Understanding cratonic flood basalts

Paul G. Silver *, Mark D. Behn ¹, Katherine Kelley ², Mark Schmitz ³, Brian Savage

Carnegie Institution of Washington, DTM, United States

Received 3 March 2005; received in revised form 23 January 2006; accepted 30 January 2006
Available online 20 March 2006

Abstract

The origin of continental flood basalts has been debated for decades. These eruptions often produce millions of cubic kilometers of basalt on timescales of only a million years. Although flood basalts are found in a variety of settings, no locale is more puzzling than cratonic areas such as southern Africa or the Siberian craton, where strong, thick lithosphere is breached by these large basaltic outpourings. Conventionally, flood basalts have been interpreted as melting events produced by one of two processes: 1) elevated temperatures associated with mantle plumes and/or 2) adiabatic-decompression melting associated with lithospheric thinning. In southern Africa, however, there are severe problems with both of these mechanisms. First, the rifting circumstances of several well-known basaltic outpourings clearly reflect lithospheric control rather than the influence of a deep-seated plume. Specifically, rift timing and magmatism are correlated with stress perturbations to the lithosphere associated with the formation of collisional rifts. Second, the substantial lithospheric thinning required for adiabatic decompression melting is inconsistent with xenolith evidence for the continued survival of thick lithosphere beneath flood basalts domains. As an alternative to these models, we propose a new two-stage model that interprets cratonic flood basalts not as melting events, but as short-duration drainage events that tap previously created sublithospheric reservoirs of molten basalt formed over a longer time scale. Reservoir creation/existence (Stage I) requires long-term (e.g. \( \gg 1 \) Ma) supersolidus conditions in the sublithospheric mantle that could be maintained by an elevated equilibrium geotherm (appropriate for the Archean), a slow thermal perturbation (e.g. thermal blanketing or large-scale mantle upwelling), or a subduction-related increase in volatile content. The drainage event (Stage II) occurs in response to an abrupt stress change in the lithosphere, which leads to the initiation and propagation of lithospheric dikes. Such a model accounts for the short eruption time of flood basalts, the evidence for lithospheric control of the eruptions, and the continued survival of cratonic lithosphere following these magmatic events. The most notable consequence of this model is that it implies the existence of large reservoirs of magma, comparable to the eruption volume of flood basalts that have been, and are still likely, present beneath stable continental lithosphere.

© 2006 Elsevier B.V. All rights reserved.

Keywords: flood basalts; craton; lithosphere

1. Introduction

From time to time, the Earth’s surface is pierced by tremendous outpourings of basalt, the primary melt product of the Earth’s mantle. These so-called flood basalts are unlike the basalts from arc volcanoes that decorate convergent plate boundaries, nor are they like...
basalts that are erupted along the globe-encircling length of mid-ocean ridges. They are indeed more akin to “floods”, characterized by eruptions of millions of cubic kilometers of basalt that can spread over a large fraction of a continent with great rapidity—often in less than one million years. Flood basalts occur in both continental and oceanic settings, where they are termed continental flood basalts and oceanic plateaus, respectively. They are often associated with continental breakup, as in the case of Paraná, Karoo, and Deccan flood basalts accompanying the breakup of Gondwanaland. Others are associated with rifting in stable cratons that does not lead to breakup, such as the Mid-Continent Rift in North America, or the Ventsdorp Rift of southern Africa. Their occurrence is not predicted by plate tectonics and consequently their origin remains enigmatic. What causes them to occur? How do they erupt so quickly? Where does the basalt come from?

Much of the thinking on the creation of flood basalts is based on the classic paper of White and McKenzie [1], which interprets flood basalts as melting events that are caused by one of two processes: (i) a high-temperature plume that serves to replace subsolidus sublithospheric mantle with supersolidus mantle, or (ii) the thinning of the lithosphere, thereby reducing the pressure of mantle material that is sitting just below the base of the lithosphere (Fig. 1A,B). The first of these is inspired by the plume hypothesis [2,3], that such outpourings are the result of deep, high-temperature plumes, while the second is an adaptation of the adiabatic-decompression melting model that so successfully accounts for the formation of mid-ocean-ridge basalts [4]. The combination of both of these mechanisms provides a reasonable explanation for flood basalts that are associated with breakup.

In cratonic environments and in the absence of breakup, the necessary lithospheric thinning could be caused by either stretching of the lithosphere [1], or through some form of convective instability and lithospheric delamination, such as that advocated for Tibet (e.g. [5,6]). The latter process would account for regions where there is little surface indication of the required lithospheric stretching needed for adiabatic decompression, although such models predict uplift following delamination. Cratonic southern Africa represents a particularly stringent test of these two mechanisms. There have been several massive outpourings or intrusions of continental basalts in the Precambrian of southern Africa, and thanks to the results from the recent Kaapvaal Project [7], it is now possible to examine the characteristics of these magmatic events in unprecedented detail. This information is in the form of new analyses of crust and mantle xenolith suites (chronology, geothermobarometry) from southern Africa’s numerous kimberlite pipes, as well as new seismic constraints on the structure of the crust and upper mantle provided by the associated Southern African Seismic Experiment.

2. Cratonic flood basalts in southern Africa

We consider the occurrence of 4 large magmatic events (Table 1, Figs. 1 and 2) spanning the time period 1.8–2.7 Ga: Ventsdorp (2.71 Ga), Great Dyke (2.57 Ga), Bushveld (2.06 Ga), and Soutpansberg Trough (1.88 Ga). The Ventsdorp is commonly regarded as a flood basalt covering much of the Archean Kaapvaal Craton [8–11]. The Bushveld, the world’s largest layered igneous intrusion, has been referred to as the intrusive equivalent of a flood basalt, given its extensive volume and short duration [8,12]. The 500-km-long Great Dyke is likely the eroded remnant of a large igneous province.
Indeed the remnant intrusion alone has an area of 60,000 km², roughly a third of the area of magmatic events referred to as flood basalts. Barton and Pretorius [13] interpreted the >3 km thickness of Soutpansberg Group basaltic eruptives as erosional remnants of a much larger Proterozoic large igneous province.

Analysis of the rifting characteristics of these 4 magmatic events shows that they were closely related to collisional orogens in two important ways. First, the actual rift orientations suggest that they exploited pre-existing mechanical anisotropy in the lithospheric mantle which was produced by strain-induced lattice preferred orientation of olivine [14,15] imparted to the mantle by a previous orogenic event. The alignment of olivine, manifested as seismic anisotropy through the fast polarization direction of split shear waves, primarily reflects the collisional deformation associated with the ca. 2.9 Ga Kimberly and Pietersburg orogens along the western and northern margins of the Kaapvaal craton, respectively [16,17], as well as earlier events on the Zimbabwe craton [15]. For all four magmatic events, the rift orientations are parallel to this pre-existing mantle fabric. Indeed, there is increasing evidence that a close relationship exists between mechanical anisotropy and seismic anisotropy in other cratonic environments such as Australia [18] and the Canadian Shield [19], as well as in oceanic environments [20].

Second and particularly striking, the timing of these rifting/magmatic events are virtually synchronous with subsequent major orogens (Table 1, Fig. 2B) [15]. The Neoarchean Ventersdorp event is temporally linked with compressional deformation and crustal thickening in the Central and Southern Marginal Zones of the Limpopo Orogen [21–23]. Similarly, the Great Dyke was emplaced during the exhumation of the Northern Marginal Zone of the Limpopo Orogen over the Zimbabwe Craton [23–26], and the Paleoproterozoic emplacement of the Bushveld–Molopo Farms Complex was nearly synchronous with the Magondi Orogeny along the NW margin of the Zimbabwe–Kaapvaal Cratons [27–29]. Significant dextral transpressional reactivation of the Triangle and Palala Shear Zones of the Limpopo Belt are also recorded at this time [30–32]. Subsequent extensional reactivation of the Palala Shear Zone and formation of the Soutpansberg Trough coincided with compressional deformation in the Kheis Belt along the western Kaapvaal Craton [33,34]. We interpret such synchronicity as evidence that these magmatic outpourings were produced by collisional rifts, i.e. rifts that formed in the stress field produced by an adjacent collisional orogen, with the actual orientation of the rift being controlled by pre-existing mechanical anisotropy. Other more recent examples of collisional rifts include the Rhine Graben, and the Baikal Rift [35–37].

This close temporal and spatial relationship between magmatic events and collisional orogens renders the plume explanation highly unlikely, as it would require that plume events just happen to occur with each major orogen. Rather, it appears that these magmatic events are caused by stress perturbations in a lithosphere weakened by mechanical anisotropy, not melting due to deep-seated, high-temperature mantle plumes.

It is also difficult to appeal to lithospheric thinning, either by lithospheric stretching or delamination, to produce adiabatic decompression melting beneath the southern African craton. The first issue is timing. Under the assumption that the million-year flood basalt time scale is a melting-event time scale (i.e. residence time of melt at depth is negligible), there are few processes that can realistically occur over such a short period of time. For example, pressure-release melting by lithospheric stretching requires stretching factors (or $\beta$ factor in the terminology of [1]) of order two. Thus, to produce large volumes of melt in 1 My would require a strain rate of $10^{-6}$/yr, which is an order of magnitude higher than the strain rates of present-day rifts ($\sim 10^{-7}$/yr or less). This includes the most prominent extensional environments active today, such as the Baikal Rift, East African Rift, and Basin and Range. (Baikal Rift: $0.3 \times 10^{-7}$/yr [38]; East African Rift: $0.5 \times 10^{-7}$/yr [39], Basin and Range: $<1 \times 10^{-7}$/yr [40]).

It is more difficult to reject delamination on the basis of timing. While [1] argued that this process is too slow, others suggest that it can occur as rapidly as a few million years [5,6]. However, even if lithospheric delamination could conceivably occur in a sufficiently short period of time, it still must be determined if delamination indeed has occurred in the case of the southern African flood
Fig. 2. (A) Map of southern Africa showing major rift systems with basaltic magmatism: Great Dyke, Venterdorp (surface exposures shown), Bushveld, and Soutpansberg. Also shown, shear-wave splitting fast polarization directions (from [82]). Abbreviations: TML: Thabazimbi Murchison Lineament, CL: Colesberg Lineament, PSZ, Palala Shear Zone, TSZ, Triangle shear zone. Green filled circles denote locations of kimberlite pipes, whose mantle xenoliths sample the mantle beneath the Venterdorp and Bushveld. Ages of nodules are denoted in blue. (B) Schematic illustrating development of collisional rifts and basaltic magmatism in southern Africa. Arrows show compression direction of orogen. Short lines show mantle fabric inferred from seismic anisotropy, which is either being formed (yellow) or fossilized (black). Red (blue) bars denote magmatic events during (after) activity. Top-left: Emplacement of the Venterdorp supergroup due to the early phase of the Limpopo orogen (2.71 Ga). Top-right: Emplacement of the Great Dyke due to the late phase of the Limpopo orogen (2.57 Ga). Bottom-left: Reactivation of Limpopo structures by Magondi Orogen and the creation of a collisional rift that produced the Bushveld Instrusion (2.06 Ga). Bottom-right: Creation of the collisional Soutpansberg Rift by the Kheis Orogen (1.88 Ga). After Silver et al. [15].
basalts. The most direct means of testing for delamination in cratonic environments is to study the geothermo-barometry and chronology of mantle xenoliths from kimberlite pipes to determine whether the mantle portion of the lithosphere was removed at the time of magmatism. These xenoliths provide an age—either of basalt depletion, or in the case of diamond growth, the age of diamond formation—and an estimate of the \( P, T \) conditions at the time of kimberlite eruption, which, in our case, was well after the magmatic events. If the mantle lithosphere was indeed removed, then there should be no mantle xenoliths with dates prior to the magmatic event in question. In the case of diamond formation ages from silicate or sulfide diamond inclusions, diamond formation should only occur 100–300 My after the hypothesized delamination event, which is the time necessary to regrow the lithosphere by simple conductive cooling, and thus return the lithospheric mantle to the diamond stability field.

Data collected in southern Africa during the Kaapvaal project permits such a test. For two of the magmatic events, the Bushveld (2.06 Ga) and Ventersdorp (2.71 Ga), there are kimberlite pipes that are near or in the source zone of the basaltic intrusions. For the Bushveld, there are two studied kimberlites on the northern and southern edge of the intrusion, Klipspringer and Premier, respectively, and one, Palmietgat, which is within the Bushveld. The most extensive data set is from the Premier mantle xenolith suite. A study of Re–Os minimum depletion ages, \( T_{RD} \) ([41], Fig. 3) reveals a range of ages, most of which in fact predate the Bushveld. The available xenoliths with well-constrained pressures do not reveal any younging with depth as might be expected from delamination. Indeed the deepest xenolith (7 GPa) is the oldest, with a \( T_{RD} \) of 3.5 Ga. For the Klipspringer pipe, sulfide diamond inclusions yield Re–Os model ages of 2.55 Ga [42], significantly earlier than the Bushveld intrusion. Finally, the Palmietgat pipe has thus far yielded 4 sulfide inclusion diamond ages of 3.0 Ga [43]. Regarding the Ventersdorp, there are numerous 2.9–3.2 Ga diamond formation ages for the nearby mines at Kimberley [44,45], which, as in the case of the Bushveld, predate the Ventersdorp event.

From the evidence available from this extensive xenolith data set, we conclude that the occurrence of delamination is highly unlikely. Indeed, it is very difficult to find any evidence for delamination from the xenolith record of any cratonic environment worldwide. The only clear evidence for delamination is for the North China Craton [46]. Delamination has also been invoked for the Wyoming craton [47] although more recent data in fact argue for lithospheric preservation [48]. We thus conclude that the presently accepted mechanisms for flood basalts—mantle plumes and lithospheric thinning—are incapable of accounting for the southern African Precambrian magmatic events discussed above.

3. A new model for cratonic flood basalts

We propose an alternative model for cratonic flood basalts that is compatible with a collisional-rifting origin, leaves the cratonic lithosphere intact, and satisfies the short 1 My time scale for these events. The basic point of departure from White and McKenzie [1] is that a flood basalt is not a melting event, but instead constitutes the drainage of a pre-formed sublithospheric molten basaltic
We hypothesize that a reservoir of molten basalt is formed at the base of the lithosphere and is maintained over extended periods of time (long compared to 1 My). The presence of such a reservoir requires two conditions: that the sublithospheric mantle is often above the solidus, and equally important, that the lithospheric mantle and underlying molten reservoir are in equilibrium. This second consequence, in turn, requires a contrast in both solidus temperature and mechanical strength between lithosphere and underlying asthenosphere. There is indeed good evidence for such a reservoir of molten basalt. The million-year time scale is then a reservoir-exhaustion time scale, rather than a melt-production time scale. The magmatic event is the result of the abrupt, stress-induced creation of lithospheric dikes. That such lithospheric dikes can form near the base of the lithosphere is dramatically demonstrated by kimberlite eruptions that can originate from depths in excess of 200 km. The initial stress perturbation can have many causes, although collision-induced rifting is probably the dominant mechanism for rifts that do not lead to continental breakup. We thus propose a two-stage model that decouples the process of magma formation from magma transport to the Earth’s surface. This decoupling has several desirable features. First, the formation and maintenance of a magma reservoir can occur over periods of time much longer than 1 My, which significantly broadens the class of allowable formation mechanisms. Second, this model leaves the lithosphere intact since narrow lithospheric dikes are unlikely to pervasively disrupt the lithosphere.

3.1. Stage I: formation/maintenance

One interpretation is that high temperature melts have infiltrated the lithosphere, effectively advecting heat into a reservoir of molten basalt is formed at the base of the lithosphere and is maintained over extended periods of time (long compared to 1 My). The presence of such a reservoir requires two conditions: that the sublithospheric mantle is above the solidus, and equally important, that the lithospheric mantle and underlying molten reservoir are in equilibrium. This second consequence, in turn, requires a contrast in both solidus temperature and mechanical strength between lithosphere and underlying asthenosphere. There is indeed good evidence for such a contrast beneath ancient cratons, since the lithosphere appears to be both melt and volatile depleted from previous basalt-extraction events. This contrast would be expected to take the form of a transition zone, rather than a sharp boundary. Since the melt-plus-depleted peridotite is approximately equal in composition to the fertile peridotite, we would expect the melt to coexist with the depleted peridotite at the bottom of the lithosphere up to the depth where the fertile-mantle solidus crosses the lithospheric geotherm. This could lead to a transition zone thickness of about 20 km at the base of the lithosphere, although the formation of a “decompaction” zone might reduce this thickness. As an example of this transition for a dry mantle, we have used the MELTS algorithm to calculate the melting behavior of both a fertile peridotite (MM3) and a depleted peridotite (DDM) representing the asthenosphere and lithosphere respectively. For a base-of-lithosphere depth of 180 km (6 GPa) and for a potential temperature of 1750 °C the fertile asthenosphere yields a >10% melt fraction while the depleted lithosphere does not melt. If the lithospheric base is at 240 km (8 GPa), then a potential temperature of 1850 °C leads to the same contrast in melt behavior. Clearly, any volatile enrichment of the asthenosphere would dramatically reduce the solidus and therefore the asthenospheric temperature required for melting. Indeed, it is well established that water addition can dramatically reduce the peridotite solidus at high pressures by more than 500 °C. Regarding the contrast in mechanical strength between lithosphere and asthenosphere, Sleep has estimated that the lithosphere must be at least 20 times more viscous than the underlying mantle in order for cratonic lithosphere to avoid entrainment over billions of years. In fact, the actual viscosity contrast of cratonic lithosphere may be much higher. One study of the effect of devolatilization suggests that this contrast can be as high as 500, apparently more than enough to stabilize the lithosphere.

Probably the most significant requirement of this molten-reservoir model is that the subcratonic mantle often remains above the solidus, and that, as a result, there are long-lived pools of melt at sublithospheric depths. This raises two questions: (i) Is there any evidence for such high base-of-lithosphere temperatures in the present or the past, and (ii) if so, what is the possible source of this heat? Some information on recent cratonic geotherms is provided by the geothermobarometry applied to mantle xenoliths from the large number of Cretaceous kimberlites that erupted in southern Africa during the breakup of Gondwanaland. A recent compilation by James et al. based on low-temperature, on-craton nodules reveals a roughly linear gradient from 0 °C at the surface to about 1200 °C at a depth of about 180 km (Fig. 4). A major source of uncertainty, however, is the depth to the base of the chemically distinct lithosphere. If it extends to 250 km, a number that is compatible with recent seismic tomography beneath southern Africa, this linear gradient implies a base-of-lithosphere temperature of about 1550 °C. The high-temperature xenoliths, however, suggest an inflected rather than linear geotherm, with a steeper thermal gradient near the base of the craton. There have been a variety of interpretations of this inflected geotherm that are permitted by the data. One interpretation is that high temperature melts have infiltrated the lithosphere, effectively advecting heat into the base of the lithosphere. Since these melts had to have come from greater depth, it implies that the sublithospheric mantle is at least partially molten despite
the apparently subsolidus temperature for dry fertile mantle at the top of the asthenosphere. A second interpretation is that this inflection represents a disequilibrium conductive geotherm resulting from a thermal perturbation applied to the base of the lithosphere. With this interpretation, the linear gradient implied by these high-temperature xenoliths can be extrapolated to 250 km depth, giving a base-of-lithosphere temperature of 1850 °C (Fig. 4). We note that this extrapolated temperature is near the fertile-mantle dry solidus at base-of-lithosphere depth. A third interpretation of this inflected geotherm is that it reflects the apparently strong temperature dependence of thermal conductivity [60], which would predict an inflected equilibrium geotherm. The nearly factor-of-two increase in thermal gradient suggested by the high-temperature xenoliths is broadly consistent with the predicted temperature dependence of this parameter. We note that all three interpretations would suggest the possible presence of melt at sublithospheric depths.

Given the evidence for higher sublithospheric mantle temperatures in the Archean, it is possible to have had an equilibrium geotherm with a semi-permanent supersolidus asthenosphere at that time. Estimates of the difference in temperature between the Archean and the present, based on the chemistry of komatiites, range from 100° to 500 °C, depending on the assumed komatiite source region ([61] and references therein). As noted above, volatiles could dramatically depress the asthenosphere solidus. Indeed, the collision that created the collisional rift presumably signified the termination of an extended phase of nearby subduction, which might be responsible for increasing the volatile content of the asthenosphere in adjacent regions. Given the range of permissible time scales for reservoir formation and maintenance, there are several phenomena that could significantly perturb the temperature at the base of the lithosphere. For example, at the long-term end, thermal blanketing [62] can increase base-of-lithosphere temperature by hundreds of degrees in about 300 My. Given the large thermal inertia of thick lithosphere, this could still lead to a perturbed geotherm similar to the ‘kink’ in Fig. 4, based on a simple analysis of one-dimensional heat flow. Similarly, the process of advective thickening of the lithosphere, thought to be the way in which thick cratons are initially formed [17,49], should produce a decrease in heat loss through the lithosphere, and a corresponding increase in the temperature of the sublithospheric mantle, assuming the heat entering the sublithospheric mantle from below is unchanged. There may be large-scale thermal anomalies in the deep mantle that ultimately reach the base of the lithosphere, such as those inferred to be associated with lower mantle low velocity anomalies beneath southern Africa [63]. Finally, there might be a continual rain of ‘failed’ plumes that deposit their heat at the base of the lithosphere because they are unable to penetrate through cratonic lithosphere.

3.2. Stage II: drainage

The actual eruption of basalt onto the surface is interpreted as the drainage of the basaltic reservoir established in Stage I, through the creation of lithospheric dikes. For the two-stage model to be viable, the overlying lithosphere must be effectively impermeable to this reservoir of magma sitting below it (Fig. 1). While it is true that island arc basaltic magmas have very short residence times, as determined by uranium series disequilibrium observations (e.g. [64]), these regions are always tectonically active and have very thin lithosphere directly beneath the location of the arc. This environment differs markedly from a stable cratonic environment, characterized by very thick lithosphere and the virtual absence of tectonic activity except for participation in occasional collisions. We assume that lithospheric dikes only form in response to perturbations in
lithospheric stress. The million-year eruption time for flood basalts is then interpreted as the time necessary to drain this sublithospheric reservoir. Drainage time can, in turn, be separated into the porous-flow migration of magma to the dike source, followed by the ascent of magma through the dike. It is likely that the latter process is geologically instantaneous (Rubin, personal communication), so that the rate limiting step is the magma migration to the dike. The problem is analogous to melt generation at spreading centers, where it appears necessary for the magma to be focused toward the spreading center by the formation of a porous, melt-rich decompaction zone of accumulating magma between a partial melt zone below and impermeable ‘cap’ above [65,66]. Given the $10^6$ km$^3$ volume of flood basalts, and assuming this represents a 10% melt of a peridotite source, then the partial melt region (melt plus residue) is $10^7$ km$^3$. Assuming the reservoir is in the shape of a disk with thickness ranging from 100 to 10 km, corresponding to a range of disk radii of 178–564 km, the required melt migration velocity to completely drain the reservoir would be 18–56 cm/yr. This number is consistent with estimates for melt migration near spreading ridges, which are at least 100 cm/yr [67], although one must keep in mind that the drainage of a sublithospheric reservoir is primarily horizontal so that the available gradient in magma pressure would be less than in the spreading-ridge case.

Finally, in the absence of continental breakup, these lithospheric dikes are unlikely to lead to the large-scale thermal disruption of the lithosphere because they are highly localized. As discussed by Rubin [66], through-going lithospheric dikes need only reach a critical width of a few meters for sublithospheric magma to travel to the surface. Below this width, the walls of the dike freeze out the magma, thus blocking the propagation of magma through the fracture. This critical thickness is small enough to ensure that the thermal perturbation to the lithosphere from the passage of basaltic magma is indeed sufficiently localized to have a negligible effect on the overall strength and stability of the lithospheric mantle. Even if the dike were continually active over 1 My, the thermal perturbation would conduct only a few kilometers into the host rock, and thus not lead to a larger-scale disruption of the lithosphere.

The model does not, however, require that every dike fully traverse the 250 km lithosphere, draining basalt directly from the base of the lithosphere to the surface. The major and trace element composition of lavas from the Ventersdorp, for example, require some amount of both low-pressure fractional crystallization and crustal assimilation, especially for the lavas that erupted later in the volcanic episode [9], suggesting that these basalts temporarily resided in a crustal magma chamber prior to eruption. The residence time of magma in such shallow crustal reservoirs must be short, however, in order to allow the eruption of primitive basaltic lava (i.e. in equilibrium with mantle olivine at Fo$_{90}$, as are the oldest Ventersdorp lavas), and indeed to allow eruption at all (the Bushveld is an example of a magma chamber that ultimately crystallized in place within the crust).

4. Consequences and tests of the model

For a model to be useful it should be testable. To that end, we consider several possible tests for the hypothesis presented, based on its observational consequences. First, the two-stage model predicts deep melting, at the base of cratonic lithosphere. Melt chemistry is a function of both temperature and pressure, and this sensitivity has been exploited to invert for the minimum and maximum depth of melting using both trace and major elements. For example [1] used the relative abundances of the heavy rare-earth elements (HREE) in erupted lavas to determine whether melting occurred in the garnet stability field, corresponding to about 90 km depth. Beyond this constraint, however, HREE modeling is less reliable, due to the present lack of partitioning and melting constraints for relevant temperatures, pressures, and compositions. This limitation is illustrated by the estimated minimum melting depth for the Ventersdorp flood basalts of $\sim 30$ km by [1], which is in direct conflict with the xenolith record, as discussed earlier. Major element inversions [68,69], based on FeO and Na$_2$O concentrations, appear to hold greater promise. FeO constrains the maximum depth of melting, while Na$_2$O constrains the total melt fraction. Combining these two constraints provides estimates of melt-column depth and potential temperature. Application to Basin and Range basalts [69] yields minimum depth estimates that are consistent with independent estimates of the depth to the base of the lithosphere. While this promising methodology could be used in cratonic environments, at present, it has not been calibrated by laboratory experiments above 4 GPa, which is a necessary step in evaluating the hypothesis for melts originating at pressures that are nearly double this value. We nevertheless note that the Ventersdorp basalts have higher concentrations of FeO than average mid-ocean ridge basalts (at 8 wt.% MgO; ave. FeO$_{Vend}$=11 wt.% [9], ave. FeO$_{MORB}$=9.1 wt.% [70]), which is consistent with a deeper origin. A more stringent test will require systematic melt experiments at higher pressure, in order to extend this approach to appropriate pressures. A related issue is the
need for better experimental constraints on the solidus of dry fertile peridotite at high pressure. We note, for example, that a recent experiment at 10 GPa found the solidus to be at approximately 1725 °C, which would imply a solidus at 250 km depth that is about 100 °C lower than shown in Fig. 4 [71]. Finally, this two-stage model predicts long residence times of perhaps 10–100 my, so that a constraint on residence time for erupted basalts would potentially be at test. While such a test presently exists for very short timescales (based on uranium-series disequilibrium ~10,000–100,000 yrs), a test for longer residence times is presently not available, although there may be ways to attack this problem in the future.

Another test is provided by the 250 Ma Siberian cratonic flood basalt, which is arguably the largest of all known flood basalts. This event thus provides another opportunity to determine whether the mantle portion of cratonic lithosphere has, or has not, been disrupted and removed by such a major outpouring of basalt. As in the case of southern Africa, there are kimberlites near the flood-basalt locations. The eruption times of these Yakutian kimberlites [72], ~360, 230 and 150 Ma, both predate and postdate the flood basalt. Thus, comparison of depletion ages between the pre- and post-flood-basalt mantle xenoliths would provide a means of determining if the craton survived the flood basalt process, as in southern Africa.

As a final test, if sublithospheric melt concentrations are still present today, then they can be imaged by seismology. While the Earth is certainly cooler today than in the 1.8–2.7 Ga time period that is the focus of this study, cratonic flood basalts (unrelated to breakup) have continued to occur, the best known being the Neo-Proterozoic (1.1 Ga) Mid-Continent Rift in North America, and the Siberian flood basalt noted above. This latter event is particularly relevant to the issue of whether such subcratonic melt reservoirs are present today, since the thermal state of the Earth at 250 Ma is virtually the same as the present, suggesting that such melt zones are likely present today. If so, they should be detected seismically as thin low-velocity zones beneath the lithosphere.

The detection of low velocity zones is a notoriously difficult problem in seismology, since downgoing waves would not bottom in the zone. A promising approach that is well suited for detecting such low velocity zones is the receiver function method [73,74], which focuses on the analysis of arrivals that have been reflected and/or converted by the top and/or bottom of such a structure. A recent modification of the receiver function method, which makes use of S waves that convert to P (traditional receiver functions use P waves that convert to S) is especially promising. These waves are particularly useful in the top 250 km, where P-receiver functions typically have large artifacts associated with Moho reverberations [75]. We note that in the last decade, there have been a variety of studies, based on reflection/conversion-based methodologies that provide evidence for low seismic velocity zones between the base of the lithosphere and the 410-km discontinuity, and suggest the presence of melt in various parts of the world: China (e.g. [76]), Siberia, southern Africa [77], the Arabian plate [78], western North America [79,80], and even globally [81]. Recently, beneath eastern North America, a high-impedance-contrast boundary has been observed which is interpreted as partial melt sitting just below the lithosphere [74]. These reports are suggestive but not yet definitive. Indeed, a global search for such sublithospheric low velocity zones would constitute a thorough test of the model we have proposed.

Acknowledgements

We thank A. Rubin and an anonymous reviewer for constructive reviews and K. Burke, R. Rudnick, A. Rubin, C.T. Lee, S. Shirey, R. Carlson, D. James, S. Gao, A. Levander, and M. Fouch, for helpful discussions. The southern African seismic observations discussed in this manuscript were made possible by the great efforts of the Kaapvaal Seismic Group, with special thanks to R. Kuehnel. Thanks to A. Clements for help with manuscript preparation. This work, as well as the new seismic and geochemical observations, were supported by the National Science Foundation (EAR9526840), the Carnegie Institution of Washington, and several southern African institutions.

References


