied example. Plume impingement beneath the craton in the Cretaceous is thought to be responsible for kimberlite magmatism at that time (Le Roex et al., 2003; Becker and Le Roex, 2006). It may also have resulted in the deep intersection of the volatile-rich peridotite solidus and production of partial melts that have pervasively metasomatised and refertilised the deep mantle root, leading to “loss” of 40 km of the Kaapvaal root by asthenospherisation. However, unradiogenic Os retained in some of the deepest and hottest samples suggests that the root has not been completely replaced (Kobussen et al., 2008; Begg et al., 2009; O’Reilly and Griffin, 2010). The deep metasomatised root has warmed substantially (Bell et al., 2003) and become in parts almost indistinguishable from asthenospheric mantle. The thinner “LAB” corresponds to a depth where the amplitude of non-thermal seismic variations beneath the Kaapvaal craton is sharply reduced, which has been interpreted as the base of the chemical boundary layer (Artemieva, 2009). In contrast, teleseismic tomography and S receiver functions have revealed a lithosphere thickness of 250-300 km (Zhang and Tanimoto, 1993; Sodoudi et al., 2013).

In summary, it appears that intact and undisturbed cratonic roots will not produce a geophysically detectible LAB signal, whereas those undergoing melt-related reworking and rejuvenation do. Therefore, we next investigate the conditions that must be met in order to generate melt beneath the cratonic LAB and to keep it there long enough to produce a discontinuity, as well as the nature of the melt that is most likely to be produced.

4.1.2 Nature of melt near cratonic LABs: Evidence from mantle samples
Convecting mantle will only melt beneath cratonic LABs in exceptional circumstances and if so, small melt fractions usually prevail because the lithospheric lid impedes further decompression melting of upwelling mantle (Foley, 2008). Nevertheless, kimberlites and related magmas were undoubtedly repeatedly produced beneath cratonic lithospheres, at least from the Proterozoic onward (Becker and Le Roex, 2006; Tappe et al., 2014). Mantle samples provide evidence for the oxidising nature of the melts interacting with cratonic roots (Frost and McCammon, 2008; Foley, 2011), as they record more oxidising conditions at their depth of entrainment than would be expected in a fertile or depleted mantle (Fig. 3C). Indeed, a general correspondence between metasomatism and oxidation has been noted (Woodland et al., 1996; Creighton et al., 2010; Yaxley et al., 2012).

Carbonatites, which are only stable at temperatures far below the dry peridotite solidus (Fig. 3A) (Dasgupta and Hirschmann, 2010; Dasgupta, 2013), may form in a volatile-bearing peridotite due to decompression in sublithospheric convection cells established as a consequence of topography beneath cratons (e.g. King and Ritsema, 2000), in a chemically buoyant plume (e.g. Simmons et al., 2007; Dannberg and Sobolev, 2015), or due to thermal insulation beneath large continental plates (Andersen and King, 2014; Porritt et al., 2015). In contrast, kimberlite formation may additionally require excess Tp (thermal plume) enabling upwelling mantle to be above the carbonated peridotite solidus over a large enough pressure interval to dissolve SiO2 (Dasgupta, 2013), and/or enriched conditions (high CO2-H2O contents; section 4.1.3). Indeed, kimberlite magmatism has been linked to episodic impingement of mantle plumes beneath cratons (Heaman et al., 2004, 2015). In addition, the vast majority of mantle peridotites have trace-element patterns consistent with having equilibrated with carbonated silicate melts rather than carbonatites (e.g. Simon et al., 2007;
Aulbach et al., 2013). This may also be supported by \( f_{O_2} \) measurements (Stagno et al., 2013), indicating that enough melting has occurred to dilute incipient carbonatite melts with silicates.

**4.1.3 Conditions for melt production below intact craton roots**

While there is no present-day magmatic surface expression of kimberlite or other magmatism in intact cratons, as discussed in section 4.1.1, there is no *a priori* reason why melt should not currently be present beneath some cratonic LABs. Several conditions must be met to cause a seismic or electromagnetic discontinuity by the presence of melt below intact cratons: (1) The melt must be of a composition and volume sufficient to produce the discontinuity, (2) it must be interconnected and (3) it must not drain, or be continuously replenished (Karato, 2012). In general, in an upwelling mantle package, the fraction and nature of partial melts is determined by the composition-dependent melting interval and melt productivity. The depth of onset of partial melting is a function of temperature (hotter = deeper), composition (such as H\(_2\)O content where wetter = deeper), and \( f_{O_2} \) (metal saturation, onset of redox melting), while the depth where partial melting ends is the cratonic LAB (e.g Novella and Frost, 2014; Green, 2014). The effect on the mantle solidus of adding H\(_2\)O to the peridotite source is much stronger than that of CO\(_2\), which, being more incompatible, nevertheless dominates in the melt composition for small melt fractions (Hirschmann, 2010).

The nature of the melt produced beneath the LAB has a strong effect on the registered seismic and electrical signature it can produce. For example, carbonated melts are more conductive and less viscous than basaltic melts (Gaillard et al., 2008), therefore having good grain-wetting properties (Minarik and Watson, 1995; Hunter and McKenzie, 1989). Very small amounts (0.03-0.2%) are required to produce an electrical anomaly, such as that observed beneath the São Francisco craton (Pinto et al., 2010; Sifré et al., 2015). The melt fraction required to produce a 6 to 9% reduction in shear wave velocity, as is typically observed at
oceanic LABs, is more debated since according to conventional wisdom 3.5 to 4% are required for isotropically distributed melt, which decreases to 0.25 to 1.25% for melt desegregating into bands due to deformation (Kawakatsu et al., 2009). It must be noted, however, that recent experimental data indicate that less than 1% of melt would explain most shear wave velocity reductions at the LAB (Chantel et al., 2016). Such melt fractions are associated with the formation of silico-carbonatite or small-volume basaltic melts (Hirschmann, 2010; Dasgupta et al., 2013), but could also be reached for incipient melts if buoyancy-driven melt migration is taken into account (Grégoire et al., 2006). Since only up to 1.5 wt% of melt can be produced at 100 km depth in a typically enriched source with 300 ppm CO₂ and 300-500 ppm H₂O, or 1% of melt at 180 km depth in a source with 1600 ppm H₂O at ambient T_P (Dasgupta et al., 2013; Novella and Frost, 2014), higher water or alkali contents and/or excess T_P are required to produce significant melt volume beneath cratonic roots at 250 km depth.

We use a predictive approach by employing an improved version of the recent thermodynamic parameterisation considering the presence of volatiles, i.e. CO₂ and H₂O (Massuyeau et al., 2015). Our calculations show that, assuming sufficiently oxidising conditions to stabilise carbonates (Rohrbach and Schmidt, 2011; Stagno et al., 2013), addition of 200 ppm H₂O and 200 ppm CO₂ to dry peridotite, for a relatively depleted mantle, produces 0.05 % of incipient carbonate-rich melt at 250 km depth (Table 1). Thus, buoyancy-driven melt migration (Grégoire et al. 2006) to accumulate melts in some compaction layers is additionally required in order to describe observed seismic anomalies. At exceptional excess T_P, as potentially obtained in plume cores, and/or in a particularly enriched mantle like the OIB source (800 ppm H₂O - 800 ppm CO₂), >0.4-0.5 % of carbonated silicate melt can be produced at about 200-250km depth. The amount of excess H₂O required in the mantle source
depends on hydrogen solubility in the minerals, which may be minimum near cratonic LAB depth (Mierdel et al., 2007). An interconnected melt may form from smaller individual melt increments if these accumulate at the rheological boundary between lithosphere and asthenosphere (Hirschmann, 2010).

4.1.4 Melt retention and resupply at the LAB

The geodynamic and compositional requirements to generate sufficient melt at the LAB have likely been met many times throughout cratonic histories, but in order to produce a geophysically detectible anomaly today, it is also necessary to either retain or replenish this melt. Once formed, compaction and topography at the LAB may cause the melt to drain even at small melt fractions (McKenzie, 1989; Faul, 2001), and in particular the low viscosity of carbonatite melt militates against retention at the LAB (e.g. Hirschmann, 2010). Because the LAB represents a permeability barrier, melts are expected to migrate upward buoyantly and focus at depressions in the LAB if the slope towards that dip is sufficiently high, until they are extracted in veins (Gregg et al., 2012). In fact, craton edges and other zones with weaker cratonic signatures (slower velocities, lower elastic thickness and resistivity compared to intact roots), such as intra-cratic boundaries, have been identified as the locus of preferred kimberlite magmatism (Poudjom Djomani et al., 2005; Malkovets et al., 2007; Artemieva, 2009; Begg et al., 2009; Faure et al., 2011). In this context, it is worth highlighting that the sharp seismic signature for oceanic LAB may be related to the fact that it overlaps the melt freezing boundary (Sifré et al 2014), which effectively represents a thermodynamic permeability boundary, but that this model cannot be extended to much deeper cratonic LABs.

Drainage might be retarded under some circumstances. Melt mobility at cratonic LAB depths may be counteracted by the small ratio of density contrast between basaltic melts and mantle
at 6 to 7 GPa (Sakamaki et al., 2013). Volatile-bearing alkali basalts may even be neutrally buoyant at cratonic LAB depths (Crepisson et al., 2014). However, it must be emphasised that more data on the physical properties of CO₂- and H₂O-rich melts that are expected at cratonic LABs are required to fully explore their stability at the LAB and thus their potential to explain seismic and electrical anomalies when present. Finally, several processes may facilitate melt replenishment below cratonic LABs. Long-lived thermochemical mantle upwellings (e.g. Zhao et al., 2015) could continuously advect peridotite that is oxidising enough to enable redox melting at 250 km depth. Alternatively, low-density methane-rich fluids continually rising from the deep upper mantle and transition zone could be oxidised at the depth of metal saturation, providing a continuous supply of reduced carbon from which carbonated melt could be generated near the base of thick lithospheres (Rohrbach et al., 2007; Frost and MacCammon, 2008; Rohrbach and Schmidt, 2011) (Fig. 4).

4.2 Mid-lithospheric discontinuities (MLDs): Geophysical expression of metasomatism, Palaeo-LABs or future LABs?

While negative velocity gradients beneath intact cratons at ~170-250 km depth are generally interpreted as the LAB (section 4.1), seismic discontinuities, with a thickness of ~30-40 km and occurring at 60-160 km depth (predominantly between 80-120 km depth) beneath continental lithosphere (Thybo, 2006; Rychert and Shearer, 2009) have been interpreted as MLDs (Fischer et al., 2010; Yuan and Romanowicz, 2010; Ford et al., 2010; Foster et al., 2014; Wirth and Long, 2014; Selway et al., 2015; for a compilation of MLD locations, depths and associated velocity decreases see Rader et al., 2015) (Fig. 3D). In some cases, cratonic MLDs are continuous with neighbouring, younger tectonic areas (Western Canada, USA, Australia; Mercier et al., 2008; Ford et al., 2010; Foster et al., 2014), whereas in others, they are truncated, indicating trans-lithospheric boundaries that depict terrane sutures at depth (Courtier et al., 2010) (Fig. 2C).
Discontinuities within the lithospheric mantle typically have negative gradients, but both negative and positive phases are observed (Selway et al., 2015; Hopper and Fischer, 2015). Negative phases are generated by waves travelling upwards from a seismically slower to a seismically faster layer, whereby the amplitude of the phase correlates positively with the size of the velocity contrast and sharpness of the velocity gradient (Selway et al., 2015). MLDs are also detected when adjacent layers show seismic anisotropy contrasts with regard to strength and orientation (Wirth and Long, 2014) and in some instances they correlate with anomalous conductivity (e.g. Jones et al., 2001). An important point that has been emphasised in recent studies is that the results of seismic studies depend critically on the energy spectrum used to filter the data. Thus, low frequencies favour the appearance of single coherent large phases, whereas higher frequencies allow the detection of multiple conversions with lower amplitudes that appear as multiple MLDs (Wirth and Long, 2014). As a consequence, single larger MLDs may be an artefact of lower frequency filtering (Wirth and Long, 2014).

Selway et al. (2015) present a comprehensive review of the various proposed origins of MLDs with regard to physical and geological plausibility, based on the evaluation of calculated S-wave velocities and of forward modelled RF, using various compositions and mineral modal abundances. In keeping with their work, we focus here on mechanisms involving the presence (1) of contrasting fabrics, such as changes in orientation of anisotropy, (2) of melts or fluids in the present-day lithosphere and (3) of seismically slow minerals (hydrous minerals, pyroxenes) deposited from melts or fluids during past tectonothermal events.

4.2.1 MLDs related to fabric
Anisotropy contrasts at the pervasive-to-focused melt flow transition

Asthenosphere-derived partial melts are thought to penetrate into and react with the deep cratonic lithosphere diffusively and pervasively, and the effects of this melt percolation are evident in all cratonic mantle sections (Gaul et al., 2000; Griffin et al., 2002, 2004). Variations in the degree of interaction of multiple batches of melt with the mantle lithosphere, with which the melt is initially out of chemical and thermal equilibrium, may lead to variable depths of freezing due to cooling and reactions with pre-existing mantle. After each “conditioning” of the lithosphere, subsequent batches of melt may percolate upwards further before experiencing chemical or thermal death, until the melt finally focuses into veins and channels and is extracted to the surface (Gregg et al., 2012) (Fig. 7A). This melt focussing into veins and deformation associated with it may cause radial anisotropy (Fig. 7B). A discontinuity with a negative velocity gradient may then ensue if radial anisotropy is associated with an abrupt transition from a layer with strong to a layer with weak anisotropy with depth, depending on the sign of the radial anisotropy (Rychert and Shearer, 2009). This is observed in the Australian continent (Yoshizawa and Kennett, 2015). A scenario involving ~10% pyroxenite and wehrlite dykes (tectonic stockworks) has been proposed by Snyder (2008, 2013) and Snyder et al. (2014) to explain seismic anisotropy observed below the Canadian Shield, although calculations suggest that the contrast in density and seismic velocity is small (Tommasi and Vauchez, 2015).

In the Kaapvaal craton, annealing in coarse peridotites weakened the fabric related to prior deformation in response to melt percolation, but 12% of the samples, possibly from the shallower, cool and little metasomatized part of the mantle, retained foliations and lineations (Baptiste et al., 2012). Since the dispersion of crystal-preferred orientation (CPO) and associated weakening of seismic anisotropy may be assisted by melt-rock interaction
(Tommasi et al., 2004), the products of which may be oversampled in magma-entrained xenoliths (section 2.2), the presence of seismic anisotropy may be underestimated based on the xenolith record.

Frozen rock fabrics: MLDs as palaeo-LABs

The frequent association of MLDs with azimuthal anisotropy has been ascribed to frozen rock fabric and can produce a discontinuity if the anisotropy contrast between adjacent layers is sufficiently large (Wang et al., 2014; Beghein et al., 2014). Azimuthal anisotropy is unlikely to result in consistently negative phases independent of the back-azimuth, which is how the majority of MLDs appear in SRF studies (Table 1 in Rader et al., 2015), and it can therefore not globally cause MLDs having negative phases (Selway et al., 2015). Nevertheless, some MLDs have been associated with formation and thickening of cratonic nuclei (Abt et al., 2010; Fischer et al., 2010; Miller and Eaton, 2010; Yuan and Romanowicz, 2010; Snyder et al., 2013; Cooper and Miller, 2014; Wirth and Long, 2014). If they represent palaeo-LABs prior to root thickening, azimuthal anisotropy could result by the same mechanisms as proposed for modern LABs, i.e. CPO of minerals deforming at the interface of the protocratonic nucleus and the palaeo-convecting mantle, provided that a sharp change in orientation or intensity of the CPO is involved (Montagner, 1998; Lebedev et al., 2009). Interestingly, initial lithosphere thicknesses estimated by Lee and Chin (2014) for peridotite xenoliths (Siberian, Slave, Kaapvaal, Tanzanian cratons) range from 30-150 km depth, which overlaps the current depths of MLDs of ~60 to 160 km (Selway et al., 2015).

Dipping or tapered fabrics associated with some MLDs in cratonic lithosphere may relate to craton root thickening involving plate tectonics (Bostock, 1997; Cook et al., 1999; Chen et al., 2007; Snyder 2008; Miller and Eaton, 2010; Ashchepkov et al., 2014; Cooper and Miller,
2014; Wirth and Long, 2014) (Fig. 2C). Alternatively, upwelling mantle underplating a palaeo-LAB, with material flowing laterally towards thinner craton margins after impinging upon the lithospheric lid (Sleep, 2002; Shervais and Hanan, 2008) may play a role (Fig. 2A). Topography developed at the palaeo-LAB prior to root thickening (Fig. 4) helps explain the seismic complexity of cratonic lithospheric mantle and the occurrence of multiple, spatially distributed MLDs with both negative and positive velocity phases that appear only when short-period filtering is applied (Selway et al., 2015). A major argument in favour of MLDs being caused by processes dating back to cratonic lithosphere construction is that these discontinuities sometimes appear to be correlated with tectonic units and truncated at craton margins (Porritt et al., 2015). They show steep-sided seismic velocity offsets that therefore have persisted for billions of years (Poupinet al., 2003).

4.2.2 Present-day ponded melts or fluids

The global existence of a discontinuity at cratonic MLD depths has been ascribed to the presence of melt (Thybo, 2006). The solidi of various mantle source rocks, combined with the thermal state of cratonic lithosphere, determine at which temperature and depth melt can be present (Fig. 3A). In disrupted cratons currently overlying anomalously hot mantle, MLDs plausibly represent the top of a live melt percolation front. For example, reduced S-wave velocities at depths of 50-100 km beneath the Tanzanian craton and adjacent rift have been correlated to the presence of phlogopite-bearing rocks and pyroxenites in mantle xenolith suites (Wölbern et al., 2012). In contrast, the cool geothermal gradients of intact cratonic mantle preclude the presence of melt at the LAB that could percolate to MLD depths, except in cases of mantle containing abnormal amounts of H₂O and/or CO₂ or at exceptionally high T_p (Fig. 3A), as discussed in section 4.1.
The effects of melt percolation on the mantle immediately prior to kimberlite eruption are evident in chemically zoned minerals and unequilibrated microstructures in mantle xenoliths that could not survive in the mantle for extended periods of time (>10s of million years) (O’Reilly and Griffin, 2010; Pearson and Wittig, 2014). The trace-element and isotopic signatures in mantle xenoliths from various cratons identify these melts as silico-carbonatites similar to kimberlites (Simon et al., 2003; Agashev et al., 2013; Aulbach et al., 2013) (Figs. 3C, 6). Saline, CO$_2$-poor fluids (brines) have also been implicated in the late (around the time of kimberlite magmatism) formation of fibrous diamonds, for which a genetic connection to kimberlites has been established (e.g. Klein BenDavid et al. 2007, 2014; Weiss et al., 2015). Given the extended time before perturbed gradients relax back to the conductive geotherm (Michaut et al. 2007) and the young age of kimberlite magmatism in many cratons (<500 Ma, see Table 2 in Grütter, 2009), the effects of melt infiltration may still contribute to abnormal present-day heat flows (e.g. Kaapvaal craton; Bell et al., 2003).

Partial melting in the cratonic lithosphere may result from sources with considerably lower solidi than fertile peridotite, such as pyroxene- and/or hydrous mineral-rich domains, the generation of which then requires some prior tectonomagmatic activity, during which they were emplaced from melt produced either in or below the lithosphere (decompression, influx of volatiles, heating) (Fig. 5). The most noted examples are assemblages of mica-amphibole-rutile-ilmenite-diopside (MARID) and related rocks, which form part of xenolith suites in the Kaapvaal craton and which have been linked to Cretaceous kimberlite activity (Dawson and Smith, 1977; Erlank et al., 1987; Konzett et al., 1998, 2000; Grégoire et al., 2002). The remobilisation of such metasomes as low-volume melts requires only a modest amount of heat, for example during plume impingement, or decompression during stretching of the continental lithosphere (Mackenzie, 1989; Niu, 2008). However, if such melts or fluids were
present at MLD depth, they should be visible as conductive layers in MT studies, which is not generally the case (Selway et al., 2015).

4.2.3 Deposition of metasomatic minerals during past metasomatism

In light of the arguments presented in the previous section, it seems unlikely that globally detected cratonic MLDs are the product of the current presence of melt or fluid. In contrast, seismically slow hydrous minerals and pyroxenes may have been deposited in the cratonic lithosphere at some time in the past and may be stable over aeons in cool cratonic roots (Selway et al., 2015). The amount of metasomatic minerals necessary to produce the measured 2-7% velocity reductions over 10-20 km are 5-10% for phlogopite or 10-15% for carbonates or >15% for amphiboles or 45-100% for pyroxenite (Rader et al., 2015). Density contrasts of such metasomatic layers with the surrounding mantle may lead to long-term gravitational instability (Rader et al., 2015). Rather than precluding their current presence at MLDs, such metasomes may imply that such accumulations are not likely to be old and that future redistributions to density-stable conditions are possible. A major argument in favour of a secondary origin of at least some MLDs, unrelated to craton construction, is their continuity across distinct tectonic units, as is the case for the Kalahari craton (Savage and Silver, 2008) and parts of the North American continent (Hopper and Fischer, 2015). The main modes of introduction of metasomatic minerals would have involved metamorphic (dehydration, partial melting) processes in underthrusting oceanic slabs or infiltration of small-volume asthenosphere-derived partial melts. Archaean geothermal gradients would have been twice as high (corresponding to surface heat flows of 45-90 mW/m²) as those recorded in xenolith entrained in kimberlites over the last Ga (Michaut et al., 2009) (Fig. 3B). Nevertheless, the old ages of diamonds (summarised in Gurney et al., 2010; Helmstadt, 2013) indicate that the roots at depth >~150 km must have cooled rapidly to near-present conditions after craton stabilisation.
Metasomatic minerals accumulated during subduction

Subduction-accretion of oceanic slabs has been a popular model for the construction of cratonic roots (Helmstaedt and Schulze, 1989; Canil, 2004; Pearson and Wittig, 2008, 2014) and has been invoked to explain dipping seismic discontinuities, and in some cases spatially coincident seismic and conductive anomalies, in some cratonic mantle samples (Cook et al., 1999; Snyder, 2008; Chen et al., 2009; Cooper and Miller, 2014). Copious amounts of water may have been introduced from seawater-altered sediment, oceanic crust and mantle into mantle overlying subduction zones during dehydration reactions (Rüpke et al., 2004; Schmidt and Poli 2014), leading to formation of phlogopite and orthopyroxene by reactions of the mantle with siliceous hydrous fluids or melts (Wyllie and Sekine, 1982; Sato et al., 1997; Konzett and Ulmer, 1999; Vielzeuf and Schmidt, 2001). Such hydrous mineral-rich lithologies may persist for Ga timescales. For example, phlogopite-rich xenoliths, though very few, with 2.7 Ga minimum ages coincident with the age of craton amalgamation, have been reported from the central Slave craton (Aulbach et al., 2007). Phlogopite in mantle xenoliths from the Kaapvaal craton have been dated to 1.25 -1.0 Ga, coincident with and linked to metasomatism in the course of the formation of the Namaqua-Natal belt at the southern craton margin (Hopp et al., 2008). The deposition of hydrous metasomatic mineral assemblages near dipping slabs may reinforce seismic velocity anomalies produced by associated shear zones (Hopper and Fischer, 2015) (Fig. 5A).

Mantle metasomatism

Seismically slow hydrous minerals, such as amphibole (at pressures <3 GPa; Fumagalli et al. 2009) or phlogopite at temperature <1150º C (Sekine and Wyllie 1982; Hansen et al., 2015; Selway et al., 2015) also precipitate from OIB-like potassic, hydrous magmas in within-plate
settings (Pilet et al., 2010, 2011; Grant et al., 2014). The composition of most kimberlites demands that they interacted with (or even formed in) metasomatised refractory cratonic mantle sources (Becker and Le Roex, 2006; Sokol et al., 2014) and that, conversely, the cratonic mantle is modified by their passage. Because small-volume melts do not transport heat, complete or partial crystallisation in the cool cratonic lithosphere is expected (McKenzie, 1989). Sleep (2009) suggests that kimberlites are rarely able to penetrate beyond mid-lithospheric depths, where the cratonic mantle is under compression and becomes cold and viscous, and instead form sills that crystallise in situ as they cool. The kimberlite production rate over the last 2.5 Ga would correspond to 0.4 km³/year, amounting to a total thickness of 2 km if it were concentrated in a layer beneath the continents, which would sustain a kimberlite eruption every 10 to 100 years; consequently, the much lower observed frequency indicates that the majority of kimberlite melts are trapped at depth (Sleep, 2009).

Since low-volume melts and fluids are enriched in incompatible elements, among them the heat-producing elements (HPE: K, Th and U), they may have significant thermal effects (O’Reilly and Griffin, 2013). Estimates of the heat produced by radioactive decay in a several km thick kimberlite layer suggest that it should result in a strong curvature in the regional geotherm around that depth that would be resolvable with xenolith data (Sleep, 2009), which is not observed (Grütter, 2009). Large accumulations of amphiboles (15% over a depth interval of 20 km) are also precluded based on heat flow arguments, as modelled by Selway et al. (2015). However, smaller volumes of hydrous minerals than those required to affect cratonic heatflow may explain MLDs which have smaller and more distributed velocity reductions than it appears in studies using low-frequency filtering (Wirth and Long, 2014). Rader et al. (2015) suggest that a metasomatic phlogopite layer, now at MLD depth, was precipitated at the pre-thickening palaeo-LAB. This appears unlikely, given that the
temperature at the LAB, where the lithospheric mantle temperature approaches that of the convecting mantle, precludes the stability of phlogopite (Fig. 3), even more so in the warmer Archaean. It is also doubtful that subsequent cratonic root thickening and gradual displacement of the LAB to greater depth is due to conductive cooling (Rader et al., 2015) because this process confers neither the dryness and viscosity nor the chemical depletion and buoyancy required for craton survival (Lendardic et al., 2003; King, 2005; Wang et al., 2014).

Observed and inferred accumulations of metasomatic minerals in cratonic mantle

A major possible problem with calling for volumes of hydrous minerals large enough to produce a seismically visible anomaly is that direct evidence for the presence of carbonate-, phlogopite- or amphibole-rich rocks in cratonic xenolith suites appears sparse, that they are reported at depths different from MLD depths, or that they are distributed over depth intervals too large to produce a discontinuity (Selway et al., 2015; Rader et al., 2015). However, this paucity may relate to: (1) Kimberlites sampling a limited and anomalous volume of the mantle that may be unusually metasomatised, but may also have disproportionately contributed to the generation of small-volume, volatile-rich magmas, with preferential consumption of assemblages rich in hydrous minerals and pyroxenes (e.g. Grégoire et al., 2002; Becker and Le Roex, 2006; Konzett et al., 2013). However, since kimberlites form at high pressures, metasomatic minerals deposited at lower pressures, where MLDs occur, should not be tapped by episodic kimberlite magmatism. (2) Amphibole- and phlogopite-rich rocks being underreported in the literature because many studies focus on obtaining constraints on the age and mode of cratonic lithosphere formation, which therefore explicitly target peridotite devoid of visibly metasomatic minerals. (3) Carbonates, moderate accumulations (10-15%) of which can produce a strong negative velocity gradient (Rader et al., 2015), rapidly decomposing upon decompression, implying that their rare occurrence in
mantle xenoliths has no bearing on their presence at mantle depth (Canil, 1990). Similarly, hydrous minerals may be preferentially disrupted during transport in the kimberlite or alteration at the surface. Indeed, even when metasomatic minerals are not directly observed, evidence for the passage of low-volume hydrous or carbonated melts, from which they would have plausibly precipitated in the cratonic lithosphere, is wide-spread (Fig. 6).

Direct evidence for the presence of metasomatic minerals in cratonic mantle comes from mantle xenoliths and xenocrysts. A tabulated summary of observations is given in Rader et al. (2015) and Selway et al. (2015). As mentioned in section 4.2.2, some of the most prominent examples of such metasomes occur in the Kaapvaal craton (MARID and related rocks). Pressure-temperature estimates locate these metasomes in the shallower lithosphere (1.5-3.7 GPa), while age dating demonstrates a close temporal link to the older Group II and subsequent Group I kimberlite magmatism (80-90 Ma) or to the preceding Karroo magmatic event (Konzett et al., 1998, 2000, 2013; Grégoire et al., 2002; Bell et al., 2005; Kobussen et al., 2008; Giuliani et al., 2013, 2014, 2015b). Phlogopite-bearing polymict breccias from 110-115 km depth have also been linked to group I kimberlite magmatism and were interpreted as failed kimberlite intrusions that paved the way for subsequent kimberlite eruption (Giuliani et al., 2014). Although the Kaapvaal MARID and related rocks are carbonate-poor, xenoliths composed of phlogopite and carbonate in lamproites from the Churchill province attest to the co-precipitation of these two minerals (Peterson and LeCheminant, 1993). In contrast, nominally anhydrous, pyroxene-rich metasomes, perhaps similar to those in western branch of the East African Rift (Davies et al., 1989), are expected at higher temperatures (greater depths).
Intense low-temperature phlogopite-related metasomatism, inferred from the TiO$_2$-Zr systematics of garnet xenocrysts, is recorded at depths of 100-120 km in the Kaapvaal craton, consistent with the aforementioned presence of MARID and similar rocks in Kaapvaal xenolith suites (O’Reilly and Griffin, 2006), and is commonly observed in shallow garnet peridotites (Fig. 6). Interaction with carbonatites is fingerprinted by characteristic strong LREE-enrichment and relative depletions in the HFSE (Rudnick et al., 1993; Ionov et al., 1993) (Fig. 6). Moreover, the chemical composition of typically cratonic magmas can be used to infer the mineralogy of their mantle source. Numerous experimental studies address the relationships between phlogopite and potassic carbonated melts, such as Group II kimberlites (e.g. Sweeney et al., 1993; Ulmer and Sweeney, 2002; Enggist and Luth, 2016). These exotic rocks have been emplaced over a considerable time span - between 65 Ma in the Indian Bastar craton (Chalapathi Rao et al., 2011) and 1.92 Ga in the Karelian craton in NW Russia (Priyatkina et al., 2014) – and occur not only in Southern Africa, but also in Australia, India, Russia and Finland (Giuliani et al., 2015b). Multiple generations of metasomes, in part refilling older melt pathways and ranging from isotopically young (asthenosphere-derived) carbonate-phlogopite assemblages to isotopically aged (lithosphere-derived) MARID-type assemblages, may be present in the deeper reaches of a single craton. This has been deduced from the chemical and isotopic composition of alkaline ultramafic melts in the North Atlantic and Congo-Tanzanian cratons (Tappe et al., 2006, 2007, 2008, 2009; Rosenthal et al., 2009). Thus, there is ample indirect evidence for the presence of seismically slow minerals in the cratonic lithosphere.

**Phase equilibria and thermodynamics: Can metasomatism explain MLDs?**

In this section, we discuss the nature of the melts responsible for metasomatism of the cratonic lithosphere following two distinct steps: (1) A review of the experimental work
carried out over the last decades, and (2) a predictive approach using the recent thermodynamic parameterisation from Massuyeau et al. (2015). Because of the limited stability of amphibole (<1070°C, <3 GPa, depending on water fugacity, Fig. 3B; Green et al., 2010), its precipitation in the cratonic lithosphere can only explain MLDs observed at shallower levels, to approximately 90 km. Thus, a different mechanism is needed to explain the presence of MLD at greater depths. As reviewed above, there is both direct and indirect evidence for the existence of phlogopite (±clinopyroxene, K richterite)-rich veins. Here, we focus on the potential “metasomatising” role of CO₂-H₂O-alkali-rich liquids (such as carbonatites, kimberlites, nephelinites and basanites) in the introduction of phlogopite in the cratonic mantle. We also probe the depth range in which such melts can be produced relative to the field of phlogopite stability and the presence of MLDs.

Numerous experimental studies have focused on the formation of melts under cratonic pressure-temperature-composition conditions, with the emerging consensus that both CO₂ and H₂O are necessary for melting to occur in cratons (Foley, 1989; Foley et al., 2009; Brey et al., 2008, 2009; Girnis et al., 2011; Novella and Frost, 2014; Sokol and Kruk, 2015; Sokol et al., 2013, 2014, 2015; Sharygin et al., 2015). In most cases, experimental melts are low to very low in silica, spanning the compositional range from carbonatites to kimberlites. Considering the present-day geotherm of intact cratons (Fig. 8A), it is clear that melting is only possible at depth exceeding 150-200 km depth (Fig. 3A). In that case, melt composition is only carbonatitic (Fig. 9). Significantly higher lithospheric heat production in the Archaean implies that geothermal gradients were elevated and that cratonic lithospheres have cooled with time (Michaut et al., 2009) (Fig. 8B). Consequently, for a cratonic geotherm close to formation (3000 Ma), kimberlitic melts are stabilised from about 150 to 300 km depth, whereas carbonatitic melts are predicted above 150 km (Fig. 9). From 2000 Myrs to now, the
melt is mostly carbonatitic at any depth. These calculations, performed for different geotherms according to the age of the craton, show that silica-poor alkali-volatile-rich melts, from carbonatites to kimberlites, can be stabilised in cratons and can introduce the “metasomatising” agents necessary for the formation of phlogopite.

More formally, phlogopite in CO$_2$-bearing peridotite is stabilised at pressures lower than about 4-5 GPa (Figs. 3B,8) according to the following reaction (Ulmer and Sweeney, 2002):

$$2\text{phlogopite} + \text{MgCO}_2 = 3\text{enstatite} + \text{olivine} + \text{garnet} + \text{K}_2\text{CO}_3(l) + 2\text{H}_2\text{O}(l)$$

with the metasomatic agents K$_2$CO$_3$(l) and H$_2$O(l) added by silica-poor alkali-volatile-rich melts previously described. As noted above, the typical shear velocity reduction of between 2 and 7 % at 60-160 km depth (Rader et al., 2015) can be obtained by accumulation of 5-15 % of phlogopite (Fig. 10) in this depth range. If we compare the cooling of the cratonic lithosphere with time and the stability domain of phlogopite (Fig. 8B), several interesting observations can be made: (1) Phlogopite can be stabilised and persist at pressures relevant to typical cratonic settings. (2) Phlogopite cannot precipitate from melts leaving a fertile mantle source that contains water as the only significant volatile, except for very hot conditions (>1200º C), shallow pressures (2 GPa) and elevated water contents (460 ppm H$_2$O). Conversely, a carbonated and hydrated convecting mantle intersects its solidus at depths overlapping the stability field of phlogopite and the reported presence of MLDs. However, in modern cool cratonic lithosphere, melts form in this source at depths greater than the stability field of phlogopite, and upward percolation is required in order for phlogopite to precipitate under these conditions. (3) As suggested by Rader et al. (2015), the cooling of lithosphere with time leads to displacement of the intersection of the carbonated peridotite solidus and to the development of a “permeability cap” where subsequent melt batches are stalled. The
cooling not only triggers a deepening of the solidus intersection, but also a decrease in the thickness of the phlogopite layer, with an increase in the maximal pressure for the precipitation of phlogopite. These different observations are illustrated in a cartoon (Fig. 8C).

5. Summary and conclusions

We have integrated the physical properties, chemical composition, mineralogy and fabric of cratonic mantle samples (xenoliths, xenocrysts, melts) with experimental and thermodynamic constraints on the formation and migration of melts below and within cratons in order to isolate the petrologically most viable mechanisms for geophysically observed cratonic mantle discontinuities, namely the LAB and MLDs. Bearing in mind that geophysical and petrologic modes of lithosphere interrogation do not operate on the same spatial or temporal scales, our main summary and conclusions are as follows:

(1) In cratons, the LAB is typified by a transition from depleted to fertile (or refertilised) peridotite, by the preponderance of sheared mantle xenoliths lying on the high-temperature limb of kinked geotherms and by reduced shear (S)- and compressional (P)-wave speeds at ~170-250 km depth, in some cases accompanied by a reduction in electrical resistivity. Because the presence of a deep lithospheric lid beneath intact cratons requires deep intersection of the solidus, only small degrees of interconnected partial melts (carbonatite, kimberlite, alkaline basalt) are expected the LAB, possibly linked to upwelling of thermochemically anomalous convecting mantle. This may explain the paucity of seismic cratonic LAB observations. Rapid drainage of melt may be counteracted by pervasive shearing and small crystal sizes in the deepest cratonic mantle, by small density contrasts between melt and lithosphere and by replenishment in long-lived mantle upwellings. Seismically detectible LABs are therefore sites of melt-lithosphere interaction that may in the future lead to severe disruption of the cratonic root.
(2) Cratonic LABs are mutable boundaries that are affected by thermochemical erosion from melting in upwelling convecting mantle, thermal insulation beneath large continental masses, craton-edge driven convection and mechanical erosion or hydration from subduction at craton margins. This leads to various states of LAB disturbance ranging from melt infiltration and densification of the deep cratonic root, which is isotopically still distinguishable from the sublithospheric mantle (Kaapvaal craton), to true asthenospherisation and partial or complete loss of deep cratonic lithosphere (Tanzania, North China craton), and finally complete destruction in the course of rifting (North Atlantic craton). Melt percolation and attendant rejuvenation explain weak, intermittent or discordant LAB signals.

(3) MLDs are recorded between 60 and 160 km depth (predominantly 80-120 km) over a range of ~30-40 km and are characterized by both positive and negative seismic velocity anomalies, often accompanied by anisotropy. Because of the cool thermal state of intact cratonic roots, the present-day global presence of melts at MDL depths is unlikely. Preferred explanations centre on the accumulation of seismically slow metasomatic minerals (metasomes rich in hydrous minerals, pyroxenes and/or carbonates) and the acquisition of fabric due to deformation during craton construction involving plume subcretion or collisional tectonics. The latter is consistent with the observation of those MLDs that appear truncated at intracratonic sutures.

(4) Azimuthal and radial anisotropy accompanied by seismic velocity reduction may result from the accumulation of metasomes as layers at or as subvertical veins above the depth at which melt flow transitions from pervasive to focussed flow at the mechanical boundary layer, above which cratonic mantle is cold and viscous. They have ages ranging from
Archaean to immediately preceding kimberlite eruption and are the expected consequence of the interaction of cratonic lithosphere with hydrous siliceous melts/fluids during subduction beneath (proto)cratons, or with asthenosphere-derived small-volume melts, such as the kimberlites that eventually transport cratonic mantle samples to the surface. The continuity of some MLDs across distinct tectonic units suggests a secondary origin of at least some MLDs, unrelated to craton construction. Thermal and density considerations imply that some of these buoyant metasomes, which are rich in heat producing elements, are young and may be the sites of future reorganisation of cratonic mantle roots.

(5) Amphibole stability being limited to <~90 km depth, our thermodynamic approach focuses on the depth range in which CO₂-H₂O-alkali-rich liquids (carbonatites, kimberlites, nephelinites and basanites) can be produced relative to the field of phlogopite stability and the presence of MLDs. Our calculations show that such melts form in the asthenosphere and penetrate upwards, possibly conditioning the lithosphere during multiple intrusions. Phlogopite is precipitated as a function of age-dependent thermal state, thus explaining variable MLD depths. Even if not directly observed, such isotopically aged metasomes have been shown to contribute to small-volume volatile-rich melts typically penetrating cratonic lithospheres. Finally, we suggest that the apparent paucity of evidence for phlogopite- or amphibole-rich assemblages in the mantle xenolith record at geophysically imaged MLD depths, if not due to preferential disaggregation or alteration, relates to undersampling by both kimberlites and humans.

Acknowledgments

SA gratefully acknowledges funding from the Deutsche Forschungsgemeinschaft under grant DFG AU386/8. Andrea Tommassi and Ingo Wölbern are thanked for providing invaluable
informal feedback on the manuscript. Insightful and constructive formal reviews by Gumer Galán and Michel Grégoire are highly appreciated, as are the seamless editorial handling by Andrew Kerr and the positivity, patience and swiftness with which Tim Horscroft has supported this work throughout the process.

References


Metasomatism from an Archean (similar to 3.5 Ga) Subduction Zone. Science 326, 1089-1091.


magmas, and relationship to the kimberlite megacrystic suite. Chemical Geology 353, 4-18.
Grant, T.B., Milke, R., Wunder, B., 2014. Experimental reactions between olivine and orthopyroxene with phonolite melt: implications for the origins of hydrous amphibole plus phlogopite plus diopside bearing metasomatic veins. Contributions to Mineralogy and Petrology 168.


upper mantle: An experimental study in a peralkaline KNCMASH system to 8.5 GPa. Journal of Petrology 38, 537-568.


Sleep, N.H., 2002. Local lithospheric relief associated with fracture zones and ponded plume
Sleep, N.H., 2009. Stagnant lid convection and carbonate metasomatism of the deep continental
lithosphere. Geochemistry Geophysics Geosystems 10, Q11010, doi:10.1029/2009GC002702
Smith, D., Riter, J.C.A., Mertzman, S.A., 1999. Water-rock interactions, orthopyroxene growth, and
Si-enrichment in the mantle: evidence in xenoliths from the Colorado Plateau, southwestern United
Snyder, D.B., 2008. Stacked uppermost mantle layers within the Slave craton of NW Canada as
defined by anisotropic seismic discontinuities. Tectonics 27.
Snyder, D.B., 2013. Imaging Archaean-age whole mineral systems. Precambrian Research 229, 125-
132.
architecture of the Slave craton, northwest Canada, as determined from an interdisciplinary 3-D
Seismic evidence for stratification in composition and anisotropic fabric within the thick lithosphere
of Kalahari Craton. Geochemistry Geophysics Geosystems 14, 5393-5412.
experiments on the Udachnaya kimberlite at 6.3-7.5 GPa: Implications for the role of H2O in
magma generation and formation of hydrous olivine. Geochimica Et Cosmochimica Acta 101, 133-
155.
kimberlite magmas: Constraints from multiple saturation experiments. American Mineralogist 99,
2292-2302.
Sokol, A.G., Khokhryakov, A.F., Palyanov, Y.N., 2015a. Composition of primary kimberlite magma:
constraints from melting and diamond dissolution experiments. Contributions to Mineralogy and
Petrology 170.
carbonation and wehrlitization of lithospheric peridotite upon interaction with carbonatitic melts.
Doklady Earth Sciences 465, 1262-1267.
Soustelle and Tommasi 2010
Interaction in the Supra-subduction Mantle: Microstructures and Water Contents in Peridotite
Soustelle, V., Tommasi, A., Demouchy, S., Franz, L., 2013. Melt-rock interactions, deformation,
hydration and seismic properties in the sub-arc lithospheric mantle inferred from xenoliths from
Dynamical constraints on kimberlite volcanism. Journal of Volcanology and Geothermal Research
155, 18-48.
and the extraction of carbon from Earth's interior. Nature 493, 84-88.
Sweeney, R.J., Thompson, A.B., Ulmer, P., 1993. Phase relations of a natural MARID composition
and implications for MARID genesis, lithospheric melting and mantle. Contributions to Mineralogy
and Petrology 115, 225-241.


Figure Captions

**Fig. 1** Global map of cratons (modified after Bleeker, 2003; Pearson and Wittig, 2008, 2014). Grey: exposed Archaean crustal rocks; white dotted lines: other fragments of composite cratons; yellow dashed lines: Archaean cratonic blocks amalgamated during the Proterozoic; stars: intact (thick) cratonic lithosphere sampled by kimberlites and related rocks; circles: disturbed (thinned) cratonic lithosphere sampled by basalt-borne xenoliths (an extensive list of mantle xenolith localities is provided by Nixon, 1987). Uncertain boundaries are denoted with question marks.

**Fig. 2** Models for craton thickening and associated development of fabric. (A) A cratonic nucleus is subcreted by upwelling plume mantle (e.g. Griffin et al., 1999, 2004; Arndt et al., 2009). Vertical lineation in the plume stem gives way to horizontal lineation when the plume impinges upon the LAB (Ito et al., 2014). Dipping lineations may result when material flows laterally towards thinner craton margins (Sleep, 2002; Shervais and Hanan, 2008). Additional radial anisotropy may be preserved if small-scale convection in the plume (Agrusta et al., 2015) is frozen in after melt extraction. Arrows below the plume head indicate mantle flow. Inset shows expected variation of olivine Mg# (molar Mg/(Mg+Fe)) with depth, which decreases due to the effects of polybaric melt extraction (1); a discrete step forms when mantle packages formed at different T_P are juxtaposed whereby the pre-existing lithospheric lid prevents high degree of depletion (high Mg#) in the subcreted mantle volume; red stippled line (2) with downward arrow illustrates that the deepest portions of cratonic mantle would have been winnowed by a variety of processes (e.g. Arndt et al., 2009) prior to thickening, whereas green horizontal arrow (3) shows effect of refertilisation (Fe addition). (B) A cratonic nucleus is accreted/subcreted by buoyant oceanic crust and mantle during subduction (e.g. Pearson and Wittig, 2008; Helmstaedt, 2009). Imbrication of oceanic slabs leads to
dipping lineations while horizontal lineations result from flat subduction. If the two packages formed at similar time by partial melting to low pressure (as is true at ridge settings), the Mg# profile of each lithosphere package would be identical, with largest degrees of depletion (high Mg#) at the top of each package. (C) Collision of two cratonic nuclei subsequent to ocean closure (Lee et al., 2011), with different thicknesses at the time of collision due to formation at different T_P (hence melting intervals; see Fig. 3A). The steep resulting boundary, such as that determined for the amalgamated eastern and western domains in the Slave craton (Poudjom-Djomani et al., 2005), truncates pre-existing MLDs that may have developed at different depths during the pre-collision stage of lithosphere thickening resulting from processes, such as those shown in (A) and (B). Near the boundary, olivine Mg# with depth may show complexity.

Fig. 3 (A) Solidi of dry and carbonated peridotite (Dasgupta et al., 2013) (solid blue lines), and combination of the solidus temperatures of Wallace and Green (1988) from 0 to ~4 GPa, and Foley et al. (2009) at higher pressures (dashed blue line), relative to a range of conductive geotherms applicable to intact (35-40 mW/m²) and disrupted cratonic lithosphere (50-60 mW/m²). The red line on the carbonated peridotite solidus highlights the ledge above which upward percolating melts would freeze; it would lie at greater depth in water-bearing or depleted mantle (Hopper and Fischer, 2015). Shown for comparison are mantle potential temperatures (T_P) at modern ridges (~1350°C; Herzberg et al., 2010), Archaean ridges (~1450°C; Davies, 2009) and at excess T_P (plumes); black stars illustrate how the dry peridotite solidus is intersected at increasing depth with increasing T_P, allowing the formation of increasingly thick melt-depleted residue (= lithosphere) due to larger melting intervals (e.g. Herzberg et al., 2010); thickness variations may further result from composition-dependent location of the mantle solidus (Hirschmann, 2000); purple star shows intersection of the
carbonated peridotite solidus at much greater depth, relevant to the generation of small
degrees of melt at modern LABs. (B) Stability of phlogopite and amphibole (pargasite): (1)
Stability of amphibole (Selway et al., 2015), (2) stability of MARID-type K richterite and (3)
solidus of dry MARID (Sweeney et al., 1993); (4) solidus of dry MARID, (5) stability of
MARID-type K richterite and (6) of K richterite in the system KCNMASH (Konzett et al.,
1997); (7) stability of phlogopite (Ulmer and Sweeney, 2002). Shown for comparison are
geothermal gradients expected in intact cratonic roots today and 3 Ga ago (Michaut and
Jaupart, 2007). (C) Oxygen fugacity relative to the fayalite-magnetite-quartz (FMQ) oxygen
buffer determined for cratonic mantle samples. Also shown are the fields for graphite,
diamond and carbonate melt stability, as well as the carbonate concentration in the melt as a
function of fO2, with increasing fO2 stabilising increasing amounts of CO3 2- in the melt (after
Stagno et al., 2013). Thick red stippled line shows fO2 as a function of pressure for pre-
metasomatic mantle from Creighton et al. (2009) for comparison, highlighting that many
kratonic mantle xenoliths are more oxidised. (D) Histogram shows number of MLD
observations per 2° x 2° bin across the globe (from Rader et al., 2015); black bar below
shows depths of intact cratonic LABs (O Reilly and Griffin, 2010).

**Fig. 4** Establishment of a palaeo-LAB with topography due to winnowing at the base of the
lithosphere. Rifting and stretching due to impingement of a thermo-chemical plume, passive
mantle upwelling, the establishment of small-scale convection cells in the sub-lithosphere due
to continental insulation, or convection at the craton edge due to a step in in the LAB may
lead to infiltration by low-degree melts and consequent thermochemical erosion (Lee et al.,
2011), densification and destabilisation (Wang et al., 2015) and foundering of denser, less
depleted and potentially cooler deep residues (Arndt et al., 2009; Michaut et al., 2009). Low-
density methane-rich fluids below the depth of metal saturation may provide a steady supply
of carbon that is oxidised just below the LAB, escaping as carbonatite (Frost and McCammon, 2008; Rohrbach and Schmidt, 2011). Melt may form and/or be focussed into depressions at the LAB.

**Fig. 5** (A) Intact craton with initial disturbance at the LAB and craton margins (left) - evidenced by magmatism involving small-volume melts, such as kimberlites (e.g. Slave and Kaapvaal cratons), which eventually may lead to loss of the deep lithospheric root in a disrupted craton (right; e.g. Tanzanian, North China and North Atlantic cratons) - evidenced by magmatism involving larger melt volumes, such as alkali basalts (Foley, 2008). Inset shows schematically the deposition of phlogopite and formation of orthopyroxene at the expense of olivine, as expected during the interaction of slab-derived hydrous, siliceous fluids/melts with the overlying mantle wedge (Wyllie and Sekine, 1982), which can strengthen the anomalous seismic signal expected from dipping shear zones (Hopper and Fischer, 2015). (B) Attendant changes in geothermal gradients as deduced by pressure-temperature arrays of mantle xenoliths and xenocrysts. An initially cool thermal state (1), corresponding to surface heat flows of 35-42 mW/m² (Grütter, 2009) in an intact cratonic root, gives way to a significantly higher geothermal gradient (3) (50-60 mW/m²; e.g. eastern North China craton; Xu, 2001). Intermediate stages (2) may be indicated by kinked or stepped geotherms, or xenolith-derived arrays showing large scatter in temperature at a given pressure (O’Reilly and Griffin, 2010).

**Fig. 6** (A) Schematic evolution of REE patterns in garnet: “Normal” chondrite-normalised patterns with LREE depletion and monotonously increasing REE contents give way to mildly sinusoidal and strongly sinusoidal patterns (Aulbach et al., 2013); the process is reversible upon interaction with high-volume melts, which also leads to refertilisation and densification
(Fe addition) in seemingly little depleted peridotites (O’Reilly and Griffin, 2013; Pearson and Wittig, 2014). (B,C) Trace-element abundances and ratios in garnet from the Kaapvaal, Siberian and Slave cratons (PetDB; http://www.earthchem.org/petdb/) betraying the passage of increasingly volatile-rich melts (decreasing melt volumes) after interaction with garnet-bearing mantle (Aulbach et al., 2013). Low Zr/Y has been identified as typical of phlogopite-bearing xenoliths and can be recognised in garnet even when phlogopite is not sampled in the xenolith volume (Griffin et al., 1999).

Fig. 7 Melt migration and solidification in cratonic mantle where a thermal boundary layer (TBL), the depth of which is movable, depending on the temperature (ambient or excess $T_p$) at the LAB, underlies a mechanical boundary layer (MBL) with a cold, viscous cratonic core. (A) Melts formed at the LAB in response to heating, decompression or addition of volatiles penetrate into the cratonic root. Melts out of chemical or thermal equilibrium with the lithosphere are expected to react out and freeze. Upward mobility of subsequent melt batches is facilitated by prior “conditioning” of the mantle through chemical reactions leading to precipitation of metasomatic minerals (O’Reilly and Griffin, 2013) and through heating. (B) Melt percolation in the TBL may be pervasive where initial deformation gives way to recovered (recrystallised) microstructures which are seismically nearly isotropic; in contrast, in the MBL, melt extraction may only be possible upon focussing and vein formation, with attendant deformation along the conduit walls (Bascou et al., 2011; Baptiste et al., 2012), leading to formation of a radially anisotropic layer. The rheological contrast between the two layers may represent a barrier to upward movement of melt, leading to ponding and fractional crystallisation of melt, with deposition of seismically slow minerals, such as phlogopite, pyroxenes and carbonate. Both the transition from an isotropic to an anisotropic layer, and partial crystallisation of the melt may cause an MLD. (C) Schematic grain-scale effects of
pervasive melt percolation with interconnected melt pockets and attendant chemical zoning in olivine and other minerals, along with deposition of texturally and compositionally unequilibrated clinopyroxene ± garnet from silica-undersaturated, alkaline melts. This gives way to melt accumulation and fractional crystallisation at the rheological boundary formed by the transition from the TBL to the MBL. In the MBL, melt is extracted in veins, with adjacent rocks showing strong fabric. Successful extraction of the melt may require plating of the conduit walls with a melt-peridotite reaction assemblage.

**Fig. 8** (A) Pressure-temperature conditions prevailing in intact and disrupted cratons (having deep and shallow lithospheric roots, respectively) and in the oceanic domain, delimited by geotherms corresponding to different surface heat flows in mW.m\(^{-2}\) (grey curves, Hasterok and Chapman, 2011). Colour curves designate the solidi for dry (dark, Hirschmann, 2000), hydrated (200 p.p.m. and 460 p.p.m. H\(_2\)O; blue, Hirschmann et al., 2009) and H\(_2\)O-undersaturated CO\(_2\) peridotites (labelled “CO\(_2\)-H\(_2\)O”; purple, combination of the solidus temperatures of Wallace and Green (1988) from 0 to ~4 GPa, and Foley et al. (2009) at higher pressures); the stability domain of phlogopite (Ulmer and Sweeney, 2002) is shown for comparison. Purple stars highlight the intersection between a continental geotherm and the solidus of H\(_2\)O-undersaturated CO\(_2\) peridotite, and blue stars point to the intersection between a geotherm and the solidi for hydrated peridotite. The depth ranges of MLDs and LABs in cratons are indicated on the left side. At higher temperatures than the dry peridotite solidus, basaltic melts are produced (red domain). At lower temperatures, incipient melting occurs due to the presence of CO\(_2\), with the production of CO\(_2\)-H\(_2\)O-rich melts (purple domain). Incipient melting is initiated by redox melting at depth, which corresponds to the change of carbon speciation from graphite or diamond (represented by diamond symbols) to carbonate. The decreasing size of the diamonds between 250 and 150 km depth symbolises the variable depth
of redox melting (Rohrbach and Schmidt, 2011; Stagno et al., 2013). (B) Evolution of
depression of the middle and lower crust. (C) Coexistence of CO2-rich and CO2-poor magmas is
favored by the presence of a CO2-rich mantle plume, which can provide a source of CO2-rich
magma. The mechanism by which CO2-rich magmas are generated involves the interaction of
CO2-bearing fluids with the mantle and crust. The presence of CO2-rich magmas can have
significant implications for the evolution of the crust and mantle, including the formation of
CO2-bearing rocks and the generation of CO2-rich magmas.

Fig. 9 Approximate melt composition is calculated for a cratonic lithosphere using a revised
thermodynamic model of Massuyeau et al. (2015), calibrated on an expanded set of
experimental data (197 new experimental points) to increase the accuracy on peridotite
melting under CO2-poor conditions at pressures > 1-2 GPa. Geotherms for the cratonic
lithosphere are a function of its age (Michaut et al., 2009), and we consider a mantle with 500
ppm H2O – 500 ppm CO2. Cratonic lithosphere is described as a two-layer structure: (1) a
depleted layer that extends from 60 km to 160 km, (2) a less depleted layer that extends from
160 to 250 km depth (Afonso et al., 2008). The mineral proportions and chemical
compositions of these layers are approximated from Walter (2005).

Fig. 10 Shear velocities (Vs) of different peridotite models relevant to cratons, from a Voigt-
Reuss-Hill average (Hacker et al., 2003). Calculations are done for 2 to 5 GPa, following the
present-day geotherm (Michaut et al., 2009; Fig. 8) and using an Excel spreadsheet developed
by Hacker and Abers (2004). Velocity for a reference model depleted cratonic lithosphere peridotite (Afonso et al., 2008), consisting of 85 % olivine (75.2 % forsterite, 9.8 % fayalite) and 15 % orthopyroxene (13.5 % enstatite, 1.5 % ferrosilite) is shown. The dark-shaded area around the reference model represents variations due to adding 6 % garnet: almandine, grossular or pyrope, for which we can observe a small variation of Vs, principally characterized by an increase. Different thin maroon curves are the reference peridotite model modified by adding different proportions of phlogopite: 2, 5, 10 and 15 %. Phlogopite is added by replacing orthopyroxene (principally enstatite followed by ferrosilite component) (Ulmer and Sweeney, 2002). The presence of phlogopite, between 2 and 15 % in proportion, is a good candidate to explain the decrease of Vs, between 2 and 7% (Rader et al., 2015) of the reference state as represented by the blue-shaded area.
Fig. 1 Aulbach et al.
Fig. 2 Aulbach et al.
Fig. 3 Aulbach et al.
Fig. 4 Aulbach et al.
Fig. 5 Aulbach et al.
Fig. 6 Aulbach et al.
Fig. 7 Aulbach et al.
Aulbach et al. Fig. 10
Table 1 SiO\textsubscript{2} content of melts and melt fractions as a function of source composition and T\textsubscript{P}

<table>
<thead>
<tr>
<th>LAB depth (km)</th>
<th>100</th>
<th>150</th>
<th>200</th>
<th>250</th>
</tr>
</thead>
<tbody>
<tr>
<td>%melt in PM\textsuperscript{1}</td>
<td>0.65</td>
<td>0.41</td>
<td>0.27</td>
<td>0.17</td>
</tr>
<tr>
<td>SiO\textsubscript{2} content (wt.%)</td>
<td>39.0</td>
<td>32.8</td>
<td>24.7</td>
<td>12.5</td>
</tr>
<tr>
<td>%melt in DM\textsuperscript{2}</td>
<td>0.21</td>
<td>0.13 (0.05)*</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>SiO\textsubscript{2} content</td>
<td>37.7</td>
<td>30.3 (6.65)*</td>
<td>5.53</td>
<td>4.98</td>
</tr>
<tr>
<td>%melt in EM\textsuperscript{3}</td>
<td>1.15</td>
<td>0.72</td>
<td>0.52</td>
<td>0.40</td>
</tr>
<tr>
<td>SiO\textsubscript{2} content</td>
<td>39.5</td>
<td>33.7</td>
<td>27.5</td>
<td>21.2</td>
</tr>
<tr>
<td>%melt in hot EM\textsuperscript{4}</td>
<td>3.13</td>
<td>1.66</td>
<td>1.26</td>
<td>1.11</td>
</tr>
<tr>
<td>SiO\textsubscript{2} content</td>
<td>45.6</td>
<td>42.1</td>
<td>39.1</td>
<td>36.7</td>
</tr>
</tbody>
</table>

*At 150 km depth, for the depleted mantle, immiscibility may occur as predicted by the model. Values in brackets represent the SiO\textsubscript{2} content and melt fraction of the immiscible carbonate liquid coexisting with the silicate-rich melt.

\textsuperscript{1}Primitive mantle (PM) at a mantle T\textsubscript{P} of 1350°C and with 500 ppm H\textsubscript{2}O – 500 ppm CO\textsubscript{2}

\textsuperscript{2}Depleted mantle (DM, MORB source) at a mantle T\textsubscript{P} of 1350°C and with 200 ppm H\textsubscript{2}O and 200 ppm CO\textsubscript{2}

\textsuperscript{3}Enriched mantle (EM, OIB source) at a mantle T\textsubscript{P} of 1350°C and with 800 ppm H\textsubscript{2}O and 800 ppm CO\textsubscript{2}

\textsuperscript{4}Enriched mantle (EM) at a mantle T\textsubscript{P} of 1530°C and with 800 ppm H\textsubscript{2}O and 800 ppm CO\textsubscript{2}

EM and DM are from Hirschman (2010), PM represents an intermediate composition.
Abstract

Geophysically detectible mid-lithospheric discontinuities (MLD) and lithosphere-asthenosphere boundaries (LAB) beneath cratons have received much attention over recent years, but a consensus on their origin has not yet emerged. Cratonic lithosphere composition and origin is peculiar due to its ultra-depletion during plume or accretionary tectonics, cool present-day geothermal gradients, compositional and rheological stratification and multiple metasomatic overprints. Bearing this in mind, we integrate current knowledge on the physical properties, chemical composition, mineralogy and fabric of cratonic mantle with experimental and thermodynamic constraints on the formation and migration of melts, both below and within cratonic lithosphere, in order to find petrologically viable explanations for cratonic mantle discontinuities.

LABs characterised by strong seismic velocity gradients and increased conductivity require the presence of melts, which can form beneath intact cratonic roots reaching to ~200-250 km depth only in exceptionally warm and/or volatile-rich mantle, thus explaining the paucity of seismical LAB observations beneath cratons. When present, pervasive interaction of these -typically carbonated- melts with the deep lithosphere leads to densification and thermochemical erosion, which generates topography at the LAB and results in intermittent seismic LAB signals or conflicting seismic, petrologic and thermal LAB depths. In rare cases (e.g. Tanzanian craton), the tops of live melt percolation fronts may appear as MLDs and, after complete lithosphere rejuvenation, may be sites of future, shallower LABs (e.g. North China craton).

Since intact cratons are presently tectonomagmatically quiescent, and since MLDs produce both positive and negative velocity gradients, in some cases with anisotropy, most MLDs may
be best explained by accumulations (metasomes) of seismically slow minerals (pyroxenes, phlogopite, amphibole, carbonates) deposited during past magmatic-metasomatic activity, or fabric inherited from cratonisation. They may accumulate as layers at, or as subvertical veins above, the depth at which melt flow transitions from pervasive to focussed flow at the mechanical boundary layer, causing azimuthal and radial anisotropy. Thermodynamic calculations investigating the depth range in which small-volume melts can be produced relative to the field of phlogopite stability and the presence of MLDs show that phlogopite precipitates at various pressures as a function of age-dependent thermal state of the cratonic mantle, thus explaining variable MLD depths. Even if not directly observed, such metasomes have been shown to be important ingredients in small-volume volatile-rich melts typically penetrating cratonic lithospheres. The apparent sparseness of evidence for phlogopite-rich assemblages in the mantle xenolith record at geophysically imaged MLD depths, if not due to preferential disaggregation in the kimberlite or alteration, may relate to vagaries of both kimberlite and human sampling.
Highlights

- Review of petrologically plausible explanations for cratonic mantle discontinuities
- Melt at cratonic LAB depth (≥200 km) due to anomalous sublithospheric X-T, retarded drainage and/or resupply
- Conflicting seismic-petrologic LAB signatures due to thermochemical-mechanical erosion
- MLDs result mostly from past deposition of metasomes at mechanical boundary layers
- Omnipresent metasomes documented by alkaline melts, underestimated due to non-representative sampling