Thermal and Magmatic Evolution of the Moon

Charles K. Shearer¹, Paul C. Hess², Mark A. Wieczorek³, Matt E. Pritchard⁴, E. Mark Parmentier², Lars E. Borg¹, John Longhi⁵, Linda T. Elkins-Tanton⁶, Clive R. Neal⁶, Irene Antonenko⁷, Robin M. Canup⁸, Alex N. Halliday⁹, Tim L. Grove¹⁰, Bradford H. Hager¹⁰, D-C. Lee¹¹, Uwe Wiechert¹²

¹Inst. of Meteorites, University of New Mexico, Albuquerque, New Mexico, U.S.A.
²Dept. of Geol. Sci., Brown University, Providence, Rhode Island, U.S.A.
³Institut de Physique du Globe de Paris, Paris, France
⁴Dept. of Earth & Atmospheric Sci., Cornell University, Ithaca, New York, U.S.A.
⁵Lamont-Doherty Earth Observatory, Palisades, New York, U.S.A.
⁶Dept. of Civil Eng. & Geol. Sci., Univ. of Notre Dame, Notre Dame, Indiana, U.S.A.
⁷University of Toronto, Toronto, ON, Canada
⁸Dept. of Space Studies, Southwest Research Institute, Boulder, Colorado, U.S.A.
⁹Dept. of Earth Sciences, University of Oxford, Oxford, United Kingdom
¹⁰Dept. of Earth, Atmospheric, & Planetary Sci., MIT, Cambridge, Massachusetts, U.S.A.
¹¹Academica Sinica, Institute of Earth Sciences, Taipei, Taiwan
¹²Eidgenossische Technische Hochschule, Zurich, Switzerland

Corresponding author e-mail: Charles K. Shearer <cshearer@unm.edu>

1. INTRODUCTION

As with all science, our continually developing concepts of lunar evolution are firmly tied to both new types of observations and the integration of these observations to the known pool of data. This process invigorates the intellectual foundation on which old models are tested and new concepts are built. Just as the application of new observational tools to lunar science in 1610 (Galileo’s telescope) and 1840 (photography) yielded breakthroughs concerning the true nature of the lunar surface, the computational and technological advances highlighted by the Apollo and post-Apollo missions and associated scientific investigations provided a new view of the thermal and magmatic evolution of the Moon.

1.1. Pre-Apollo view of the thermal and magmatic evolution of the Moon

Many of the early views of the Moon manifested in mythology and art throughout the world were primarily tied to lunar and terrestrial cycles and the relationships between the Sun and the Moon. Prophetically, myths involving the lunar deities Mwuetsi from Zimbabwe and Coyolxauhqui from Mexico told of rather violent or catastrophic events in which the Moon was expunged from the Earth. Numerous ancient scientific observations were made about the nature of the Moon ranging from those uncovered in early Neolithic sites that correctly identified mare Crisium and mare Humorum to the insights made by Greek philosophers such as Anaxagoras (ca. 500-428 B.C.) and Democritus (ca. 460-370 B.C.), who attached terrestrial analogues to its character (stone, mountains).

With the advent of the telescope (1610) and photography (1840) as scientific tools for lunar exploration, semiquantitative data could be collected that would provide an intellectual foundation for scientific interpretation. Initially, modern terrestrial geological analogs were extended to the Moon (lunar highlands, volcanic craters, seas). Combined with the rigor of
computational modeling, these observational data were extended to predict the original thermal state of the Moon and its thermal and magmatic history. Its proximity to the Earth made the Moon a prime candidate for the source of a wide range of meteorites (chondrites, eucrites) (Urey 1962, 1965; Duke and Silver 1967) and potential extraterrestrial materials (tektites) (Verbeek 1897). Numerous models were made for the ejection of material from the lunar surface (Arnold 1965). The possibility that the Moon was the source for these materials erroneously added “lunar sample” mineralogical and geochemical observations to the pre-Apollo computational models for its thermal and magmatic history.

1.1.1. Pre-Apollo view of the initial thermal state of the Moon. The conditions and processes under which the Moon formed had profound implications for its initial thermal state and its subsequent thermal and magmatic evolution. Pre-Apollo models for its origin fall within three groups (1) accretion or condensation along with the Earth as a double planetary system, (2) fission from a rapidly rotating Earth, and (3) capture by the earth of a fully formed body that was assembled elsewhere in the solar system.

The co-accretion of the Earth and Moon had been a fairly popular model for the origin of this double planet system. In its simplest form it suffered from two major problems. It failed to explain the orbital relationship between the two bodies and their contrasting densities. More elaborate models appealed to fractionation during accretion or slightly different accretional environments to account for the differences in density. In the 19th century it was anticipated that planetary bodies were formed by the accretion of incandescent matter. Therefore, the Moon and the Earth were once molten and differentiated during the crystallization of these molten oceans (Thomson 1864). In the co-accretion model, hot or cold accretion of the Moon dictated its initial thermal state.

In the fission model initially proposed by Darwin (1879), the Moon separated from the Earth by solar tidal forces and the Pacific Ocean basin was the resulting scar. The rationale for this model was that it accounted for the density differences between the Earth and the Moon because it implied that this event occurred following core formation and involved material derived from the upper mantle and crust of the Earth. The similarity of the composition of tektites to the terrestrial crust and the erroneous conclusion that were derived from the Moon (Verbeek 1897; Nininger 1943, 1947; O’Keefe 1963) added inaccurate geochemical evidence for this model. The fission model was initially criticized because it did not account for the orbital dynamics of the Earth-Moon system and because tides of the magnitude required to extract the Moon were thought to be mechanically impossible. The thermal consequences of the fission model for the Moon were not widely explored. The incorporation of a terrestrial crustal component enriched in U, Th, and K into the Moon would have resulted in melting of the lunar interior over a substantial period of time due to the release of radioactive heat. Although abandoned by many prior to Apollo missions, the primordial extraction of the Moon from the Earth by other mechanisms proved to be attractive following the Apollo missions.

The perceived similarities of moons in the solar system, the uniqueness of the Earth-Moon system, and the dynamics of the Moon’s orbit around the Earth were the philosophical basis for the capture model. Capture of the Moon by the Earth was long advocated by Urey (1952, 1957, 1959) and the theoretical basis was provided by Gerstenkorn (1955), MacDonald (1964), and Goldreich (1966). The capture model in its pre-Apollo incarnations implies that a relatively undifferentiated and cool Moon was produced elsewhere in the solar system with a primitive composition similar to chondritic meteorites. The Moon as a source for chondritic meteorites was advocated by many (Urey 1959, 1962, 1965; Arnold 1965; Öpik 1966) until the return of the first Apollo mission to the Moon. Most of the capture models suggested that the Moon was captured during the very early history of the Earth. However, in a pamphlet that was privately printed in 1908, F.B. Taylor suggested that the Moon was captured as late as the Cretaceous. Fitting within this model, Urey (1952) calculated that the initial interior of the
Moon was less than 600 °C and more likely was 300 °C. Consequently, magmatic evolution of the Moon primarily involved surface melting caused by impacts and basin formation.

1.1.2. Pre-Apollo view of the thermal and magmatic evolution of the Moon. The first observations related to the thermal and magmatic evolution of the Moon were made by Renaissance scientists in the 17th century. Drawing analogies to crater-producing volcanic processes on Earth, these early observers understandably attached a volcanic origin to the lunar craters. Dana (1846) bolstered these earlier views of a volcanic origin for lunar craters. Gilbert (1893) and Baldwin (1949) challenged this view by concluding that the large lunar craters were of impact origin. The debate continued until the Apollo missions.

Besides the debate on the origin of lunar craters, the interpretation of the early magmatic history of the Moon hinged upon the nature of the lunar maria and highlands. Originally considered seas by the first lunar explorers during the Renaissance, 20th century observers prior to Apollo speculated that they were asphalt lakes (Wilson 1962), dust (Gold 1955), sedimentary rocks (Gilvarry 1968), impact derived melts (Urey 1952), and flood basalts (Baldwin 1949; Kuiper 1954; Fielder 1963). Each of these origins implies a distinctive thermal and magmatic history for the Moon. Proposed ages for the mare ranged from 4 billion years to tens of millions of years (Baldwin 1949; Hartmann 1965; Gault 1970). Drawing on the initial terrestrial analogy made by Galileo in 1610 for the lunar terra, the brighter reflectivity and the presumed lunar origin for tektites, most observers prior to Apollo equated the lunar highlands to terrestrial continental masses. Although their composition and origin were far less debated than the maria, the highlands were suspected to be largely volcanic or to represent more sialic rocks such as granites and rhyolites.

Gilbert (1893) predicted that the Moon formed cold and remained cold throughout its history. The mathematical problem of the cooling of a sphere radioactively heated was solved by Lowan (1933) and first applied to the thermal history of the Moon in the 1950s (Urey 1952, 1955) and MacDonald (1959). There was increasingly a perception (Runcorn 1962; Tozer 1967; Schubert et al. 1969) that solid-state convection has been important in controlling the thermal history of the Moon and other planetary bodies. Urey’s early work on the thermal history of the Moon suggested the possibility that internal temperature achieved by radioactive heating could be sufficient to melt at least a part of the lunar interior. Urey (1960) remained skeptical of the extent of lunar melting predicted by this model and anticipated that the Moon accreted cold and perhaps the interior experienced only moderate heating. The melt produced in the lunar interior through radioactive decay was predicted to behave as a highly viscous liquid. Its convection in the lunar interior was modeled by Kopal (1961) and Runcorn (1962, 1963). Prior to the Apollo missions, the basic assumptions underlying Runcorn’s approach, namely that the Moon is internally hot enough for a large part of its mass to behave as a molten globe, was challenged by several observations. A rigid crust and a mostly solid interior were implied by relief of the lunar surface and a significant free libration of the Moon in longitude (Koziel 1964).

Looking back from the vantage point of one of mankind’s genuine scientific adventures, many of these observations and models for the Moon appear naive and ambiguous. The status of understanding the thermal and magmatic history of the Moon is perhaps best depicted in a lament made by Cook (1967) only two years prior to Apollo and quoted by Taylor (1975): “No conclusions can be drawn other than that the interpretation of the existing experimental data leads to many ambiguities.”

1.2. Apollo view of the thermal and magmatic evolution of the Moon

The Apollo missions and associated science had a profound impact on our understanding of the thermal and magmatic history of the Moon. A detailed analysis of the view of the Moon immediately following the Apollo missions is presented by Taylor (1975). Critical to this immense intellectual step forward was the collection of lunar samples and their integration
into both Earth-, orbital-, and surface-based observations made prior to and during the Apollo missions. Prior vanguard missions (Luna, Ranger, and Surveyor) provided invaluable guidance for planning the Apollo missions, but the remotely sensed data were mostly ambiguous at answering important questions concerning the thermal history of the Moon.

The observational data derived from the Apollo missions resolved the nature and age of many of the basic lithologies and landforms making up the Moon that added critical constraints in interpreting the magmatic and thermal evolution of the Moon. Some of the fundamental observations made by Apollo-related science were:

1. Large impacts produced all of the large lunar craters and a variety of impact lithologies.
2. Lunar magmatism was essentially basaltic in nature.
3. The maria represent impact basins that were flooded with fluid basalts, produced by melting of the lunar mantle and a majority were erupted to the lunar surface in the interval between 3.1-3.9 billion years ago. The mare basalts exhibit a wide range in composition and are depleted in europium compared to other rare earth elements.
4. The terra are complex, ancient (>3.9 billion years old) and dominated by plagioclase-rich lithologies such as ferroan anorthosite that is enriched in refractory elements. Several other plutonic lithologies were also identified. The trace element characteristics of the primary lunar crust are complementary to the mare basalts and their mantle sources.
5. The Moon is depleted in volatile and siderophile elements and enriched in refractory elements compared to the Earth, yet oxygen isotopes indicate that the Earth and Moon have an ancestral relationship.
6. The Moon is highly differentiated and therefore is not the source for primitive meteorites such as chondrites. This initial differentiation event occurred soon after lunar accretion and was completed by \( \approx 4.4 \) Ga.
7. The coefficient of moment of inertia and seismic data are consistent with nearly uniform density with depth and a small lunar core (diameter <700 km). Conductivity, heat-flow, and tectonic observational data added additional constraints for the internal temperature and the bulk composition of the Moon.

The synthesis of these observational data led to numerous fundamental models for the thermal and magmatic evolution of the Moon immediately following Apollo. The Moon also provided a window to the early thermal evolution of terrestrial planetary bodies that is inaccessible due to eradication of early planetary surfaces by tectonics or by the remoteness of other planetary surfaces. Early Apollo-era results led to the realization that immediately following accretion the Moon underwent extensive, Moon-wide melting. The resulting expression of this event was eloquently referred to as the lunar magma ocean (LMO). Models involving a LMO held that the early lunar crust of ferroan anorthosite was produced by plagioclase flotation and that the mantle source for the mare basalts was a product of the sinking and accumulation of more dense mafic silicates. Early crystallization sequences were calculated for the LMO from olivine-dominated early cumulates to ilmenite-rich late stage cumulates. The final dregs of this Moon-wide melting event produced a widespread, uniformly distributed, incompatible-element-rich and isotopically uniform material referred to as KREEP (Potassium, Rare Earth Elements, Phosphorous). Crystallization of the LMO was prior to 4.4 Ga. and models for the depth of the LMO ranged from 100 km to whole-Moon melting. Models that advocate melting of only the outer portion of the Moon imply that primitive lunar mantle lies below the cumulate pile. The extent of early lunar melting was tied to the potential sources for the primordial heat. The composition of a given mare basalt reflected the composition of the cumulate horizon that was melted. Therefore, it was expected that basalt
compositions reflected depth of melting. Estimates for the depth of origin for the mare basalts ranged from 100 to 400 km. The mantle sources for the mare basalts were volatile-poor and most were thought to be very reduced. Even so, pyroclastic deposits found at the Apollo 15 and 17 landing sites were interpreted to be products of volatile driven fire-fountaining. The existence of basaltic volcanism older than the basin-filling basalts was speculative. Most of the highland basaltic rocks such as the Fra Mauro basalts were concluded to have been produced by impacts. The magmatic rocks in the lunar highlands were characterized into several distinct plutonic suites. It was concluded that these plutonic rocks were not partial melts from the same sources as the mare basalts. Whether all the highland plutonic rocks were products of the same primordial lunar differentiation or a broader range of lunar processes (LMO, post-LMO magmatism, impacts) was a point of debate.

On the basis of observations from the Apollo missions, numerous thermal history models for the Moon appeared in the literature prior to 1976. The LMO concept was prominent in many of these early models. However, most avoided treating this episode of lunar differentiation in detail. The early history of the Moon immediately following initial differentiation was dominated by an internal thermal environment that resulted in significant magmatic and tectonic activity. Thermal models both with and without solid convection were suggested and were equally able to satisfy many of the observations made during the Apollo missions. Many of the models concluded that convection played an important role in lunar differentiation, dynamics, and the introduction of lateral heterogeneity. There was reasonable agreement that the range of present-day temperatures in the lunar interior is rather low to a depth of 500 km and rises to about 1000 °C at 700 km depth. Below this depth, there was less agreement in the models, with estimates for temperatures at the center of the Moon ranging between 1000-1600 °C. Some models predicted a partially molten lower mantle and partially molten core. Most of these thermal models were dependent upon the mechanism (primary accretional zoning or early differentiation) for the distribution of heat-producing elements in the outer part of the Moon and the existence of an undifferentiated lunar mantle.

The observations and the resulting science related to the Apollo missions established a robust observational and intellectual framework to provide context for future lunar exploration. Since the high-water mark of the 1970s, new inroads have been made owing to new analytical-computational-technological advances, new missions during the last decade of the twentieth century, and a retrospective view of the pioneering science attributed directly to Apollo. In this chapter, we present new insights into the thermal and magmatic evolution of the Moon afforded us by these new approaches and views.

2. LUNAR ORIGIN AND FORMATION

The origin of the Moon is of fundamental importance in providing a context for lunar and terrestrial studies in a variety of fields, and has received much attention, particularly over the past 50 years. Numerous theories of lunar origin have been advanced, and have been well summarized in a book (Origin of the Moon 1985) and several review papers (Boss 1985; Wood 1986; Stevenson 1987). Almost 20 years ago, at a Conference in Kona, Hawaii, (that led to the publication of the Origin of the Moon book), a single hypothesis of lunar origin came to the forefront: the giant impact (e.g., Brush 1996). In the intervening time, scenarios that might lead from a giant impact to the Moon as we known it have been explored and summarized in another volume, Origin of the Earth and Moon (2000). A variety of physical properties of the Moon must be explained, and in this section, we summarize how consistent the giant impact model is with the real Moon. We first review why the giant impact model is currently the favored hypothesis, and then explore how well it explains observed lunar characteristics, such as early lunar dynamics, composition, initial thermal state, and timing of formation.
2.1. Mechanisms of lunar formation

There are four general types of lunar origin models: fission, capture, co-accretion, and giant impact. Characteristics of these models, as well as their strengths and weaknesses, are covered by Boss (1985), Wood (1986), and Stevenson (1987). The fission model postulates that the early Earth was spinning so rapidly that it became rotationally unstable, causing material in the equatorial regions to be flung into orbit to yield the Moon. In the capture model, the Moon formed elsewhere in the solar system, but it eventually suffered a close encounter with the Earth and was able to enter a bound and stable circum-terrestrial orbit. During Earth accretion, a circum-terrestrial disk might have formed through a variety of mechanisms, which could have led to the formation of the Moon. The giant impact model supposes that such a disk formed from the ejecta of a giant collision of a projectile and the Earth, while the co-accretion model supposes that the circum-terrestrial disk was present throughout Earth’s accretion, and was built up of smaller heliocentric projectiles.

The key properties of the Earth/Moon system that serve to separate the models from each other are the ability of each to explain the large angular momentum of the Earth-Moon system and the observation that the Moon is depleted in iron. There is no single observation that rules out the capture model, but there is only a very narrow range of orbital parameters that allow capture to occur (Wood 1986). The dynamical problem with the capture model can be summarized as follows: a dissipative process near the Earth (e.g., extended atmosphere or accretion disk) must operate to allow the object to be captured into Earth orbit, but such a process must disappear soon after the object is in orbit so that it will not continue to lose orbital energy and crash into the Earth. Compositionally, it is also difficult to explain why a captured Moon would not have a bulk iron content similar to the terrestrial planets. The fundamental objection to the fission and co-accretion models is the same; they cannot readily explain the angular momentum of the Earth-Moon system. The fission model requires much more angular momentum than the system currently has, and there is no viable mechanism to remove a sufficient amount of angular momentum. For example, the amount of angular momentum lost to solar tides is likely small (e.g., Canup et al. 2001).

It is difficult, but not impossible, for co-accretion to produce a sufficiently high angular momentum Earth-Moon system. Thus, on the basis simply of the angular momentum, co-accretion must be considered a viable alternative to the giant impact (Weidenschilling et al. 1986; Morishima and Watanabe 2001). Another challenge for the co-accretion model is to explain how the Moon could form side-by-side with the Earth but be deficient in iron. Weidenschilling et al. (1986) proposed that the iron deficiency can be created during co-accretion if the circum-terrestrial disk acts as a compositional filter, where mechanically weak silicates break up and are retained in the disk, while strong iron objects pass through the disk and do not become accreted to the proto-Moon. Recent work has strengthened the co-accretion model by showing that the proto-Earth could have gravitationally altered the orbits of heliocentric impactors such that their trajectories would become spatially non-uniform and the right amount of angular momentum could be deposited into the Earth-Moon system if accretion occurred in the presence of nebular gas (Ohtsuki and Ida 1998). The recent co-accretion model of Morishima and Watanabe (2001) makes use of the spatially non-uniform distribution of heliocentric impactors to get the right angular momentum, but requires the co-accretion to happen before dissipation of the solar nebula or early in Earth accretion. As the solar nebula probably dissipated within the first 10 m.y., it is unlikely that such an early formation is consistent with the proposed formation age of the Moon within the first 50 m.y. (see below). Furthermore, co-accretion during the early formative stages of Earth requires a delicate balance between tidal evolution and mass addition to prevent the Moon from crashing into the Earth. The required balance between tidal evolution and mass addition is physically unlikely (Sleep and Fujita 1998).
The Moon’s lack of a large iron core together with planet accretion models that predict that large impacts would be common in the final stages of terrestrial accretion led Hartmann and Davis (1975) to propose that the Moon formed from iron-depleted mantle material ejected from the Earth during a large-scale impact. They also hypothesized that material subject to such conditions might be depleted in volatile elements relative to the Earth. An independent and contemporaneous investigation by Cameron and Ward (1976) recognized that the oblique impact of a roughly Mars-sized planet could account for the rapid initial terrestrial rotation rate implied by the current angular momentum of the Earth-Moon system, and suggested that vaporization might provide a physical mechanism to emplace material into bound orbit. The concepts described in these two works form the basis for what is now the leading theory for lunar origin, the giant impact hypothesis.

A viable model for lunar formation must create a Moon that is consistent with myriad observations, from studies of lunar composition provided by the Apollo and Luna missions, to detailed properties of the lunar orbit. Before exploring how well models of the giant impact can explain these observations, a definition is in order. What makes a giant impact different than a common large impact? Non-giant impacts would have ejected material on ballistic trajectories that would re-impact the proto-Earth, or on hyperbolic ones that would leave the system altogether, but not on trajectories that would place the ejecta in stable orbits that would allow them to accrete into the Moon. A giant impact is not only different in size than other impacts, it involves fundamentally different physics that allows material to enter orbit through impact jets, vapor clouds, gravitational torques, or gravitational interactions among the ejected debris (Stevenson 1987). One might ask whether a giant impact is likely? Were there vagrant protoplanets roaming the early solar system big enough to cause such a collision? The answer appears to be yes. Many models of accretion of the terrestrial planets indicate that there could be objects of the necessary size formed near the Earth (e.g., Wetherill 1986; Agnor et al. 1999; also the review by Canup and Agnor 2000). The results of calculations used to determine the size distribution of the protoplanets in the terrestrial neighborhood must, however, be considered cautiously because they neglect the effects of fragmentation (Stewart 2000) and still struggle to reproduce the low inclination and eccentricity orbits of Earth and Venus (e.g., Chambers and Wetherill 1998; Agnor et al. 1999).

Accepting that a giant impact is possible, other questions arise: can a giant impact eject enough mass into Earth orbit such that it can accrete into a single Moon with the observed properties? With recent computational advances, it has become possible to more fully explore the parameter space of plausible impact scenarios and determine which are most consistent with the Moon. Dynamical models of lunar formation, discussed in the next section, are important because they not only address whether a giant impact could lead to formation of the Moon, but also provide key information for geochemical models, such as the likely mass of the impactor and proto-Earth at the time of the giant impact, the relative amount of material from each that goes into forming the Moon, and the thermal state of the proto-Earth and ejecta following the impact.

2.2. Dynamical models

2.2.1. Overview. Beginning in the mid-1980’s (e.g., Benz et al. 1986, 1987; Melosh and Kipp 1989), numerical hydrocode simulations of planet-scale collisions sought to determine 1) whether an impact could eject sufficient iron-depleted material into a bound orbit to form the Moon, and 2) what specific type of impact would be required. Although early simulations were hindered by coarse resolutions and very long computational run times (see also Chapter 5), results generally supported the impact hypothesis (Benz et al. 1986, 1987, 1989; Cameron and Benz 1991). Calculations relating to the subsequent evolution of a protolunar disk composed of solids suggested that the disk could spread extremely rapidly via disk particle collisions, with a characteristic time scale of only about a month (Ward and Cameron 1978). More detailed models
that considered the thermodynamic behavior of the disk suggested that disk material would be a mixture of fluid and vapor, whose overall evolution time would be regulated by the radiative cooling time of the disk, ~ 100 years (Stevenson 1987; Thompson and Stevenson 1988).

In the mid-1990’s, models describing the accumulation of impact-ejected material were developed and provided new constraints on the distribution of orbiting material required to yield a lunar-size moon (Canup and Esposito 1996; Ida et al. 1997; Kokubo et al. 2001). In particular, N-body accretion models predicted the rapid accretion of a single moon at a typical radial distance of 3-4 Earth radii ($R_\oplus$) and revealed a relationship between the angular momentum and mass of the protolunar disk and the size of the resulting moon (Ida et al. 1997). In cases where multiple moons accumulated initially, analytic and numerical models suggested that further evolution of the system due to tidal interaction with the Earth and satellite-satellite interactions would still yield a single moon system, as the moons collided with each other, or as inner moons collided with the Earth (Canup et al. 1999).

Recent works have attempted to identify specific impact scenarios that can account for the current Earth-Moon system, in essence attempting to link the predictions of impact simulations with those of the lunar accumulation models (Cameron 1997, 2000, 2001; Cameron and Canup 1998; Canup and Asphaug 2001). A common difficulty has been producing a massive enough protolunar debris disk with an impact that leaves the Earth-Moon system with its final mass and angular momentum. An inability to produce simultaneously the main dynamic features of the Earth and Moon with a single impact led to proposals that multiple impacts might have been required, or that the Moon-forming impact occurred when Earth was only about half-accreted (Cameron and Canup 1998; Cameron 2000, 2001). While both of these possibilities are not inconsistent with the character of impacts predicted during the final stages of terrestrial accretion (Agnor et al. 1999), they are more restrictive and problematic than the original single-impact hypothesis. However, recent work (Canup and Asphaug 2001) suggested that such restrictions may be unnecessary, as high-resolution simulations predict that an impactor containing about 10% of the Earth’s mass appears capable of both producing the Moon and an Earth-Moon system with approximately the current mass and angular momentum.

Below we discuss recent impact simulation results; for more detail, the reader is referred to the chapters by Cameron and Kokubo, and Canup and Ida in the “Origin of the Earth and Moon” (2000) volume.

2.2.2. Constraints on the Moon-forming impact. The angular momentum contained in both the lunar orbit and the Earth’s spin has been nearly conserved throughout the history of the Earth-Moon system, with some loss due to tidal interaction with the sun. Tidal interaction between the Earth and Moon causes the lunar orbit to expand and the Earth’s rotation to slow, implying that the Moon initially formed very close to a more rapidly rotating Earth. A minimum orbital distance for lunar formation would be the Earth’s Roche limit ($a_{\text{Roche}} = 2.9 R_\oplus$ for lunar density material); significant accretion interior to this distance would be frustrated by planetary tides (e.g., Canup and Esposito 1995).

The current angular momentum of the Earth-Moon system, $L_{E-M} = 3.5 \times 10^{41}$ g-cm$^2$/s, is quite large, and implies that the initial length of the terrestrial day was only about 4-5 hours when the Moon was located near the Roche limit. A lunar mass satellite ($M_1 = 7.349 \times 10^{25}$ g or $0.012 M_\oplus$) orbiting on a circular orbit with $a = a_{\text{Roche}}$ would contain an orbital angular momentum of $M_1 (GM_\oplus a_{\text{Roche}})^{1/2} = 0.18 L_{E-M}$, providing a lower limit on both the mass and angular momentum of orbiting material necessary to yield the Moon assuming completely efficient accumulation.

In reality, much of the orbiting material initially within $a_{\text{Roche}}$ will re-impact the Earth in the course of angular momentum exchange within the protolunar disk and through interaction with the forming Moon; some may also be ejected from the system entirely. Lunar accumulation
simulations (Ida et al. 1997; Kokubo et al. 2001; also review by Kokubo et al. 2000) predicted a characteristic initial satellite orbital radius $\sim 1.2a_{\text{Roche}}$, or $a \sim 3.5R_\odot$. A basic conservation argument can be made to estimate the mass of a satellite, $M_s$, forming with $a = 1.2a_{\text{Roche}}$ from an initial protolunar disk containing mass $M_D$ and angular momentum $L_D$ (Ida et al. 1997):

$$M_s = \frac{1.9L_D}{\sqrt{GM_Da_{\text{Roche}}}} - 1.1M_D - 1.9M_{\text{esc}}$$

(4.1)

where $M_{\text{esc}}$ is the amount of escaping material during accretion.

In the simplest impact scenario, the Moon-forming impact would produce an Earth-Moon system consistent with the current one, requiring no significant subsequent dynamic modification. In this case, a candidate impact should place material containing sufficient mass and angular momentum to yield the Moon, typically at least 1.5-2 lunar masses, and produce an Earth with approximately its final mass and an Earth-Moon system with somewhat more than its current angular momentum in order to account for later loss due to interactions with the sun. In addition, the ejected material must be appropriately iron-depleted to account for the Moon’s low bulk density. The exact fraction of iron contained in the Moon is uncertain. The upper limit on the mass fraction of the lunar core is estimated to be $\sim 3\%$ (e.g., Hood and Zuber 2000); additional iron is incorporated in silicates in the lunar mantle and crust (e.g., Lucey et al. 1995). From the bulk lunar density, an upper limit on the total percent by mass of iron in the Moon is estimated to be about 8% in the limit that all of the iron is contained in low-density silicates (Wood 1986).

2.2.3. Numerical impact simulation method. The general approach in performing numerical impact experiments has been to consider four primary impact variables: impact parameter, $b$, [defined here as $b = \sin(\xi)$, where $\xi$ is the impact angle], impact velocity, $v_{\text{imp}}$, [whose minimum value is the mutual escape velocity of the colliding objects, $v_{\text{esc}} = \{2G(M_I + M_{\text{Tar}})/(R_I + R_{\text{Tar}})\}^{1/2}$], the mass ratio of the impactor to the total mass, $\gamma$, and the total colliding mass, $M_T$. A series of simulations are examined to determine what set of impact conditions yields the most favorable results; however, the possible parameter space is large and multi-dimensional, and individual impact simulations are computationally intensive.

Most works modeling potential Moon-forming impact have utilized a method known as smooth particle hydrodynamics, or SPH (e.g., Lucy 1977; Benz 1990). The SPH technique requires no underlying grid and is well suited to intensely deforming systems evolving within mostly empty space. Moreover, its Lagrangian nature allows for tracking of material history during a simulation. In SPH, an object is represented as a large number of spherical overlapping ‘particles’ whose evolution is tracked as a function of time. Each particle represents a quantity of mass of a given composition, whose 3-dimensional spatial extent is specified by a probability density function known as the kernel, and the characteristic spatial width of the particle, known as the smoothing length. The functional form of the kernel does not change during a simulation, but the smoothing length of each particle is adjusted to maintain an overlap with a minimum number of other particles, thus insuring that even low-density regions are smoothly resolved. For modeling planet-scale impacts (i.e., those involving objects 1000’s of kilometers in radius or larger), the evolution of each particle’s kinematic (position and velocity) and state (internal energy, density) variables are evolved due to 1) gravity, 2) compressional heating and expansional cooling, and 3) shock dissipation. The chosen form for the equation of state relates a particle’s specific internal energy and local density to pressure as a function of specified material constants. Commonly used equations of state range from simple analytic forms (e.g., Tillotson 1962) to complex semi-analytic methods (e.g., ANEOS; Thompson and Lauson 1972).

A critical element in the accuracy of SPH is numerical resolution. In early works, simulations with $N = 3000$ particles were done in which a lunar mass (and thus the typical
amount of ejected material) was represented by only a few tens of SPH particles. Computational advances now allow for much greater resolution (Cameron 2000, 2001; Canup and Asphaug 2001), and recent simulations have $N = 10^4$-$10^5$. With $N > 10,000$ particles, comparisons (Cameron 2000; Canup and Asphaug 2001) indicate that predictions for the amount of orbiting material are consistent to within about 10-15% across simulations with varying resolution. However, resolving the amount of orbiting iron remains a challenge, as modern resolutions describe a lunar mass using hundreds to thousands of SPH particles and the upper limit on the amount of lunar iron with typically several tens of particles.

2.2.4. Simulating lunar-forming impacts. SPH simulations performed in the 1990s by Cameron were reviewed in detail in Cameron (2000, 2001). Since the work of Cameron and Benz (1991), increasingly larger impactors relative to the targets were considered in an effort to increase the yield of material placed into bound orbit, with Cameron (2000, 2001) considering collisions that all involved impactors containing 30% of the total colliding mass, or $\gamma = 0.3$.

A prevailing trait of these simulations has been an apparent difficulty in placing sufficient mass into orbit to yield the Moon for a total system mass and impact angular momentum close to that of the current Earth-Moon system (e.g., Cameron 1997, 2000, 2001; Cameron and Canup 1998). Two classes of impacts, both with $\gamma = 0.3$, were identified as being capable of producing the Moon. The first involved an impact by an object with about $3 \times$ the mass of Mars, with an impact angular momentum, $L$, much greater than $L_{E-M}$, typically by a factor of 2. A significant dynamic event—such as another giant impact—would then be required to decrease the Earth-Moon system angular momentum by this large amount subsequent to the Moon-forming event. The second class of impacts involved an impact of an object with roughly twice the mass of Mars with $L \sim L_{E-M}$, but with a total mass (impactor plus target) of only $M_T \sim 0.65M_{\oplus}$. In this “early-Earth” scenario, the Earth is only partially accreted when the Moon forms and must subsequently gain $\sim 0.35M_{\oplus}$, with the later growth involving sufficiently small and numerous impacts so that the system angular momentum is not significantly altered.

Both of these impact scenarios are more restrictive than the original single impact-hypothesis. A specific problem with the early-Earth scenario is the potential for the Moon to become contaminated by iron-rich material during the period when the Earth was accumulating the final ~35% of its mass (e.g., Stewart 2000). If, during this period, the Moon also accumulated even an approximately proportionate amount of material, it would have gained excessive amounts of iron. The ratio of impact rate onto the Moon vs. that with the Earth is $N_I = f_M R_M^2/R_{\oplus}^2$, where $(f_M/f_{\oplus})$ is the ratio of the gravitational focusing factor of the Moon to that of the Earth, and $R_M$ and $R_{\oplus}$ are lunar and terrestrial radii. Reasonable values of impactor velocity yield $0.03 < N_I < 0.074$ (Stewart 2000); assuming an impacting population with a terrestrial abundance of iron implies that the Earth could accrete $0.10M_{\oplus}$ before impacts with the Moon delivered an amount of iron equal to the upper limit on the mass of lunar iron. There are several factors that might serve to mitigate the accretion of iron by the Moon, including a less than perfect accretion efficiency (e.g., Morishima and Watanabe 2001). However, in general, as the amount of material which must be added to the Earth after the Moon-forming impact in a given impact scenario increases, difficulties with the Moon becoming compositionally more similar to the other terrestrial planets during the period of subsequent terrestrial growth increase. The simplest explanation for the Moon’s unusual compositional characteristics is that it is the result of an impact that occurred near the very end of terrestrial accretion.

Work has been ongoing to identify other impact scenarios that can more closely produce the Earth-Moon system. An important step in this regard has been the realization that the results of SPH simulations display common trends when viewed in terms of scaled quantities. Canup et al. (2001) re-examined the results of Cameron’s (2000) simulations and found that the ratio of the mass of the orbiting material to the total colliding mass, $(M_D/M_T)$ is a function of the impact angular momentum ($L$) scaled by a quantity $L^*$. Here $L^*$ is the critical angular
momentum for rotational stability of a spherical body with mass $M_T$. The quantity $(L/L^*)$ for a $v_{imp} = v_{esc}$ impact is given by

$$\frac{L}{L^*} = \frac{\sqrt{3}}{K} \gamma (1 - \gamma) \sqrt{(1 - \gamma)^{1/3} + (1 - \gamma)^{1/3} \sin \xi}$$

where $K$ is the gyration constant (2/5 for a uniform sphere), $\gamma$ is the impactor-to-total mass ratio, and $\xi$ is impact angle. The protolunar disk masses and angular momenta produced by SPH impact simulations performed by Cameron (2000) with $\gamma = 0.3$ were found to be well approximated by power-laws of the form

$$\frac{M}{M_L} = C_M \left[ \frac{L}{L^*} \right]^{s_M} ; \quad \frac{L}{L^*} = C_L \left[ \frac{L}{L^*} \right]^{s_L}$$

with $C_M = 0.054$, $s_M = 3.9$, and $C_L = 0.38$, $s_L = 4.4$. For $\gamma = 0.3$, the maximum yield of orbiting material resulted for an impact angular momentum about 70-80% of that of a grazing impact, independent of the total colliding mass $M_T$.

2.2.5. Recent simulations involving smaller impactors. Using the scaling analysis in Canup et al. (2001), Canup and Asphaug (2001) estimated what other impactor sizes might be better able to produce the Earth-Moon system. They predicted that the maximum yield for an $L \approx L_{E-M}$ and $M_T \approx M_\oplus$ impact should be achieved when $b \approx L_{E-M} L_{graz} \approx 0.8$, with the angular momentum of a grazing impact, $L_{graz}$, given by (again assuming $v_{imp} = v_{esc}$):

$$L_{graz} = [\frac{3}{(4\pi)}]^3 \sqrt{\frac{2G}{\gamma} M_T^{5/3}}$$

where $\rho$ is the average target/impactor density, and $f(\gamma) = \gamma (1 - \gamma)[(1 - \gamma)^{1/3} + (1 - \gamma)^{1/3}]^{1/2}$. Impactors with $\gamma < 0.12$ had been ruled out as lunar-forming candidates in early low-resolution studies (Benz et al. 1987), because they appeared to produce overly iron-rich disks. However, for those $N = 3000$-particle simulations, a single iron particle contained a mass comparable to the upper limit on the lunar core.

In a survey of 36 simulations, Canup and Asphaug (2001) found a variety of successful candidate impacts involving impactor-to-target mass ratios $\gamma \sim 0.1$, or an impactor mass of $\sim 6 \times 10^{26}$ g. This is essentially equivalent to the mass of Mars. Such impacts are attractive in that they appear to require little or no dynamical modification of the system after the Moon-forming impact, thus avoiding problems associated with a period of extended terrestrial growth after lunar formation. Nominally, a smaller impactor than those proposed by Cameron (2000, 2001) would also be more likely, because in collisional populations small objects are more common than large ones (the number of objects, $dN$, in a mass range $dm$ is typically proportional to $dN \propto m^{-q} dm$, with $q \sim 1.5-1.8$). The Mars-mass impactor proposed as optimal by Canup and Asphaug (2001) is essentially the same as that originally proposed a decade prior to the first Moon-forming impact simulations by Cameron and Ward (1976).

2.2.6. Discussion and open issues. Recent work simulating potential Moon-forming impacts indicate that a wide variety of oblique, low-velocity impacts are capable of placing material into bound orbit where it would then accumulate into a satellite(s). Two specific impact scenarios have recently been promoted: one involves the impact of an object with about 2× the mass of Mars with an Earth than is only about 65% formed after the impact (Cameron 2000, 2001), while the other invokes the impact of a Mars-mass object with a fully formed Earth (Canup and Asphaug 2001). The former involves a very massive impactor relative to the target protoearth, with the impactor containing 30% of the total mass; the latter involves a somewhat smaller impactor containing about 12% of the total mass. Although not yet investigated with simulation, we predict that there will be a class of impacts intermediate to these two cases involving progressively decreasing impactor-to-total mass ratios combined with increasing
total colliding masses, which would also yield a lunar-sized Moon and an Earth-Moon system with an angular momentum close to $L_{E-M}$. The range of these impacts that could be viable Moon-forming candidates would then depend upon on how much the Earth could accrete in the presence of the Moon without the Moon becoming contaminated with iron or volatile-rich material. Other types of impacts with $v_{\text{imp}} > v_{\text{esc}}$ may also be promising candidates.

To date, impact simulations predict that the majority of the material placed in orbit originates from the impactor, with contributions from the target proto-Earth of up to tens of percent by mass (e.g., Canup and Asphaug 2001). Recent work suggests that this may be compatible with the lunar and terrestrial siderophile-element profiles (Righter 2002), although the identical O-isotope signatures in the Moon and Earth (see Section 4.2.5) then require the impactor and the target to have very similar compositions.

The choice of equation of state (EOS) on impact simulation outcome has an important effect and has yet to be fully assessed for potential Moon-forming impacts. Both early simulations (Benz et al. 1986, 1987) and Canup and Asphaug (2001) utilized the Tillotson EOS (Tillotson 1962), while those of Benz et al. (1989), Cameron and Benz (1991), and Cameron (1997, 2000, 2001) used ANEOS (Thompson and Lauson 1972). Tillotson is a simple analytical EOS that lacks an actual treatment of phase changes and mixed phase states. Unlike Tillotson, ANEOS handles mixed phases in a thermodynamically consistent manner; however, in its standard rendition, ANEOS treats all vapor as monatomic species (e.g., Melosh and Pierazzo 1997). The entropy required for vaporization of molecular species such as mantle material is therefore greatly overestimated by this feature of ANEOS, which may be responsible for the apparent lack of vapor production in the simulations of Cameron (e.g., 2000, 2001). An extension to ANEOS to allow for vaporization of molecular species was undertaken by Melosh (2000), and simulations have now been done with this EOS (Canup et al. 2002). Open questions include the relative role of gas pressure gradients vs. gravitational torquing in placing material into bound orbit, and the predicted initial physical state (solids-melts-vapor) of the ejected material.

2.3. Lunar accretion time scales and the early lunar orbit

Assuming that a giant impact occurred and ejected sufficient mass into Earth orbit, the question then becomes whether the material would accrete into a single Moon that is consistent with observed physical and orbital properties. Recent numerical simulations that track the evolution of a swarm of $10^3$-$10^4$ moonlets in a circum-terrestrial disk using direct N-body simulation indicate that a single Moon would form under a variety of initial conditions on a timescale of order 1 month to 1 year (Ida et al. 1997; Kobuko et al. 2000, 2001). Lunar accretion seems to be somewhat inefficient (only 10-55% of the ejected material goes into forming the Moon) so an important conclusion from these simulations is that more than 1 lunar mass must be ejected into orbit from the giant impact (Canup and Esposito 1996; Kobuko et al. 2000, 2001). Many different initial conditions (size and radial distribution of particles, total mass in the disk, etc.) are put into the N-body simulations because the exact initial disk conditions are unknown. There is uncertainty both because the distribution of ejecta in the circum-terrestrial disk from the giant impact remains an active area of research (see Section 2.2) and because the amount of time the disk evolves between the giant impact and the onset of accretion is not well constrained. It is unlikely that accretion could begin immediately following the giant impact because the ejected material is mostly melt and vapor, and significant cooling would be required before accretion could occur. In fact, N-body simulations assume that all particles in the circum-terrestrial disk are solid and behave like infinitely strong particles during collisions (e.g., no fragmentation). Thompson and Stevenson (1988) estimated that 10-100 years might elapse between the giant impact and when the material is cool enough to accrete, but coupled thermal and dynamic models will be necessary to constrain the exact time (e.g., Stewart 2000; Kokubo et al. 2000). Using a hybrid N-body/liquid disk model, Canup and Ward (2000) suggested that accretion exterior to the Roche limit could occur rapidly, while the Roche interior disk remains...
Shearer et al. Thermal & Magmatic Evolution of the Moon

hot and evolves on the longer, 10-100 year time scale. Thus one important unanswered question about the giant impact is that of timing. How soon after the giant impact did accretion begin and over what time interval did the Moon form? These issues are important for understanding whether the Moon was ever completely molten, with major implications for the lunar thermal evolution and composition (see Sections 2.4 and 2.6).

Simulations suggest that the Moon would accrete near the Earth’s Roche limit, at around 3-4 Earth radii, and would subsequently tidally evolve outward to its present location, where it is still moving slowly away from the Earth at about 3 cm/yr (Dickey et al. 1994). However, lunar accretion simulations also predict that the initial inclination of the lunar orbit relative to the Earth’s equator would be of order 1 degree, while integrations of the lunar orbit back in time indicate that this initial inclination was at least 10 degrees (Goldreich 1966; Touma and Wisdom 1994). There are currently two different proposed explanations for how the Moon’s orbit could have become significantly inclined to the Earth’s equator. (1) If a remnant of the protolunar disk continued to exist interior to the Roche limit for some time after the accretion of a Moon outside \( a_{\text{Roche}} \), resonant interaction between the Moon and this disk through the 3:1 inner vertical resonance could increase lunar orbital inclination to the required value (Ward and Canup 2000). A bending mode of the disk is excited, which coherently perturbs the lunar orbit at each midplane crossing, progressively increasing the tilt of the lunar orbit. This mechanism would occur during the earliest evolution of the Moon’s orbit, within a few hundred years of the Moon’s accretion. Its effectiveness is dependent upon the Roche interior disk containing at least \(~0.5\) lunar masses and having a characteristic viscous lifetime \( \geq 50-100 \) years; for such values, a small initial inclination would be increased to \(~10-15\) degrees. The recoil of the Moon destroys the disk by causing it to decay and be accreted by the Earth. (2) If a Roche interior disk did not interact with the early Moon, later resonant interactions of the Moon with the Sun might have been able to increase the lunar inclination to appropriately high values instead (Touma and Wisdom 1998; Touma 2000). Assuming the protolunar disk has been removed, the Moon’s passage through a secular resonance with the Sun, known as the evection resonance (where the apsidal precession rate of the Moon equals the orbital period of the Earth, 1 year), would occur at about 4.6 Earth radii (Kaula and Yoder 1976; Touma and Wisdom 1998). Capture into this resonance would lead to an increase in the lunar eccentricity, which if followed by capture in a second inclination-eccentricity resonance (termed the eviction resonance by Touma and Wisdom) can produce the needed inclination. This mechanism’s effectiveness depends upon having the tidal dissipation “Q” factor of the Moon change between passage through evection and eviction, such that an early low rate of dissipation allows for eccentricity growth at the evection resonance, whereas a later enhanced rate of dissipation allows for capture into the eviction resonance. Such a change could occur if an initially cold Moon was tidally heated during the eccentricity excitation phase (Touma and Wisdom 1998).

In summary, the difficulty of reconciling the large initial lunar inclination with many lunar origin models, including the giant impact model, appears to have been mitigated by plausible dynamical scenarios that could increase the inclination of an initially equatorial Moon to \(~10\) degrees. The next section explores whether other observations about the Moon (particularly from geochemistry and geology) can constrain whether the Moon formed hot or cold, and might be used to infer the timescale between the giant impact and accretion.

2.4. The initial lunar thermal state

A giant impact is a major thermal event that could have consequences for numerous properties of the Moon that we observe today. Recent N-body simulations of lunar formation (Section 2.3) predict that once accretion began, it proceeded very rapidly, leaving little time to lose heat of accretion. If accretion occurred less than 100 years after the giant impact, it is likely that the Moon formed almost completely molten (Pritchard and Stevenson 2000). Uncertainties in the physical state of the ejecta from a giant impact (Section 2.2) and the
interval between the giant impact and accretion (Section 2.3) make it difficult to state precisely the range of initial lunar temperatures. Therefore in this section we consider the extreme case: do any observations rule out a completely molten Moon? For more complete analysis see Pritchard and Stevenson (2000).

The most widely cited observational constraint on the initial temperature profile within the Moon is the radius constraint, which relates surface observations of faulting on the Moon to internal thermal evolution (MacDonald 1960; Solomon and Chaiken 1976). If the Moon formed almost completely molten and cooled monotonically throughout geologic history, then the lunar radius would contract, building up stress in the rigid outer layers of the Moon through the relation (Solomon 1986),

\[
\frac{\Delta R}{R} = \frac{(1 - \nu)\sigma}{E}
\]

where \(\Delta R/R\) is the fractional change in radius, \(\nu\) is Poisson’s ratio, \(E\) is Young’s modulus, and \(\sigma\) is the horizontal stress in the lithosphere. If the stress exceeds a certain level, faulting would occur. The fact that there are no global scale contractional or extensional features on the Moon constrains the amount of radius change (and therefore temperature change) within the Moon since the creation of the stable lunar surface at 3.5 Ga, after the end of heavy bombardment. By comparison, global thrust faults on Mercury have been used as evidence of global contraction (e.g., Strom et al. 1975). That no such features are seen on the Moon indicates that there has been a delicate balance between heating of the lunar interior and cooling of the exterior throughout lunar history, such that the radius has remained nearly constant (Solomon and Chaiken 1976; Solomon 1977, 1986; Kirk and Stevenson 1989). The implication for lunar origin scenarios is that the interior of the Moon must have started cold, and certainly not molten. However, there are several scenarios that complicate the relation between internal thermal evolution and surficial tectonics such that hotter conditions deep within the Moon would be possible (Pritchard and Stevenson 2000). First, as Equation (4.5) indicates, the Young’s modulus must be used to relate stress to strain. The outer layers of the Moon are highly cracked due to impact processes, as indicated by Apollo seismic experiments (e.g., Toksöz et al. 1974), and this type of material can withstand more strain before failing than solid rock (He and Ahrens 1994; Pritchard and Stevenson 2000). Mathematically, this means that by using a lower \(E\), a higher strain is allowed before the stress exceeds failure criteria. Another possible factor that could obscure the relation between initial temperature and faulting is if the lunar thermal state has not decreased monotonically. For example, a change in convective regime within the interior could lead to episodes when the interior heated up. Other factors that complicate the relation between initial temperatures and lunar faulting are mentioned in Pritchard and Stevenson (2000) and below. The conclusion of these works is that a hot initial Moon cannot be ruled out.

An initially molten Moon has geochemical consequences. One of the first and most unexpected results of the Apollo missions was that the outer several 100 km of the Moon were probably initially completely molten (see Section 4.3). The giant impact and subsequent rapid accretion of the Moon naturally leads to a LMO, whereas the slower accretion envisioned in co-accretion models may or may not (e.g., Wood 1986). Early proponents of the giant impact also claimed that vaporization during the giant impact could lead to the observed volatile depletion of the Moon (Wood 1986), and more recently some have noted that accretion of the Moon within 1 month could itself lead to elemental fractionation (Abe et al. 1999). Alternatively, the existence of volatiles in samples thought to be from deep within the lunar interior has been cited as evidence that the Moon was never totally molten (Dickenson et al. 1988). Another line of evidence against vaporization of proto-lunar material is the relative abundance of alkali elements. Although the Moon is depleted in absolute abundances of volatile alkalis, their relative abundances (e.g., Li/Na, Na/K, K/Rb, Rb/Cs) do not show a volatile-depletion pattern
(i.e., Li < Na < K < Rb < Cs) relative to the Earth (McDonough and Sun 1995; Jones and Drake 1993; Taylor 1982). Turcotte and Kellogg (1986) also claimed that the Moon could not have begun molten because of the inferred lack of mixing and vigorous convection in the lunar interior. Although the issues of lunar composition are discussed in more detail in Section 2.6, it is emphasized here that the geochemical data have multiple interpretations, and there is not necessarily a simple relation between a hot/molten state and a complete depletion in volatiles or homogenization. The existence of volatiles deep in the lunar interior could be consistent with a molten Moon if the process of volatile extraction was inefficient. Similarly, the existence of undepleted xenoliths from the terrestrial mantle (Jones and Palme 2000) cannot rule out a terrestrial magma ocean and the existence of a giant impact (Stewart 2000). Furthermore, the volatile depletion of the Moon might be inherited from the giant impactor and have little to do with the giant impact event itself (Humayan and Cassen 2000). It is also possible that the Moon could have begun nearly molten and be poorly mixed if the mantle heterogeneity was formed during differentiation, and if mantle convection was weak or non-existent.

Two sets of geophysical data could provide information on the initial thermal state of the Moon, but they require further investigation. First, is the possibility that a record still exists from the time when the Moon was close to the Earth and it rotated faster. The fundamental shape of the lunar surface and gravitational field (as recorded in the low order spherical harmonics of those fields) might be interpreted as a “fossil” of the earlier rotation state, which might be inconsistent with a totally molten Moon (Hood and Zuber 2000). The other set of geophysical data is paleomagnetism of the lunar samples (discussed in Chapter 3), which might indicate the presence of an early convecting lunar dynamo, although there are theoretical difficulties with this interpretation. If there really was a lunar dynamo, then core formation occurred early in the Moon, and this has implications for the early thermal state, but a totally molten Moon is not required (Section 6.3.9).

2.5. Age of the Moon

The age of the Moon represents more than just a point in time in early solar system history. The various theories for the origin of the Moon as well as those for the accretion of the terrestrial planets more generally are directly testable by the precise determination of the Moon’s age. Furthermore, without a properly tested theory for how the Moon formed we cannot evaluate the reasons for its chemical composition and what that composition reveals about the early stages of planetary development. The age of the Moon is now known extremely well; consequently, we can confirm certain theories and refute others. In this section the background to these issues is briefly explained and the state of understanding at the end of the Apollo missions is summarized. Then the powerful new constraints from hafnium-tungsten (Hf-W) cosmochemistry are discussed.

The earliest dynamic and chemical evolution of the terrestrial planets is now relatively well understood in terms that are consistent with both isotopic and chemical data as well as theoretical modeling. The process of runaway gravitational growth should build Moon-sized objects over periods of roughly 10^5 years (Lissauer 1987). Therefore if the Moon formed as such a planetary embryo there are obvious implications for its expected age. The importance of protracted accretionary processes involving collisions between such embryos were not widely appreciated until the 1970’s (Safronov 1974; Wetherill 1976, 1986), and the concept of the origin of the Moon as a result of a late collision between two planets came about in the 1980’s (Cameron and Benz 1991). The most serious alternate theory was capture. However, this theory makes no predictions that the Moon itself should form late. The Giant Impact theory predicts that the age of the Moon should post-date the origin of the solar system by some considerable time period; probably tens of millions of years if Wetherill’s calculations are correct.

Attempts to date the Moon were initially focused on dating the oldest rocks and therefore providing a lower limit. These studies mainly involved precise Sr, Nd and Pb isotopic constraints
(Tera et al. 1973; Wasserburg et al. 1977; Hanan and Tilton 1987; Carlson and Lugmair 1988; Shih et al. 1993; Alibert et al. 1994). At the end of the Apollo era, Wasserburg and colleagues wrote “The actual time of aggregation of the Moon is not precisely known, but the Moon existed as a planetary body at 4.45 Ga, based on mutually consistent Rb-Sr and U-Pb data. This is remarkably close to the $^{207}$Pb/$^{206}$Pb age of the Earth and suggests that the Moon and the Earth were formed or differentiated at the same time.” (Wasserburg et al. 1977). Such constraints on the age of the Moon still leave considerable scope (>100 m.y.) for an exact age.

Some of the most precise and reliable early ages for lunar rocks are given in Chapter 5 (Table 5.7). They provide considerable support for an age of >4.42 Ga. Perhaps the most compelling evidence comes from ferroan anorthosite 60025, which defines a relatively low first-stage $\mu$ and an age of about 4.5 Ga. Carlson and Lugmair (1988) reviewed all of the most precise and concordant data and concluded that the Moon had to have formed in the time interval 4.44-4.51 Ga, which is consistent with the estimate of 4.47 ± 0.02 Ga of Tera et al. (1973). The anomalous ages and initial Nd isotopic compositions of selected ferroan anorthosites is discussed in Section 3.4.

Model ages can provide upper and lower limits on the age of an object. Halliday and Porcelli (2001) reviewed the Sr isotope data for early solar system objects and showed that the initial Sr isotopic compositions of early lunar highlands samples are all slightly high relative to the best estimates of the solar system initial ratio (Fig. 4.1). Even the most conservative assessment of the Sr isotope data would conclude that the difference between the $^{87}$Sr/$^{86}$Sr of the bulk solar system initial at 4.566 Ga = 0.69891 ± 2 and the Moon at (roughly) 4.515 Ga = 0.69906 ± 2 is resolvable. A Rb-Sr model age for the Moon can be calculated by assuming that objects formed from material that separated from a solar nebula reservoir with the Moon’s current Rb/Sr ratio. As the Rb/Sr ratios of the lunar samples are extremely low, the uncertainty in formation age does not affect the calculated initial Sr isotopic composition, hence model age, significantly. The CI chondritic Rb/Sr ratio ($^{87}$Rb/$^{86}$Sr = 0.92) is assumed to represent the solar nebula. This model provides an upper limit on the formation age of the object, because the solar nebula is thought to represent the most extreme Rb/Sr reservoir in which the increase in Sr isotopic composition could have occurred. The Sr isotopic composition, however, probably evolved in a more complex manner over a longer time period. The calculated time required to generate the difference in Sr isotopic composition in a primitive solar nebula

![Figure 4.1](image.png)
environment is $11 \pm 3$ m.y. (Halliday and Porcelli 2001). Unless the Rb/Sr of the bulk solar system or solar nebula is not that of CI chondrites, this is the earliest point in time at which the Moon could have formed. This alone establishes that the Moon acquired its Rb/Sr ratio (or chemical composition) at a stage that is later than predicted from modeling the accretion dynamics of planetary embryos by runaway gravitational growth (Lissauer 1987).

A similar model-age approach can be used with Hf-W cosmochemistry. The W isotopic compositions of bulk rock lunar samples range from chondritic to $\varepsilon_W > 10$ (Lee et al. 1997, 2001). This was originally interpreted as the result of radioactive decay of formerly live $^{182}$Hf within the Moon, which has a variable but generally high Hf/W mantle (Lee et al. 1997; Halliday 2000). We now know that a major portion of the $^{182}$W excess in lunar samples is cosmogenic and the result of the reaction $^{181}$Ta($n,\gamma$) $\rightarrow^{182}$Ta($\beta^-$) $\rightarrow^{182}$W while these rocks were exposed on the surface of the Moon (Leya et al. 2000; Lee et al. 2002). This can be corrected using (1) estimates of the cosmic ray flux from Sm and Gd compositions, (2) the exposure age and Ta/W ratio or (3) internal isochrons of W isotopic composition against Ta/W (Lee et al. 2002). The best current estimates for the corrected compositions are shown in Figure 4.2. The spread in the data is reduced and the stated uncertainties are greater relative to the raw W isotopic compositions (Lee et al. 1997). Most data are within error of chondritic but a small excess $^{182}$W is still resolvable for some samples.

From these data, the most obvious and clear implication of the chondritic W isotopic compositions is that the Moon, a high Hf/W object, must have formed late. Otherwise the very high Hf/W would have resulted in a large excess in $^{182}$W produced from the relatively large amount of (still) live $^{182}$Hf. On this basis Halliday (2000) pointed out that the chondritic W isotopic composition was hard to explain if the Moon formed before about 50 Myrs after the start of the solar system. Subsequent to these studies it was shown that the present-day W isotopic composition of the average solar system, as measured in chondrites is offset to $\varepsilon_W = -1.9 \pm 0.1$ (Kleine et al. 2002; Schönberg et al. 2002; Yin et al. 2002). This change, as well as the corresponding change in the initial $^{182}$Hf/$^{180}$Hf of the solar system shortens the model timescales. Kleine et al. (2002) and Yin et al. (2002) argued that the Moon had to have formed within the first 30 Ma of the start of the solar system. More recent modeling of the magnitude of the radiogenic effect on the Moon indicates a somewhat younger age of 40 to 50 Ma (Halliday 2003, 2004).

Of course the radiogenic excesses found in some lunar samples could be either the result of inherited isotopic heterogeneity or radioactive decay of with formerly live $^{182}$Hf. High precision oxygen isotope data (Wiechert et al. 2001) provide no evidence of inherited heterogeneities. Although significant meteoritic material might have been incompletely admixed into the Moon after it formed, (perhaps equivalent to the late veneer on Earth) and early large impacts clearly struck the Moon, there is no indication of incompletely admixed meteoritic components in the lunar interior. This does not exclude addition of meteoritic material completely but limits the amount to a few percent. Admixing 3% H-chondritic material, for example, would be detectable when 0.016‰ deviation (in $\Delta^{17}$O) is considered significant (Fig. 4.3). Even less material from L, LL or carbonaceous chondrites would be detectable because these groups are further displaced from the terrestrial fractionation line (TFL). Although CI-chondrites overlap with the TFL, they characterized by high $\delta^{18}$O values above 16‰. Any proportion larger than 5% would increase the $\delta^{18}$O value by at least 0.5‰.

Furthermore, the oxygen data indicate that the Earth and Moon have compositions that are identical ($\Delta^{17}$O$_{\text{MOON}} = 0.000003 \pm 0.000005$, 99.7% confidence). Therefore, if the giant impact theory is correct, the proto-Earth and the impacting planet were constructed from an identical mix of inner-solar-system material and provenance. Only if the $\Delta^{17}$O values of the proto-Earth and impactor were identical to within 0.03‰ would it be possible that the average $\Delta^{17}$O value of the Moon plots within 0.005‰ on the TFL. This calculation is conservative, assuming the Moon is made of equal proportions from the proto-Earth and the impactor planet. Improved
models of the Moon-forming giant impact indicate that a large portion (i.e., 70-90%) derived from the impactor (Asphaug and Canup 2001). If the material from impactor and the proto-Earth were incompletely mixed, it would not be detectable in the oxygen isotopes. However, all giant impact models involve an extraordinarily energetic start to the Moon and preservation of pre-accretion heterogeneities is unlikely.

2.6. Compositional constraints on the origin of the Moon

Estimates for the composition of the Moon have been made in numerous studies (Anders 1977; Wänke et al. 1977; Taylor 1982; Ringwood 1986; Jones and Delano 1989; O’Neill 1991; summarized by Newsom 1995) and in Chapter 3 (Section 5.3). A comparison between the estimated bulk composition of the Earth’s mantle and the Moon indicates that there are both similarities and differences, but the differences are such that a lunar origin by fission can be ruled out because it predicts that the two should have identical compositions (Drake 1986). Estimating planetary bulk composition is difficult because samples of the terrestrial and lunar lower mantles are lacking. Furthermore, the compositions of the objects(s) that
would form a Moon through co-accretion, capture, or the giant impact are not known. A particular problem for the giant impact is that the exact proportion of material from the Earth and impactor that form the Moon is unknown, although recent numerical simulations of the giant impact indicate that 70-90% of the Moon would consist of material from the impact (see Section 2.2). Still, identical oxygen isotopic compositions of the Earth and Moon, and differences from other solar system materials provides a strong constraint (Wiechert et al. 2000). The similarity in the oxygen isotopes suggests that the material that formed the Moon came from nearly the same location within the solar nebula as Earth, and although consistent with current simulations of the giant impact (Wiechert et al. 2000), it is also consistent with other lunar formation scenarios such as co-accretion and capture.

2.7. Summary

We cannot yet answer the question of whether the Moon as we know it is completely consistent with formation by a giant impact into the early Earth, because realistic models of the giant impact, evolution of the protolunar disk, and lunar accretion are only beginning to be explored. However, during the past 20 years, sophisticated (albeit incomplete) numerical models have revealed that a giant impact can eject sufficient mass into Earth orbit to yield a Moon consistent with the observed orbital and compositional properties, providing strong support for the giant-impact hypothesis. These models predict that the Moon can accrete very quickly (within a year or so), and further work that couples thermal and dynamical evolution is needed to determine plausible initial thermal conditions for the Moon. The long-standing constraint on initial lunar temperatures from the lack of contractional faults on the Moon should be re-examined using a wider range of plausible lunar thermal histories and rheological conditions for the lunar lithosphere. In terms of geochemistry, some now argue that because a large contribution to lunar composition might come from the (mostly unknown) composition of the impactor, lunar formation from the giant impact might not be “testable” in extreme cases (Spudis 1996; Jones and Palme 2000). Along with more sophisticated and complete models of the giant impact, co-accretion models that could be consistent with the angular momentum constraint, bulk iron deficiency, and the timing of terrestrial accretion, should be explored. In the end, we may only be able to say that one scenario is more plausible than another, but that many possibilities are allowed.
3. INITIAL MAGMATIC DIFFERENTIATION

3.1. The magma ocean: rationale, origin, size and final outcome

3.1.1. The rationale for a lunar magma ocean. Ever since petrographic studies revealed anorthositic lithologies in Apollo 11 soils (Wood et al. 1971), researchers have appealed to buoyant segregation of relatively low-density plagioclase in an extensive, primordial melting event to explain the plagioclase-rich crust of the Moon (e.g., Solomon and Toksoz 1973; Taylor and Jakes 1974). That event, which came to be known as the lunar magma ocean, has received additional affirmation through the chemistry and petrology of mare basalts and geophysical modeling.

The abundance, distribution, characteristics, and apparent antiquity of the plagioclase-rich lunar crust remain the most compelling arguments for a globe-encircling magma ocean. Remote-sensing data from the Apollo missions for Fe, Mg, Al/Si, and Mg/Si (Adler and Trombka 1977) were interpreted as indicating that the highlands are at least as anorthositic as the soils and polymict breccias (75% plagioclase) from the Apollo 16 site (Warren 1985). Estimates for the bulk composition of the veneer of megaregolith on the surface of the Moon range from anorthositic gabbro (Spudis and Davis 1986) to ferroan-anorthositic norite (Spudis et al. 1999). The bulk composition of the deep lunar crust remains obscure, but material in large basin floors and materials ejected from large impact basins indicate that the lower crust is basaltic in composition. Estimates for its average composition range from norite to KREEP basalt. Clearly, the bulk composition of the lunar crust is not that of anorthosite (Al$_2$O$_3$ ≈ 35 wt%) as predicted by most LMO models, but approaches that of anorthositic norite (Al$_2$O$_3$ ≈ 25 wt%). This is not, however, an impediment to the LMO concept. The lunar crust does not represent simply a primordial lithology produced during early differentiation, but a composite planetary feature consisting of LMO-produced anorthosites, post-LMO igneous assemblages derived from melting of the lunar mantle and emplaced into the crust, and megaregolith mixtures of these petrogenetically distinct magmatic assemblages. Longhi (1982) estimated that the lunar crust (assuming a mean thickness of 60 km) contains the equivalent of at least 5-6 km of anorthosite. Wood (1986) calculated that a minimum thickness of 20 km of pure anorthite must reside in the lunar crust. Crustal models by Spudis et al. (1999) and Hawke et al. (2003) suggested that this primordial anorthosite crust resides primarily in the middle crust (15-35 km) and is exposed in some areas of the northern far side and within inner rings of multiring basins (see also Chapter 3).

The antiquity of the ancient lunar crustal rocks has been authenticated in numerous studies (reviewed by Carlson and Lugmair 1981; Nyquist 1982; Nyquist and Shih 1992; Snyder et al. 2000; Nyquist et al. 2002). Very old ages were postulated initially for the ferroan anorthosites because of their primitive $^{87}$Sr/$^{86}$Sr ratios. Unfortunately, ferroan anorthosites are extremely poor in Rb and REE, so no Rb-Sr or Sm-Nd internal isochron ages were measured prior to the late 1980s (Carlson and Lugmair 1988). The new results for the ferroan anorthosites (Carlson and Lugmair 1988; Alibert et al. 1994; Borg et al. 1997, 1999) confirm ancient ages, although their implied crystallization ages point to a much more complex picture of the initial stages of lunar differentiation or an early disturbance of the isotopic systematics. This will be discussed later in this section.

Other chemical attributes of the ferroan anorthositic crust provide further evidence for their origin as LMO flotation cumulates. Differences in mineral compositions, such as molar ratios of Ca/(Ca+Na) (or An content) in plagioclase and Mg/(Mg+Fe) in mafic minerals (Warner et al. 1976) and trace element signatures, such as Sr, Eu/Al, Eu/Sm, Sc/Sm and Ti/Sm, between the ferroan anorthosites and spatial associated mafic lithologies (Norman and Ryder 1979; Raedeke and McCallum 1980; Warren and Wasson 1980; James and Flohr 1983; Warren 1986) are not consistent with a complementary petrogenetic relationship. Thus, there appears to be an absence
of mafic cumulates complementary to the ferroan anorthosites in the lunar crust. This implies that the differentiation event that produced the early lunar crust was of a large scale and that the bulk of mafic cumulates associated with the ferroan anorthosites remain in the lunar mantle.

The notion of a complementary, plagioclase-depleted lunar interior did not gain wide acceptance until after studies of the Apollo 12 mare basalts. Initially, the nearly ubiquitous negative Eu-anomalies in rare-earth-element (REE) plots of mare basalt analyses were interpreted as indicating residual plagioclase in the basalt source region. Experimental work by Green et al. (1971a,b) demonstrated that plagioclase was absent from the liquidi of primitive mare basalts at both low and high pressure. The inability of plagioclase to crystallize in the early stages of low-pressure differentiation or to remain a residual phase in the mare basalt source region meant that the widespread signal of plagioclase fractionation in mare basalts was indicated that an earlier event depleted the mare-basalt source region of plagioclase (Helmke et al. 1972; Philpotts et al. 1972). In that the ferroan anorthosites have a substantial positive Eu anomaly, the simple inference is that the mare basalt source regions were the complementary mafic cumulates of the same magma that produced the ferroan-anorthositic crust.

As reviewed by Warren and Wasson (1979) and Warren (1985), the ratios of incompatible elements among lunar rocks rich in such elements virtually all conform to a uniform pattern throughout the sampled portion of the Moon. The key characteristic of this pattern is that it is high in concentration of incompatible elements such as K, Rare Earth Elements, P (KREEP), Th, U, Zr, Hf, and Nb, and exhibits a uniform fractionation among the incompatible elements. For example, it is uniformly enriched in the light REE relative to the heavy REE as indicated by precisely determined La/Lu and Sm/Nd elemental ratios and \(^{143}\text{Nd}/^{144}\text{Nd}\) isotopic ratios. The widespread uniformity and the extreme incompatible-element enrichment of the KREEP signature relative to bulk Moon can not be produced by small-scale fractional crystallization processes (Taylor 1975 1982; Warren and Wasson 1979; Warren 1985). Model ages for KREEP also suggest its origin during the early stages of lunar differentiation (Nyuquist and Shih 1992). This uniformity and enrichment was viewed as being produced from the residuum (urKREEP) of a single, global magma during primary lunar differentiation. Although recent remotely sensed data indicate that this geochemical signature is localized in the Procellarum-Imbrium region on the Earth-facing side of the Moon, it is still a major planetary signature that requires a large-scale planetary process for its origin. Warren and Wasson (1979) and Walker et al. (1979) cautioned that most rocks with a KREEP signature are too mafic in major-element composition and mineralogy to be simple, unaltered samples of LMO residuum. More likely, the KREEP signature was incorporated into the lunar crust through remobilization, assimilation, and mixing.

Numerous attempts have been made to demonstrate the former existence of a LMO, gauge its size, and predict its effects. Warren (1985) provided a thorough review and analysis of the problem. Here, we focus on the formation of the LMO and its effect on lunar thermal history, the extent of early differentiation as a gauge of its size and duration of crystallization, and the relation between ferroan anorthosites and the LMO.

3.1.2. Thermal constraints for a magma ocean. Although several heating mechanisms have been proposed (see summaries in Warren 1985, Shearer and Papike 1999), the only plausible way to melt a large portion of the Moon in a short time is by rapid (< 10\(^3\) years) accretion. So when Wetherill’s (1976, 1980) dynamic modeling of the early solar system showed that the planets accreted on a time scale of 10\(^2\) to 10\(^8\) years, serious doubts arose about the plausibility of the LMO. Several alternate hypotheses that did not involve an LMO were advanced to explain the Eu-anomaly in mare basalts (e.g., Walker 1983) and the abundance of ferroan anorthosite without cogenetic mafic rocks in the lunar crust (e.g., Longhi and Ashwal 1984). However, the development of the giant impact hypothesis alleviated many of the concerns about rapid heating. As discussed earlier, models of the dynamics of debris orbiting Earth after the giant impact predicts very short accretion times ranging from several months to
a year. Furthermore, collisional heating of particles within the disc might maintain the inner portion of the disc in a liquid to partially vaporized condition (Cameron and Benz 1991). Melting a substantial portion of the Moon no longer seems to be a serious obstacle.

Perhaps, the most insightful constraints on the size of the LMO came from the work of Solomon and Chaiken (1976), who modeled the change in planetary volume through time as a function of initial temperature distribution (Section 4.2). Their thermal modeling was consistent with the outer 200 km of the Moon being initially molten. Solomon (1977) revised the depth of the molten zone to 300 km. Kirk and Stevenson (1989) repeated the calculations and found that if one allowed for heating and melting of the part of the Moon below the LMO that the volume increase produced by partial melting through time could offset the contraction produced by solidification of a 400 km deep LMO. To provide sufficient expansion, the amount of melt required was ~100× the volume of the mare basalts. This magma need not have erupted or even intruded the crust; however, the production of this magma does coincide with the eruption of the mare basalts. So unless the Moon supported two coeval, but distinct melting systems throughout its history, it is reasonable to equate the melting predicted by Kirk and Stevenson (1989) with mare basalt magmatism. Head and Wilson (1992) calculated that mare basalt eruption geometries and rates implied a system of dikes and sills 50× the volume of the erupted mare basalts. Thus the 100:1 ratio of melt to erupted mare basalt envisioned by Kirk and Stevenson (1989) is certainly plausible. As shown below, however, the mare basalts could not have formed from melting the undifferentiated lunar interior. Moreover, there is no evidence of melts of primitive lunar material ever erupting as lavas. Thus it appears from thermal modeling that 400 km is an upper limit to the thickness of the LMO.

It may be that some additional thermal process, not incorporated in the thermal models described above, counteracted the contraction expected if the initial depth of melting exceeded 400 km. One such process is asymmetric heating of the crust. Haskin (1998) called attention to the enrichment of heat-producing, incompatible elements in a region of thinned crust on the lunar near side, the Procellarum-KREEP terrane (PKT; see Chapter 3). Perhaps higher temperatures in the PKT expanded and/or weakened the crust sufficiently to mask the effects of global contraction of a LMO deeper than 400 km. Another such process is deep burial of heat-producing elements. Revising a suggestion of Kesson and Ringwood (1975), Hess and Parmentier (1995) proposed that overturn in the late-stage cumulates of an LMO would have sent plumes of ilmenite and ferropyroxene sinking into the lunar interior, carrying with them trapped melt enriched in heat-producing elements. Even though the buildup of heat released by radioactive decay would eventually lead to melting, the presence of ilmenite would constrain the melts to be denser than their surroundings at depths >500 km (Delano 1990). Such a deep-seated heat source was not considered in the thermal calculations described in Section 2.4. It is unclear whether sufficient abundances of incompatible elements could have been sequestered in the deep interior to accommodate a deep (>400 km) LMO; this remains a fruitful area of investigation.

3.1.3. Duration of magma-ocean crystallization. The duration over which the LMO crystallized is uncertain. Part of the problem is the uncertainty in the age of the ferroan-anorthositic crust. The anorthosites are among the oldest lunar rocks, yet their crystallization ages are difficult to interpret (see Section 3.4). Low $^{87}\text{Sr}/^{86}\text{Sr}$ ratios for ferroan anorthosites approach those of primitive meteorites, implying that they have been separated from any significant concentrations of Rb for a length of time comparable to the age of the Moon (Nyquist and Shih 1992). Model ages for nonpristine rocks using U-Pb isotopes indicate a major differentiation event occurred at about 4.47-4.56 Ga (Oberli et al. 1979; Carlson and Lugmair 1988; Hanan and Tilton 1987; Premo and Tatsumoto 1991). Relatively young ages for some ferroan anorthosites (Borg et al. 1998) may be explained by prolonged subsolidus cooling or open system behavior, but their derivation from light REE-depleted sources (Borg et al. 1998) is still open to interpretation (Section 3.4).
The idea that the KREEP component in the lunar crust was the residuum from the crystallization of the LMO has played a key role in estimating the duration of lunar differentiation (Papanastassiou et al. 1970; Tera and Wasserburg 1974; Lugmair and Carlson 1978). The average model age for KREEP is $4.42 \pm 0.07$ Ga for Sm/Rb = 2.3 (Nyquist and Shih 1992). The interpretation of this model age is open to speculation in that it may represent a true crystallization age, a cooling or disturbance age, or a mixing event. It does, however, represent the maximum duration over which the LMO crystallized. Sm-Nd model ages for the mare basalts are interpreted as indicating that their source region (i.e., mafic cumulates from LMO crystallization) were isotopically closed by 4.4 Ga.

Thermal models for the duration of LMO crystallization are highly dependent on the size of the LMO and the nature and thermal conductivity of the early crust. Thermal models predict that before plagioclase became a stable liquidus phase, parts of the LMO may have crystallized over a period of time ranging from $10^5$ years to less than $10^7$ years (Solomon and Longhi 1977; Binder 1976; Minear 1980). The discrepancy between these durations result primarily from assumptions about the chilled margin or proto-lunar crust (thickness of proto-crust and boundary layers, viscosity, and composition). Assuming development of a thick (50-100 km), stable and continuous plagioclase crust following substantial LMO crystallization, Solomon and Longhi (1977) and Minear (1980) predicted total LMO crystallization times of $6 \times 10^7$ to $2 \times 10^8$ years. Shorter time scales for crystallization are anticipated for a thinner lunar crust or an early crust destabilized by large impacts. Minear (1980) estimated that the LMO crystallized over a period of $3 \times 10^7$ to $8 \times 10^7$ years in a multi-stage crustal growth model that anticipated an early lunar proto-crust and a later ferroan-anorthosite dominated crust disrupted by impacts and regrown by conductive cooling. In models in which the ferroan-anorthositic crust formed after LMO crystallization (Longhi 2003), the LMO could have crystallized over a timeframe of less than 200 years to 10 m.y. (Longhi 2003).

Although the excess $^{182}$W produced by the early decay of $^{182}$Hf to $^{182}$W is significantly smaller than initially reported (Lee et al. 1996), there still appear to be mantle reservoirs for the mare basalts that have excess $^{182}$W ($\varepsilon_{^{182}W} > 0$). As the high $\varepsilon_{^{182}W}$ appears to be associated with the mare basalts that were generated by melting of a clinopyroxene- and ilmenite-bearing mantle, Shearer and Newsom (1998, 2000) suggested that the radiogenic W sources were produced during LMO crystallization and not core formation. From the extent of Hf/W fractionation by ilmenite and high-Ca clinopyroxene determined by Righter and Shearer (2003) and Shearer and Righter (2003), over 95% of the LMO crystallized less than 60 million years after solar-system formation. Depending on the accretion age of the Moon, this inference suggests that most of the LMO crystallized over a period of 10-40 million years. Such a short period of time to crystallize the LMO contrasts with models suggesting durations from 100 m.y. to greater than 200 m.y. (Solomon and Longhi 1977; Longhi 1980). Rapid LMO crystallization also implies that the transition from LMO (characterized by anorthosite formation) to serial magmatism (with the emplacement of Mg-suite plutonic rocks) occurred earlier, in agreement with age constraints from other studies (Shih et al. 1993).

### 3.1.4. Chemical and petrological constraints on early lunar differentiation

The concept of the LMO began with the recognition of an anorthositic crust, so it is useful to examine some mass-balance constraints involving the crust. Figure 4.4 illustrates the results of some simple calculations that determine the amount of Al$_2$O$_3$ residual to the formation of crusts of different thickness and composition for 3 potential whole-Moon compositions. Following the gravity modeling of Haines and Metzger (1980), Taylor (1982) adopted 2.93 and 73.4 km for the average density and thickness, respectively, of the lunar highlands crust; he also derived 26 wt% Al$_2$O$_3$ as the average crustal composition. The 50:1 ratio for intrusive to extrusive mare basalts implies approximately 20 wt% of the crust consists of mare basalt (Head and Wilson 1992). Taking the average Al$_2$O$_3$ concentration of mare basalts to be 10 wt%, then removing the
mare basalt component from the crust increases the Al$_2$O$_3$ to 30 wt% and decreases the average thickness to ~58 km. These values are likely to be extremes because most of the intrusive mare magmas are likely to be situated beneath maria not modeled by Haines and Metzger (1980). Global crustal mass balance based on global FeO and using basin ejecta as probes of the middle to lower crust indicate average crustal Al$_2$O$_3$ of 24-25 wt% (see also Chapter 3, Section 5.3). According to this type of calculation, the Al$_2$O$_3$ concentration of the pre-mare-basalt lunar crust probably lies within the yellow-shaded area bounded by 26 and 30 wt% Al$_2$O$_3$ (Fig. 4.4). This total Al$_2$O$_3$ includes LMO flotation cumulates as well as post-LMO plutonic rocks.

Figure 4.4 also shows that the chondritic (C) composition cannot satisfy the Al$_2$O$_3$ requirement of the Moon unless nearly the entire Moon was initially molten and the extraction of Al$_2$O$_3$ was nearly perfect. In Figure 4.4, the TWM and LPUM compositions both contain sufficient Al$_2$O$_3$ for total melts of the Moon, but LPUM does not for an LMO <500 km deep. We may constrain the composition-depth relations further by considering constraints imposed by mare-basalt petrogenesis. For example, the pressures of multiple saturation on the liquidus of the picritic green glasses provide a crude estimate of the minimum depth of the mare basalt source region and, therefore, the minimum depth of early differentiation. Elkins-Tanton and Grove (2000) showed multiple saturation pressures from 1.3 to 2.5 GPa, which correspond to depths of 260 to 700 km. If decompression provided most of the heat for mantle melting on the Moon, as it does on the Earth, then pressures of multiple saturation represent average depths of melting (Klein and Langmuir 1987). Longhi’s (1992) modeling of green-glass petrogenesis by polybaric melting, which is updated below, constrained the green glass source region to lie at depths ≥1000 km and to contain ~1.4 wt% Al$_2$O$_3$ (horizontal blue shading). Figure 4.4 indicates that the TWM composition contains too much Al$_2$O$_3$ to produce the depth and composition of the green glass source region, whereas LPUM agrees well for an LMO depth between 500 km and the entire Moon.
3.1.5. Petrologic constraints on the extent of early lunar melting and differentiation. The compositions of basalts reflect the compositions and depths of their source regions, although not as directly as we would like. Not only do magma compositions change prior to eruption in response to low-pressure crystallization and assimilation, but also the melting process intrinsically obscures itself. Melting in the Earth’s mantle occurs primarily by the release of pressure during convective upwelling. Adiabatic decompression permits approximately 10% melting per GPa (Hess 1992). Low porosity and density differences lead to separation of melt from matrix during the melting process (McKenzie 1984). Consequently, most primitive magmas are blends of low-degree melts accumulated over a range of depths from a progressively depleted source. An important corollary is that a magma composition cannot be related to a specific depth by major elements. Nonetheless, there have several reasonably successful calculations of primitive magma composition based upon parameterization of the P-T-X data obtained from melting experiments (Kinzler and Grove 1992a; Klein and Langmuir 1987; Longhi 1992).

In order to understand the implications of mare basalt compositions for lunar structure, it is useful to compare them to terrestrial basalts in the same format that the melting calculations are presented. Accordingly, Figure 4.5 compares the compositions of low-Ti mare basalts with those of the most abundant terrestrial basalts—mid-ocean ridge (MORB), ocean island (OIB), and continental flood basalts (CFB). Each set of compositions is projected against appropriate olivine-saturated, low-pressure liquidus boundaries calculated by the algorithms presented by Longhi (1991). The terrestrial basalts have higher proportions of plagioclase (Pl) to pyroxene (Opx, Di) components reflecting higher concentrations of both Al₂O₃ and alkalis. The high relative proportions of the Pl component lead to early crystallization of plagioclase and augite and the relatively late appearance of low-Ca pyroxene as pigeonite. This is the typical tholeiitic trend. The lower proportion of the Pl component in the mare basalts leads to the early appearance of pigeonite. The differences in the moderately volatile alkalis between lunar

![Figure 4.5. Comparison of the compositions of low-Ti mare basalts with those of the most abundant terrestrial basalts—mid-ocean ridge (MORB), ocean island (OIB), and continental flood (CFB). Each set of compositions is projected against appropriate olivine-saturated low-pressure liquidus boundaries calculated by the algorithms presented by Longhi (1991).](image-url)
and terrestrial basalts reflect different accretion histories, whereas the differences in Al₂O₃ are more likely to reflect different magmatic histories.

Figure 4.6 shows a series of batch and polybaric fractional melts of a source with the composition of depleted terrestrial upper mantle from Kinzler and Grove (1992b) calculated with the programs BATCH and FFTHERM. BATCH (Longhi 2002), which calculates equilibrium crystallization and partial melting, is a new version of MAGPOX (Longhi 1991, 1992). FFTHERM is a modification that balances the amount of melting in each pressure step (0.1 GPa) against the temperature drop along an adiabat of 15°/GPa, the calculated temperature drop along the solidus, and the temperature equivalent of the heat of fusion. Any melt that exceeds the assumed background porosity fraction of 0.008 is withdrawn and accumulated. In Figure 4.6 the calculated batch and accumulated fractional melts overlap a significant portion of the natural compositions in both projections. The more extreme natural compositions with negative Opx components probably reflect the influence of CO₂ in the melting process. The batch melts all represent liquids that are saturated with or nearly saturated with an aluminous phase (plagioclase, spinel, or garnet). Saturation with an aluminous phase at high pressure ensures that a liquid will have a high ratio of plagioclase to pyroxene components.

The whole-Moon composition of Taylor (1982) contains 6 wt% Al₂O₃—1.5× the estimated concentration of the Earth’s upper mantle (Hart and Zindler 1986). It should not be surprising, therefore, that 10% batch melts of TWM have relatively high ratios of plagioclase to pyroxene components, a generally tholeiitic aspect, and do not overlap the compositions of the mare volcanics (Fig. 4.7). At 30% melting at 3.0-3.5 GPa, TWM melts begin to approach the component proportions of the mare volcanics, because the garnet is melted out, but Mg/(Mg+Fe) (Mg') values are much higher than those of the green glasses (0.72 vs. 0.61). Also, there is no plausible way to generate such high degrees of melting at depths (>700 km) corresponding to these pressures. Similar arguments apply to batch melts of LPUM.

Figure 4.8 illustrates the track of accumulated polybaric partial melts of TWM, LPUM, and GGS (a modified version of the green glass source region of Longhi (1992) that was constructed

Figure 4.6. A series of calculated batch and polybaric fractional melts of a source with the composition of depleted terrestrial upper mantle (Terrestrial “dPUM”) from Kinzler and Grove (1992b).
by adding forsteritic olivine, ferroan clinopyroxene, and ilmenite in various proportions). Each composition has garnet at the solidus at 4.0 GPa, but GGS has only a trace, and following the removal of 0.2% liquid from the source, garnet does not appear at lower pressures. The high-\(P\) melts of TWM have markedly lower quartz coordinates than melts of LPUM or GGS because of different solidus assemblages. Orthopyroxene is not present at the TWM solidus above 3.2 GPa because of a peritectic reaction with the liquid. At these pressures, melting of TWM begins in the presence of olivine, augite, garnet, and, possibly spinel. At \(\sim 7\%\) melting, orthopyroxene begins to crystallize, but the polybaric melting process never gets beyond 1-2\% melting at any
time. The orthopyroxene-free liquids lie on a boundary curve that is not tightly constrained by experimental data. However, the data that do exist constrain the olivine + clinopyroxene + garnet curve to trend toward low silica (Longhi 1995). The LPUM accumulated melts cut across the green glass array in the Wo (wollastonite, CaSiO$_3$) projection in Figure 4.8a, but are far removed in Figure 4.8b. The GGS melts do cut across the field of the green glasses in both projections. Some sample compositions are listed in Table 4.1. These compositions are by no means unique—it is possible to spread the track of the calculated liquids out by varying the initial composition or the porosity (lower porosity means more melt is extracted at each step, so the source is depleted faster and the melts become less aluminous).

The fact that neither batch nor polybaric melting of bulk lunar compositions can produce the picritic green glasses or mare basalts has several implications for lunar differentiation. The simplest are that: a) the primitive isotopic and trace-element patterns identified in the green glasses (e.g., Delano 1986; Neal 2001) are not intrinsic, but result from some sort of hybridization or assimilation; b) the post-LMO deep melting of the Moon envisioned by Kirk and Stevenson (1989) could not have produced the mare volcanics, at least not directly.

The most significant implications relate to the scale and extent of the primordial differentiation. To the extent that the model is accurate, the results imply that the differentiation event must have been sufficiently fluid to deliver differentiated material to depths of ~1200 km. Conceivably, this transport could have been accomplished by crystallization of an LMO of modest depth (400-500 km) followed either by plumes of dense Fe-rich pyroxene ± ilmenite that penetrated the deep interior or by large-scale overturn that placed the initially unmelted portion of the Moon above part or all of the crystallized LMO. Indeed, the presence of primitive components in some of the volcanic glasses (Delano 1986) suggests that some undifferentiated material survived at least until the time of mare volcanism whereas the $P$-$T$ path of the depleted green glass source implies that only modest amounts of undifferentiated Moon could have been present during mare volcanism.

Figure 4.9 shows the $P$-$T$ path of the GGS that produced the compositional trend in Figure 4.8 in relation to its own solidus and the solidii of TWM and LPUM. Clearly, the picritic glass magmas would have melted and partially assimilated any primitive material with which they came into contact. More important, however, is the observation that the temperatures along the GGS solidus exceed the solidus temperatures of undifferentiated material. Thus it is not likely that any significant fraction of undifferentiated lunar material was present above ~1200 km at the time of the green glass magmatism, otherwise it would have melted preferentially.

<table>
<thead>
<tr>
<th></th>
<th>TWM</th>
<th>LPUM</th>
<th>GGS</th>
<th>FF11*</th>
<th>Ap15G</th>
<th>Ap14 VLT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO$_2$</td>
<td>44.4</td>
<td>46.1</td>
<td>43.9</td>
<td>45.8</td>
<td>44.1</td>
<td>46.0</td>
</tr>
<tr>
<td>TiO$_2$</td>
<td>0.31</td>
<td>0.17</td>
<td>0.07</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>6.14</td>
<td>3.93</td>
<td>1.44</td>
<td>7.24</td>
<td>7.81</td>
<td>9.30</td>
</tr>
<tr>
<td>Cr$_2$O$_3$</td>
<td>0.61</td>
<td>0.50</td>
<td>0.42</td>
<td>0.43</td>
<td>0.33</td>
<td>0.58</td>
</tr>
<tr>
<td>FeO</td>
<td>10.9</td>
<td>7.62</td>
<td>16.15</td>
<td>20.7</td>
<td>21.0</td>
<td>18.2</td>
</tr>
<tr>
<td>MnO</td>
<td>0.15</td>
<td>0.13</td>
<td>0.18</td>
<td>0.28</td>
<td>0.10</td>
<td>0.21</td>
</tr>
<tr>
<td>MgO</td>
<td>32.7</td>
<td>38.3</td>
<td>36.26</td>
<td>17.9</td>
<td>16.7</td>
<td>15.9</td>
</tr>
<tr>
<td>CaO</td>
<td>4.60</td>
<td>3.18</td>
<td>1.55</td>
<td>7.12</td>
<td>8.41</td>
<td>9.24</td>
</tr>
<tr>
<td>Na$_2$O</td>
<td>0.09</td>
<td>0.05</td>
<td>0.03</td>
<td>0.16</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>K$_2$O</td>
<td>0.009</td>
<td>0.003</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>Mg'</td>
<td>0.84</td>
<td>0.90</td>
<td>0.80</td>
<td>0.61</td>
<td>0.59</td>
<td>0.61</td>
</tr>
</tbody>
</table>

*a accumulated melt composition: GGS source, 4.0 GPa initial melting, 0.1 GPa steps, 0.008 porosity*
Such melts would have been lunar tholeiites in which high-Ca clinopyroxene was the dominant pyroxene. Such rocks have not been found among the samples. Therefore, ~1200 km is a reasonable approximation of the depth of the primordial differentiation.

The recent interpretation of seismic data contrasts with the petrologic interpretation of the extent of the LMO to ≈ 1200 km depth. Several of the Apollo-era seismic models have been analyzed to infer constraints on the structure and composition of the lunar mantle (Buck and Toksöz 1980; Hood and Jones 1987; Mueller et al. 1988; Hood and Zuber 2000; Khan et al. 2000). In general, the preferred interpretation of the increase in seismic velocity at 500 km depth is a change in composition to more aluminous and Mg-rich silicates. This inferred change in composition and increase in the degree of heterogeneity has been interpreted as indicating that the Moon was initially melted and differentiated only to a depth of 500 km. If this interpretation is correct, it implies that an Al- and Mg-rich primitive mantle exists below 500 km, that all the Al in the primordial lunar crust must have been extracted from the upper mantle, and that much of the mare basaltic magmatism must have involved melting above 500 km.

A final consideration is the composition of the green glass source itself. Figure 4.8 illustrates the bulk compositions TWM, LPUM, and GGS. The three symbols are widely spaced, reflecting their different Al$_2$O$_3$ concentrations, 6, 4, and 1.4 wt%, respectively. Estimates of the difference in Al$_2$O$_3$ between primitive and deleted upper mantle on the Earth are on the order of 0.1 wt% (Kinzler and Grove 1992b). This difference is smaller than the size of the symbols. If the Moon has a bulk composition similar to that of the Earth’s upper mantle, then differentiation by extraction of basalt would imply that roughly twice the mass of the GGS had been extracted. Such material is not in evidence on the surface, and if buried would have to have a remarkably cryptic history, inasmuch as any basalt would be much less refractory than either GGS or undifferentiated Moon. Although the GGS composition was developed to fit the green glasses, it nonetheless is consistent with mixing of early- and late-stage components of an LMO. These various lines of evidence are thus consistent not only with an LMO, but with one that involved at least 2/3 of the Moon’s radius. From the mass balance relations in Figure 4.8 and the need for a low-Al$_2$O$_3$ source for the green glasses, the bulk composition of this ocean was very likely similar to an alkali-depleted version (LPUM) of the Earth’s upper mantle.

3.1.6. Final outcome. The LMO model predicts several outcomes regarding the current nature of the Moon. An appealing aspect of these LMO models is that they provide (1) processed
mantle cumulate assemblages that are appropriate sources for subsequent periods of lunar magmatism and (2) a primary lunar crust consisting of plagioclase-rich flotation cumulates.

Different cumulate rock types have been identified as potential sources for a wide range of parental magmas for both pre-mare and mare basalts. For example, the highly fractionated KREEP component in the cumulate pile is considered to have been remobilized and emplaced into the lunar crust. Many of the pre-mare basaltic magmas are thought to have acquired a KREEP signature through assimilation or mixing. The late-stage, ilmenite-bearing cumulates have been considered as possible sources for the high-Ti mare basalts (Taylor and Jakes 1974; Taylor 1982; Snyder et al. 1992), whereas early olivine-orthopyroxene dominated cumulates have been proposed as the mantle source for the very-low-Ti picritic magmas (Taylor and Jakes 1974; Taylor 1982).

Chemical signatures within the LMO cumulate stratigraphy are illustrated in Figure 4.10. For example, early cumulates should be enriched in elements such as Ni and perhaps Co, and have REE patterns with virtually no Eu anomaly (Shearer and Papike 1989). Later cumulates should develop an REE pattern with negative Eu anomalies, be systematically enriched in incompatible elements and depleted in compatible elements such as Ni.

The crystallization history of the LMO has been summarized by Schnetzler and Philpotts (1971), Taylor and Jakes (1974), Taylor (1982), Longhi (1977, 1981), and Snyder et al. (1992). Schematic diagrams for the cumulate pile produced by LMO crystallization are presented in Figures 4.10 and 4.11. The sequence of crystallization is highly dependent on LMO bulk composition and the pressure and flow regimes under which crystallization occurred, and is therefore difficult to predict. In a dynamically simple LMO, the crystallization sequence advocated by most of these models is olivine → orthopyroxene ± olivine → olivine + clinopyroxene ± plagioclase → clinopyroxene + plagioclase + ilmenite. On the basis of estimated bulk compositions of the LMO (Taylor and Bence 1975; Ringwood and Kesson 1976; Buck and Toksöz 1980; Warren 1986; Hughes et al. 1988, 1989), olivine crystallizes first until the olivine-orthopyroxene boundary line is reached. The extent of the olivine cumulate assemblage in the LMO cumulate pile is estimated to be between 30 and 40%. The effect of the olivine-orthopyroxene field boundary on cumulate assemblages depends on both the pressure of crystallization and the efficiency of crystal accumulation (equilibrium versus fractional crystallization). As pressure increases, this boundary moves toward the olivine apex of the olivine-anorthite-SiO$_2$ pseudoternary, which has two effects. First, this shift effectively decreases the volume of the monomineralic olivine assemblage in the cumulate pile. Second, the olivine-orthopyroxene boundary becomes a cotectic surface, resulting in coprecipitation of olivine + orthopyroxene regardless of the efficiency of crystal separation. Within the context of the flow regimes proposed by Spera (1992), olivine and orthopyroxene precipitated from magmas at fairly high-pressure near the base of the cumulate zone. At the upper LMO boundary, significant radiative heat loss resulted in rapid olivine and orthopyroxene crystallization. These early crystallization products of the upper boundary layer should sink and be incorporated into the inertial inner region of the LMO (a nearly isothermal region dominated by high convective velocities detailed in Section 3.2). It appears unlikely that these early phases from the upper boundary would have been incorporated into the basal cumulate layer. The appearance of low-Ca clinopyroxene and high-Ca clinopyroxene followed the precipitation of orthopyroxene and olivine. The exact sequence of the appearance of these pyroxenes is dependent upon bulk composition. The relationship among clinopyroxenes and plagioclase in the LMO crystallization sequence is discussed in detail by Longhi (1980). In the bulk composition used by Snyder et al. (1992), high-Ca clinopyroxene crystallized after plagioclase.

The appearance of plagioclase in the crystallization sequence is fundamentally important to understanding the development and evolution of the early lunar crust. In simple dynamic models, the lunar crust is thought to have formed by plagioclase crystallization and flotation
after substantial amounts of the LMO had crystallized. Using the initial bulk composition suggested by Warren (1986) in which the $\text{Al}_2\text{O}_3$ was 7 wt%, Snyder et al. (1992) calculated that plagioclase would be a liquidus phase after 57% LMO crystallization. Using estimates of bulk compositions containing <5 wt% $\text{Al}_2\text{O}_3$, Snyder et al. (1992) estimated that plagioclase would appear in the crystallization sequence after 70-80% crystallization. The ability of plagioclase to float and accumulate in “rockbergs” at the surface also depends on the bulk Mg’ of the LMO.
Flotation of anorthositic rocks (plagioclase + minor mafic minerals) may have been impossible until the "magmaosphere" became more Fe-rich after substantial crystallization of mafic silicates (Warren and Wasson 1980). High degrees of crystallization prior to plagioclase flotation are also implied by the trace-element characteristics of the ferroan anorthosites (Palme et al. 1984, 1997). On the basis of these models, a large portion of the cumulate pile has either no Eu anomaly or only a minor one (Taylor 1982; Shearer and Papike 1989; Snyder et al. 1992).

In more complex models for LMO dynamics that will be discussed in the subsequent section, it could be expected that plagioclase was a liquidus phase during a substantial period of crystallization (Longhi 1980). Due to rapid heat loss of the upper LMO boundary, plagioclase might have become a liquidus phase earlier in the upper boundary layer than in the lower one. This may have resulted in formation of a protocrust prior to 60% LMO crystallization. In addition, this early plagioclase crystallization may have imposed a negative Eu anomaly on early basal cumulates (olivine + orthopyroxene) by sinking of residual melt into the central zone of the LMO.

Ilmenite-bearing cumulates precipitated after 90% crystallization of olivine + pyroxene + plagioclase. Ilmenite made up between 3-12% of the cumulus minerals. Incompatible trace elements excluded from the crystal structures of olivine, pyroxene, and plagioclase, were concentrated in the residual melt (KREEP). The signature of this KREEP component was later incorporated into the feldspathic highland crust through remobilization, assimilation, and mixing.

Longhi (2003) suggested a stratified, post-LMO lunar mantle that is different from that illustrated in Figures 4.10 and 4.11. This difference is not due to a dramatic difference in crystallization sequence of the LMO, but the efficiency by which plagioclase could be removed and accumulated from the LMO. As discussed below, the efficiency of plagioclase removal is dependent upon the evolution of different flow regimes in the crystallizing LMO. Under an end-member model in which the ferroan-anorthosite crust is not a product of LMO crystallization, Longhi (2003) proposed that the post-LMO Moon consists of four density-stratified layers: (1) an undifferentiated zone beneath the LMO cumulate pile that is initially denser than the overlying olivine-rich cumulates because of its lower Mg' and the presence of garnet, (2) an intermediate ultramafic zone of dunites grading upward into harzburgites, (3) an upper mafic zone composed of norites and gabbro-norites, and (4) a thin protocrust consisting of dense LMO residua, quenched LMO liquids, and perhaps a small amount of anorthosite. These differences in densities and thermal characteristics would drive subsequent cumulate overturn resulting in the post-LMO formation of the ferroan-anorthosite crust, formation of mare-basalt mantle sources by fertilizing the lower LMO cumulate pile with upper LMO cumulates, and the derivation of mantle plumes from the undifferentiated zone to form Mg-suite magmas.

Complementary to the LMO cumulates that constitute at least a portion of the lunar mantle are the flotation cumulates that make up the primordial lunar crust. They represent the only remnant in the lunar sample collection that is presumably a product of the first stage of lunar differentiation. Unlike most of the other terrestrial planets, the primary lunar crust is preserved and provides significant information concerning the initial differentiation of the Moon. A summary of the petrography, geochemistry, and distribution of the ferroan anorthosites that comprise this primary crust is given by Papike et al. (1998) and in Chapter 3.

Apart from high $\text{Al}_2\text{O}_3$ and $\text{CaO}$, and the low Mg' documented in Chapter 3, the ferroan anorthosites are characterized by low incompatible element abundances with the trivalent REE generally less than chondritic. Their REE patterns are typified by a large positive Eu anomaly with $(\text{Eu}/\text{Sm})_N$ as high as 100. This large anomaly is attributed to the highly reducing conditions during the early stages of lunar differentiation (= IW−1) which resulted in a high proportion of divalent Eu relative to terrestrial magmatic environments. Estimates of parental magma compositions from bulk rock and plagioclase core compositions suggest that the
parent magmas had refractory incompatible-element ratios close to chondritic (e.g., Ti/Sm, Sc/Sm) (e.g., Hubbard et al. 1971; James et al. 1991). These parental melts are estimated to contain REE concentrations 10-50× chondritic abundances with fairly flat REE patterns (Floss 1991; Floss et al. 1991, 1998; Jolliff and Hsu 1996; Papke et al. 1997b). The calculated parent melts with lower abundances of REE lack distinct Eu anomalies and are presumably earlier and more primitive, whereas those with higher REE have small negative Eu anomalies. This is consistent with a model magma ocean originally with 5 wt% Al₂O₃ crystallizing plagioclase from ~78-90% solid (Papke et al. 1997b). However, the general lack of correlation between An contents of the plagioclase with the Mg’ of coexisting mafic minerals does not fit the simple crystallization model implied by the REE. In general, this lack of correlation between mineral compositions is attributable to variable amounts of adcumulus growth (Morse 1982; Ryder 1982) or more complex processes (e.g., Haskin et al. 1981; Jolliff and Haskin 1995).

Although in the general model of anorthosite petrogenesis the anorthosites are complementary to the mare-basalt source regions, the relationship is complex. Some of the mare sources are too magnesian, contain too much high-Ca pyroxene, and too much Ni to be directly complementary to ferroan anorthosites (e.g., Ryder 1982, 1991). These lines of evidence suggest modification of the LMO cumulates by mixing during solid-state convection. More details are presented in Section 3.3.

3.2. Physics of the lunar magma ocean

Most models for the thermal evolution of the LMO are based on the assumption that the outer portion of the Moon was totally molten (Wood et al. 1970; Wood 1972, 1975; Taylor and Jakes 1974; Hubbard and Minear 1975; Walker et al. 1975; Longhi 1977, 1981; Solomon and Longhi 1977; Minear and Fletcher 1978; Tonks and Melosh 1990; Spera 1992). Early models focused on the geochemical and petrologic implications of this early differentiation event.

Among the more sophisticated of these early models were those developed by Solomon and Longhi (1977). Their thermal models were based on the numerical solutions of the finite-difference analog of the equation for energy conservation with spherical symmetry, as originally described by Toksöz et al. (1972). Although these thermal models are one dimensional and did not describe convective flow in the LMO, they accounted for heat loss from the base and top of the LMO and approximated the average thermal state as a function of time. The models proposed by Solomon and Longhi (1977) shared several parameters. The thermal conductivity was assumed to be constant at 4 × 10⁶ erg/cm-sec-K and the specific heat was set to 1.2 × 10⁹ erg/g-K. The surface temperature was fixed at −20 °C. Initial density was set at the mean lunar density of 3.34 g/cm³. In more complex models the cumulate layers were assigned appropriate densities on the basis of mineral assemblage. The heat of fusion (4 × 10⁹ erg/g) was taken to be distributed uniformly with temperature over the solidus-liquidus interval. Convective heat transport with the liquid part of the LMO was simulated by the assumption that heat in excess of that necessary to maintain a given layer along a specific temperature profile is transferred to the overlying layer at each time step. The average abundances of heat-generating elements were taken from Langseth et al. (1976) and these elements were modeled to be strongly partitioned into the liquid phase. Models proposed by Solomon and Longhi (1977) included both crystal-liquid fractionation and the expected variation of the physical properties of the liquid and solids with the extent of LMO crystallization. The starting solidus and liquidus were taken from the data of Ringwood (1976) for the whole-Moon composition of Taylor and Jakes (1974) and the effects of pressure were ignored.

In the more sophisticated models of Solomon and Longhi (1977), cooling and crystallization would occur primarily at the top of the LMO in response to radiative cooling. The base of the LMO consists of primitive lunar mantle. This model assumes a protocrust that is a Moon-wide chilled margin at least a few kilometers thick (Solomon and Longhi 1977). The protocrust of this scale forms a cold thermal boundary layer that affects the overall heat
loss from the LMO. A thinner protocrust on the order of meters predicted by other models (e.g., Hofmeister 1983) would provide less impedance to heat loss. Individual blocks of this chilled margin are recycled into the LMO by a combination of mini-plate tectonics and impact disruption. The early crystallizing minerals, dominated by olivine, sink to the base of the LMO. Initial cooling is efficient, crustal growth is slow, and the lower mafic cumulate layer grows rapidly. Once plagioclase becomes a liquidus phase and is buoyant, crustal growth accelerates and the LMO cooling rate decreases. Herbert et al. (1977) envisioned the formation of the initial plagioclase crust as surface expressions of convection-cell subsidence. On the basis of these models, the time for the crystallization of approximately 90% of the LMO would have been 100 to 200 m.y. The duration of crystallization of the LMO is virtually independent of initial depth as long as it is covered by a 70 km thick insulating layer of flotation cumulates.

More complex models of the dynamics of a totally molten LMO have been explored by Tonks and Melosh (1990) and Spera (1992). A workshop on the physics and chemistry of magma oceans from 1 bar to 4 mbar contributed significantly to understanding the complexities of the LMO (1992, LPI Tech. Report 92-03). Studies by Solomatov and Stevenson (1993a,b,c), Abe (1997), and Solomatov (2000) have application for the LMO as well as for other terrestrial planets. At the high Rayleigh numbers relevant to the LMO, convection changes to a regime referred to as “hard” turbulent convection. At even higher Rayleigh numbers, convection is expected to enter a new regime of turbulent convection (”ultrahard”), but recent experimental studies by Glazier et al. (1999) suggested that hard turbulence is probably the pertinent regime for the LMO. Tonks and Melosh (1990) used the turbulent mixing-length theory of Kraichnan (1962) to describe a hard-turbulent LMO regime that initially convected vigorously with an initial Rayleigh number (Ra) on the order of $10^{25-27}$. Circulation in a hard-turbulent regime is characterized by small correlation scales for velocity, temperature, and composition, both spatially and temporally (Spera 1992). The evaluation by Spera (1992) used experimental and simulation studies combined with a scale analysis. Spera adopted the experiments and arguments of Castaing et al. (1989) to support a $\text{Nu} \sim \text{Ra}^{2/7}$ rather than the traditional $\text{Nu} \sim \text{Ra}^{1/3}$ scaling. The Nusselt number ($\text{Nu}$) is the ratio of the actual heat flux at the surface relative to the conductive flux giving the same drop in temperature across the horizontal boundaries of the system. The difference in scaling adopted by these models is critical. As shown by Spera (1992), calculation of the LMO surface temperature and cooling rate from the two different scaling laws leads to significantly different results. Adopting the 2/7 law yields an LMO surface temperature ($T_s$) of 580 °C, whereas the traditional 1/3 law yields an LMO $T_s$ of 1100 °C. Since the radiative cooling rate is proportional to $T_s^4$, cooling is nearly a factor of 10 faster for the 1/3 law scaling. A better understanding of the relationship between Nu-Ra relations for high Ra convection of fluids with high Prandtl number ($\text{Pr} = \text{kinematic viscosity/thermal diffusivity}$) is required to better constrain models for the duration of LMO cooling and crystallization (Spera 1992).

In their comparison of the nature of magma oceans on the Earth and Moon, Tonks and Melosh (1990) concluded that the effect of gravitational acceleration on the slopes of the solidus, liquidus, and adiabatic temperature profiles was one factor in the divergent geochemical evolutionary paths of the two bodies. They proposed that the adiabat of the LMO lies entirely between the liquidus and solidus. Therefore, nearly the entire LMO would have been at subliquidus conditions. An implication of this is that crystals suspended in the LMO would remain below the liquidus temperature. They concluded that individual crystals would continue to grow no matter where they were transported in the LMO and would eventually become large enough to overcome convective stirring and settle into the cumulate pile. This conclusion is a simplification of the relationship between the adiabat and the liquidus. It would appear that crystals that are liquidus phases in cooler portions of the LMO could be transported to and absorbed in hotter regions that are still below the liquidus.
In high \( Ra \) convection regimes where \( Ra^{1/3} > 35 \sigma^{1/2} \) (Eqn. 4.6) (\( \sigma = \) Prandtl number), Tonks and Melosh (1990) and Spera (1992) defined a complex LMO system that could be divided conceptually into several thermal-mechanical regimes on the basis of rheological properties. Immediately adjacent to either the floor or top of the LMO is a thin thermo-mechanical boundary layer where heat is transported by conduction. The thickness of this layer is defined by Spera (1992) as follows:

\[ \delta_T = 4L Ra^{-1/2} \]

where \( L \) is the time dependent depth of the LMO. The surface heat flux supported by the high Rayleigh number heat flux is defined by Spera (1992) as follows:

\[ q_s \approx \frac{k \Delta T}{4L} Ra^{1/2} \]

where \( \Delta T \) is the temperature difference across the LMO from its base, at or near the solidus, to its surface and \( k \) is the thermal conductivity (\( k \approx 3 \text{ Wm}^{-1}\text{K}^{-1} \)). The thickness of the surface conductive crust and the basal cumulate pile increase through time.

Adjacent to both boundary layers are sublayers where heat is transferred primarily by convection but in which viscosity dominates. The thickness of the layer is as follows (Spera 1992):

\[ \delta_v = 4L Ra^{-1/2} \]

and the rms convective velocity within the layer is:

\[ u_c = \left( \frac{k}{L} \right) Ra^{1/2} \]

The separation of crystals from melt is restricted to this flow regime near the floor region or along the base of the crust. Melt trapped within the basal cumulates communicates very slowly with the adjacent layers because of the large Darcy friction within the crystal mush (Oldenburg and Spera 1990, 1991; Spera 1992).

The interior layer of the LMO is dominated by inertial forces. The temperatures of this part of the LMO are nearly isothermal and typical convective velocities are on the order of:

\[ u_c = \left( \frac{\nu}{L} \right) \sigma^{-2/3} Ra^{1/2} \]

This layer consists of any material above its liquidus, along with partly crystalline regions where crystal fraction is below a value \( \Phi_{critical} \) too low for crystals in the liquid to form networks. One of the important features of hard turbulent convection in this zone is the existence of large-scale turbulent eddies (Tonks and Melosh 1990; Solomatov 2000). Tonks and Melosh (1990) evaluated crystal suspension and settling in this zone using empirical studies of sediment transport. Settling and flotation only take place for Rouse numbers (terminal velocity of the crystal/turbulent friction velocity) greater than 1. Results show that crystals up to 1-2 cm in diameter can be suspended from the onset of cooling. As viscosity increases in the later stages of cooling, crystals of 10 cm to >1 m in diameter may be suspended in the inertial layer. The critical crystal size for equilibrium crystallization to be an important process during the crystallization of early magma oceans was estimated by Solomatov (2000) as:

\[ D_{crit} = 10^{-3} \left( \frac{\eta_t}{0.1 \text{ Pa}\cdot\text{s}} \right)^{1/3} \left( \frac{F}{10^5 \text{ W}\cdot\text{m}^{-2}} \right)^{1/2} \text{ m} \]
The dimensions and depths of the various rheological zones during the evolution of the LMO depend on the liquidus and solidus slopes, and the intersections of the adiabat with the liquidus and solidus and the line of $\Phi_{\text{critical}}$ (see Solomatov 2000; Fig. 4.9). The structure of a magma ocean may thus depend significantly on planet size, a relationship also noted and discussed by Tonks and Melosh (1990). Determination of $\Phi_{\text{critical}}$ is significant for the modeling process. An example of the thermal structure and evolution implied by many of the above idealizations is shown in Figure 4.12, which also shows two potential relationships between the lunar adiabat and the solidus and liquidus (Tonks and Melosh 1990; Spera 1992).

Spera (1992) reconstructed the thermal and physical history of the LMO using the above parameters. During the early stages of the LMO with $Ra \approx 10^{25}$, $\sigma \approx 10^2$, and $L \approx 500$ km, the conductive and discontinuous crust had a thickness of $\delta_T \approx 0.1$ m through which enormous heat flux, $q_s \approx 2 \times 10^4$ W/m$^2$, is carried. Solomatov (2000) estimated a heat flux of $\sim 10^6$ W/m$^2$ during the initial stages of MO crystallization. Hofmeister (1983) argued that the thickness of the chilled crust would reach a steady-state thickness of no more than a few meters. As long as frequent impacts and shear stresses generated by convection kept the surface discontinuous, heat transfer would remain high and be governed by the balance between heat brought to the surface by convection and radiative heat transferred from the hot LMO surface. During this time, the viscous sublayer has a limited thickness of $\delta_y \sim 0.5$ km. Based on the above equations, the magma convection velocities within this layer are on the order of $10^{-2}$ m/s.

Within the inertial layer, convective velocities are extremely high ($\sim 10^{-4}$ m/s; Spera 1992; Solomatov 2000). The temperature gradient across the layer is small due to the very rapid advective transport of heat. Cooling rates are extraordinarily high during early stages of LMO crystallization because of the significant radiative heat loss through the discontinuous and thin

---

**Figure 4.12.** Schematic diagrams illustrating LMO structure for high Rayleigh number convection (after Spera 1992) and relationships between two possible lunar adiabats and the solidus and liquidus. Adiabat 1 is constructed based on the regime diagram of Spera (1992). A portion of this adiabat lies above the liquidus. In this relationship crystals nucleate in relatively narrow zones near the top and bottom of the LMO. Crystals that are swept into the main body of the LMO remelt, thus keeping the mean crystal size small. During cooling the adiabat will shift to lower temperatures. Adiabat 2 is from Tonks and Melosh (1990) and lies between the liquidus and the solidus for the Moon. In this scenario, crystals continue to grow in size until they are large enough to settle out or float. Tonks and Melosh (1990) suggested that the LMO crystallized predominantly under the condition of the adiabat lying entirely between the solidus and liquidus.
Estimates of crystal size during the early stages of LMO crystallization are very close to the critical size separating fractional and equilibrium crystallization (Solomatov 2000), which means that both equilibrium and fractional crystallization are equally acceptable within the uncertainties of the physical parameters.

As the LMO continued to cool, the thicknesses of the boundary layers and the viscous sublayers continued to grow at the expense of the interior region. The development of the former results in the substantial decrease in radiative heat loss. Tonks and Melosh (1990) examined the longevity of the inertial flow layer using Equation (4.6). They calculated that the inertial flow zone disappeared at $\Phi_{\text{critical}} = 46\%$ in a 400 km LMO and 49\% in a 1000 km terrestrial magma ocean. This is similar to the disappearance of the inertial flow zone estimated by Murase and McBurney (1973). Most theoretical and experimental investigations based on spherical particles estimate a $\Phi_{\text{critical}}$ value between 50 and 60\% (e.g., Campbell and Forgacz 1990; Lejeune and Richet 1995; Marsh 1995). Philpotts et al. (1999) found that in natural basaltic systems at low pressures, plagioclase laths form three-dimensional networks at crystal fractions as low as 25\%. Therefore, the liquidus phases and their typical crystal forms may be a highly significant parameter. The propagation of this rheological front from the base of the terrestrial magma ocean toward the surface may take as little as $10^3$ years (Solomatov 2000). Melosh (1992) suggested that at this crystal “lock up” point the effective viscosity increases by many orders of magnitude and melt moves by percolation through the crystals. In this environment, chemical differentiation of the magma ocean takes place as the liquid percolates through the solid and may be driven by density differences (Melosh 1992). At this stage, fractional crystallization may be the more dominant process. Magma-ocean cooling and further crystallization within this rheological regime may be as fast as $10^7-10^8$ years (Davies 1990). However, if surface recycling was inefficient, thereby slowing convection beneath the lunar crust, the crystallization time would be longer (Solomatov and Moresi 1996; Solomatov 2000).

Using initial bulk compositions of 5-7 wt\% $\text{Al}_2\text{O}_3$, Snyder et al. (1992) calculated that plagioclase would be a liquidus phase after 57\% to 80\% of the LMO crystallized. On the basis of these estimates, the onset of plagioclase crystallization would occur when the crystal fraction is well above any reasonable value of $\Phi_{\text{critical}}$ and the inertial zone of the LMO has disappeared. Therefore, plagioclase flotation depends not only on the buoyancy of plagioclase in highly evolved liquids, but also on the ability of plagioclase to separate buoyantly from a crystal mush. Upon the appearance of plagioclase on the liquidus and the development of a well-defined plagioclase crust, the rate of LMO cooling would drop precipitously owing to thermal blanketing effects of stable crust. The crust and cumulate zone continue to grow until they eventually merge at the base of the crust.

Tracking incompatible and heat-producing elements trapped in interstitial liquids is one critical step to understanding the magmatic source regions that would result from LMO crystallization. Interstitial liquids are trapped in solidifying crystal mushes when solidification proceeds more quickly than compaction and percolation. The bottom of the partially solidified zone moves to shallower depth due to cooling by heat loss at the surface. Therefore, the upward velocity $V$ of this solidification front will influence the amount of trapped, interstitial liquid in the LMO cumulate pile. Buoyant percolation of liquid that is less dense than solid would occur at a pore velocity given by:

$$u = \frac{K}{\Phi \mu} \Delta \rho g$$  \hspace{1cm} (4.13)

where $K$ is permeability, $\Delta \rho$ is the density difference between the crystals and the liquid, $\mu$ is the liquid viscosity and $\Phi$ is the melt fraction. If liquid percolation occurs through tube-like channels along mineral grain edges, then
where $b$ is the grain size (cf. Wark and Watson 2002). The amount of melt trapped is given by the value of $\Phi$ for which $V \approx u$. Escape of interstitial melt requires compaction of the solid. This estimate neglects the resistance of deformation of the solid that would further increase the trapped melt fraction (cf. Shirley 1986). A surface heat flux by radiation through an $H_2$-$H_2O$ iron-wüstite buffered atmosphere (Abe et al. 2000), at a surface temperature equal to the potential temperature of the liquid, would give trapped liquid fractions on the order of 1-10% for a grain size of 1 mm and a liquid viscosity of 1-10 Pa·s. If a conductive lid forms at the surface, the resulting much lower rates of cooling would give a smaller trapped melt fraction. Even a trapped melt fraction less than 1% may be significant in controlling the incompatible-element content of LMO cumulates.

3.3. Physics of cumulate overturn

3.3.1. Introduction and relevance. As discussed in previous parts of Section 3, crystallization of the LMO produced a wide compositional range of cumulates that eventually served as mantle sources for subsequent periods of lunar basaltic magmatism. Models for the stagnant crystallization of an LMO predict a cumulate stratigraphy (Fig. 4.10) that is inconsistent with the observation that there is a limited relationship between basalt composition and its depth of origin (Section 4.5). Also, it appears that basaltic magmas represented by the volcanic glasses, and probably most of the mare basalts, were a product of melting initiated deep in the lunar mantle (Delano 1986; Shearer and Papke 1993, 1998; Elkins-Tanton et al. 2003a). Since Ti-bearing cumulates and plagioclase do not crystallize until near the top of a solidifying magma ocean (Fig. 4.10) and assimilation of Ti by low-Ti magmas appears not to be a viable mechanism (Ringwood and Kesson 1976; Van Orman and Grove 2000; Elkins-Tanton et al. 2001), the source materials needed for mare magmatism requires some mechanism of overturn in which late magma-ocean cumulates are mixed downward into the underlying mantle of earlier cumulates to a depth of at least 400 km.

Solidification of a magma ocean would define initial planetary compositional differentiation, but could lead to an unstable cumulate density stratification, which overturns to a stable configuration. The nature of the final, stable compositional stratification can have important implications for a planet’s subsequent evolution by delaying or suppressing thermal convection and by controlling the initial temperature distribution and the distribution of heat-producing elements in the mantle. Processes occurring during the first few hundred million years of magma-ocean solidification and overturn surely shaped the evolution of the planet on much longer time scales.

For the Moon, recent progress on LMO crystallization (e.g., Snyder et al. 1992; Hess and Parmentier 1995) point to unexplored aspects of cumulate formation and overturn that may be critical to understanding lunar evolution. The origin of the source materials and the processes responsible for mantle melting that generated lunar basalts (mare basalts, Mg-suite rocks) and perhaps the temporal evolution of lunar magnetic field may all have been controlled by the initial formation and disruption of the LMO cumulate pile.

As discussed above, LMO crystallization models suggest that highly fractionated, late-stage liquids were trapped between deep mafic cumulates and a flotation anorthositic crust. Depending on the LMO composition used, Ti-oxides (i.e., ilmenite) would begin to crystallize from the late-stage liquid with clinopyroxene and plagioclase when 89-95% of the LMO had solidified. Depending on the thickness of anorthosite crust assumed, crystallization of ilmenite would begin at a depth between 150 and 100 km at a temperature between 1180 and 1125 °C (Hess and Parmentier 1995; Van Orman and Grove 2000). Assuming efficient separation from
plagioclase, the high-Ti cumulates would have a density of 3700-3800 kg/m$^3$, compared to the underlying olivine + pyroxene mantle density of about 3300 kg/m$^3$. Ringwood and Kesson (1976) proposed that because the solid ilmenite + clinopyroxene cumulate layers were denser than the underlying, less evolved olivine- and pyroxene-bearing cumulates, the ilmenite + clinopyroxene cumulates sank into the underlying cumulates. Ringwood and Kesson (1976) and other investigators (e.g., Shearer et al. 1992, 1993; Hess and Parmentier 1995; Zhong et al. 2000) hypothesized that the sunken high-Ti cumulates subsequently remelted and contributed to the formation of high-Ti mare basalts. Other investigators have challenged the hypothesis of deeply foundered, remelted high-Ti cumulates on the basis that high-Ti melts are not buoyant below 200-300 km depth in the Moon (Circone and Agee 1996; Wagner and Grove 1997), or on the basis that the high viscosities in the cold planetary lid would prevent or delay their sinking (Elkins-Tanton et al. 2002).

**3.3.2 Evolution of the mantle immediately following magma ocean crystallization.** As discussed above, crystallization of the LMO would generate an unstably stratified mantle that should overturn to a gravitationally stable state. Dense cumulates that were added to the top of the already solidified mantle would have been gravitationally unstable and would tend to sink. The relative timescales of LMO solidification and cumulate overturn also influence the final cumulate stratigraphy. If overturn occurred while the LMO was crystallizing, deeper already solidified mantle would be brought to the solidification front thus affecting the composition of the remaining liquid. Overturn therefore involves a competition between the rate of thickening of the solidified layer and its Rayleigh-Taylor time scale (Elkins-Tanton et al. 2003b).

The time scale of Rayleigh-Taylor instability and overturn of an unstably stratified viscous layer of thickness $d$ and solid-state viscosity $\eta$ with stress-free top and bottom boundaries is

$$t_{\text{overturn}} = \frac{4\pi^2 \eta}{\gamma gd^2}$$

where $\gamma$ is the compositional density gradient, assumed constant with depth, and $g$ is gravity (Hess and Parmentier 1995). As the thickness of the solidified layer $d$ increases, its overturn time decreases. Gravity at the solidification front also increases as solidification proceeds, further decreasing overturn time as the solidified layer thickens. At the beginning of solidification near the bottom of a 1000 km deep LMO on the Moon, $g$ is about 0.7 m/s$^2$, overturn of a 100 km thick layer with $\gamma = 10^{-4}$ kg/m$^3$/m in olivine and olivine-orthopyroxene cumulates (below a depth of about 100 km in Fig. 4.13) and $\eta = 10^{19}$ Pa·s, is about 2 million years. This value of viscosity is taken to be comparable to values near the bottom of the Earth’s oceanic lithosphere, in the range $10^{19}$-$10^{20}$ Pa·s. The creep rate of mantle silicates, however, depends strongly on the presence of small amounts of water. In the absence of water, laboratory deformation studies (Hirth and Kohlstedt 1996) indicate that the viscosity is larger by at least a factor of 100. In this case the overturn time of a 100 km thick layer is more than 200 m.y. longer than estimates for the solidification time of the LMO. Near the end of solidification when $d$ is 10× thicker, the overturn time is only a few million years.

Overturn rates therefore do not become less than thickening rates until near the end of LMO solidification, thus the LMO is expected to crystallize almost completely before density instability initiates cumulate overturn. Based on the simple model of very rapid, turbulent convective heat transfer to the surface, assuming that a solidified layer at the surface participates in this convection rather than forming a conductive lid, solidification may be more rapid than this estimated overturn time for the olivine-orthopyroxene cumulates. Once a plagioclase flotation crust forms a conductive lid, the last stages of solidification will be much longer. Therefore overturn of the mafic cumulates may occur while the evolved portion of the LMO is still liquid.
The simplified overturn scenario is where the entire cumulate pile is simply reordered with a monotonically decreasing density upwards and assuming that plagioclase separates perfectly by flotation. The resulting compositional profiles before and after overturn are shown in Figure 4.14. Prior to overturn, primitive olivine-orthopyroxene cumulates with progressively increasing Fe/Mg are overlain by an ilmenite-rich layer. After overturn, this dense layer forms the base of the cumulate pile with a high-Ti and low-Si composition (Fig. 4.14).

During overturn, hot material rising from depth could cross its solidus through decompression and melt (Fig. 4.14). The hottest cumulates that rise to just below the plagioclase flotation crust are olivine with an Mg' of 95-96. The cumulates just below are a mixture of high Mg' olivine and orthopyroxene. The approximate solidus for these materials is shown as a dashed line, indicating that the overturned cumulates in this model should not remelt through decompression. In other models, overturn can create up to 50 km of mafic material that could form a lower lunar crust. The very old ages of the Mg-suite (Papike et al. 1998, and references therein) suggest that a mechanism such as remelting high-Mg cumulates during overturn may be required to explain their formation.

Despite gravitational instability, however, the rate and occurrence of overturn depend on complex rheologies of solids deforming by thermally activated creep. The calculations presented in Elkins-Tanton et al. (2002) suggest that, with a non-linear stress-dependent rheology appropriate for a pyroxene-dominated cumulate, the late-forming high-Ti cumulates may creep at rates too low to allow descent via Rayleigh-Taylor instabilities on a time scale appropriate to form the high-Ti basalts. Wieczorek and Phillips (2000) presented a model with a ten-km thick layer of material of a KREEP-basalt composition at the base of the anorthositic crust that predicts that melting should eventually extend to depths of about 600 km and persist through the entire evolution of the Moon to the present day. Hess and Parmentier (2001) pointed out that such a high concentration of heat sources might prevent the crystallization of the LMO over much of the history of the Moon, but in particular they question whether the

![Figure 4.13. Solidus and liquidus for a lunar bulk mantle composition from Ringwood and Kesson (1976), with approximate adiabats (thin lines) for the liquid and partly crystalline magma oceans to demonstrate the possibility of crystalline, crystalline network, and liquid and crystal mush zones in the solidifying mantle. The adiabats have slopes such that during almost the entirety of magma ocean crystallization the adiabat lies between the solidus and liquidus, as discussed in Tonks and Melosh (1990). Plagioclase begins to crystallize when the magma ocean is about 80% solid, at which point the magma ocean is likely to have a higher than critical crystal fraction, making complete plagioclase flotation impossible (modified after Tonks and Melosh (1990).)](image-url)
Shearer et al. Thermal & Magmatic Evolution of the Moon

high temperatures at the base of the crust generated in such a model would allow the presence of an elastic lithosphere thick enough to support the mascons at the time of mare basalt eruptions (Solomon and Head 1980).

Longhi (2003) suggested a somewhat different outcome resulting from density and temperature stratification of an LMO cumulate pile that contained an upper, plagioclase-bearing mafic zone, formed in the absence of a thick, floating anorthositic crust (Section 3.1). He proposed that because of unstable density profiles in the mafic and ultramafic zones (lower, olivine- and pyroxene-bearing cumulates) produced by the upward decrease in Mg, two-layer convection occurred. Overturn in each zone not only brings lower-density mineral assemblages upward but also advects heat, resulting in transportation of olivine cumulates at 1800 °C to the base of the mafic cumulate zone that contains mineral assemblages with solidus of 1000-1250 °C. Melting in the mafic cumulate zone ensues and produces thick suspensions of plagioclase in ferroan-anorthosite-like liquid that ascend buoyantly. The denser, complementary pyroxenitic component (both Ti-poor and Ti-rich) sinks into the lunar interior to form the source region for the mare basalts through mixing of early LMO cumulates. Overturn steepens and eventually reverses the density profile in the LMO cumulate pile. This change in the density profile along with radioactive and conductive heating of undifferentiated mantle beneath the LMO cumulate pile may activate plumes. These plumes may eventually melt to produce additional post-LMO magmas such as the Mg-suite.

3.3.3. Spatial scales of mantle overturn. The spatial scale of gravitational instability is especially important for testing the overturn hypothesis and for understanding the subsequent
Shearer et al. (Thermal & Magmatic Evolution of the Moon)

The simplest outcome of the above idealized overturn would be for all the ilmenite cumulates to move to the bottom of the LMO cumulates, or perhaps the base of the mantle. This dense material, rich in heat-producing elements, would heat up with time allowing thermal expansion to offset compositional density. If the deep layer is pure ilmenite cumulate, a temperature increase of several thousand °C is required for thermal expansion to offset the high intrinsic density, which would result in a completely liquid layer. This melt would be expected to be denser than the solid (Circone and Agee 1996), and so might remain at depth in a molten state. Thus deep, compositionally dense layers may also be difficult to remobilize through convective instabilities, and may form a reservoir sequestering material from the rest of the mantle (Alley and Parmentier 1998; Elkins-Tanton et al. 2003b). Alternatively, thermally driven convective upwelling in the overlying mantle might entrain small amounts of such dense melt creating a possible mare-basalt source. If the ilmenite cumulates mixed with olivine-pyroxene cumulates during overturn, then a smaller increase in temperature might allow the solid mixture to become buoyant at a more modest temperature increase, rise, decompress, and melt (Zhong et al. 2000; Stegman et al. 2003).

The final stages of this heterogeneous overturn would thus involve the simultaneous solidification and sinking of the ilmenite-rich cumulates. The ilmenite-rich layer before overturn is relatively thin but contains the densest mantle material (Fig. 4.15). The wavelength or spacing of dense diapirs would depend on the thickness of the layer from which they form and thus on the balance between the rate of solidification and the rate of diapiric sinking. Estimated diapir sizes for reasonable ranges of solidification rate are invariably less than 10 km (Hess and Parmentier 1995). The low sinking velocity of small diapirs means that they would sink to depths only on the order of 100 km over the time for final solidification of the LMO, thus generating a layer consisting of a mixture of ilmenite-poor cumulates with ilmenite-rich inclusions (Beard et al. 1998 suggested deeper sinking of the pyroxene-bearing cumulate). This mixed layer is denser than underlying cumulates and its thickness may be large enough that it would be most unstable at long wavelengths (cf. Parmentier et al. 2002). The observed hemispheric asymmetry in the distribution of mare basalts on the surface and Th in surface materials (Lawrence et al. 1998) might thus be explained. Alternatively, ilmenite cumulates that initially sink to the bottom of the mantle might rise, decompress, and melt in a spherical-harmonic mode-one pattern (Zhong et al. 2000) generating mare basalts in only one region of the surface. Stegman et al. (2003) suggest that such a process may also explain the generation and subsequent decay of a lunar core magnetic dynamo based on available paleomagnetic data.

3.4. Challenges to the LMO hypothesis

3.4.1. Introduction. The LMO model has been the cornerstone of lunar petrology since the examination of the first samples returned by the Apollo missions (Smith et al. 1970; Wood et al. 1970). One of the first challenges to the LMO model was the identification of energy sources capable of producing extensive planetary-scale melting. Initial models for lunar accretion and other energy-producing processes were found to generate inadequate amounts of heat. As illustrated earlier in this chapter (Section 2), with the acceptance of new lunar accretion models that predict the storage of enough energy in the Moon to induce extensive lunar melting, the necessity for alternative, low melting differentiation models is diminished. The potential problems with forming a primordial lunar crust through the separation and flotation of plagioclase in an LMO following moderate degrees of crystallization (40-60%) were discussed in detail in Secton 3.2. Another potential problem that needs to be factored into LMO models are new isotopic observations for the ferroan anorthosites.

3.4.2. Ages of early crustal rocks. Isochron ages have been determined for a growing number of lunar crustal rocks belonging to the alkali, magnesian, and ferroan-anorthositic suites. Notable inconsistencies exist between the measured ages and initial isotopic compositions of many lunar crustal rocks and the ages and isotopic compositions expected
According to LMO models. Such inconsistencies suggest that the LMO model, in its simplest form, cannot explain all of the isotopic systematics observed in the lunar rocks.

Ages for alkali-suite rocks range from ~3.80 to 4.37 Ga, for Mg-suite rocks, from ~4.18 to 4.46 Ga, and for ferroan anorthosites, from 4.29 to 4.56 Ga (Fig. 4.16; see also Chapter 3, Fig. 3.5 and Chapter 5, Table 5.6). Numerous Rb-Sr ages determined on Mg-suite rocks range from 3.90 to 4.61 Ga and also overlap with the Sm-Nd ages determined on the ferroan anorthosites (see summaries in Nyquist and Shih 1992; Nyquist et al. 2001). Sm-Nd isotopic
Contemporaneous crystallization of the ferroan anorthosites and Mg-suite rocks is not necessarily contemporaneous with the magmatism that produced at least some ferroan anorthosite. Data from Lugmair et al. (1976), Nakamura et al. (1976), Carlson and Lugmair (1981), Nyquist et al. (1981c), Compston et al. (1984a; 1984b), Carlson & Lugmair (1988), Shih et al. (1993), Alibert et al. (1994), Meyer et al. (1996), and Norman et al. (2000) suggest that the ages of the ferroan anorthosites and Mg-suite rocks post-date urKREEP formation. This figure demonstrates that the ages of the ferroan anorthosites and Mg-suite rocks may overlap, suggesting that by about 4.4 Ga both types of magmas were being produced on the Moon. Even the oldest ages of the ferroan anorthosites and Mg-suite rocks are within error of one another requiring the magma ocean to have existed for only a short period of time.

Analyses of the ferroan anorthosites are very difficult owing to their monomineralic, highly shocked, and commonly polymict character, and their very low concentrations of REE. Thus the ferroan anorthosites that have been dated tend to have higher abundances of mafic minerals than typical ferroan anorthosites (in order to obtain a wider range of parent/daughter ratios to give a more precise isochron) and may therefore not be representative of the ferroan-anorthositic suite as a whole. Furthermore, the Sm-Nd isotopic systematics of the ferroan anorthosites could somehow be disturbed (see discussion in Borg et al. 1999), and the Sm-Nd ages of the ferroan anorthosites may not represent crystallization ages. For the purpose of this initial discussion, the ferroan anorthosite ages are taken at face value. We then explore the implications of plausible mechanisms that have been proposed to disturb the Sm-Nd isotopic systematics of the ferroan anorthosites.

There is significant overlap between the ages of the alkali, magnesian, and ferroan-anorthositic suites of rocks, and particularly the latter two. The overlap is not dependent solely on the relatively young ages of ferroan anorthosites such as 62236. In fact, Mg-suite clast 15445,247 (Shih et al. 1993) is within analytical uncertainty of the age of the oldest ferroan anorthosite (Fig. 4.16). The ages of the ferroan anorthosites are younger than the age of LMO crystallization, estimated from the short-lived chronometer $^{182}$Hf$\rightarrow^{182}$W ($t_{1/2} = 9$ Ma). Although this chronometer only provides model ages for parent/daughter fractionation it suggests that lunar differentiation occurred between 4.50 and 4.52 Ga (Lee et al. 1997; Shearer and Newsom 1999; Righter and Shearer 2003; Shearer and Righter 2003). Furthermore, the age of urKREEP formation (i.e., the late-stage residual melt of LMO crystallization) has also been estimated by a variety of isotopic methods. Nyquist and Shih (1992) suggested that an average model age of 4.42 ± 0.07 Ga is probably the most representative of the age of urKREEP formation. This model age is as old or older than the Sm-Nd age of ferroan anorthosite 62236 (4.29 ± 0.06 Ga; Fig. 4.16), suggesting that the age of 62236 could post-date urKREEP formation.

Taken at face value, these age relationships suggest that the earliest Mg-suite magmatism was contemporaneous with the magmatism that produced at least some ferroan anorthosite. Contemporaneous crystallization of the ferroan anorthosites and Mg-suite rocks is not
consistent with the standard LMO model in which ferroan anorthosites are the oldest crust and are intruded by younger Mg-suite plutonic rocks. The crystallization ages of at least some ferroan anorthosites post-date the age estimates for crystallization of the LMO, therefore the age relationship between the ferroan anorthosites, Mg-suite rocks, and LMO crystallization are more consistent with lunar crustal genesis by intrusion of multiple magma bodies, i.e., serial magmatism (e.g., Walker 1983; Longhi and Ashwal 1985; Longhi 2003).

3.4.3. Initial Nd isotopic composition. Six ferroan anorthosites have been dated by the Sm-Nd isotopic system: 60025, 67016, 62236, 67215, and clasts from lunar meteorites MAC88105 and Yamato 86032 (Carlson and Lugmair 1988; Alibert et al. 1994; Borg et al. 1999; Norman et al. 2000, 2003; Nyquist et al. 2002). It is apparent that all of the ferroan anorthosites from the Apollo 16 site analyzed thus far have positive initial \( \varepsilon_{\text{Nd}}^{143} \) values (Fig. 4.17), which is consistent with derivation from LREE-depleted sources. However, the ferroan-anorthosite clasts from the lunar meteorites have negative initial \( \varepsilon_{\text{Nd}}^{143} \) values. Data on the MAC88105 clast is inconclusive for negative initial \( \varepsilon_{\text{Nd}}^{143} \) because it may be contaminated by Mg-suite materials (Nyquist et al. 2002). Thus, the presence of an Mg-suite component could lower the initial Nd value of MAC, making it negative. Conversely, the clast from Y86032 appears to be a ferroan-anorthosite breccia without contamination from other lithologies.

The \( ^{147}\text{Sm}^{144}\text{Nd} / ^{143}\text{Nd} \) ratio of potential ferroan-anorthosite sources are estimated from their ages and initial \( \varepsilon_{\text{Nd}}^{143} \) values, assuming that their sources formed at ~4.56 Ga. The model illustrates potential relationships between the four analyzed ferroan anorthosites from Apollo 16. All of the ferroan anorthosites lie within error of a single growth line with a \( ^{147}\text{Sm}^{144}\text{Nd} \) ratio of 0.279 ± 0.17. Thus, although differences in crystallization and exposure ages indicate that these four ferroan anorthosites are unlikely to be comagmatic, they maybe derived from the same source (Borg et al. 1999). Conversely, they may simply be derived from different sources with similar \( ^{147}\text{Sm}^{144}\text{Nd} \) ratios. In any case the fact that the Apollo 16 ferroan anorthosites have positive initial \( \varepsilon_{\text{Nd}}^{143} \) values, and consequently appear to be derived from LREE-depleted sources, is not consistent with the simple LMO model. This inconsistency stems from the fact that if the initial conditions are chondritic, the LMO is not expected to dramatically fractionate the Sm/Nd ratio, and hence \( \varepsilon_{\text{Nd}} \) values, throughout crystallization.

The preceding discussion presumes that the Sm-Nd isotopic systematics of the samples have not been disturbed by shock processes. Although most Sm-Nd isotopic systematics appear fairly well behaved, this is not a forgone conclusion (Norman et al. 2003).

![Figure 4.17. Time-\( \varepsilon_{\text{Nd}}^{143} \) diagram of ferroan-anorthosites and magnesian-suite rocks (sources of data same as Fig. 4.16). Note that all ferroan anorthosites (FAN) have positive initial \( \varepsilon_{\text{Nd}}^{143} \) values, consistent with derivation from LREE-depleted sources. The \( ^{147}\text{Sm}^{144}\text{Nd} \) ratio estimated for the sources of FAN, assuming the source formed 4.56 Ga, is 0.279 ± 0.017. This value is typical of highly depleted sources from terrestrial planets.](image)
3.4.4. Comparison between modeled anorthositic cumulates and ferroan anorthosites.

Snyder et al. (1992) estimated the mineralogy and major and trace element compositions of the LMO cumulates using a combination of phase equilibria and equilibrium/fractional crystallization models. Using these models as a basis, the Sm-Nd abundances of anorthositic cumulates may be calculated in a similar fashion, but using the modal mineralogy observed in actual samples. This approach permits the modeled compositions to be directly compared to analyzed ferroan anorthosites. The ferroan anorthosites are assumed to have formed with their present modal mineralogy during the interval of crystallization from 78-86% (Snyder et al. 1992). The results of the model are not affected if the 86-95 interval is used instead. From these data and model results, the \(^{147}\text{Sm}/^{144}\text{Nd}\) ratios of the anorthositic cumulates can be calculated.

The isotopic composition of the cumulates depends on when they formed. One estimate of the time of LMO crystallization comes from the 4.42 ± 0.07 Ga model age for KREEP (Nyquist and Shih 1992). This age is comparable, but slightly younger than the age estimated from the Hf-W isotopic system of 4.50 and 4.52 Ga (Lee et al. 1997; Shearer and Newson 1999). For the models presented below the LMO is assumed to have occurred at 4.50 Ga. The \(\varepsilon_{\text{Nd}}\) values of the anorthositic cumulates that are modeled are not, however, strongly dependent on this age because of the long half life of the \(^{147}\text{Sm} \rightarrow ^{143}\text{Nd}\) (\(t_{1/2} = 106\) Ga) chronometer.

Figure 4.18 is a plot of the \(\varepsilon_{\text{Nd}}\) of whole-rock values for ferroan anorthosites 60025, 62236, 67215, and 67016 versus the \(\varepsilon_{\text{Nd}}\) calculated for anorthositic cumulates with the modal mineralogy of the ferroan anorthosites. It is apparent from this figure that the measured \(\varepsilon_{\text{Nd}}\) whole-rock values are substantially below the 1:1 line predicted by the model. The \(\varepsilon_{\text{Nd}}\) calculated for the anorthositic cumulates are dependent on the time of LMO crystallization, the Sm/Nd ratio of the bulk Moon, and the Nd isotopic composition of the bulk Moon. However, the crystallization of the LMO must be significantly younger than 4.35 Ga (3.85 Ga) and the bulk Moon is required to have a \(^{147}\text{Sm}/^{144}\text{Nd}\) ratio significantly higher than the chondritic value of 0.1967 (0.207) and \(\varepsilon_{\text{Nd}}\) value significantly higher than 0 (+4.1) to account for the differences between the modeled and measured \(\varepsilon_{\text{Nd}}\) whole rock values of 60025. The situation is worse

---

**Figure 4.18.** Present-day \(\varepsilon_{\text{Nd}}\) of modeled anorthositic cumulates versus present-day whole-rock values. Source of whole-rock values same as Figure 4.17. Whole rock value for 60025 calculated from data from mineral fractions assuming a mode of plag:ol:pig = 80:10:5. The composition of the modeled anorthositic cumulates are calculated using the LMO model of Snyder et al. (1992) and the modal mineralogy of the individual ferroan anorthosites assuming formation at 4.50 Ga. The present-day \(\varepsilon_{\text{Nd}}\) of values of FAN are more radiogenic than the calculated values suggesting that they are not simple floatation cumulates of a magma ocean.
for samples that lie farther from the 1:1 line on Figure 4.18. Another possibility is that the ferroan anorthosites that have been analyzed isotopically contain an LREE-depleted component that was not considered in the models. Although the presence of this component would shift a sample to the right on Figure 4.18, it is unlikely because it would have to be a major source of REEs in the ferroan anorthosites. A final possibility is that the analyzed ferroan anorthosites are mixtures of materials derived from a LREE-depleted source. This also seems unlikely because most mineral phases from the ferroan anorthosites lie on isochrons. The simplest explanation for the differences between the modeled and measured $\varepsilon_{143}^{Nd}$ whole-rock values is that these ferroan anorthosites are derived from a source that is more LREE-depleted than is hypothesized for the LMO.

The LREE-depleted source region postulated for the ferroan anorthosites is confirmed, at least for sample 62236, by the presence of a small positive $\varepsilon_{142}^{Nd}$ anomaly of $+0.25 \pm 0.11$ (Borg et al. 1999). Positive $\varepsilon_{142}^{Nd}$ anomalies, such as that observed in 62236, are the result of the decay of $^{146}\text{Sm} \rightarrow ^{142}\text{Nd}$ ($t_{1/2} = 103$ Ma), in an LREE-depleted source region (i.e., Sm/Nd ratio that is greater than the chondritic value). Although the size of the $\varepsilon_{142}^{Nd}$ anomaly observed in 62236 is relatively small, it is significantly larger than that predicted by the models. For example, a modeled anorthositic cumulate forming at 4.50 Ga is expected to have a negative $\varepsilon_{142}^{Nd}$ value of $-0.45$. As with whole-rock $\varepsilon_{143}^{Nd}$ values, the $\varepsilon_{142}^{Nd}$ is not dependent on isochron determinations and is expected to be unaffected by post-crystallization shock metamorphism of the sample.

3.4.5. Ramifications of isotopic data for the LMO. Both new and resurrected models have been employed to account for the production of ferroan anorthosites with younger ages and derived from LREE-depleted mantle sources (positive $\varepsilon_{142}^{Nd}$). Wetherill (1975) envisioned a differentiation model in which accretional melting of the Moon was restricted, creating isolated magma chambers that crystallized to form large layered intrusions consisting of mafic cumulates (source for mare basalts) capped with anorthosites (lunar highlands). An extensive primitive lunar mantle remained unprocessed during these events. Walker (1983) and Longhi and Ashwal (1985) suggested that subsequent heating of the Moon due to the decay of long-lived radionuclides initiated global convection that heated the stack of intrusions envisioned by Wetherill (1975). This resulted in melting and allowed the separation, growth and ascent of anorthositic diapirs. The coeval mafic component was buried in the mantle at depths greater than the sampling depth of multiring impact basins (30-60 km). At such depths, these mafic cumulates could have mixed with primitive lunar mantle assemblages that did not participate in the localized melting. This model accounts for the overlapping ages of ferroan anorthosites and Mg-suite rocks, the young age of some of the ferroan anorthosites, and perhaps even the LREE-depleted mantle sources for the ferroan anorthosites.

This type of model, however, has several weaknesses (Hess 1989; Shearer and Papike 1993). First, it is clear that substantial early melting occurred rather than insignificant melting. Second, this model does not adequately explain the diversity of mantle sources for the mare basalts (e.g., TiO$_2$ variability). Third, in comparison with terrestrial analogs, most of the larger layered intrusions do not have such large anorthosite units. Fourth, there is little evidence for the existence of a primitive lunar mantle that was involved in lunar magmatism. Fifth, the apparent consistency of the KREEP-like interelement ratios in almost all nonmare materials is not compatible with basalt crystallization in relatively small, isolated magma chambers (Warren 1985).

One way to reconcile the isotopic results discussed above is to have a chondritic LMO crystallizing early and very quickly and for the ferroan anorthosites from the Apollo 16 landing site not be floatation cumulates of the LMO. Rather they have to be derived from mafic sources characterized by LREE-depletion. Multi-stage growth modeling of $\varepsilon_{142}^{Nd}$ and $\varepsilon_{143}^{Nd}$ values in 62236 by Borg et al. (1999) suggest that the source region of this sample formed at $4.45 \pm 0.01$ Ga. This age is younger than some estimates of LMO crystallization of (e.g., 4.50-4.52 Ga; Lee et al. 1997; Shearer and Newsom 1999) and therefore does not require the
source of 62236 to be a cumulate of the LMO. A post-LMO melting model was advanced by Longhi (2003) to explain the isotopic observations made by Borg et al. (1999) for the Apollo 16 ferroan-anorthosite. His calculations implied that the LMO was capable of producing plagioclase-bearing mafic cumulate sources as a result of the inefficient flotation-separation of plagioclase during crystallization. Remelting of these sources at relatively low-pressures (5-10 kb) is capable of generating ferroan-anorthosite-like magmas. These post-LMO magmas would be derived from LREE-depleted sources and would have ages that post-date ferroan anorthosites produced during LMO crystallization and overlap with the Mg-suite.

3.4.6. Crystallization and impact resetting ages of ferroan anorthosites. An alternative to distinct periods of ferroan-anorthosite production is that post-crystallization, open-system processes affected the Nd isotopic composition and/or Sm/Nd ratios of the ferroan anorthosites. These subsolidus processes have been explored by Shearer et al. (2002), Nyquist et al. (2002), and Norman et al. (2003). Numerous studies have demonstrated that the ferroan anorthosites have undergone significant post-crystallization cooling, reheating to the point of remelting, brecciation, and mixing (i.e., Nord and Wandless 1983). Shearer et al. (2002) showed that 62236 consists of several presumably ferroan-anorthositic lithologies that have distinct trace-element signatures and that have been metamorphosed to fairly high temperatures (800-1000 °C). They suggested that during these high temperature events, isotopic reequilibration may have occurred among the different lithologies in the breccia. Nyquist et al. (2002) showed that the plagioclase matrix surrounding the Yamato 86032 ferroan-anorthosite clast had a distinctly different Nd isotopic composition and Sm/Nd. The higher REE diffusion rates in plagioclase compared to pyroxene and the significantly different REE patterns of the two phases suggest that open system processes may have affected plagioclase more severely than primary pyroxene. Using isotopic data derived from the mafic minerals from the Apollo 16 ferroan anorthosites, Norman et al. (2003) derived an age of $4.456 \pm 0.04$ Ga and an $\varepsilon_{Nd}^{142}$ of $+0.8 \pm 1.4$. Using isotopic data derived from mafic minerals from the Apollo 16 ferroan anorthosites and bulk data from an anorthositic clast from Yamato 86032, Nyquist et al. (2002) derived an age of $4.49 \pm 0.09$ Ga and an $\varepsilon_{Nd}^{142}$ of $0.0 \pm 0.8$. The scale and mechanism of REE distribution among phases and their effect on the Nd isotopic composition and Sm/Nd of plagioclase are still in need of validation. Further isotopic measurements of ferroan anorthosites are required to better understand the observations and diametrically opposing conclusions made by Borg et al. (1999) and Norman et al. (2003).

4. LUNAR (PRE-MARE) MAGMATISM 4.5 TO ~3.85 Ga

4.1. Introduction

4.1.1. Definition of pre-mare magmatism. Based on Earth observations and a cursory view of the Apollo and Luna sample suites, to a first approximation the lunar crust consists predominately of an ancient highland crust of ferroan anorthosite, younger basalts that fill many impact basins, and impact-generated lithologies. However, a closer examination of the sample suite indicates the existence of a wide compositional breadth of plutonic and volcanic rocks that represent lunar magmatic events that presumably followed ferroan-anorthosite formation and preceded the eruption of basin-filling basalts (Fig. 4.19). Although referred to as pre-mare magmatism, it is highly likely that mare magmatism is a continuation of early episodes of basaltic magmatism. For example, on the basis of lunar meteorites, KREEP basaltic magmatism is known to extend to at least 2.9 Ga (Borg et al. 2004). In addition, the sampling of the plutonic rocks discussed below is biased because their availability is a product of impact events that were curtailed significantly after 3.9 Ga. These plutonic and volcanic lithologies have been arranged into several petrologic groups: Mg-suite plutonic rocks, alkali-rich plutonic rocks, and highly evolved highland plutonic rocks, KREEP basalts, high-Al basalts, and high-K basalts (Hubbard et al. 1971; Brown et al. 1972; Warner et al. 1976; Meyer
1977; Warren and Wasson 1977, 1979, 1980; Norman and Ryder 1979; Shervais et al. 1984, 1985; Warren 1989). Many of these lithologies occur as clasts in highland soils and breccias, and have been modified by post-crystallization processes to varying degrees. Of the 384 kg of samples returned from the Moon, less than 0.75% (2.88 kg) are coherent fragments of these pre-mare lithologies (see Chapter 3, Section 2.1). Of this mass, approximately 85% is represented by only 7 rock fragments.

The existence of large volumes of pre-mare basaltic lithologies has also been postulated based on remotely sensed data (Head 1974; Metzger and Parker 1979; Schultz and Spudis 1979; Bell and Hawke 1984; Davis and Spudis 1985, 1987; Hawke et al. 1990; Head and Wilson 1992). A series of low-albedo surface units of varying distribution and areal extent have been interpreted as pre-mare volcanic deposits (Head 1974). These deposits are concentrated in and near upland areas adjacent to maria, appear to have a smooth surface that gives the appearance of mantling underlying topography, and have an age that is post-Imbrium and pre-mare (Head 1974). Head and Hawke (1992) have suggested that perhaps up to a third of the erupted basalts at the lunar surface were this type of “cryptomaria volcanism.” The distribution of plutonic rocks has been examined through windows to the lower lunar crust provided by central peaks of impact craters (Pieters and Wilhelms 1985; Hawke et al. 1986; Lucey et al. 1986; Pieters 1991, 1993; Tompkins and Pieters 1999; see also Chapter 3). The relationship of older dark-mantle deposits and central peak lithologies to the volcanic and plutonic lithologies represented by the sample suite is equivocal.

4.1.2. New approaches to exploring episodes of lunar magmatism. The relatively old magmatic lithologies are critical for reconstructing the magmatic and thermal history of the Moon between 3.85 and 4.5 Ga. Are these lithologies products of melting resulting from internal (mantle) or external heat (impact)? If endogenous, are these lithologies related to one another? Is younger mare magmatism really a continuation of these early stages of magmatism? Understanding their composition, distribution, and origin is important for reconstructing the structure of the lunar crust and calculating the bulk composition of the crust and the whole Moon.

New data on these lithologies have focused on three sets of interpretive problems: lack of new samples, small sample size, and lack of geological context in which to place the sample suite. “New” samples of these early stages of lunar magmatism have been “mined” as clasts and mineral fragments from breccias. Because of their complex history (magmatic, extended cooling, impact, potential mixing) and in many cases small, potentially unrepresentative
sample size, recent work on these samples has focused upon isotopic and microbeam studies. Many of these samples have little or no geological control. In order to better use the sample data to reconstruct the magmatic, crustal, and mantle evolution of the Moon, their relationship to remotely sensed lithologies must be established and they must be put within the context of the global data sets.

4.2. Plutonic rocks, magnesian suite

4.2.1. Introduction. Rocks that belong to the Mg-suite (also referred to as the Mg-suite, magnesium-rich plutonic rocks, highland Mg-suite or HMS) can be differentiated from other plutonic-hypabyssal lithologies making up pristine lunar highland rocks with a variety of mineralogical and chemical criteria (Chapter 3). They are distinguished from ferroan anorthosites and alkali-suite rocks by the An content of plagioclase vs. the Mg/(Mg+Fe) of coexisting mafic minerals (Fig. 4.20). Compared to the ferroan anorthosites, they generally have higher abundances of mafic minerals and elevated incompatible-element concentrations (Chapter 3 and references therein). The REE contents of rocks from the Mg-suite vary widely (≈0.4-400× chondrite) and in most cases are greater than those of ferroan anorthosites (Fig. 4.21). The mineralogical and chemical differences between ferroan anorthosites and the Mg-suite have been interpreted as indicating formation of these two crustal rock types under different petrogenetic circumstances (i.e., they cannot be simply related along a common liquid line of descent).

Commonly, it is assumed that the Mg-suite plutonic rocks intruded preexisting LMO-derived ferroan anorthositic crust and that they formed intrusive complexes analogous to terrestrial layered mafic intrusions. The evidence for the former is that early age dates suggested that the ferroan anorthosites are older than the Mg-suite rocks. The evidence for the latter is tied to the sampled lithologies that are analogous to terrestrial layered intrusions and mimic a liquid line of descent (dunite → troctolite → norite) predicted in the olivine-anorthite-SiO$_2$ pseudoternary system. However, orbital and field evidence, although far from complete, calls for a much more complex scenario. In fact, as noted above, the interpretation of some recent isotopic data cast some doubt on the temporal relationship between the ferroan anorthosites and the Mg-suite. Also, there is a spatial dissociation between ferroan anorthosite (Apollo 16) and Mg-suite (Apollo 14, 15, and 17) samples suggesting that lateral and vertical separations dominate the distributions rather than the intrusion of one into another (Ryder et al. 1995; Papike et al. 1998).

![Figure 4.20](https://example.com/figure420.png)

Figure 4.20. Anorthite content (An) in plagioclase vs. Mg' of coexisting mafic silicates for lunar highland plutonic rocks (modified after Warren 1986 and Snyder et al. 1995).
As with the ferroan anorthosites, the petrologic history of the Mg-suite is obscured to varying degrees because of the effect of intense meteorite bombardment and prolonged subsolidus cooling. The first isotopic data were interpreted as indicating that Mg-suite magmatism followed the generation of the ferroan-anorthositic lunar crust and extended over a period of 400 m.y. (Carlson and Lugmair 1981a; Carlson 1982; Dasch et al. 1989) (Figs. 4.16 and 4.19). This is also implied by the initial Sr isotopic compositions of the Mg-suite plutonic rocks that are higher than those of ferroan anorthosites (Carlson and Lugmair 1981a,b). The duration of Mg-suite magmatism is also problematical. Some of the younger ages for the Mg-suite may not be crystallization ages at all, but may be more closely related to subsolidus closure (Carlson and Lugmair 1981a,b; Carlson 1982). Conversely, the decrease in the intensity of meteorite bombardment after ~3.9 Ga limited the natural sampling of these deep crustal lithologies.

4.2.2. Recent petrologic observations. Many of the basic petrologic observations concerning the Mg-suite were made during the early- to mid-1970s. Since that time numerous petrologic and geochemical observations have been made that reflect on the origin of the Mg-suite and their place in the evolution of the lunar crust.

In an attempt to reconstruct the geology of the lunar highlands crust, Spudis et al. (1991) and Ryder et al. (1997) collected electron-microprobe data on mineral fragments in poikilitic melt breccias produced by the Serenitatis and Imbrium impact events. They held that the nature and distribution of the mineral fragments could be used to reconstruct the stratigraphy of the highland crust and thereby provide additional insight into crust-building episodes related to the emplacement of the Mg-suite. Ryder et al. (1997) concluded that the highland lithologies at this site were extremely complex, consisting of numerous separate igneous intrusions that crystallized at various crustal depths. Mg-suite plutons were emplaced at depths of 40-50 km in the lunar crust (McCallum and Schwartz 2001). Some of the mineral fragments were from previously unknown Mg-suite lithologies (i.e., troctolites with olivine with even higher Mg'). The deepest crust sampled by this impact event was basaltic in composition, consisting of KREEPy gabbroic rocks with limited olivine. Mg-suite lithologies consisting of norites and troctolites overlaid the gabbroic rocks. The intrusive relationship between the Mg-suite and ferroan anorthosites generally predicted by more simple crust-building models was lacking in the Serenitatis region. Ferroan anorthosites are essentially absent in the mineral fragments.
The studies of Jolliff et al. (1993) and Snyder et al. (1995) are examples of a similar approach for investigating “new” Mg-suite plutonic rocks. They compiled major-trace element and isotopic data for relatively pristine rocks that occurred as small (<100 g) clasts from Apollo 14 breccias. Using these data, they calculated parental melts and placed these samples within the context of large Mg-suite samples. The study by Snyder et al. (1995) concluded that these remnants of the Mg-suite were cumulates derived from the crystallization of 0 to 43% of pristine KREEP basalt (Fig. 4.20). They also suggested models for the incorporation of the KREEP component and elevated large-ion lithophile elements into the parent magmas.

One of the more insightful petrologic-experimental observations concerning the origin of the Mg-suite was made by Longhi (1981) and was reemphasized by Warren (1985) and Hess (1998). They concluded that the crystallization sequence of the Mg-suite, like that of the mare basalts (orthopyroxene followed by clinopyroxene), indicated that lunar basaltic magmatism was non-tholeiitic. Longhi (1981) and Hess (1998) interpreted this observation as indicating that the mantle source for the parental magmas for the Mg-suite had a sub-chondritic Ca/Al ratio and consisted of mineral assemblages dominated by harzburgites rather than lherzolites. Alternatively, this could be a product of very high degrees of partial melting (Longhi 1981) or assimilation of a ferroan-anorthosite crust (Warren 1985).

4.3.2. Isotopic studies. Nyquist and Shih (1992) and Snyder et al. (2000) compiled ages for the Mg-suite rocks. These ages ranged from 4.61 ± 0.07 to 4.17 ± 0.02. This range is perhaps a little deceiving in that it includes 40Ar-39Ar and Rb-Sr ages that are clearly disturbed by impact-related processes. Still, from these data many of the early studies concluded that the Mg-suite rocks post-dated the ferroan anorthosites. U-Pb and Sm-Nd data have more precisely defined the duration of Mg-suite magmatism. A compilation of U-Pb and zircon ages for the Apollo 14 Mg-suite rocks yields a range of 4.136 to 4.320 Ga. Shih et al. (1993) determined Sm-Nd ages for two norites from the Apollo 15 site of 4.46 ± 0.07 and 4.28 ± 0.03 Ga. It is interesting to note that these two ages come from the same clast in sample 15445 from locations separated by one cm (Nyquist and Shih 1992). These data seem to indicate that Mg-suite magmatism was initiated soon after the formation of the primitive ferroan-anorthosite crust and extended for 100s of millions of years. Snyder et al. (1995) suggested that Mg-suite magmatism on the lunar near side occurred first in the northeast and then swept to the southwest over a period of 300 to 400 million years. However, it is possible that in reality all of these intrusions were solidified before 4.3 Ga and represent a significant episode of crust building (Papike et al. 1998).

Nyquist and Shih (1992), Shih et al. (1993), and Snyder et al. (1995) have shown that many of the Mg-suite rocks exhibit enriched Nd isotopic signatures (i.e., negative ε143Nd) and a 147Sm/144Nd ratio of 0.169 that approximates that of KREEP (Fig. 4.22). This signature contrasts with what would be expected for early Mg-rich LMO cumulates, which should have a depleted Nd isotopic signature (i.e., LREE depletions with a positive ε143Nd). The similarity of 147Sm/144Nd ratio implies that the KREEP basalts are potentially the volcanic equivalent to the Mg-suite and Mg-rich KREEP basalts could be analogous to the parental magmas for the Mg-suite (Snyder et al. 1995).

A dunite from the Mg-suite has a very subchondritic 187Os/188Os value of 0.1045 (Walker et al. 2004). This means that the mantle source of its parent melt could not have been highly fractionated with regard to Re/Os prior to its formation. This and other siderophile-element data indicate that the lunar mantle (at least the sources for Mg-suite magmas) had much lower abundances of highly siderophile elements relative to the terrestrial and martian mantles.

4.2.4. Trace element mineralogy. One of the difficulties of interpreting bulk rock geochemical data from the Mg-suite is that they represent cumulates, not melts, and many of the samples are small clasts that are non-representative of the cumulate lithologies from which they were derived. Because of this, numerous studies have approached the interpretation of the Mg-
suite by trace-element analysis of individual phases using the ion microprobe. These studies have used this approach to interpret magmatic history by comparing and inverting elemental data of individual phases and for interpreting age dates by investigating on a mineralogical scale the effects that mixing, cooling, and reheating have on isotopic systematics. To date, trace-element analyses by ion microprobe have been reported for pyroxene, olivine, plagioclase and phosphates for most of the lithologies making up the Mg-suite (Jolliff et al. 1993; Papike et al. 1994, 1996; Shervais and McGee 1997, 1999; Shearer and Floss 2000; Shearer and Papike 2000; Shearer et al. 2001). Based on these analyses and the assumption that the trace-element compositions of mineral cores could be inverted to determine melt compositions, these investigators determined the trace element characteristics of Mg-suite magmas.

As implied by the bulk rock data, the plutonic rocks of the Mg-suite appear to have crystallized from magmas with REE and incompatible trace-element concentrations equal to or greater than the high-K KREEP estimated by Warren (1985) (Fig. 4.23). In addition, relative to high-K KREEP, the calculated parental magmas for the norites are depleted in Eu (Papike et al. 1994, 1996). The trace-element data provide additional insight from the distribution of the incompatible elements in the lithologies that cannot be extracted from the whole-rock data. That is, phases that crystallized early from the melt (Figs. 4.23 and 4.24) (cores of olivine, orthopyroxene) are enriched in incompatible elements indicating that this enrichment was a characteristic of the parental magmas, not a product of late-stage magmatic processes or subsolidus mixing or mingling of a KREEP component.

Several other insights are offered by the ion-probe data regarding the characteristics of the Mg-suite parental magmas and the petrogenetic relationships among highland lithologies. A puzzling characteristic of the Mg-suite is that most lithologies have primitive major-element characteristics (high Mg') yet are enriched in incompatible elements. With such a high Mg', compatible elements such as Ni should also be high. For example, Ni in olivine (Mg' ≈ 0.92) from terrestrial komatiites is approximately 3700 ppm (e.g., Karner et al. 2001) and the Ni in olivine (Mg' ≈ 0.75) from the Apollo 12 olivine basalts ranges from 400 to 500 ppm (Papike et al. 1999; Karner et al. 2001). Surprisingly, Ni in olivine from the Mg-suite ranges from 300 to 100 ppm for the ultramafic lithologies (Mg' for olivine ≈ 0.90 to 0.85 ) and 160 to 40 ppm

Figure 4.22. Initial εNd values vs. crystallization ages for Mg-suite rocks, evolved highland plutonic rocks and KREEP basalts. These rocks form a linear array corresponding to an LREE-enriched source with 147Sm/144Nd = 0.169 (modified after Nyquist and Shih 1992).
Figure 4.23. Chondrite-normalized REE plots for plagioclase and pyroxene in magnesian-suite norites and pyroxenes (analyzed by ion microprobe) and parental melts calculated from them by inversion using appropriate distribution coefficients (modified from Papike et al. 1998).

Figure 4.24. Ni vs. Co for olivine from the various lithologies making up the magnesian-suite, ferroan-anorthosites and Select mare basalts (modified from Shearer and Papike 2005).
for the troctolites and norites (Mg’ for olivine ≈ 0.94 to 0.71) (Shearer et al. 2001) (Fig. 4.24). Therefore, the only “primitive” signature of these magmas is their high Mg’.

Among the Mg-suite lithologies thus far analyzed, there is clearly a relationship between Ni and Co in the olivine (Fig. 4.24). Nickel and to some degree Co decrease from the ultramafic rocks → troctolites, spinel troctolites, Mg anorthosites → gabbronorites. Yttrium in olivine and incompatible elements in plagioclase and the bulk rock increase with decreasing Ni. However, the olivine-bearing norites plot at higher Co. Compositions of the olivine in these lithologies are offset toward ferroan-anorthosite olivine that plot at much higher Co.

4.3. Plutonic rocks, evolved highland rocks

4.3.1. Introduction. Here, included in this discussion of evolved lunar highland rocks are the alkali suite, consisting of anorthosites, norites, gabbros, and monzodiorites, a suite of quartz monzodiorites (monzogabbro), and a range of siliceous plutonic rocks referred to as granites, rhyolites, and felsites (Hubbard et al. 1971; Brown et al. 1972; Warren and Wasson 1980; Jolliff 1991; Snyder et al. 1995). The sample population and the petrography of these evolved highland rocks are described in Chapter 3 and by Papike et al. (1998). The small size and potential non-representative nature of these small samples is an obvious obstacle to their interpretation.

Early studies of the lunar highlands identified plutonic rocks with anomalously high alkali-element contents (Hubbard et al. 1971) and highly evolved mineral and chemical signatures (Brown et al. 1972). These alkali-rich lunar rocks are generally defined as containing greater than 0.1 wt% K₂O and 0.3 wt% Na₂O (Hubbard et al. 1971; Brown et al. 1972; Warren and Wasson 1980). Although this is low compared with terrestrial magmatic rocks, the bulk Moon is highly depleted in alkali and volatile elements (Taylor 1982). Although the alkali rocks have an elevated incompatible-element contents (La = 20 to >1000× chondrite), they rarely have the typical KREEP incompatible-element abundance pattern (Fig. 4.25) nor the high K₂O contents observed in the other evolved plutonic rocks such as the quartz monzodiorites and granites/felsites (Jolliff 1991; Papike et al. 1998). Characteristic features of these rock types are discussed in Chapter 3.

The evolved highland lithologies appear to make up a much smaller portion of the lunar crust relative to the ferroan anorthosites and Mg-suite rocks. However, this conjecture is based on their limited abundance in the lunar sample suite and speculation concerning their origin. Most of the early ages of tiny fragments of these evolved rocks indicate that this style of magmatism followed the production of ferroan-anorthosite and overlapped with episodes of Mg-suite magmatism.

4.3.2. Isotopic studies. Nyquist and Shih (1992) and Snyder et al. (2000) compiled ages for the alkali-suite rocks. Ages tabulated at that time range from 4.370 ± 0.03 to 3.796 ± 0.01. Like the Mg-suite rocks, this range is deceptive in that it includes ⁴⁰Ar-³⁹Ar and Rb-Sr ages that are clearly disturbed. Zircon (U-Pb) ages and Sm-Nd data have more precisely defined the duration of alkali-suite magmatism. A compilation of zircon ages yields a range of 4.028 to 4.370 Ga for alkali anorthosite (Apollo 14), gabbronorite (Apollo 16), and quartz monzodiorite (Apollo 15). Zircon ages for the lunar granites (Apollo 12, 14, 17) range from 3.883 to 4.360 Ga. Sm-Nd ages for an alkali anorthosite (Snyder et al. 1995) and a granite (Shih et al. 1985) are 4.108 ± 0.053 Ga and 4.110 ± 0.200 Ga, respectively.

Shih et al. (1985) and Snyder et al. (1995) have shown that granite and alkali anorthosite exhibit enriched initial-Nd isotopic signatures (εNd = −0.6 and −1.0, respectively) and a ¹⁴⁷Sm/¹⁴⁴Nd ratio of 0.169 that approximates that of KREEP (Fig. 4.22). The similarity of ¹⁴⁷Sm/¹⁴⁴Nd ratio implies that KREEP, KREEP basalts, Mg-suite rocks, and the evolved plutonic rocks have a common petrologic link, which is consistent with the inference from mineral compositions (e.g., Fig. 4.20).
4.3.3. Trace element mineralogy. Snyder et al. (1994) and Shervais and McGee (1997, 1999) presented analyses of plagioclase and pyroxene from norites and anorthosites. Based on these analyses, it appears that the alkali norites crystallized from magmas with REE and incompatible trace-element concentrations equal to or slightly greater than “high K-KREEP” estimated by Warren (1985). Relative to high K-KREEP, the calculated magmas for the anorthosites and norites of the alkali suite are slightly depleted in HREE and enriched in Eu (Shervais and McGee 1997, 1999). From ion microprobe analyses in phases from an alkali gabbro (in 14318), Shervais and McGee (1997) concluded that the alkali suite samples they analyzed were derived from parental magma with trace-element characteristics similar to those for many of the Mg-suite lithologies. Snyder et al. (1994) deduced that the parent magma was similar to the composition of the quartz monzodiorite lithologies and more evolved than those for the Mg-suite.

Jolliff et al. (1993) analyzed the REE in apatite and merrillite in a suite of samples including alkali-suite lithologies. Although the phosphates are likely to represent products of very late-stage trapped melt, the coexisting silicate mineral assemblages were used to model the geochemical evolution of residual melt, constrained by bulk trace-element analyses by INAA, and thus to back out likely parent-melt compositions. Parent-melt REE concentrations inferred in this way were found to be consistent with fractionation of KREEP-like parent melts. Jolliff et al. (1999) reported ion-microprobe analyses of pyroxene and plagioclase in sample 14161,7373, a well-preserved phosphate- and alkali-rich monzogabbro that likely crystallized in a shallow intrusive body. Although this sample has the highest bulk
REE concentrations of any lunar rock, the trace-element concentrations of its minerals are consistent with fractionation of a KREEP-like parent magma. Segregation of phosphates from late-stage melts, especially RE-merrillite, coupled with the effects of late-stage silicate-liquid immiscibility to produce felsic segregations, imposes a trace-element signature that leads to the unusual enrichment pattern seen in granite/felsite.

### 4.4. Remotely sensed data on lunar plutonic rocks

If the Mg-suite and alkali-suite lithologies crystallized at depths of 10-30 km (Herzberg and Baker 1980) or deeper, they should rarely be exposed at the surface, unless excavated by large basin-forming events or uplifted in the central peaks of large craters. Their identification through remote sensing methods is further complicated because characteristics such as Mg', which is used to distinguish Mg-suite rocks, cannot be determined from current spectral data (e.g., Clementine) that are available at high spatial resolution. However, the relative abundance of mafic minerals can be obtained from such data and may be used to distinguish candidate Mg-suite rocks from anorthosites and basalts (see Chapter 2).

The study of rock types exposed in the central peaks of large craters by Tompkins and Pieters (1999) and described in Chapter 3 classified central peaks according to their major mineralogy. Exposures containing less than 80% plagioclase were identified as gabbro, norite, troctolite, gabbro-norite, anorthositic troctolite, anorthositic gabbro, and anorthositic gabbro-norite following the classification scheme of Stöffler et al. (1980). These were inferred to be Mg-suite lithologies because these rock types are known from the samples. Central peaks containing anorthositic norite, on the other hand, could indicate either Mg-suite rocks or ferroan-anorthositic rocks (Jolliff and Haskin 1995; Norman et al. 1995; Tompkins and Pieters 1999; Hawke et al. 2003; Norman 2003) (Fig. 4.26). Because Mg’ is not determined, however, the assignment to the ferroan, magnesian, or alkali suites remains an inference.

In one location, Tycho, the suggestion of Mg-suite plutonic material in the central peaks is perhaps more conclusive. The identification of gabbroic material in the peaks of Tycho (Tompkins and Pieters 1999) is supported by previous studies, including the analysis of

![Figure 4.26](image.png)

**Figure 4.26.** Locations of proposed ancient volcanic units. Hidden mafic deposits are shown with circles, and the names corresponding to Table 4.3 are given. Potential Mg-suite materials are indicated with squares. Possible locations of KREEP volcanism are identified by triangles. Basemap is a downsampled version of the Shaded Relief Map of the Lunar Near Side and Far Side Hemispheres, L 10M, I-2769, Lambert azimuthal equal area projection, US Geological Survey 2002.
telescopic hyperspectral data (Hawke et al. 1986; Lucey and Hawke 1989; Pieters 1993). The central peaks of Tycho exhumed material from a depth of approximately 15 km (Tompkins and Pieters 1999), which is too deep to represent buried extrusive materials. Thus, it is likely that Tycho crater tapped into an Mg-suite pluton or a gabbroic intrusion petrogenetically related to the mare basalts. From the composition of surrounding crater peaks, and from other studies (Hawke et al. 1986), the deposit is thought to be vertically and laterally extensive, on the order of tens of kilometers (Tompkins and Pieters 1999).

The mafic central-peak lithologies are not constrained to the western lunar near side, i.e., the Procellarum KREEP terrane (PKT). If these are indeed exhumed Mg-suite plutonic rocks, this is an important observation that would constrain petrogenetic models for both the origin of the Mg-suite and the nature of asymmetric lunar magmatism prior to mare volcanism. More studies are needed to determine the volumetric significance of these types of deposits and the role they played in early lunar magmatism. In particular, it is important to combine UVVIS spectral data (mineralogy) with Lunar Prospector compositional data (KREEP signature: Th, Sm) to fingerprint the nature of Mg-suite lithologies both inside and outside the PKT, although in practice this is difficult because of the coarse spatial resolution of the LP data (Chapter 2).

4.5. Models for early lunar plutonism

Age dates and reconstructions of the lunar crust using remotely sensed data and mineral fragments paint a complex picture of crust-building by the emplacement of post-LMO magmas. The Mg-suite appears to mark the transition between magmatism associated with the LMO and serial magmatism. This transition period may have occurred as early as 30 m.y. (Shearer and Newsom 2000) to as late as 200 m.y. after LMO formation (Solomon and Longhi 1977; Longhi 1980). The duration of Mg-suite emplacement into the crust is also unclear. The absence of Mg-suite rocks with ages that correspond to the younger episodes of KREEP-basaltic magmatism is possibly a result of a lack of deep sampling because of a decrease in impact flux after ~3.9 Ga. Although it seems logical to assume that the Mg-suite was emplaced into the older ferroan-anorthosite crust, neither crustal reconstructions such as those of Ryder et al. (1997) nor analysis of central peaks confirm this relationship. The study by Ryder et al. (1997) seems to suggest a near-side crust consisting of multiple episodes of Mg-suite emplacement.

Several models for the generation of the Mg-suite have been dismissed based on more recent data: impact origin, LMO origin, remelting and remobilization of late-stage LMO cumulates, and metasomatism of Mg-suite rocks by an incompatible-element-enriched fluid. An impact origin for the Mg-suite had been proposed to explain the chemical paradox of primitive and evolved chemical signatures in the same rocks (Wänke and Dreibus 1986; Taylor et al. 1993). The model proposed by Taylor et al. (1993) combined late accretion impactors (~bulk Moon) as a source for the primitive component with LMO crystallization products (anorthosite, KREEP) as a source for the evolved component. The impact of this material into the Moon during the end of LMO crystallization would have mixed this primitive material with remelted ferroan-anorthosite and residual KREEP liquid. The resulting magmas pooled beneath the ferroan anorthositic crust and subsequently intruded the crust. There are several difficulties with this model. The trace-element fingerprints for impactors, such as elevated siderophile-element abundances, are not found in the Mg-suite. For example, the Cr/Ni ratios for the Mg-suite show a typical lunar value (Cr/Ni > 5) in contrast with primitive cosmic abundances (Cr/Ni = 0.25). There is also a mass-balance problem with this process. The ratio of the mass of the impactor to the mass of the impact melt is too small (O’Keefe and Ahrens 1977) to make a substantial contribution to the high Mg′ observed in the Mg-suite. Formation of the Mg-suite by impact also circumvented the problem of a heat source capable of producing large volumes of primitive, high-Mg magmas following LMO crystallization (Taylor et al. 1993). However, several studies have identified viable processes that could have triggered melting in the deep
Shearer et al. Thermal & Magmatic Evolution of the Moon

lunar mantle (Ryder 1991; Spera 1992; Warren and Kallemeyn 1993; Shearer and Papike 1993). Hess (1994) explored the possibility that the Mg-suite was generated by impact melting of the plagioclase-rich lunar crust and olivine-rich cumulates of the LMO. This eliminates the mass-balance problem in the model of Taylor et al. (1993). Hess (1994) postulated that large impact-melt sheets that were superheated and sufficiently insulated could cool slowly and differentiate to produce the troctolite-norite-gabbro sequence observed in the Mg-suite. Variable incorporation of KREEP and crustal components during shock melting could explain the variation in the evolved component in the Mg-suite plutonic rocks. The siderophile-element signature of the impactor would have been significantly diluted during these processes. If an impact origin for the Mg-suite is correct, one would not expect that the ferroan-anorthosite and Mg-suite chemical trends to be distinctly different (Fig. 4.20). The observed trends cannot be rationalized in terms of mixing of crustal and LMO cumulate components. Hess (1994) also pointed out that impact melts involving a substantial upper-mantle component should have olivine on the liquidus and olivine-rich mineral assemblages are expected. However, dunites are not well represented in the Mg-suite, at least in the returned sample suite and on the lunar surface. Finally, if a substantial upper-mantle component was incorporated into surface impact-melt sheets, it should be expected that upper mantle lithologies would have been excavated and incorporated into the lunar regolith. No such samples have been found.

Wood (1975), Longhi and Boudreau (1979) and Raedeke and McCallum (1980) proposed models in which the Mg-suite was produced during LMO crystallization and evolution. Wood (1975) suggested that the Mg-suite and ferroan anorthosites were simply contemporaneous products of crystal accumulation and trapped melts. In this scenario, the Mg-suite intrusions consist of cumulus olivine/pyroxene plus plagioclase, whereas the ferroan anorthosites consisted of cumulus plagioclase with mafic crystals produced from intercumulus melts. In addition to the differential incorporation of cumulates and trapped melts, Longhi and Boudreau (1979) proposed that the cumulus minerals in both rock types were produced by different styles of crystallization. The plagioclase in the ferroan anorthosites were products of equilibrium crystallization. The cumulate mafic silicates of the Mg-suite precipitated at approximately the same time, but under conditions of fractional crystallization. Raedeke and McCallum (1980) demonstrated that minerals from the banded zone in the Stillwater Complex showed two fractionation trends remarkably similar to the fractionation of Mg/(Mg+Fe) of mafic minerals versus An of plagioclase exhibited by the lunar highland plutonic rocks (see Figure 4.20). They attributed the bimodality of the Stillwater rocks to differences in the style of crystallization (fractional crystallization accompanied by crystal accumulation versus equilibrium crystallization of trapped intercumulus liquid in a plagioclase-rich crystal mush). A contemporaneous relationship between members of these suites is also suggested by the overlapping ages for some of the Mg-suite rocks with the ferroan anorthosites (Shih et al. 1993).

The above models are not consistent with all observations made for the Mg-suite. First, the ages for many Mg-suite rocks postdate the ferroan anorthosites that may be derived from LMO crystallization (Papanastassiou and Wasserburg 1975; Carlson and Lugmair 1981a,b; Nyquist and Shih 1992; Shih et al. 1994). Some of the apparent overlap of ages between the ferroan anorthosites and the Mg-suite is the combined result of an early transition between the LMO and serial magmatism (Shearer and Newsom 2000) and an inaccurate interpretation of the crystallization ages for these early crustal rocks (Papike et al. 1998). Second, detailed geochemical characteristics are not consistent with a near contemporaneous origin of the two magmatic suites from the LMO. For example, the Mg-suite parental magmas have high initial Sr, Sm, and Eu/AI, but low Sc/Sm and Ti/Sm relative to the melts parental to the ferroan anorthosites (Norman and Ryder 1979; Raedeke and McCallum 1980; Warren and Wasson 1980; James and Flohr 1983; Warren 1986). These KREEPy signatures indicate that the Mg-suite cumulates crystallized after the development of KREEP.
The above difficulties prompted the development of numerous models that advocate a post-LMO igneous origin for the Mg-suite that involves the melting of LMO cumulates or even an undifferentiated lunar mantle. These models call upon either the melting of shallow, evolved LMO cumulates, or melting of the deep lunar mantle (LMO cumulates, or primitive lunar mantle) with incorporation of KREEP or crustal component through assimilation, mixing or metasomatism. Hess et al. (1978), Hess (1989), Snyder et al. (1995), Papike et al. (1994, 1996, 1997) and Korotev (2000) explored the possibility that the parental magmas for the Mg-suite originated by partial melting of LMO cumulates at a shallow depth. Although they did not draw a genetic connection between the Mg-suite and KREEP basalts, Hess et al. (1978) suggested that the KREEP-like highland basalts were generated by partial melting of LMO cumulates that crystallized soon after extensive ilmenite crystallization, but prior to the formation of KREEP (between 95 and 99% crystallization of the LMO). These magmas may have assimilated KREEP. Alternatively, Hess (1989) explored the possibility that the KREEP-rich magmas were a product of partial melting of a lower lunar crust that had been altered metasomatically by KREEP. Snyder et al. (1995) and Papike et al. (1994, 1996, 1997) suggested that pressure release melting and remobilization of these rock types may be related to catastrophic impacts on the lunar surface. However, pressure-release melting is unlikely to be important for the deep lunar crust and very shallow mantle as the pressures in those environments are rather low (i.e., 5-10 kb).

Neal and Taylor (1991) suggested that a metasomatic event occurred after emplacement and crystallization of the Mg-suite magmas in the lunar crust. Korotev (2000) proposed that the Mg-suite magmas are the product of dissolution of anorthosite and the most primitive olivine cumulates into the late-stage, KREEP-enriched liquid of the LMO. The KREEP-rich liquid would become multisaturated with both plagioclase and olivine. Although Hess and Parmentier (2001) discounted this model based on geochemical rationale, they did provide a thermal model for this dissolution process. They suggested that because the KREEP layer contains more than 200× chondritic heat-producing elements, it would undergo reheating and would grow by dissolving adjacent lithologies.

In most cases, these models do not account for the substantial observational data for the Mg-suite. Models that advocate remobilization of the KREEP horizon generally cannot account for the extraordinarily high incompatible-element content and the high Mg' of the Mg-suite (Shearer and Papike 1999; Hess and Parmentier 2001). Nor can they account for higher Ni abundance and lower Cr/Ni of the Mg-suite. The mechanism in which a large melt layer is produced just below the crust would retard the transport of later mare magmas to the lunar surface and impede the preservation of mascons associated with large impact basins (Hess and Parmentier 2001). The subsolidus addition of a “KREEP-like” component to the Mg-suite lithologies is not likely because ion probe studies indicate that the cores of large olivine and orthopyroxene grains in the Mg-suite have a KREEP signature. There is no evidence that these components diffused across these grains from a metasomatic fluid.

Models that require the Mg-suite to represent crystallization products of high Mg' magmas generated in the deep lunar mantle were developed to resolve some of the problems with a shallow cumulate source (James 1980; Warren and Wasson 1980; Longhi 1981, 1982; Morse 1982; Smith 1982; Hunter and Taylor 1983; James and Flohr 1983; Shirley 1983; Shervais et al. 1984; Warren 1986; Ryder 1991; Hess 1994). Hess (1994) demonstrated that magmas with Mg' appropriate for Mg-suite parent magmas could be generated by melting of early LMO cumulates. An LMO with an Mg' value equivalent to the bulk Moon (80 to 84; Jones and Delano 1989; O’Neill 1991) upon crystallization at high pressures would produce early cumulates of olivine having Mg' greater than 91. Subsequent melting of these cumulates would produce magmas with Mg' equivalent to that of the parental Mg-suite magmas. Because of the pressure dependence of the FeO-MgO exchange equilibrium between olivine and basaltic melt, crystal-
lization of these high-pressure melts near the lunar surface would result in liquidus olivine that is slightly more magnesian than residual olivine in the mantle source. Higher Mg’ values for the melt may result from the reduction of small amounts of FeO to Fe in the source (Hess 1994).

Melting of the early LMO cumulates initially could have been triggered by either radioactive decay (Hess 1994) or cumulate overturn (Hess and Parmentier 1994), or both. The deep lunar mantle materials would have been less dense than the overlying cumulates and tend to move upward and be subjected to pressure-release melting. This would have resulted in relatively high degrees of melting (>30%; Ringwood 1976; Kesson and Ringwood 1977; Herbert 1980). Partial melting of early LMO cumulates could have produced primitive melts with high Mg’, but these magmas would not have the same geochemical characteristics as the Mg-suite. For example, these primitive magmas would not possess the high incompatible-element enrichments, fractionated Eu/AI and Na/(Na+Ca), low Ni and plagioclase as a liquidus phase until the Mg’ of the melt was <42. Two types of processes have been proposed to resolve this problem: assimilation of KREEP and melting of a hybrid KREEP-early cumulate sources.

For KREEP assimilation, Warren (1986) calculated that if these high-Mg magmas assimilated ferroan-anorthosite and KREEP, they would have reached plagioclase saturation at values of Mg’ appropriate for Mg-suite magmas. Such magmas also would have inherited a fractionated incompatible-element signature (high REE, fractionated Eu/AI). Hess (1994) explored the thermal and chemical implications of anorthosite melting and plagioclase dissolution by high-Mg basaltic magmas. In his analysis of anorthosite melting and mixing as a mechanism to drive high Mg’ magmas to plagioclase saturation, Hess concluded that the resulting crystallization of olivine and the mixing of relatively Mg-poor coticet melts produced from the anorthosites would lower the Mg’ of the hybrid melt below that expected for the Mg-suite parental magmas. In addition, diffusion rates for Al₂O₃ in basaltic melts are extremely slow (Finnila et al. 1994) and indicate that the time scales to dissolve even a small amount of plagioclase are of the same order as the characteristic times of solidification of a large magma body. Similar thermal constraints are less severe for the shallow melting-KREEP assimilation by primitive Mg-suite magma. Although KREEP assimilation does not dramatically drive the Mg-rich magmas to plagioclase saturation, this mechanism may account for the evolved trace-element signatures in the Mg-suite. However, mixing of relatively viscous melts of KREEP composition with more-fluid Mg-rich magmas could be prohibitive (Finnila et al. 1994). In addition, melt compositions for the Mg-suite norites that were calculated from pyroxene trace element data have a KREEP component equivalent to or slightly higher than KREEP (Papike et al. 1997). Simple mass-balance calculations indicate that it may be impossible for primitive Mg-suite magma to assimilate such an abundant amount of KREEP (Shearer and Floss 1999). However, it is possible that the lower Mg’ of the norites along with their elevated REE may reflect assimilation of KREEP and fractional crystallization of a more primitive basalt that was parental to the troctolites. Thus-far unevaluated aspects of assimilation models are the assimilation of a KREEP horizon with a temperature elevated by heat-producing elements, a lunar mantle capable of producing parental magma with a higher Mg’, and the effect of crustal processes on the subsolidus redistribution of elements in the Mg-suite lithologies (i.e., pyroxene reequilibration with intercumulus melt, subsolidus reequilibration of primary olivine or pyroxene compositions; McCallum and Schwartz 2001).

As an alternative to assimilation at shallow mantle levels, is it possible that the KREEP component was added to the deep mantle source for the Mg-suite? Hess (1994) proposed that the source for the Mg-suite may be a hybrid mantle consisting of early LMO cumulates (dunite) and a bulk Moon component that may be either primitive lunar mantle or an early quenched LMO crust. Polybaric fractional melting of a 50-50 mixture would produce melts with appropriate Mg’ and reasonable Al₂O₃. However, these mixing components would not produce some of the incompatible-element signatures exhibited by the Mg-suite. Shearer et
Shearer et al. (1991) suggested a cumulate overturn mechanism to transport a KREEP component to the deep lunar mantle, thus explaining the evolved KREEP signature imprinted on selected picritic glasses associated with mare basaltic magmatism. A similar process may have produced the KREEP signature in the Mg-suite. A potential pitfall of this model is the Ni content of the olivine in Mg-suite rocks. Although the Mg′ of olivine (Fo$_{92-88}$) in the Mg-suite is high and suggests a primitive magma derived from deep within the LMO cumulate pile, the Ni content of the olivine determined by electron microprobe (Ryder 1982) and ion microprobe (Shearer and Papike 2000; Shearer et al. 2001) ranges from 10 to 350 ppm. This is fairly low for primitive lunar basaltic magma. In comparison, ion microprobe analyses of olivine cores (Fo$_{75-72}$) from Apollo 12 olivine basalts yield Ni concentrations of 450-515 ppm (Papike et al. 1999). The calculated olivine composition in equilibrium with mare basalts more primitive than the Apollo 12 olivine basalts (i.e., Apollo 15 green glass) is approximately Fo$_{85}$ and Ni = 1500 ppm. This poses a dilemma for interpretation of the high Mg′ of the Mg-suite and hints at the involvement of a metal phase either during melting of the source or during evolution of the basaltic magma (Papike et al. 1997). Alternatively, it could suggest that the parental magmas for the Mg-suite are not derived from deep within the LMO cumulate pile, but from a lunar mantle that was not part of the LMO cumulate lithologies.

The most reasonable models for the generation of the parental magmas for the Mg-suite and other highland plutonic and volcanic rocks involve the incorporation of a KREEP component. How the KREEP component is incorporated into these magmas is the point of debate. How would the asymmetrical distribution of KREEP on the Moon, as indicated by remotely sensed data, influence the formation and distribution of the Mg-suite magmas under these two petrogenetic scenarios (shallow assimilation versus deep melting of a hybrid mantle)? If magmas parental to the Mg-suite where produced by a two-stage process involving mantle melting followed by KREEP assimilation, there should be two types of high-Mg plutonic rocks: those with a KREEP signature spatially associated with the PKT on the near side of the Moon and those without a KREEP signature emplaced into ferroan anorthosites in non-KREEPy lunar terranes. In the scenario in which KREEP in the mantle source drives melting, it should be anticipated that only KREEP-enriched Mg-suite rocks will exist and they will be closely associated with the PKT on the near side of the Moon. Combining the remote sensing approach taken by Tompkins and Pieters (1999) with gamma-ray chemical signatures of KREEP may provide additional insight into distinguishing between these two petrogenetic models.

The presence of a trace-element KREEP signature and Nd isotopic systematics implies a link not only among the rocks making up the Mg-suite, but also among the Mg-suite, alkali suite, evolved plutonic rocks, and the KREEP basalts. As a first approximation, Warren (1988) and Snyder et al. (1995, 2000) suggested that fractional crystallization of a KREEP basalt could potentially produce the cumulate lithologies making up the Mg-suite. The KREEP basalts in the lunar collection have bulk rock Mg′s of approximately 61 to 66. This range in Mg′ is significantly lower than would be expected for parental magmas of the Mg-suite cumulates (mafic minerals with Mg′ ≥ 90). However, it is possible that more Mg-rich KREEP basalts exist but have not yet been sampled.

The Mg-suite and evolved plutonic rocks may be either part of a continuum of crystallization products of parental basaltic magmas similar to the KREEP basalts (Ryder 1976; James 1980; Warren and Wasson 1980; James et al. 1987; Snyder et al. 1995; Shervais and McGee 1999) or separate episodes of basaltic magmatism (Warren and Wasson 1980). Using a KREEP basalt (from the Apollo 15 site) as a starting parental magma, Snyder et al. (1995) demonstrated that fractional crystallization and the accumulation of mineral phases and trapped KREEP-like residual liquid (2-15%) could produce the range of mineral and rock compositions observed in the highland Mg- and alkali suites. The sequence of crystallization and accumulation that they proposed is Mg-suite (0-43% crystallization) → alkali anorthosites,
alkali norites (43-74% crystallization) → alkali gabbros, alkali norites (74-90% crystallization) → quartz monzodiorites (90-99.8% crystallization).

Shevrais and McGee (1997, 1999) envisioned a more complex process to account for the REE characteristics of the alkali-suite magmas and the difference in the albite content of the plagioclase between the calculated liquid lines of descent for KREEP magmas and alkali-suite magmas. They concluded that in addition to fractional crystallization, assimilation of ferroan anorthosites, magma mixing, and local equilibrium crystallization were important processes relating these magmas.

An alternative model is one in which the Mg-suite and alkali-suite magmas represent contemporaneous, but separate, episodes of basaltic magmatism (James 1980; Warren and Wasson 1980; James et al. 1987). There is some compositional evidence to suggest genetically distinct highland rock types. For example, James et al. (1987) subdivided many of these highland rock types into various groups on the basis of their mineral chemistry and mineral associations. Whether these subdivisions are artificial or petrologically significant is unclear. Within this scenario, however, differences between the two suites may be attributed to the depth of initial melting prior to assimilation. For example, Mg-suite magmatism would be a product of deep mantle melting followed by KREEP assimilation just below the lunar crust, whereas the more evolved magmas would involve initial melting at shallower mantle levels followed by assimilation. Hunter and Taylor (1983) suggested that the parental magmas of the alkali suite could represent the “dregs” of the LMO.

Materials of granitic composition occur as immiscible glasses within the mesostasis of mare basalts, in melt inclusions, as glasses in selective lunar soils, as crystalline clasts in breccias, and as segregations in monzogabbro (Papike et al. 1998; Jolliff et al. 1999). In all petrologic models, the granite, rhyolite, and felsite clast are considered to be products of extensive fractional crystallization of alkali-suite parental magma. Marvin et al. (1991) proposed that these evolved granitic rock types were produced through the crystallization and removal of phosphates and zircon from parental magmas equivalent to the quartz monzodiorite of the alkali suite. This fractional crystallization mechanism accounted for the enrichment of REE, Zr, and Hf in the quartz monzodiorites relative to the more evolved lunar granites. It has been also proposed that the lunar granites are a product of liquid immiscibility following extensive fractional crystallization of a KREEP basalt. Experimental studies by Hess et al. (1975) illustrated that after 90-95% of crystallization of a variety of lunar basaltic magmas, spherules of granitic magma exsolved from the residual ferrobasaltic liquid. Chemical signatures of these lunar granites have been attributed to primarily immiscibility (Quick et al. 1977; Taylor et al. 1980; Warren et al. 1983; Neal and Taylor 1989; Snyder et al. 1993). On the other hand, Jolliff (1991), Jolliff et al. (1993, 1999) and Shearer et al. (1998) similarly argued that many of these chemical signatures were a result silicate-liquid immiscibility, but after phosphate crystallization.

4.6. Volcanic rocks, KREEP basalts

4.6.1. Introduction. KREEP basalts were returned primarily by the Apollo 15 and Apollo 17 missions. Small and rare fragments of KREEPy basalts from the Apollo 14 and 16 and Luna 20 sites have been documented (Salpas et al. 1987; Jolliff et al. 1991; Nyquist and Shih 1992), but only the Apollo 15 and 17 basalts lack elevated siderophile-element abundances attributed to meteorite contamination. The KREEP basalts occur primarily as numerous small fragments and particles among the rake material, fines and as clasts in breccias. Only a few samples of the KREEP basalts are over 1 g, 72275 (clast 91, 2.73 g), 15382 (3.2 g), and 15386 (7.5 g). Most are subophitic to intersertal in texture. The mineralogy is dominated by pyroxene and plagioclase. In the Apollo 15 basalts many pyroxenes have cores of orthopyroxene, overgrown with pigeonite and then augite. The pyroxene in the Apollo 17 basalts is primarily pigeonite with minor augite. Olivine is extremely rare. The KREEP basalts are distinguished from the mare
basalts by their higher plagioclase abundances and their highly enriched concentrations of incompatible minor and trace elements that have elemental ratios similar to the “type” KREEP samples (soils and breccias from Apollo 14) (Fig. 4.27). Compared to the Apollo 15 KREEP basalts, those from Apollo 17 are more iron-rich, lower in incompatible elements and more depleted in heavy REE. A list of KREEP basalts in the Apollo collection and more detailed description of the petrography are presented by Papike et al. (1998) and in Chapter 3.

The KREEP basalts sampled from the Apollo missions have crystallization ages of 3.84 to 4.08 Ga (Nyquist and Shih 1992). The Apollo 15 KREEP basalts (3.84 ± 0.02 Ga) are younger than the KREEP basalts from the Apollo 17 site (≈ 4.01 to 4.08 Ga). Based strictly on sample mineralogy, geochemistry and isotopic data, the relationship among KREEP basaltic volcanism, mare basaltic volcanism, and basin formation is equivocal. The Apollo 15 KREEP basalts may represent volcanic eruptions nearly contemporaneous with the formation of the Imbrium basin and predated the earliest stages of Imbrium-filling mare magmatism. The KREEP volcanism represented by the Apollo 17 samples predates the formation of Serenitatis basin and associated basin-filling basalts. The oldest mare basalts overlap with the ages of some of these KREEP basalts, High-Ti basalts as old as 3.88 ± 0.06 Ga are found at the Apollo 11 site (Papanastassiou et al. 1977).

More recently, an olivine-gabbro cumulate clast with a significant KREEP signature was identified in lunar meteorite Northwest Africa 773 (Fagan et al. 2001; Jolliff et al. 2003, Borg et al. 2004). The Ar-Ar age of the cumulate lithology is ~2.7 Ga (Fernandes et al. 2002). The Sm-Nd age for the clast is ~2.9 Ga (Borg et al. 2004). This sample indicates that the duration of KREEP basaltic magmatism continued for at least another 1 billion years after the eruption of the Apollo 15 KREEP basalts.

4.6.2. Isotopic studies. Isotopic data may be able to answer questions concerning the age relationship between KREEP basalts and plutonic highland rocks and mare basalts and the petrogenesis of the KREEP basalts. As was stated earlier, KREEP basalts erupted to the lunar surface between 3.84 and 4.08 Ga (Nyquist and Shih 1992). This overlaps with the eruption of mare basalts. Papanastassiou et al. (1977), Snyder et al. (1996) and Shih et al. (1999) confirmed that some of the Apollo 11 high-Ti basalts were as old as 3.88 Ga. Misawa et al. (1993) determined a Sm-Nd age of 3.87 ± 0.06 Ga on basaltic lunar meteorite Asuka 881757 that had a composition that are is similar to the low-Ti to very low-Ti basalts. Compared to the ages for KREEP basalt volcanism, emplacement of Mg-suite magmas into the lunar crust is considerably older (Sm-Nd ages range from 4.18 ± 0.07 Ga to 4.46 ± 0.07 Ga). Figure 4.22 shows the initial εNd for KREEP basalts compared to highland Mg-suite rocks. Data for most of the KREEP basalts plot along the same evolution line as do the data for the Mg-suite (Nyquist and Shih 1992). NWA 773 deviates from this trend. It has an initial εNd of −7.84 and a 167Sm/144Nd ratio for its source region of 0.161 (Borg et al. 2004). This indicates that its source is more LREE-enriched than the sources of the other KREEP basalts and Mg-suite rocks.

4.6.3. Trace-element mineralogy. Trace-element data of mineral phases in KREEP basalts are essentially absent in the literature. Most of the available ion microprobe data is
from the olivine-gabbro clast in NWA 773 (Jolliff et al. 2004; Shearer and Borg 2004). The mineral phases such as plagioclase and pyroxene exhibit substantial incompatible-element enrichments relative to similar phases in mare basalts. This is consistent with the isotopic measurements that suggest that a substantial KREEP component has been added to this basalt. Olivine in the clast has Ni and Co concentrations exceeding that found in the Mg-suite and high-Al basalts (Fig. 4.24). The Ni and Co concentrations in olivine overlap or exceed that found in olivine in low-Ti mare basalts.

4.7. Volcanic rocks, high-Al and K-rich basalts

4.7.1. Introduction. The high-Al basalts (>11 wt% Al₂O₃) are enriched in Al relative to other lunar basalts, but are not “high-Al” relative to terrestrial basalts. They were returned by the Apollo 14 and Luna 16 missions and bear a close resemblance to low-Ti mare basalts based on similar contents of TiO₂ (2-3 wt%), FeO (14-18 wt%) and CaO (9-11 wt%) (Table 4.2). Feldspathic basalts from other collection sites (e.g., the one Apollo 12 feldspathic basalt 12038 and the evolved Luna 24 VLT basalts) are not considered in this discussion because they may be either unrepresentative samples, fractional crystallization products of melts that originally contained lower abundances of Al₂O₃, or may not have been derived from the location where they were collected. As with other vestiges of older magmatism, the study of these rocks has been hampered by sample size and the lack of geological context. The importance of both the Apollo 14 and Luna 16 rocks is that they perhaps represent the oldest record of lunar volcanism thus far recognized. The high-Al basalts clasts from the Apollo 14 site have ages that range from 4.0 to 4.3 Ga (Taylor et al. 1983; Dasch et al. 1987; Nyquist and Shih 1992; Shih et al. 1992), up to 200 m.y. older than the oldest known KREEP basalt. It is still a point of debate whether the high-Al basalts represent non-mare basaltic magmatism (Hubbard and Gast 1972), high-Al mare basaltic magmatism (Ridley et al. 1975), or impact melts (Snyder et al. 2000).

In addition to the high-Al basalts, Shervais et al. (1985b), Goodrich et al. (1986) and Neal et al. (1988b, 1989a,b) described basaltic clasts with relatively high potassium contents (i.e., > 0.5 wt% K₂O) in polymict lunar breccias from the Apollo 14 site. These were described as high-K (Goodrich et al. 1986) or very high-K basalts (Shervais et al. 1985b) which we will collectively call “K-rich” basalts. A description of these basalts is presented in Chapter 3.

4.7.2. Trace element studies. Although high-Al basalts were first reported during the early days of lunar sample studies (i.e., Helmke et al. 1972; Hubbard et al. 1972), it was not

---

**Table 4.2.** Comparison of bulk compositions of high alumina basalts and low-TiO₂ mare basalts.

<table>
<thead>
<tr>
<th></th>
<th>High-Alumina Basalt</th>
<th>Low-TiO₂ Mare Basalt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14053</td>
<td>14072</td>
</tr>
<tr>
<td>SiO₂</td>
<td>46.4</td>
<td>45.2</td>
</tr>
<tr>
<td>TiO₂</td>
<td>2.64</td>
<td>2.57</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>13.6</td>
<td>11.1</td>
</tr>
<tr>
<td>Cr₂O₃</td>
<td>0.40</td>
<td>0.51</td>
</tr>
<tr>
<td>FeO</td>
<td>16.8</td>
<td>17.8</td>
</tr>
<tr>
<td>MnO</td>
<td>0.26</td>
<td>0.27</td>
</tr>
<tr>
<td>MgO</td>
<td>8.48</td>
<td>12.2</td>
</tr>
<tr>
<td>CaO</td>
<td>11.2</td>
<td>9.8</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.44</td>
<td>0.32</td>
</tr>
<tr>
<td>K₂O</td>
<td>0.10</td>
<td>0.08</td>
</tr>
<tr>
<td>Total</td>
<td>100.3</td>
<td>99.9</td>
</tr>
</tbody>
</table>

Note that the low Ti mare basalts have higher FeO but lower Al₂O₃ contents. From Papike et al. 1997.
until detailed studies of clasts in lunar breccias (i.e., 14303, 14304, 14305, and 14321) during the 1980s that the importance and complexity of these possible remnants of early basaltic volcanism were first recognized (Taylor et al. 1983). The elemental data from these Apollo 14 clasts reported during the mid-1980s to the early 1990s indicated an extremely wide variation in incompatible elements both in overall concentration and REE slope with limited variation in major element whole rock chemistry (i.e., Mg' = 57-51) and mineral composition (Dickinson et al. 1985; Shervais et al. 1985a,b). The database for these basalts was expanded considerably by Neal et al. (1988, 1989, 2003). The compilation of all the analyses of Apollo 14 high-Al basalts confirm the earlier studies showing a strong positive correlation between La and Hf over an extremely wide range in composition (Fig. 4.28a).

The K-rich basalts have the same major-element composition as the high-Al basalts, but with elevated contents of K, Ba, and Rb (Shervais et al. 1985b). They exhibit the same general positive correlation between La and Hf as the high-Al basalts, although in some samples La and Hf is significantly greater. The concentration of Sm and Th overlap or exceed the range observed in the high-Al basalts and generally have higher Th/Sm ratios (Fig. 4.28b). The K/La ratios for most mare basalts and even KREEP basalts typically fall between 30 and 100 whereas the VHK basalts have ratios that exceed 1000. Chondrite-normalized trace-element patterns show a negative Eu anomaly, a flat to negatively sloped REE pattern with REE abundances up to ~40× chondritic, and with low Sc (≤60 ppm).

4.7.3. Isotopic studies. Isotopic data for the high-Al and K-rich basalts have been summarized by Nyquist and Shih (1992), Shearer and Papike (1998), and Snyder et al. (2000). Since the recognition that the high-Al basaltic clasts may represent old episodes of basaltic volcanism (Taylor et al. 1983), subsequent studies have confirmed the initial observations (Dasch et al. 1987; Neal and Taylor 1990). The work by Dasch et al. (1987) extended the age of the high-Al basalts back to 4.29 ± 0.13 Ga. The combination of all age dates up to that time led Neal and Taylor (1990) to conclude that distinct periods of high-Al basaltic volcanism occurred at 4.3, 4.1, and 3.95 Ga.

Figures 4.29 and 4.30 show initial $^{87}\text{Sr}/^{86}\text{Sr}$ and $\varepsilon_{\text{Nd}}$ for the Apollo 14 high-Al basalts as a function of their ages. Dasch et al. (1987) noted that the data for the basalts defined two distinct

![Figure 4.28](https://example.com/figure428.png)

**Figure 4.28.** (a) Hf vs. La for the high-Al and high-K basalts. Groups outlined are from Dickinson et al. (1984). Also plotted are the high-K basalts (compiled by Clive Neal) (b) Sm versus Th for the high-Al and high-K basalts.
reservoirs with different Rb/Sr. Groups 1, 2, 3, and 5 were consistent with radiogenic growth of $^{87}\text{Sr}$ in a reservoir of uniform Rb/Sr $= 0.021$, whereas groups 4 and $5'$ are consistent with radiogenic growth in a reservoir with Rb/Sr a factor of two lower. In terms of Sm-Nd isotopes, the high-Al basalts plot at $\varepsilon_{\text{Nd}}$ values greater than 0 and do not plot along the $^{147}\text{Sm}/^{144}\text{Nd}$ evolution line of 0.169 defined by the Mg-suite, the evolved plutonic rocks, and the KREEP basalts.

The K-rich basalts are characterized by $\varepsilon_{\text{Nd}}=0$ implying derivation from a long-lived chondritic reservoir (Shih et al. 1986). Because of large uncertainties, the ($^{87}\text{Sr}/^{86}\text{Sr}$)$_1$ and the implied Rb/Sr ratios are not well defined, but appear to be roughly comparable to that of Apollo 14 high-Al basalts. The initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio is low ($0.6995 \pm 0.0005$) and significantly lower than KREEP basalts (>0.700).

4.7.4. Trace-element mineralogy. One of the obvious problems with understanding the petrogenesis of these rocks is whether the compositional variations observed and modeled are a product of magmatic or impact processes, or a function of unrepresentative sampling. To this end, Hagerty et al. (2002, 2003, 2004) analyzed selected olivine-bearing high-Al basalt clasts representing incompatible-element enriched and depleted basaltic lithologies as defined by Shervais et al. (1985a), Dickinson et al. (1985, 1989), Neal et al. (1988a, 1989a,b), and Hughes et al. (1990). Plagioclase in these clasts is enriched in REE (Fig. 4.31), Ba, and Sr in the same whole-rock enrichment sequence observed by Dickinson et al. (1985) (Fig. 4.28a). In addition,
the Ce/Yb ratios for plagioclase increase (∼5 → 7 → 19) with increasing REE content (Fig. 4.31a). These indicate that the basaltic melts from which these plagioclase grains crystallized exhibited different REE contents and Ce/Yb (i.e., LREE enrichments). Assuming $D's$ for the plagioclase do not significantly change, the REE concentration in the melt increased by $\sim 10\times$ between the depleted and enriched basaltic clasts. This incompatible-element enrichment is on the same scale as predicted by whole-rock analyses of the individual clasts.

Although whole-rock analyses show limited variation in compatible elements (i.e., Co, Ni), the olivine from clasts show that each fragment exhibits different Ni-Co concentrations and behavior (Fig. 4.32). The basaltic clasts with the lowest incompatible element concentrations in the bulk rock and in plagioclase contain olivine with cores exhibiting the highest Ni abundances (360 ppm). Conversely, the clasts with the highest incompatible-element concentrations contain olivine with cores exhibiting the lowest Ni concentrations (80 ppm). Nickel decreases from the olivine cores to their rims.

The ion-microprobe analyses of olivine and plagioclase in Apollo 14 high-Al basalt clasts indicate that both compatible and incompatible elements exhibit variability consistent with compositional differences observed in the bulk samples. This indicates that the 10- to 20-fold increase in incompatible elements observed in the bulk clasts is real and not simply a product of unrepresentative sampling (i.e., higher proportions of incompatible-element-enriched mesostasis). Simple fractional crystallization of a single basaltic magma or melting of a single mantle source cannot account for the range of incompatible-element enrichments observed in the olivine-bearing high-Al basalts. The range of ages exhibited by these basalts also argues against simple fractional crystallization. Instead, the range observed in the high-Al clasts represents distinct pulses of basaltic magma (Neal and Taylor 1990; Hagerty et al. 2002, 2003; Kramer and Neal 2003) that were produced by the melting of distinct mantle source or different degrees of assimilation.

**Figure 4.31.** Chondrite-normalized rare earth element patterns for plagioclase from the high-Al basalts (Black = sample 14321, 1245; Gray = sample 14321, 1246; White = sample 14321, 1261). (Hagerty et al. 2002).

**Figure 4.32.** Ni vs. Co for olivine from the high-Al basalts (Hagerty et al. 2002).
4.8. Remotely sensed data on pre-mare volcanic rocks

4.8.1. Introduction. The global distribution of pre-mare lithologies on the Moon is difficult to describe, because to date, no contiguous deposits of non-mare volcanism have been unequivocally identified. This may result from several factors. Current remote sensing techniques are not able to distinguish the pre-mare rock types from highland or mare materials, or mixtures thereof. Additionally, these ancient lithologies may correspond to deposits that have been buried or obscured by subsequent events or that are intrusive in nature, making them difficult to observe remotely. Despite these obstacles, some advances have been made.

4.8.2. Hidden mafic deposits. The presence of ancient basalt clasts in the Apollo sample record, some as old as 4.23 Ga (Taylor et al. 1983), contrasts sharply with the age range of observed mare deposits, which are only 3.9-2.5 Ga (Head 1976). When Schultz and Spudis (1979) first recognized the existence of hidden mafic deposits, they proposed that these represented the contiguous bodies from which such ancient basalt clasts originated. Since then, many hidden mafic deposits, termed “cryptomaria” by Head and Wilson (1992) have been proposed and identified.

By definition, cryptomaria are strictly mare basalt deposits whose low albedo signature has been hidden or obscured by superposed high albedo material (Antonenko et al. 1995; Fig. 4.33). The thickness of the overlying ejecta can range from a few kilometers near large impact basins to only several meters near small craters. Thus, the smooth surface topography of hidden deposits is not always guaranteed, since thick basin ejecta may be very textured (Head 1974). Criteria for the detection of cryptomare deposits include the presence of dark-haloed craters (craters that penetrate through the overlying ejecta to excavate underlying mafic material) and the identification of spectral or geochemical anomalies (Antonenko et al. 1995). However, mare basalts can be difficult to distinguish from other mafic materials when only remote sensing tools are available, and especially if the deposits in question are obscured by other materials. As a result, the term “cryptomare” has been used to include all hidden mafic deposits, regardless of specific chemical or mineralogical composition (Antonenko et al. 1995).

A global census of all previously identified and suspected hidden mafic deposits, based on a survey of existing literature, was conducted by Antonenko (1999). A rough estimate of the potential surface area and volume was calculated for each deposit, and a possible age range determined (Fig. 4.26 and Table 4.3). This list provides a good start for understanding the importance of these deposits and their importance for early lunar volcanism. Not all such deposits are necessarily included in the list of Table 4.3. Additional hidden mafic deposits may be found upon further investigation. Also, Nectarian or pre-Nectarian ejecta deposits

Figure 4.33. Illustration showing the formation of a cryptomare. High albedo basin ejecta is emplaced on top of a pre-existing low albedo mafic deposit. These hidden mafic may be sampled by dark halo impact craters, which penetrate the regional ejecta to excavate mafic material, emplacing it in a halo around the crater.
that obscure even older mafic units may lie beneath the circular maria (Antonenko 1999). The totals given in Table 4.3 are therefore considered minimum estimates for the area and volume of all hidden mafic deposits on the Moon.

A comparison between these hidden mare deposits and known maria can be made on the basis of such totals. Maria have a total surface area of approximately $6.3 \times 10^6$ km$^2$ (Head 1975b) and a total volume of approximately $5 \times 10^6$ km$^3$ (Hörz 1978; Budney and Lucey 1998; Antonenko 1999). Hidden mafic deposits may, therefore, cover as much of the lunar surface as half of the known maria and have a volume that is equal to 30% of the known mare volume (Antonenko 1999). It is clear that hidden mafic deposits represent a significant contribution to global lunar volcanism.

The age ranges given in Table 4.3 can be used to assess the volcanic flux of these mafic materials. Considering only units that are unambiguously older than 3.8 Ga, the data indicate that volcanic flux prior to 3.8 Ga may have been as high as 60% of the flux after 3.8 Ga (Antonenko 1999). It is possible that early volcanic fluxes were comparable to those in the late Imbrian period when volcanism is believed to have peaked (Head and Wilson 1992).

The specific nature of these hidden mafic deposits is not well understood. Owing to their proximity to mare deposits, they are generally assumed to represent an earlier phase of mare volcanism. Because of their presence close to the surface, they are generally believed to be extrusive. The likelihood of shallow intrusive sills is often ruled out because the current models of mare emplacement do not allow for them (Head and Wilson 1992). However, although no

### Table 4.3. List of suspected cryptomare deposits and their characteristics.

<table>
<thead>
<tr>
<th>Cryptomare Deposit</th>
<th>Area ($\times 10^5$ km$^2$)</th>
<th>Thickness (m)</th>
<th>Volume ($\times 10^5$ km$^3$)</th>
<th>Age (Ga)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schiller-Schickard</td>
<td>3.6</td>
<td>400</td>
<td>1.5</td>
<td>3.8 - 3.92</td>
</tr>
<tr>
<td>Humorum</td>
<td>2.3</td>
<td>450</td>
<td>1.0</td>
<td>3.8 - 3.92</td>
</tr>
<tr>
<td>Procellarum</td>
<td>0.35</td>
<td>400</td>
<td>0.15</td>
<td>3.8 - 3.92</td>
</tr>
<tr>
<td>Austral</td>
<td>6.4</td>
<td>500</td>
<td>3.2</td>
<td>3.8 - 3.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.92 - 4.0</td>
</tr>
<tr>
<td>Balmer</td>
<td>1.7</td>
<td>400</td>
<td>0.68</td>
<td>3.87 - 4.1</td>
</tr>
<tr>
<td>East Frigoris</td>
<td>3.7</td>
<td>400</td>
<td>1.5</td>
<td>3.8 - 3.92</td>
</tr>
<tr>
<td>Langemak</td>
<td>4.0</td>
<td>400</td>
<td>1.6</td>
<td>3.92 - 4.1</td>
</tr>
<tr>
<td>Smythii</td>
<td>2.0</td>
<td>400</td>
<td>0.80</td>
<td>3.8 - 3.84</td>
</tr>
<tr>
<td>Taruntius</td>
<td>0.87</td>
<td>600</td>
<td>0.52</td>
<td>~3.8</td>
</tr>
<tr>
<td>South Pole-Aitken</td>
<td>2.5</td>
<td>400</td>
<td>1.0</td>
<td>3.63 - 3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.9 - 4.1</td>
</tr>
<tr>
<td>Mendel-Rydberg</td>
<td>0.31</td>
<td>400</td>
<td>0.12</td>
<td>3.8 - 3.92</td>
</tr>
<tr>
<td>Marginis</td>
<td>0.29</td>
<td>400</td>
<td>0.12</td>
<td>3.8 - 4.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.92 - 4.1</td>
</tr>
<tr>
<td>Lomonosov-Flemming</td>
<td>0.52</td>
<td>400</td>
<td>0.21</td>
<td>3.8 - 4.0</td>
</tr>
<tr>
<td>Hercules</td>
<td>0.62</td>
<td>400</td>
<td>0.25</td>
<td>&gt;3.2</td>
</tr>
<tr>
<td>Tsiolkovsky</td>
<td>0.50</td>
<td>400</td>
<td>0.20</td>
<td>3.2 - 4.0</td>
</tr>
<tr>
<td>Korolev</td>
<td>0.95</td>
<td>400</td>
<td>0.38</td>
<td>3.4 - 4.04</td>
</tr>
<tr>
<td>Humboldtianum</td>
<td>0.21</td>
<td>400</td>
<td>0.084</td>
<td>3.8 - 3.87</td>
</tr>
<tr>
<td></td>
<td>1.12</td>
<td>400</td>
<td>0.45</td>
<td>3.87 - 3.92</td>
</tr>
<tr>
<td>Maurolycus</td>
<td>1.6</td>
<td>400</td>
<td>0.64</td>
<td>3.87 - 3.92</td>
</tr>
<tr>
<td>Milne</td>
<td>0.92</td>
<td>500</td>
<td>0.46</td>
<td>3.92 - 4.0</td>
</tr>
<tr>
<td>Mendelev</td>
<td>0.53</td>
<td>400</td>
<td>0.21</td>
<td>3.2 - 3.9</td>
</tr>
<tr>
<td>Totals</td>
<td>35.7</td>
<td>15.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
clasts of intrusive mare basalts have been unambiguously identified in the returned lunar samples (e.g., Heiken et al. 1991), there appear to be examples of intrusive equivalents of mare basalts among the lunar meteorites (e.g., NW A 773, Asuka 881757, and clasts in QUE 94281 and EET 87521 and 96008; see Chapter 2). Such deposits could represent any of the more mafic materials that have been identified in the sample record, not just basalts. Some hidden mafic deposits could, for example, represent shallow plutons of Mg-suite rocks or KREEP basalt flows.

4.8.3. KREEP basaltic volcanism. Two related instances of potential KREEP volcanism have been proposed. The Apennine Bench Formation near the Apollo 15 landing site has long been suspected of being composed of extrusive KREEP basalts (Hawke and Head 1978; Spudis 1978; Metzger et al. 1979; Spudis and Hawke 1986). This formation may be underlain by an intrusive equivalent of the KREEP basalts (Hawke and Head 1978; Blewett and Hawke 2001). The Apennine Bench Formation was initially mapped as a light-plains deposit (Wilhelms and McCauley 1971). However, photogeologic studies showed evidence of volcanic features, suggesting that extrusive volcanism produced the deposit (Hawke and Head 1978; Spudis 1978). Furthermore, Apollo gamma-ray spectrometer data (Davis 1980; Metzger et al. 1977) indicated that the FeO, TiO$_2$, and Th values for the Apennine Bench were consistent with the composition of Apollo 15 KREEP basalt samples (Hawke and Head 1978; Spudis 1978; Metzger et al. 1979; Spudis and Hawke 1986). FeO and TiO$_2$ abundances derived from Clementine remotely sensed data (Lucey et al. 2000) and Lunar Prospector Th abundances (Lawrence et al. 2000) are in good agreement with those derived from the Apollo data, and are consistent with Apollo 15 KREEP basalts (Blewett and Hawke 2001). Telescopic hyperspectral reflectance spectra indicate the presence of low-Ca pyroxenes such as those found in KREEP basalts (Spudis et al. 1988; Blewett et al. 2001).

Adjacent to the Apennine Bench, the craters Aristillus and Autolycus have excavated KREEP-like material. Clementine data shows the ejecta of Aristillus to have FeO values that are lower than the nearby mare basalts, while TiO$_2$ values are greater than the surrounding mare, highlands and even the Apennine Bench (Blewett and Hawke 2001). Lunar Prospector and Apollo Th data from Aristillus also show elevated Th concentrations (Metzger et al. 1979; Gillis and Jolliff 1999). Similar FeO, TiO$_2$, and Th concentrations are noted in the ejecta of Autolycus crater (Metzger et al. 1979; Blewett and Hawke 2001). Such evidence suggests that these two craters excavated KREEP-rich material (Metzger et al. 1979; Gillis and Jolliff 1999). Considering the depth of excavation and the diameter of the crater, Aristillus ejecta originate from a depth of <5 km, which may be too deep to represent extrusive volcanism that was buried by later events, but which could be Imbrium impact melt. Blewett and Hawke (2001) suggested that the Aristillus and Autolycus craters excavated a KREEP plutonic complex. This interpretation of the provenance of Aristillus and Autolycus ejecta supports an earlier stratigraphic model suggested for the region (Hawke and Head 1978). In that model, the lowest unit consists of the “highland” basement, which was impacted to form the Imbrium basin. Thorium-rich KREEP plutonic materials and/or KREEP basalts were emplaced following basin formation, and then covered by Imbrium mare basalts. Lastly, the Aristillus and Autolycus impacts penetrated into the KREEP-rich substrate. Detailed studies of stratigraphic relations exposed by these and other craters are needed to confirm the nature of the proposed KREEP magmatic units and to determine their extents and volumes.

4.9. Models for early lunar volcanism

4.9.1. KREEP basalts. The early partial melting models for KREEP basalts were inadequate because they did not explain the constancy of relative trace-element abundances and ratios in KREEP. Therefore, most of the current models for the petrogenesis of KREEP basalts are tied to the role played by the residual dregs ("urKREEP") of the LMO during lunar magmatism. Two general types of models are currently considered: assimilation of urKREEP
by magmas undergoing fractional crystallization and generation of a melt from an urKREEP-enriched cumulate source.

Using trace-element data generated in the early 1980s and assimilation models that had been previously proposed for the Mg-suite plutonic rocks, Warren (1988) modeled the chemical compositions of Apollo 15 and 17 KREEP basalts as a product of AFC (assimilation and fractional crystallization) of urKREEP by Mg-rich parental magmas. Snyder et al. (1995) quantitatively modeled the AFC generation of parental KREEP basalts using a model and end-member components similar to that of Warren (1988).

Several recent models have suggested that perhaps the KREEP basalts and the high-Al basalts were products of melting of a hybrid lunar mantle consisting of mixtures of evolved (KREEP) and early LMO cumulates. This type of model is partially based on the LMO cumulate-overturn models of Ringwood and Kesson (1976), Spera (1992), and Hess and Parmentier (1995), and perhaps demonstrated by chemical and experimental data for the lunar pyroclastic glasses (Hughes et al. 1988, 1989, 1990; Shearer et al. 1989, 1990, 1991; Shearer and Papike 1993). Hess and Parmentier (1995), Hess (1994) and Shearer and Papike (1999) suggested the possibility that the KREEP component was transported into the deep (and Mg-rich) lunar mantle where it was incorporated into the source for magmas parental to the KREEP basalts and Mg-suite magmas. An alternative mechanism to produce a hybrid source and to initiate melting was proposed by Nyquist and Shih (1992) who suggested that mixing and melting was the product of large impacts. Clearly, the interrelationships among KREEP volcanism, highland plutonism, mare volcanism, and basin formation are complex. Are the KREEP basalts the volcanic equivalents of both the highland Mg-suite and the highland alkali suite? Is there a transition between mare basalts and the KREEP basalts?

As discussed earlier, Warren (1988) and Snyder et al. (1995) suggested that primitive KREEP basaltic magmas were parental to the plutonic cumulates making up the Mg- and alkali rock suites. This is consistent with the fractional crystallization modeling of Snyder et al. (1995) and the Nd isotopic data compiled by Nyquist and Shih (1992). There is a significant difference in ages between Mg-suite plutonic rocks (4.14-4.46 Ga) and KREEP basalts (2.9-4.08 Ga) that suggests they represent different episodes of lunar magmatism. However, this difference could be the result of sampling bias attributed to changes in the meteorite impact flux that would change with time both the efficiency of sampling of the lower crust and the preservation of old KREEP basalt flows. On the basis of lunar meteorite NWA 773 there appears to be an age overlap between the production of KREEP basalts and that of the most of the mare-filling basalts. The lower initial $\varepsilon_{Nd}$ for this one example of rather young KREEP basaltic magmatism suggests that a substantial KREEP component was in the mantle source and that KREEP in the lunar mantle played an important part in extending the duration of lunar magmatism in the Procellarum basin on the near side of the Moon (Jolliff et al. 2003; Borg et al. 2004).

4.9.2. High-Al basalts. Models for the generation of the high-Al basalts and the VHK basalts range from melting of hybridized and nonhybridized mantle sources (Reid and Jakes 1974; Taylor and Jakes 1974; Binder 1976; Kurat et al. 1976; Ma et al. 1979; Dickinson et al. 1989; Hughes et al. 1990; Shervais and Vetter 1990) to assimilation of an evolved LMO component (KREEP, lunar granite rock types) by mantle-derived basaltic magmas (Dickinson et al. 1985; Shervais et al. 1985a; Neal et al. 1988a,b, 1989a,b) to impact melting of crustal lithologies (Snyder et al. 2000). Many of these various models are similar to those proposed for the KREEP basalts, but involve different sources and assimilants.

Are the basalt represented by these fragments a product of mantle melting or impact melting? Snyder et al. (2000) suggested that the mixing of components to produce the high-Al basalts was a surface process resulting from impact. According to this hypothesis, the high-Al basalts are not true basalts but are instead impact-melt rocks. The low siderophile contents (Dickinson et al. 1985; Shervais et al. 1985a), however, are not consistent with addition of a
meteorite component to the target or melt by the potential impactor. Moreover, the narrow range of major-element bulk compositions reinforces constraints provided by phase equilibria (e.g., Neal et al. 1989b). Impact melts might be expected to show a wider range of compositions that do not respect phase boundaries, unless the impact melt has subsequently differentiated. Irving (1975) used the cotectic composition of KREEP basalts to argue for a magmatic origin. The major-element compositions of the high-Al basalts overlap those of low-Ti mare basalts, except that the CaO/Al$_2$O$_3$ weight ratios (0.8-0.9) of the former are slightly less than those of the latter (CaO/Al$_2$O$_3$~1). However, given the small samples available, slight discrepancies such as these might be expected. The convex REE patterns and the initial $\epsilon_{Nd}$ vs. age plots are both consistent with magmas having been derived from cumulate mantle sources. We tentatively agree that these basalts are related to mare volcanism and have significant petrogenetic implications.

Early studies, based primarily on high-Al basalts returned by the Luna 16 mission, placed their mantle source just below the crust at depths of 40 to 100 km (Reid and Jakes 1974; Taylor and Jakes 1974; Binder 1976; Kurat et al. 1976). In these studies, the source region was thought to be a late-stage LMO cumulate that was rich in clinopyroxene and plagioclase (Reid and Jakes 1974; Taylor and Jakes 1974; Binder 1976; Ma et al. 1979) and that was LREE-enriched (Ma et al. 1979). Ridley et al. (1975) used mineralogical and trace-element justifications to argue that plagioclase was not a required phase in the source and that a slight increase in silica activity in an olivine + pyroxene source would result in a stronger partitioning of Al into the melt. Alternatively, he suggested that the high-Al basalts could simply be residual liquids. Clearly, some of the younger and first-discovered plagioclase-rich basalts (e.g., the Apollo 12 feldspathic basalt, evolved Luna 24 VLT basalts) may have been produced by this latter process.

The wide range in incompatible trace-element abundances of high-Al basalt clasts from the Apollo 14 site was shown to be inconsistent with these models (Shervais et al. 1985a; Dickinson et al. 1985, 1989; Neal et al. 1988a,b, 1989a,b; Hughes et al. 1990). Although this variation could be a function of different amounts of mesostasis in the individual small fragments, ion microprobe data imply that the variation in the fragments is related to significant differences in melt composition. This variability primarily in incompatible trace elements was ascribed to different degrees of partial melting and KREEP assimilation (Dickinson et al. 1985; Shervais et al. 1985; Neal et al. 1988a,b, 1989a,b; Kramer and Neal 2003). The best-fit model by Neal et al. (1988, 1989a,b) and Neal and Taylor (1990) employed cyclical assimilation of KREEP by a LREE-depleted olivine basalt. They concluded that assimilation of up to 15% KREEP and 70% fractional crystallization would produce the array of compositions observed in a suite of high-Al basalts from the Apollo 14 site. Based upon Sr and Nd isotopic data first reported by Dasch et al. (1987), Nyquist and Shih (1989) modified this model by suggesting that there were multiple batches of olivine basaltic magmas produced by different degrees of partial melting of at least two mantle sources. They also concluded that the parental olivine basalt assimilated KREEP prior to extensive fractional crystallization. The latter is consistent with the ion-microprobe data of Hagerty et al. (2001-2003) that show that the cores of olivine in several samples are relatively enriched in incompatible elements such as Kramer and Neal (2003) concluded, on the basis of trace-element data, that three groups of Apollo 14 high-Al basalts existed representing different degrees of partial melting of a common source at three separate times (3.95, 4.1, 4.3 Ga). Each magma experienced variable degrees of KREEP assimilation and fractional crystallization.

Although assimilation-fractional crystallization (AFC) models involving basaltic magmas and KREEP are consistent with major- and trace-element data, they are not entirely compatible with the isotopic data. Nyquist and Shih (1992) interpreted the Sr and Nd isotopic data as follows: (1) Source regions for the high-Al basalts with low incompatible-element concentrations were established early in lunar evolution (LMO cumulates) and had depleted LREE abundance patterns. (2) Coincident with the onset of earliest episodes of high-Al magmatism, these sources were contaminated with REE-rich, LREE-enriched magmas having
The K-rich basalts with their elevated K, Ba, and Rb concentrations require a complex petrogenesis because simple partial-melting models are difficult to reconcile with the phase equilibria; primitive partial melts must be in equilibrium with plagioclase but high pressure experiments on compositions similar to K-rich basalts have only olivine and orthopyroxene on the liquidus (Walker et al. 1972). AFC models have been proposed for the origin for the K-rich basalts; however, rather than assimilating a KREEP component, their petrogenesis may have involved the assimilation of a very small amount (1-3%) of a “granitic component” (Shervais et al. 1985b; Shih et al. 1986, 1987; Neal et al. 1988b, 1989b). Initial Sr and Nd isotopic compositions of these basalts lie on the growth curves defined by the high-Al basalts indicating that Sr and Nd were not assimilated significantly from a granitic component. These results are not unique since our granite sample data base is sparse, although such a process does account for the elevated K, Rb, and Ba concentrations and elevated K/La.

One important conclusion inferred from these older basalts is that the crystallization of the LMO, convective-overturn of the LMO cumulate pile, and reheating/melting of cumulate sources were achieved by 4.3 Ga. These results are consistent with constraints provided by the short-lived isotopes (e.g., Shearer and Newsom 2000; Righter and Shearer 2003) and the physics of the cumulate overturn model (e.g., Hess and Parmentier 1999).

5. LUNAR (MARE) MAGMATISM 3.85 TO ~1.0 Ga

5.1. Introduction

Successful models for the internal evolution of the Moon must consider the volume, distribution, timing, composition and, ultimately the petrogenesis of mare basalts. Indeed, given the paucity of geophysical data, the evolution of the lunar interior can be gleaned only by unraveling the petrogenesis of the mare basalts. Mare basalts as represented by the dark smooth areas on the Moon, cover roughly 17% of the surface but are primarily relegated to fills of multiringed impact basins and irregular depressions (Head 1976). The basalts form
low albedo, generally smooth surfaces with characteristics that give evidence for highly effusive, flood-type eruptions (Head and Wilson 1992). Since these flood basalts completely fill most of the near-side basins, the early stages of basin-filling volcanism are largely erased. Partly filled basins, however, such as Orientale and Smythii afford a glimpse at the beginning stages of basin volcanism (Gillis and Spudis 2000). Even earlier volcanic episodes appear as dark-mantling deposits or “dark haloes” around impact craters that excavated pre-basin deposits (Head and Wilson 1992), but their volumes and links to mare volcanism are not yet fully documented (Antonenko 1999). Thickness of the mare flood basalts are estimated to vary widely, ranging from a few hundred meters in Smythii (Gillis and Spudis 2000) to 4.5 km in the central portions of multiringed basins (Head and Wilson 1992; see also Chapter 3). In addition to the eruption of large volumes of lava, variously colored volcanic glass beads (or picritic glasses) with mare basalt affinities were produced by fire-fountaining (Head and Wilson 1992 and references therein). These glass beads provide some of the most important constraints to models of mare-basalt petrogenesis.

Mare volcanism on the lunar far side is limited to a few volcanic deposits in comparatively small and young craters and lacks the extensive basaltic plains so characteristic of the near-side crust. This asymmetry in mare volcanism has generally been attributed to the thicker anorthosite crust on the lunar far side, where it is roughly 70 km thick (Zuber et al. 1994, 1999), on average and may act as a low density barrier preventing the eruption of the denser mare magmas. However, the largest and possibly the most ancient lunar basin, the South Pole-Aitken Basin, occurs on the lunar far side. Despite its great size, it contains only sporadic patches of mare volcanic deposits (Pieters et al. 2001; Petro and Pieters 2004). The basin is floored by lunar crust that appears to be no thicker than crust flooring the multiringed, basalt-filled craters on the lunar near side (Wieczorek and Phillips 1999). It is clear that the creation of a major multiringed basin over thinned anorthosite crust is not directly correlated to the appearance of large volumes of mare volcanism.

Radiometric crystallization ages of common mare-basalt compositions range from 3.6-3.9 Ga for the high-Ti basalts at the Apollo 11 and Apollo 17 sites to 3.16-3.4 Ga for low and very-low-Ti basalts at the Apollo 12, Apollo 15 and Lunar 24 sites (Nyquist and Shih 1992; Snyder et al. 2000). Hiesinger et al. (2000), using crater size frequency ages on well defined volcanic units, found that mare volcanism on the lunar near side was active from about 3.9-4.0 Ga to about 2.0 Ga, a span of almost 2 Ga. Younger patches of mare volcanism apparently are as young as 1.2 Ga (Hiesinger et al. 2003). Most magma production, however, occurred in the late Imbrian at about 3.6-3.8 Ga (Fig. 4.34). A second weaker pulse of volcanism between 3.3-3.5 Ga is observed in the younger multiring basins. No single basin records 2.0 Ga years of volcanism (Hiesinger et al. 2000, 2003), and the older basins such as Tranquillitatis contain only a “single” pulse of volcanism between 3.3-3.9 Ga. If all the volcanic effusive activity is considered together, the 3.5-3.8 Ga period marks the peak in volcanic output. Mare volcanism on the lunar far side, while comparatively sparse, was also a long-lived process (Pieters et al. 2001).
5.2. Composition of the mare basalts and source mineralogy

Mare basalts and picritic glasses are unique to the Moon; neither the Earth, Mars, or the basaltic meteorites have basalts that even approach the major-element composition of some of the lunar basalts. The major-element composition of mare volcanic rocks have been summarized in numerous publications (e.g., Papike et al. 1976, 1997, 1998; BVSP 1981; Delano 1986; Neal and Taylor 1992; Shearer and Papike 1993, 1999). Our aim here is to show how the composition of the mare volcanics require that the mantle source is not part of the primitive Moon but instead is a highly differentiated product of the LMO.

Relative to terrestrial basalts, the major-element compositions of the mare volcanic suite have widely varying and extreme compositions (Table 4.4). The mare volcanics and particularly the picritic glasses have significantly lower CaO, Al\(_2\)O\(_3\), and Na\(_2\)O contents but higher FeO and MgO contents than terrestrial basalts from the ocean floor and ocean islands. The low Na\(_2\)O contents probably reflect, in part, the paucity of Na\(_2\)O and other volatile elements on the Moon (Taylor 1982), but the low CaO and Al\(_2\)O\(_3\) contents record an unusual mantle source and/or petrogenesis. The low CaO and Al\(_2\)O\(_3\) contents are not characteristics of magmas in equilibrium with lherzolite mantle. The bulk compositions of picritic glasses are more similar to the compositions of komatiites than basalts; the mare “basalts” are not basalts but products of ultramafic magmas. We will, nevertheless, continue the tradition and use mare “basalts” to characterize both crystalline basalts and the volcanic picritic glasses. The reader should be cautioned not to jump to conclusions; unlike the terrestrial experience, ultramafic magmas on the Moon do not necessarily imply that such magmas are products of large degrees of melting.

Another major difference in bulk composition between terrestrial and lunar mafic/ultramafic magmas are revealed in the very high FeO contents of the mare basalts and picritic glasses. The FeO contents of the picritic glasses are roughly twice those of primitive ocean-floor basalts (Delano 1986; Hess 1992). Only highly differentiated ocean-floor basalts and rare Archean basalts or komatiites attain FeO values of the high-Ti picritic glasses. The MgO contents of the lunar picritic glasses are more typical of terrestrial picritic basalts or komatiites but given their high FeO values, the Mg' values of the lunar mare suite are significantly lower. Komatiite/picrite terrestrial magmas, for example, coexist with olivine with Mg' values in the low 90's (Hess 1989, 1992) but the most magnesian picritic glasses crystallize olivine with Mg' values in the low 80's (Delano 1990). These differences are very significant and speak to material differences in the compositions of the lunar and terrestrial mantles.

| Table 4.4. Compositions of (1) very low-Ti picrite glass (2) high-Ti picrite glass (Shearer and Papike 1993) (3) ocean floor basalt (4) ocean island basalt and (5) komatiite (Hess 1989). |
|---|---|---|---|---|
| Picrite Glasses | Terrestrial Magmas |
| **SiO\(_2\)** | 47.6 | 34.1 | 49.7 | 48.8 | 47.3 |
| **TiO\(_2\)** | 0.3 | 17.0 | 0.7 | 2.5 | 0.4 |
| **Al\(_2\)O\(_3\)** | 7.8 | 3.8 | 16.5 | 11.9 | 7.8 |
| **Cr\(_2\)O\(_3\)** | 0.6 | 0.7 | 0.1 | 0.1 | 0.4 |
| **FeO** | 16.3 | 23.5 | 8.3 | 11.4 | 10.8 |
| **MnO** | 0.3 | 0.4 | 0.2 | 0.2 | 0.2 |
| **MgO** | 18.6 | 13.6 | 9.0 | 9.7 | 25.8 |
| **CaO** | 8.0 | 6.4 | 13.8 | 10.5 | 7.6 |
| **Na\(_2\)O** | 0.2 | 0.5 | 1.9 | 2.0 | — |
| **K\(_2\)O** | 0.0 | 0.2 | 0.1 | 0.5 | — |
The range of TiO$_2$ contents in the picritic glasses is most noteworthy, varying from very low values of 0.22 wt% in Apollo 15 green glass to 17.0 wt% TiO$_2$ in Apollo 14 black glass (Shearer and Papike 1993). Terrestrial basalts, in comparison, typically have between about 0.5% and 6.0% TiO$_2$. Primitive MORB, for example, have TiO$_2$ contents between 0.5-1.2% TiO$_2$ (Presnall and Hoover 1987). Only highly magnesian alkaline lavas, as for example the meimechites, which, based on their elevated incompatible trace element content, are products of small degrees melting of a primitive or an enriched mantle source, have more extreme TiO$_2$ values that reach only 4-6% (Arndt et al. 1995). Indeed, it is easy to show as is done below that even very small degrees of melting of a typical primitive terrestrial mantle cannot give rise to the very high-Ti contents of the most enriched lunar picrite mare glasses.

The bulk partition coefficient for TiO$_2$ between lherzolite mineral assemblages and basalt is strongly affected by the high-Ca clinopyroxene and garnet content. The high-Ca clinopyroxene-basalt TiO$_2$ partition coefficient varies from about 0.3-0.5 (Skulski et al. 1994; Jones 1995). The partition coefficients for garnet and orthopyroxene are estimated from the TiO$_2$ partition coefficients between these phases and coexisting high-Ca clinopyroxene (e.g., Seitz et al. 1999). While there is considerable scatter, the garnet-basalt partition coefficient is roughly twice that of high-Ca clinopyroxene, whereas that of orthopyroxene is roughly half that of clinopyroxene. In support of these estimates, experiments done by Xirouchakis et al. (2001) indicate a $D_{(TiO_2)} \approx 0.1-0.2$ for orthopyroxene. Using $D = 0$ for olivine, a typical fertile spinel lherzolite would have a bulk $D > 0.1$. A depleted harzburgite, in contrast, would have $D \leq 0.1$.

Terrestrial fertile mantle contains roughly 0.2 wt% TiO$_2$, depleted mantle significantly less (Taylor 1980). Basalts derived by small degrees of melting of fertile lherzolite cannot have more than 2 wt% TiO$_2$ assuming $D \geq 0.1$ (the maximum enrichment of TiO$_2$ in the melt is given by $1/D = 10$). Depleted mantle approaching harzburgite in composition has a smaller bulk partition coefficient for TiO$_2$ but also a much smaller TiO$_2$ content. It follows that terrestrial basalts are constrained to have less than about 2 wt% TiO$_2$; those highly alkaline primitive basalts with higher TiO$_2$ concentrations probably reflect the influence of CO$_2$ on the melting process. It comes as no surprise that the TiO$_2$ enriched mare basalts and glasses cannot be derived from primitive mantle on a volatile-depleted Moon, but must originate from mantle that is both differentiated and TiO$_2$ enriched.

The sources for the very low-Ti mare basalts (VLT) and glasses are also unlikely to be primitive mantle. Consider, as an illustrative example, the petrogenesis of a VLT glass with 0.4 wt% TiO$_2$ (Table 4.5). Assuming a mantle with 0.2 wt% TiO$_2$, such glasses can be generated by roughly 50% melting of fertile terrestrial lherzolite. At this extent of melting all the accessory phases and high-Ca clinopyroxene are consumed leaving harzburgite in the mantle. Indeed, even 20-25% melting is sufficient to create a harzburgite residue (Hess 1992). Under these conditions, both CaO and Al$_2$O$_3$ are incompatible and their contents should be enriched, roughly doubled, in the basalt melt relative to the contents in the mantle. At first glance, these expectations are satisfied as the VLT glass has roughly 8 wt% each of Al$_2$O$_3$ and CaO or double that in fertile mantle (Table 4.5). However, for this explanation to be generally accepted, the TiO$_2$ contents of other VLT glasses should correlate with their Al$_2$O$_3$ and CaO contents. Such correlations do not exist. Whereas VLT glasses have TiO$_2$ contents that range from 0.22 to 1.0 wt%, the corresponding CaO and Al$_2$O$_3$ each are within about 10% from their mean values (Delano 1986; Shearer and Papike 1993). The highest TiO$_2$ contents, moreover, are not necessarily associated with the highest CaO and Al$_2$O$_3$ contents. These arguments show that primitive lherzolite mantle is not an appropriate source for VLT basalts and glasses (See also Section 4.3.1). Furthermore, large degrees of melting of a peridotite source are not a requisite to account for the low-TiO$_2$ contents. A more reasonable hypothesis is that the source for the VLT mare magmas is differentiated and depleted in the incompatible oxides TiO$_2$, Al$_2$O$_3$ and CaO.
With the exception of one or two VLT picritic glass compositions, it is noteworthy that the contents of both Al$_2$O$_3$ and CaO in VLT glasses appear to be buffered to near constant values (Table 4.5). What these results suggest is that the Al$_2$O$_3$ and CaO are not strongly incompatible and/or the extent of melting is roughly constant. The mantle phases that are capable of buffering the CaO and Al$_2$O$_3$ contents are plagioclase, high-Ca clinopyroxene, garnet and orthopyroxene. The first two phases listed are poor candidates since the CaO and Al$_2$O$_3$ contents of the VLT picritic glasses are too low—the VLT melts are undersaturated with respect to plagioclase and high-Ca clinopyroxene at all lunar pressures. The same is true for garnet except at depths greater than 500 km. In addition, the HREE contents of most of the picritic glasses are not consistent with significant quantities of garnet occurring as a residual phase (Papike et al. 1997). Of the possible phases listed above, only orthopyroxene has the characteristics of the desired residual phase.

The VLT picritic glasses have orthopyroxene as a liquidus phase only at high pressures (Chen et al. 1982; Chen and Lindsley 1983; Elkins-Tanton et al. 2000) so that the source lies deep within the Moon. The orthopyroxene-basalt partition coefficient for Al is pressure dependent; this dependency is approximated by $D_{Al} = 0.11 + 0.10\times$GPa. Assuming that the pressures are 2-3 GPa it follows that $D_{Al} = 0.3-0.4$. The bulk partition coefficient, of course, is a function of the olivine content in the source; assuming that the sources for mare basalts are cumulates to the LMO, the source could range from an olivine-bearing orthopyroxenite to an orthopyroxene-bearing dunite (e.g., Hughes et al. 1989; 1990; Delano 1980; Longhi 1982; Hess 2000 among others). Therefore, for the source to buffer the Al$_2$O$_3$ and CaO contents to approximately constant values for varying amounts of melting, an orthopyroxene-rich source is clearly favored. This conclusion follows from the fact that the contents of more compatible elements in partial melts are less sensitive to varying degrees of partial melting than are incompatible elements. Both CaO and Al$_2$O$_3$ in an olivine-rich mantle would behave much more incompletely than in a pyroxene-rich mantle.

The FeO contents of VLT picritic glasses range from 16-22 wt% (Table 4.5). This range in FeO contents requires either that the FeO contents of the source region vary a comparable amount and/or the mineralogy of the source is dominated by orthopyroxene rather than olivine. The FeO olivine/basalt partition coefficient for VLT to high-Ti mare basalts is approximately unity (Delano 1980; Elkins-Tanton et al. 2000). It follows that partial melting of an olivine-rich source to produce mare basalts or fractionation of olivine from such basaltic magmas will

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO$_2$</td>
<td>48.68</td>
<td>44.10</td>
<td>44.41</td>
<td>46.53</td>
<td>44.5</td>
<td>46.3</td>
<td>45.15</td>
<td>41.92</td>
<td>39.8</td>
<td>39.28</td>
<td>36.69</td>
</tr>
<tr>
<td>TiO$_2$</td>
<td>0.22</td>
<td>0.40</td>
<td>0.63</td>
<td>0.78</td>
<td>0.47</td>
<td>0.55</td>
<td>1.00</td>
<td>4.82</td>
<td>6.84</td>
<td>9.07</td>
<td>13.69</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>7.06</td>
<td>7.79</td>
<td>7.23</td>
<td>8.30</td>
<td>7.65</td>
<td>9.82</td>
<td>6.87</td>
<td>6.52</td>
<td>8.50</td>
<td>7.55</td>
<td>6.87</td>
</tr>
<tr>
<td>Cr$_2$O$_3$</td>
<td>0.60</td>
<td>0.40</td>
<td>0.50</td>
<td>0.37</td>
<td>0.45</td>
<td>0.46</td>
<td>0.64</td>
<td>0.40</td>
<td>0.54</td>
<td>0.84</td>
<td>0.80</td>
</tr>
<tr>
<td>FeO</td>
<td>16.30</td>
<td>21.82</td>
<td>19.27</td>
<td>18.54</td>
<td>21.3</td>
<td>19.5</td>
<td>23.00</td>
<td>24.10</td>
<td>21.7</td>
<td>23.35</td>
<td>20.59</td>
</tr>
<tr>
<td>MnO</td>
<td>0.34</td>
<td>0.25</td>
<td>0.28</td>
<td>0.32</td>
<td>0.32</td>
<td>0.30</td>
<td>0.33</td>
<td>0.32</td>
<td>0.36</td>
<td>0.27</td>
<td>0.23</td>
</tr>
<tr>
<td>MgO</td>
<td>18.66</td>
<td>17.00</td>
<td>18.75</td>
<td>16.42</td>
<td>17.6</td>
<td>14.1</td>
<td>15.68</td>
<td>13.74</td>
<td>11.8</td>
<td>11.90</td>
<td>11.84</td>
</tr>
<tr>
<td>CaO</td>
<td>8.43</td>
<td>8.55</td>
<td>8.05</td>
<td>8.48</td>
<td>8.37</td>
<td>9.41</td>
<td>7.93</td>
<td>7.90</td>
<td>8.60</td>
<td>8.02</td>
<td>7.68</td>
</tr>
<tr>
<td>Na$_2$O</td>
<td>0.17</td>
<td>0.10</td>
<td>0.25</td>
<td>0.14</td>
<td>0.28</td>
<td>0.16</td>
<td>0.21</td>
<td>0.68</td>
<td>0.31</td>
<td>0.35</td>
<td>0.69</td>
</tr>
<tr>
<td>KO$_2$</td>
<td>0.00</td>
<td>0.00</td>
<td>n.a.</td>
<td>0.08</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>0.03</td>
<td>n.a.</td>
</tr>
<tr>
<td>Total</td>
<td>100.5</td>
<td>100.4</td>
<td>99.4</td>
<td>100.0</td>
<td>100.9</td>
<td>100.8</td>
<td>100.5</td>
<td>98.5</td>
<td>100.5</td>
<td>99.1</td>
<td>99.8</td>
</tr>
</tbody>
</table>
leave the FeO contents of the melt largely unchanged. The FeO orthopyroxene/basalt partition coefficient is approximately 0.6-0.7 for VLT (Elkins-Tanton et al. 2000) and hi-Ti mare basalt (Delano 1980). Partial melting of an orthopyroxene-rich source from 0 to 40% is required to vary the FeO contents by the observed amounts. Based on our discussions detailing the near constancy of Al$_2$O$_3$ and CaO in VLT glasses, such a large range in the degree of melting is unlikely. It follows that much of the variation in FeO contents must be ascribed to variations of the source composition.

We conclude from this analysis that VLT basalts were derived from heterogeneous mantle dominated by orthopyroxene but with more than trace amounts of olivine. The need to include olivine in the source is argued more fully in forthcoming sections. The mantle source was depleted in CaO and Al$_2$O$_3$ relative to primitive terrestrial lherzolite; the mantle residue to the VLT cannot have contained high-Ca clinopyroxene, plagioclase, ilmenite or garnet because the melts are undersaturated with respect to these phases at lunar pressures below 3.0 GPa (Elkins-Tanton et al. 2000; Chen et al. 1982; Shearer et al. 2003; Draper et al. 2004). This conclusion does not exclude the possibility that such phases were originally in the source region albeit in small amounts and were consumed by melting or that the mantle source containing such phases was in contact with the VLT during some stage of their evolution (Longhi 1992; Elkins-Tanton et al. 2000). In addition, high pressure studies by Shearer et al. (2003) and Draper et al. (2004) hint at the possibility that garnet may be a liquidus phase at pressures above 3.0 GPa.

The high-Ti picritic glasses must also come from a similar source given that the CaO and a lesser degree the Al$_2$O$_3$ contents are similar to those of VLT glasses. What is different, of course, are the extraordinary levels of TiO$_2$ and the concomitant increase in FeO contents (Table 4.5). It is apparent from the preceding discussions that the higher TiO$_2$ and FeO contents must in large part reflect the higher TiO$_2$ and FeO contents of the mantle source. The most extreme compositions as represented by the Apollo 14 black glasses contain the highest TiO$_2$ (17 wt%) but the lowest Al$_2$O$_3$ contents of only 3.8 wt%. These low Al$_2$O$_3$ contents are roughly half of those of the VLT picritic glasses yet the CaO contents (6.5 wt%) of the high-Ti glass are not reduced a comparable amount. Whereas the Al$_2$O$_3$ contents in high-Ti glasses vary widely from 3.8-8.5 wt%, the spread in CaO contents is much more subdued (6.3-8.2 wt%). Note that the CaO/Al$_2$O$_3$ ratio in VLT is much less variable than that of the high-Ti picritic glasses. Perhaps the explanation centers on the pressure dependence of the Al$_2$O$_3$ partition coefficient of orthopyroxene. In comparison, the partition coefficient for CaO does not have strong pressure dependency.

Picritic mare glasses are the most primitive mare liquid compositions yet returned from the Moon and are the best candidates for primary magmas (Delano 1986). Crystalline mare basalts, on the other hand, are more evolved by comparison and their compositional variations within each suite of high-Ti, low-Ti and even VLT basalts appear to have been controlled by low-$P$ crystal fractionation of olivine and possibly other phases (BSVP 1981). It is not clear, however, whether the more primitive parental liquids are themselves products of fractional crystallization or approximate primary compositions. We tentatively make the assumption that the most primitive mare basalts are comparable to picritic glasses in petrogenetic importance. An instructive comparison between the two groups of magmas is given in Figure 4.35 where the TiO$_2$ and Al$_2$O$_3$ contents are plotted versus the Mg’ value. According to Papike et al. (1976) the Mg’ values of the most primitive high- and low-Ti basalts are roughly 0.50. Those basalts with larger values appear to contain cumulates of olivine. Certainly, the low-Ti olivine basalts with such elevated Mg’ have suffered significant additions of olivine (Walker et al. 1975). If this analysis is correct, then the Mg’ values of the most primitive high- and low-Ti mare basalt liquids and the picritic glasses are very similar. The same can be said of the VLT basalts and their glasses. If we now compare the major-element compositions of the picritic glasses with mare basalts with the same Mg’ values, it is evident that the mare basalts have higher CaO and Al$_2$O$_3$ contents than the picritic glasses (Fig. 4.35 and Table 4.6). High-Ti picritic glasses, for example, have Al$_2$O$_3$
contents in the range from 4-8 wt% whereas the primitive high-Ti mare basalts with Mg’ \( \cong \) 0.5 have Al₂O₃ contents from 7-12 wt%. Of course, more evolved mare basalts have even higher Al₂O₃ contents because Al₂O₃ behaves incompatibly with olivine and such basalts typically are undersaturated with respect to plagioclase.

Low-pressure experimental studies show that the most primitive mare basalts and picritic glasses have only olivine (± minor chromite) as their liquidus phase (Longhi 1992). Calculated low-pressure fractionation paths for a representative set of picritic glasses, show that the differentiates of these more primitive magmas are similar to typical crystalline mare basalts (Longhi 1987). This fact suggests that picritic glasses are parental magmas to the more evolved mare basalts. What is a puzzling result for some workers is that none of the 24 picritic glass compositions listed by Delano (1986) has exactly the right composition to be a parent to one of the known varieties of mare basalts with the one exception of the Apollo 17 green glass (Longhi 1987). The simplest explanation is that the collection of mare basalts is incomplete—we simply have a sampling problem. Remote-sensing data suggest that there is indeed a continuum of mare basalts across the TiO₂ spectrum, and the lunar meteorites are providing new and somewhat random samples with which to test the global distribution (see Chapters 2 and 3).

Delano (1990) suggested, however, that a “sampling problem” may not be a sufficient explanation for the lack of direct-lineage relationships between the picritic glasses and mare basalts. Picritic glasses, including those analyzed subsequently by Shearer and Papike (1993), have FeO/Sc weight ratios that range from about 4000 to 7000, whereas crystalline mare basalts have ratios from about 2000 to 6000 (Fig. 4.36) (Delano 1990; Shearer and Floss 2000). Fractionation of olivine leaves the FeO contents largely unchanged and thereby decrease the FeO/Sc ratios of differentiated melts. Fractionation of pyroxene, however, increases the FeO content and decreases Sc. Delano (1990) argued, however, that crystal fractionation processes could not reconcile the two data sets.
The conclusion that picritic glasses are distinct from mare basalts is supported also by the contrasting concentrations of large-ion lithophile elements in these lithologies (Shearer et al. 1994). Specifically, the Li/Be ratios in mare basalts (3-5) are consistently lower and are offset from the value of these ratios in picritic glasses (14-34) (Fig. 4.37). Since Li is moderately incompatible in olivine and pyroxene (partition coefficients are 0.15 in clinopyroxene and 0.33 in olivine) and Be is strongly incompatible ($D \approx 0$ in olivine and $D \approx 0.06$ in clinopyroxene), the Li/Be ratio should decrease in the cumulate pile with fractionation of the LMO. Indeed, KREEP-like components have the lowest Li/Be ratios. The sharp differences in Li/Be between mare basalts and picritic glasses probably record distinct source regions for the two groups of mare magmas. Neither partial melting nor fractional melting models can reconcile these data (Shearer et al. 1994).

### Table 4.6
Comparison of the bulk composition of high TiO$_2$, low TiO$_2$ and VLT mare basalts (B) with corresponding glass compositions (G) (from Papike et al. 1997).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO$_2$</td>
<td>39.72</td>
<td>37.79</td>
<td>37.88</td>
<td>39.28</td>
<td>46.40</td>
<td>43.05</td>
<td>48.1</td>
<td>46.0</td>
<td>44.10</td>
<td>45.71</td>
</tr>
<tr>
<td>TiO$_2$</td>
<td>10.50</td>
<td>12.97</td>
<td>12.50</td>
<td>9.07</td>
<td>3.28</td>
<td>3.58</td>
<td>0.36</td>
<td>0.75</td>
<td>0.40</td>
<td>0.89</td>
</tr>
<tr>
<td>Al$_2$O$_3$</td>
<td>10.43</td>
<td>8.85</td>
<td>5.73</td>
<td>7.55</td>
<td>10.16</td>
<td>8.58</td>
<td>11.2</td>
<td>10.1</td>
<td>7.79</td>
<td>8.50</td>
</tr>
<tr>
<td>Cr$_2$O$_3$</td>
<td>0.25</td>
<td>0.41</td>
<td>0.70</td>
<td>0.84</td>
<td>0.52</td>
<td>0.52</td>
<td>0.60</td>
<td>0.42</td>
<td>0.40</td>
<td>0.50</td>
</tr>
<tr>
<td>FeO</td>
<td>19.80</td>
<td>19.66</td>
<td>20.83</td>
<td>23.35</td>
<td>20.15</td>
<td>21.52</td>
<td>18.2</td>
<td>22.4</td>
<td>21.82</td>
<td>18.90</td>
</tr>
<tr>
<td>MnO</td>
<td>0.30</td>
<td>0.27</td>
<td>0.31</td>
<td>0.27</td>
<td>0.27</td>
<td>0.52</td>
<td>0.26</td>
<td>0.26</td>
<td>0.25</td>
<td>0.20</td>
</tr>
<tr>
<td>MgO</td>
<td>6.69</td>
<td>8.44</td>
<td>13.90</td>
<td>11.90</td>
<td>8.22</td>
<td>13.06</td>
<td>11.0</td>
<td>10.5</td>
<td>17.00</td>
<td>16.16</td>
</tr>
<tr>
<td>CaO</td>
<td>11.13</td>
<td>10.74</td>
<td>6.94</td>
<td>8.02</td>
<td>10.80</td>
<td>8.35</td>
<td>10.2</td>
<td>10.8</td>
<td>8.55</td>
<td>8.74</td>
</tr>
<tr>
<td>Na$_2$O</td>
<td>0.40</td>
<td>0.36</td>
<td>0.48</td>
<td>0.35</td>
<td>0.27</td>
<td>0.34</td>
<td>0.15</td>
<td>0.21</td>
<td>0.10</td>
<td>0.14</td>
</tr>
<tr>
<td>K$_2$O</td>
<td>0.06</td>
<td>0.05</td>
<td>0.15</td>
<td>0.03</td>
<td>0.07</td>
<td>n.a.</td>
<td>0.01</td>
<td>0.016</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Total</td>
<td>99.3</td>
<td>99.5</td>
<td>99.4</td>
<td>100.5</td>
<td>100.1</td>
<td>99.2</td>
<td>100.1</td>
<td>101.5</td>
<td>100.4</td>
<td>100.0</td>
</tr>
</tbody>
</table>

The conclusion that picritic glasses are distinct from mare basalts is supported also by the contrasting concentrations of large-ion lithophile elements in these lithologies (Shearer et al. 1994). Specifically, the Li/Be ratios in mare basalts (3-5) are consistently lower and are offset from the value of these ratios in picritic glasses (14-34) (Fig. 4.37). Since Li is moderately incompatible in olivine and pyroxene (partition coefficients are 0.15 in clinopyroxene and 0.33 in olivine) and Be is strongly incompatible ($D \approx 0$ in olivine and $D \approx 0.06$ in clinopyroxene), the Li/Be ratio should decrease in the cumulate pile with fractionation of the LMO. Indeed, KREEP-like components have the lowest Li/Be ratios. The sharp differences in Li/Be between mare basalts and picritic glasses probably record distinct source regions for the two groups of mare magmas. Neither partial melting nor fractional melting models can reconcile these data (Shearer et al. 1994).

### 5.3. Petrography

The petrographic characteristics of mare basalts have been thoroughly reviewed in BVSP (1981), Longhi (1992), Papike et al. (1997), Chapter 3 of this book and references therein; only the briefest summary is presented here. We focus mainly on those features that are most relevant to our understanding of mare-basalt petrogenesis. Many of the textural features and especially the modes are controlled by the vagaries of cooling and crystal differentiation. In addition, many studies (e.g., Papike et al. 1976; Rhodes et al. 1976) have concluded that the wide range of modal and chemical compositions is a result of both magmatic differentiation and the small non-representative sample size of many of the basalt lithologies.

Figure 4.37. Be vs. Li for picrite glasses (Symbols), crystalline mare basalts (rectangle) and KREEP (square). Trajectories produced by melting cumulate source by various degrees (pm) (modified after Shearer et al 1994).
The average modes of high-Ti Apollo 11 and Apollo 17 mare basalts are given in Table 4.7. The crystallization sequence inferred from petrochemical analysis and experimental investigations (Longhi 1992; Papike et al. 1997) is

\[ \text{olivine} \rightarrow \text{armalcolite} \rightarrow \text{pyroxene} \rightarrow \text{plagioclase} + \text{ilmenite} \]

where pyroxene is derived from a reaction relation between olivine and melt and ilmenite from a reaction relation between armalcolite ([Fe,Mg]Ti$_2$O$_5$) and melt. The appearances of plagioclase and to a lesser extent pyroxene are strongly affected by cooling rate (see Longhi 1992) so that their position in the crystallization sequence is variable but, in general, plagioclase, augite, and pigeonite appear after armalcolite. As such, the most magnesian samples have olivine ± armalcolite as a liquidus phases whereas more evolved magmas typically have only ilmenite, pyroxene, and plagioclase as early crystallizing phases. Experiments have shown that high-Ti mare basalts are multisaturated with 2 or more phases within 25 °C of their one bar liquidus (Longhi 1992). These are not characteristics of primary or even primitive magmas unless the mantle sources are shallow and differentiated.

Low-Ti mare basalts exhibit a wide range of textures and crystal sizes ranging from vitrophyric to porphyritic gabbros (BVSP 1981). The crystallization sequence is similar to that of high-Ti mare basalts, except that ilmenite, not armalcolite, is the primary TiO$_2$ phase and ilmenite is generally the last major phase to appear (the natural consequence of lower TiO$_2$ contents). Pigeonite typically forms cores to augite mantles so that the general crystallization sequence is

\[ \text{spinel} + \text{olivine} \rightarrow \text{pigeonite} \rightarrow \text{augite} \rightarrow \text{plagioclase} \rightarrow \text{ilmenite} \]

Iron metal, phosphates, cristobalite and quenched immiscible melts are all late appearing phases (Roedder and Weiblen 1970). A feature that distinguishes mare basalts from terrestrial basalts is the prominence of low-Ca clinopyroxene, mainly pigeonite, in the crystallization sequence. Pigeonite precedes or is largely coeval with augite in the crystallization sequence whereas in terrestrial MORBs augite is by far the most dominant pyroxene (Longhi 1981). Indeed, the tholeiitic silicate crystallization sequence on Earth is

\[ \text{olivine} \rightarrow \text{plagioclase} \rightarrow \text{augite} \rightarrow \text{pigeonite} \]

whereas in mare basalts, and in other lunar magmas, the sequence is

\[ \text{olivine} \rightarrow \text{plagioclase} \rightarrow \text{pigeonite} \rightarrow \text{augite} \]

where pigeonite and plagioclase may switch positions in mare basalts (Longhi 1992). These petrographic data alone clearly distinguish the mantle source on Earth, typically fertile lherzolite, from that of the sources for magmas on the Moon.

5.4. Trace elements

The REE patterns of mare basalts and picritic glasses provide useful but, in some cases, somewhat confusing constraints on their petrogenesis. We focus first on some features most basalts have in common. A persistent feature is that nearly all mare basalts and picritic glasses

<table>
<thead>
<tr>
<th>Apollo</th>
<th>11</th>
<th>17</th>
<th>12</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olivine</td>
<td>2.2</td>
<td>4.5</td>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>Pyroxene</td>
<td>49.0</td>
<td>47.0</td>
<td>53</td>
<td>63</td>
</tr>
<tr>
<td>Plagioclase</td>
<td>31.0</td>
<td>23.0</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>Ilmenite</td>
<td>14.0</td>
<td>24.0</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>
have negative europium (Eu) anomalies (Fig. 4.38) (Papike et al. 1997). The negative Eu anomaly is typically ascribed to the nature of the source region because primitive mare basalts and picritic glasses are far from being saturated with respect to plagioclase at any pressures (BVSP 1981; Longhi 1992). At the liquidus temperatures appropriate for basalts at 1 bar, plagioclase saturated melts would contain more than 15% Al₂O₃ (Hess et al. 1977); in contrast, most picritic glasses contain less than 10 wt% Al₂O₃ (Delano 1986; Table A3.12) and primitive mare basalts typically less than 12 wt% (Neal and Taylor 1992; Table A3.11). Mare basalts with greater Al₂O₃ contents are differentiated, have suffered accumulation of plagioclase, and/or reflect a characteristic of a mantle source (high-Al mare basalts). The Eu anomaly is therefore intrinsic to the source and is commonly attributed to the fractionation of plagioclase to form the floated anorthosite crust. (BVSP 1981). Pyroxene-bearing cumulates complementary to the anorthosite would bear the negative Eu signature. Small Eu anomalies recorded in some VLT glasses and basalts may not require plagioclase fractionation but reflect the very low distribution coefficient for Eu⁺² in low-Ca pyroxene (McKay et al. 1990). It is also possible that the Eu anomalies are products of source hybridization from Eu depleted liquids. Europium anomalies in high-Ti magmas are, on average, much larger than those in VLT and low-Ti magmas (BVSP 1981; Papike et al. 1997) and can be generated only from liquids that have lost significant quantities of plagioclase. That the depth of the Eu anomaly increases with total HREE contents is consistent with a source hybridized by various amounts of trace element-rich components that formed from liquids which had lost substantial plagioclase.

Most high-Ti and low-Ti mare basalts have a distinctive chondrite-normalized REE pattern that distinguishes lunar from terrestrial basalts. The mare basalts typically possess positive LREE slopes and negative HREE slopes, together forming a convex-up REE pattern (Fig. 4.38) (Papike et al. 1997, 1998). Terrestrial ocean-floor basalts (N-MORB), which cover 70% of the Earth, contain positive LREE slopes but generally flat HREE (e.g., Hess 1989). The N-MORB are products of partial melting of somewhat depleted lherzolite, peridotites that still contain roughly 10% (±) of high-Ca clinopyroxene. The primary melts parental to N-MORB were equilibrated with olivine, orthopyroxene, high-Ca clinopyroxene, spinel and/or garnet. The REE patterns of N-MORBs appear to be largely controlled by high-Ca clinopyroxene; in fact, the shape of DREE patterns of high-Ca clinopyroxene is very similar to the REE patterns of N-MORBs (Fig. 4.39). The distinctive convex-upward REE patterns of mare basalts, therefore, give us a clue to their petrogenesis.
The convex-upward REE patterns observed in mare basalts cannot be derived by partial melting of primitive lherzolite nor of highly depleted harzburgite mantle. Partial melts from fertile lherzolite have negatively sloped REE patterns if garnet is in the residue and negatively sloped LREE and flat HREE patterns if a garnet-free lherzolite assemblage is in the residue. Melting of high-Ca clinopyroxene-free harzburgite produces melts with positively sloped REE patterns, whereas if the harzburgites are high-Ca clinopyroxene-bearing, the HREE pattern will tend to be flat. A source formed of harzburgitic cumulates from a LMO with a flat chondrite-normalized REE pattern largely mimic the REE patterns of harzburgite created by large degrees of melting of primitive mantle. High-Ca clinopyroxene-free harzburgites have positively sloped REE patterns (e.g., BVSP 1981; Hess 1989). Partial melts generated from these cumulates will also have positively sloped REE patterns for relatively small amounts of melting.

The distinctive mare-basalt convex REE pattern must, in part, reflect a fractionated non-chondritic source, most likely the mafic cumulates of the LMO (BVSP 1981). The REE content of the cumulates is controlled by the REE pattern of the magma, the nature of the crystallizing phases and the amount of trapped intercumulus liquid. Multiplication of the bulk distribution coefficient by the abundances of the REE in the LMO existing at the time the source accumulated gives the REE content of the adcumulate. The source REE is a weighted average of the REE content of the adcumulate and intercumulus liquid. Nyquist et al. (1979) completed this exercise and concluded that the negatively sloped HREE pattern was also a feature of the LMO parental to the cumulates (i.e., the LMO did not have a flat chondrite-normalized HREE pattern). The REE patterns of low-Ti mare basalts could be reproduced by small degrees of melting of an adcumulate that retained significant quantities of high-Ca clinopyroxene in the mantle residue.

Hughes et al. (1988) argued instead, that the REE patterns in low-Ti mare basalts were produced by small degrees of melting of sources in which olivine and orthopyroxene alone were residual phases. The adcumulates were crystallized from a chondritic LMO (i.e., flat REE pattern). The negatively sloped HREE pattern was created through the hybridization of the adcumulate source through addition of late-stage LMO cumulates and their intercumulus liquid. Hybridization resulted from the convective overturn of the gravitationally unstable cumulates. Similar models were devised to produce an adequate source for the high-Ti mare basalts (Hughes et al. 1989). Unfortunately, such models are not always internally consistent when major-element constraints are included in the analysis.

Hughes et al. (1989) in their attempts to model the petrogenesis of 74220 orange glass, for example, concluded that the trace elements could be duplicated by melting 4-7% of a hybridized source formed from 95.8% of cumulate olivine and 4.2% late-stage cumulates composed of 24% augite, 40% plagioclase, 35% ilmenite, plus small but critical amounts of apatite and highly evolved KREEPy intercumulus liquid. Whereas the trace elements in these melts were similar to those of the orange glass, the major elements were not. A very powerful...
constraint on the petrogenesis is that ilmenite, clinopyroxene, and plagioclase (or spinel) are not found on the high-pressure liquidus of the orange glass (Green et al. 1971) and must therefore be consumed and eliminated from the source. The liquid produced by 5% melting and leaving only olivine in the residue has about 16 wt% TiO$_2$, 11 wt% Al$_2$O$_3$, and 9 wt% CaO. In comparison, the orange glass has about 9 wt% TiO$_2$, 6 wt% Al$_2$O$_3$, and 7 wt% CaO. The melt produced by 7% melting is a better fit for the major elements (11 wt% TiO$_2$, 7 wt% Al$_2$O$_3$, and 6 wt% CaO) but is a poorer fit for the trace elements.

The chondrite-normalized REE patterns of the picritic glasses are more varied and more clearly express the effects of source hybridization and/or assimilation (Fig. 4.40) (Shearer and Papike 1993). Whereas a few of the high-Ti picritic glasses have the characteristic convex-upward mare basalt REE pattern, the Apollo 15 and 17 orange glasses for example, some REE patterns are negatively sloped. Apollo 14 and 15 high-Ti picritic glasses have such negatively sloped REE patterns. Some VLT picritic-glass REE patterns are also negatively sloped in direct contrast to the REE patterns in crystalline VLT basalts, which generally have a positive slope. These VLT glasses have LREE abundances that are more than 30× chondritic values whereas the VLT mare basalts have LREE abundances at less than 10× chondritic. The negatively sloped patterns appear to record some incorporation of KREEP-like components in the source (Shearer and Papike 1993). Apollo 16 green glass has low total REE, almost no negative Eu anomaly and a flat REE pattern. Only the VLT Apollo 14 green glass has the convex upward REE pattern.

The Ni and Co contents of VLT glasses vary within a factor of 2 from their mean values (Shearer and Papike 1993). More significantly, the Ni/Co ratio does not vary widely and, with a few exceptions, is constrained to values between 2 and 3 (Shearer et al. 1996; Table A3.12). Both Ni and Co are compatible in olivine and pyroxene; the $D_{Ni}$ is roughly 10 and 3 and $D_{Co}$ is 1.7 and 1.1 for olivine and orthopyroxene, respectively. Using the batch melting equation, the ratio of Ni/Co in melts derived by relatively modest degrees of melting ($\leq 20\%$) is approximately the ratio of the bulk distribution coefficients, $D_{Ni}/D_{Co}$, multiplied by the initial Ni/Co ratio, $C°_{Ni}/C°_{Co}$ of the source.

It follows that the constrained variations in the Ni/Co content of the VLT glasses required that the product of bulk distribution coefficients and the initial Ni/Co contents of the source vary only slightly. It also requires that the most primitive VLT glasses experienced only very little olivine fractionation subsequent to their separation from the mantle. As little as 10% fractional crystallization of olivine will change the Ni/Co ratio in the melt by a factor of more than 2.

Similarly, the Ni/Co contents of the cumulates of the LMO will vary not only with the extent of fractionation but also as a function of whether olivine or orthopyroxene is the dominant liquidus phase. Forty percent fractionation of olivine, for example, lowers the Ni/Co ratio by more than an order of magnitude (Fig. 4.41) (Shearer et al. 1996). The effects of pyroxene fractionation are directly related to the degree of fractionation ($F$) and the Ni/Co ratio decreases with $F$.

The limited range of Ni/Co ratios in VLT picritic glasses sets important limits on the Ni/Co contents of the source region. Since Co and Ni are both strongly compatible, the Ni/Co ratio is a robust indicator of the cumulate mantle and is little affected by subsequent hybridization from highly evolved cumulates and KREEP-enriched liquids, provided that the latter occur in relatively small amounts. First, the sources must be distributed over a limited “stratigraphic section” of the cumulate pile (Shearer et al. 1996). The stratigraphic section represents roughly 10% of the volume of the cumulate pile. Second, this same ratio characterizes VLT glasses from Apollo 11, 14, 15 and 17 sites, implying that the source mantle for VLT picritic glasses is widely distributed across the lunar near-side mantle.

The Co and especially the Ni contents of high-Ti picritic glasses are lower than those in the VLT glasses. Nickel contents typically are, with a few exceptions, less than 50 ppm and some
Figure 4.40. Chondrite normalized REE in lunar picrite glasses (modified after Papike et al. 1998).
glasses have Ni below detection limits. Cobalt contents are within the same range or somewhat lower than those in VLT glasses. The Ni/Co ratios are, therefore, comparatively low and range from about 0.5-1.1 (Shearer and Papike 1993; Table A3.12). Taken at face value, these data suggest that the cumulate source of high-Ti picritic glasses is located in a somewhat more evolved part of the cumulate stratigraphy. The same conclusion may apply to the crystalline mare equivalents. The Ni/Co ratios in crystalline high-Ti mare basalts range to values as low as 0.1 with Ni contents of a few ppm (Table A3.11). Both Co and Ni contents are lower in crystalline high-Ti mare basalts than in the corresponding glasses and possibly reflect the fractionation of olivine from the parent magmas.

Accumulation of ilmenite in the high-Ti mare and picritic glass source(s) can be examined from the concentration of certain trace elements that partition into ilmenite over olivine and pyroxene. The transition elements Hf, Zr, Nb, and Ta tend to favor Ti-rich phases like ilmenite, armalcolite, sphene and rutile (McCallum and Charette 1978; Jones 1995; Horng and Hess 2000). Ilmenite/basalt partition coefficients for Zr ($D_{Zr} = 0.33$) and Nb ($D_{Nb} = 0.8$) are at least an order of magnitude greater than those for olivine and pyroxene (Table 4.8). During most of the crystallization of the LMO where olivine and orthopyroxene are the liquidus phases, inter-element ratios between the above elements should remain constant, presumably at chondritic values. However, when ilmenite and, to a lesser extent, high-Ca pyroxene join the crystallization sequence, interelement ratios between the high-valence transition elements and trace elements

Table 4.8. Approximate crystal/melt partition coefficients for basalts at 1200 °C.

<table>
<thead>
<tr>
<th>Element</th>
<th>OLV</th>
<th>OPX</th>
<th>CPX</th>
<th>ILM</th>
<th>GNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>La</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.05</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>Nd</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.15</td>
<td>0.007</td>
<td>0.03</td>
</tr>
<tr>
<td>Sm</td>
<td>0.001</td>
<td>0.002</td>
<td>0.2</td>
<td>0.01</td>
<td>0.20</td>
</tr>
<tr>
<td>Lu</td>
<td>0.02</td>
<td>0.04</td>
<td>0.4</td>
<td>0.07</td>
<td>7.0</td>
</tr>
<tr>
<td>Co</td>
<td>1.7</td>
<td>1.1</td>
<td>1.5</td>
<td>4.3</td>
<td>9.0</td>
</tr>
<tr>
<td>Ni</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>3.5</td>
<td>—</td>
</tr>
<tr>
<td>Ti</td>
<td>0.01</td>
<td>0.1</td>
<td>0.3</td>
<td>—</td>
<td>0.3</td>
</tr>
<tr>
<td>Hf</td>
<td>0.01</td>
<td>0.05</td>
<td>0.2</td>
<td>0.4</td>
<td>0.07</td>
</tr>
<tr>
<td>Zr</td>
<td>0.01</td>
<td>0.05</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Nb</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
<td>0.8</td>
<td>0.002</td>
</tr>
<tr>
<td>Ta</td>
<td>0.005</td>
<td>0.005</td>
<td>0.01</td>
<td>4.0</td>
<td>0.02</td>
</tr>
<tr>
<td>W</td>
<td>0.01</td>
<td>0.02</td>
<td>0.2</td>
<td>0.1</td>
<td>0.003</td>
</tr>
<tr>
<td>Sc</td>
<td>0.2</td>
<td>0.3</td>
<td>1.0</td>
<td>1.5</td>
<td>2.0</td>
</tr>
</tbody>
</table>

The $D$ values are not constants but are functions of $T$, $P$, melt composition and crystal composition, particularly the wollastonite content of pyroxene (Forsyth et al. 1994; Skulski et al. 1994; Shearer and Papike 1986; Shearer et al. 1995; Green et al. 2000; McCallum and Charette 1978; Jones 1995).
strongly incompatible with these phases, such as Ce, will become fractionated. Shearer et al. (1996) concluded that the source regions for both VLT and high-Ti picritic magmas were fractionated with respect to lunar abundances for Nb/Ce and Zr/Ce. The high-Ti source region, in particular, was interpreted to contain small additions of ilmenite, but no more than a few percent. Neal (2001) found that the Hf/Th and Ta/Th are decidedly suprachondritic for high-Ti mare basalts, supporting the hypothesis that ilmenite was added to the source region. Perhaps even more compelling is the suprachondritic ratio of Ta/Nb in high-Ti mare basalts because the partition coefficient for Ta is about twice that for Nb in TiO$_2$-rich minerals (Horng and Hess 2000).

5.5. Isotopic observations

Isotopic data for mare basalts are presented as initial $\varepsilon_{\text{Nd}}$ vs. crystallization age in Figure 4.42, and initial $\varepsilon_{\text{Hf}}$ vs. initial $\varepsilon_{\text{Nd}}$ diagrams in Figure 4.43 (Nyquist and Shih 1992; Snyder et al. 1994; 1997; 2000; Beard et al. 1998). The discussion will focus first on the Nd data. All mare basalts have positive $\varepsilon_{\text{Nd}}$ bounded by the chondritic evolution line, by definition $\varepsilon_{\text{Nd}} = 0$, and an $\varepsilon_{\text{Nd}}$ growth curve corresponding to a source with $^{147}\text{Sm}/^{144}\text{Nd} = 0.21-0.28$ (Nyquist et al. 1995). It follows that mare basalts are products of melting of sources with long-term LREE-depleted signatures as would be expected of pyroxene/olivine cumulates from an LMO with near chondritic Sm/Nd. Furthermore, the isotopic systematics record varying degrees of Sm/Nd fractionation in the source mantle. Apollo 12 ilmenite basalts and Apollo 17 high-Ti basalts derived from the most depleted sources, whereas Apollo 15 mare basalts came from the least depleted sources, and one sample is almost chondritic. These characteristics are difficult to duplicate in an LMO because the Sm/Nd ratio is roughly constant over more than 90% of its crystallization (Snyder et al. 1992; Beard et al. 1998). The source cannot be composed solely of primitive cumulates from the LMO.

The cumulate source is assumed to consist of two components. One component is an orthopyroxene-rich adcumulate; the cumulate would have a fixed Sm/Nd, as long as it derived from the main stages of crystallization of the chondritic LMO. A second component must have smaller Sm/Nd ratios and be enriched in both Sm and Nd relatively to the adcumulate so only small masses of this component are required. Snyder et al. (1994) suggested that this is a trapped-liquid component derived from more advanced stages of crystallization of the LMO. In their model, only a few percent of this component is needed to explain the observed data. Other enriched components are possible candidates but some form of hybridization is required to recreate the isotopic systematics of mare basalts (e.g., Hughes et al. 1989).

Interestingly, some of the Apollo 12 low-Ti mare basalts and Apollo 17 high-Ti mare basalts lay along the growth curve corresponding to the most extreme LREE depletion in the mantle source (Nyquist and Shih 1992). Other Apollo 12, Apollo 17 and Apollo 11 high-Ti mare basalts define different LREE depleted growth curves (Snyder et al. 1996). This commonality of sources from widely separated landing sites suggests that the source regions

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.42.png}
\caption{Initial $\varepsilon_{\text{Nd}}$ vs. crystallization ages of mare basalts and array for a depleted cumulate mantle. TL represents trapped liquid component (modified after Snyder et al. 2000).}
\end{figure}
extend Moon-wide and the timing of the depletion event was synchronous. These features imply a global process, no doubt the solidification of a global LMO (Nyquist and Shih 1992).

It is dangerous, however, to carry these conclusions much further. A commonality of sources is not as apparent in the Sr isotope data (Fig. 4.43). Initial $\varepsilon_{\text{Hf}}$ and $\varepsilon_{\text{Nd}}$ values for low-Ti and high-Ti mare basalts imply that the source regions for these lunar magmas are distinct (Fig. 4.44). Low-Ti mare basalts have higher initial $\varepsilon_{\text{Hf}}$ values, at a given $\varepsilon_{\text{Nd}}$, than the initial $\varepsilon_{\text{Hf}}$ values for high-Ti mare basalts (Beard et al. 1998). Indeed, the data reflect sources for low-Ti basalts that have chondrite-normalized Lu/Hf ratios approximately 4× greater than chondrite-normalized (Sm/Nd) ratios; high-Ti sources had roughly equal chondrite-normalized ratios of Lu/Hf and Sm/Nd. The two source arrays appear to intersect at KREEP basalt, perhaps indicating that these arrays are mixing lines emanating from the KREEP component (Snyder et al. 2000).

As was discussed earlier (Sections 4.2 and 4.3) it appears some of the mare basalts on the Moon have a significant radiogenic $\varepsilon_{\text{W}}$ signature, i.e., $\varepsilon_{\text{W}} > 0$ (Fig. 4.2). It is generally agreed that mare basalts were produced by melting of the cumulates from the LMO, thus the $\varepsilon_{\text{W}}$ signature is unlikely to have been inherited from the accreting bodies that formed the Moon (Shearer and Newsom 2000; Righter and Shearer 2003; Shearer and Righter 2003). Now if the lunar reservoirs received their W signature during core formation, it would be difficult to imagine a physical process that would not imprint all crystallization products with the same W depletion event. It is more likely that if this radiogenic W signature truly exists in mare basalts, it was imparted through the fractionation of high-Ca clinopyroxene and ilmenite. Those sources that have non-radiogenic W simply did not contain sufficient concentrations of the late-stage cumulates. High-Ti source regions
must be enriched in ilmenite and therefore should have $\varepsilon_W > 0$ as is observed. This places constraints on the age (Section 4.3) and composition of mare-basalt sources.

One difficulty with this model is that crystallization of ilmenite and high-Ca pyroxene from late-stage magmas would have created KREEP-rich liquids with depleted Hf/W. However, there are no igneous rocks on the Moon with $\varepsilon_W \leq 0$; all measured rocks have $\varepsilon_W \geq 0$. One possible resolution of this paradox is that fraction of Hf/W is due to fractionation of small amounts of metal and as well as late-stage high-Ca clinopyroxene and ilmenite. In this scenario, the early LMO would have $\varepsilon_W \geq 0$ so that late-stage depletion of Hf into high-Ca clinopyroxene and ilmenite would not produce $\varepsilon_W < 0$. Resolution of these questions awaits more data.

5.6. Experimental phase equilibria

Phase equilibria experiments on mare basalts and picritic glass compositions are ably reviewed by BVSP (1981) and Longhi (1992) so only the most relevant highlights and more recent experimental results are presented in this section. Low-pressure phase equilibria on VLT glass compositions have only olivine as a liquidus phase for a temperature interval of about 200 °C (Fig. 4.45) (Chen et al. 1982; Chen and Lindsley 1983; Elkins-Tanton et al. 2000). Liquidus temperatures are very high, reaching almost 1400 °C, comparable to those of some of the less magnesium terrestrial komatiites (Hess 1989). Pigeonite is the second liquidus phase, followed by plagioclase and then augite (Longhi 1992). The late appearance of plagioclase and augite contrary to their comparatively earlier crystallization in terrestrial ocean-floor tholeiites is a consequence of low CaO and Al$_2$O$_3$ contents. Indeed, VLT and also high-Ti picritic glasses are strongly undersaturated with respect to these phases at all pressures (Longhi 1992).

High-Ti picritic glasses have olivine on or near the 1-bar liquidus but the most Ti-rich compositions (TiO$_2$ > 18 wt%) also include minor amounts of Ti-rich chrome-spinel as a liquidus phase (Delano 1980; Wagner and Grove 1997). Olivine is the sole liquidus phase in the less Ti-rich Apollo 74220 orange glass (TiO$_2$ ~ 9 wt%), however (Green et al. 1975). Liquidus temperatures are somewhat lower those of the VLT glasses but are still high ($T \sim 1300$ °C). Armalcolite ([Fe,Mg]Ti$_2$O$_5$) is the next major phase to appear but develops a reaction relation with liquid after ilmenite appears at only slightly lower temperatures. Pigeonite, plagioclase and high-Ca clinopyroxene appear in that order at temperatures around 1200-1150 °C.

In summary, the common features of the 1-bar phase diagrams of the picritic glasses are that they are characterized by (1) high liquidus temperatures, (2) olivine (± Cr spinel) as the sole liquidus phase, and (3) the remaining silicate and oxide phases are relegated to much lower temperatures. These features suggest that the picritic glasses are primitive magmas generated at considerable depths in the lunar mantle. These conclusions are even more robust
if the source of these magmas contained both olivine and orthopyroxene. The expansion of the olivine field with decreasing pressure brings only olivine to the 1-bar liquidus: the greater the reduction in pressure, the larger the olivine liquidus field. The size of the olivine liquidus field in picritic glasses is more like those of terrestrial komatiites than of basalts, again an indication of the depths at which such magmas were generated.

The high-pressure phase equilibria are relatively simple. Olivine is the liquidus phase until orthopyroxene joins and then replaces olivine at high pressures. The pressures and temperatures where the picritic glasses are multisaturated with these two phases are given in Table 4.9. It is possible that garnet appears on the liquidus at higher pressures (Chen et al. 1982; Shearer et al. 2003; Draper et al. 2004). High-Ca pyroxene, plagioclase or ilmenite are not stable on the liquidus. The ilmenite liquidus, for example, is more than 50-100 °C below the temperatures of multiple saturation even in the most Ti-rich glasses (Delano 1980; Wagner and Grove 1997). Chrome spinel is a high-pressure liquidus phase in some picritic melts but the stability of this phase is very sensitive to Cr$_2$O$_3$ and oxygen fugacity, so the petrogenetic significance of this phase is unclear. Although low CaO pyroxene is the liquidus phase at high pressure, the pyroxene becomes progressively more calcic with decreasing temperatures or with increasing pressures at fixed temperatures (Green et al. 1971).

The phase-equilibria data for crystalline mare basalts are more variable because we are dealing with magmas of varying crystallinity and at different states of fractionation. The 1-bar liquidus of high-Ti mare basalt 70215, for example, is multiply saturated with olivine, armalcolite, and spinel, a feature not characteristic of primitive magmas generated at depth (Kesson 1975). If we compare these results with those of the high-Ti picritic glasses, then 70215 could be the derivative melt of a picritic glass evolved to some 100 °C below the liquidus (Longhi 1992). Other mare basalts have only olivine on the liquidus but appear to be compositions that accumulated excess olivine (Green et al. 1971; Walker et al. 1976). The petrogenetic significance of the phase equilibria of crystalline mare basalts is considered in the next section.

In the subsolidus, high-Ti mare basalts are ilmenite gabbros containing ilmenite, High-Ca clinopyroxene, and plagioclase as major phases (O’Hara et al. 1970; Ringwood and Essene 1970). With increasing pressure, the gabbro assemblage undergoes a continuous phase transition where garnet gradually replaces plagioclase as one of the major phases, rutile replaces part or all of the ilmenite. An eclogite or garnet-pyroxenite assemblage is stable at pressure greater than 1 GPa depending on the temperature and bulk composition. The eclogite assemblage is garnet + high-Ca clinopyroxene + rutile ± ilmenite ± quartz (O’Hara et al. 1971; Ringwood and Essene 1971).

### Table 4.9. The pressures and temperatures of olivine + orthopyroxene multi-saturation of phase equilibria experiments on picritic glass compositions.

<table>
<thead>
<tr>
<th>Picrite Glass</th>
<th>TiO$_2$ (wt%)</th>
<th>Pressure (GPa)</th>
<th>Temp (°C)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apollo 14B</td>
<td>0.45</td>
<td>2.4</td>
<td>1560</td>
<td>(a)</td>
</tr>
<tr>
<td>Apollo 14 VLT</td>
<td>0.66</td>
<td>1.9</td>
<td>1500</td>
<td>(b)</td>
</tr>
<tr>
<td>Apollo 17 VLT</td>
<td>0.71</td>
<td>1.8</td>
<td>1500</td>
<td>(b)</td>
</tr>
<tr>
<td>Apollo 74220</td>
<td>8.9</td>
<td>2.0</td>
<td>1490</td>
<td>(c)</td>
</tr>
<tr>
<td>Apollo 15 Red</td>
<td>13.8</td>
<td>2.5</td>
<td>1460</td>
<td>(d)</td>
</tr>
<tr>
<td>Apollo 14 Black</td>
<td>16.4</td>
<td>1.5</td>
<td>1430</td>
<td>(e)</td>
</tr>
</tbody>
</table>

5.7. Depth of the mare basalt source

High-pressure phase-relation experiments on mare basalts or their analogs record pressures of olivine plus orthopyroxene multiple saturation, which range from about 0.5 to 2.5 GPa (Longhi 1992). Multiple saturation pressures for picritic glasses, in contrast, are generally higher and define the range from 1.5-2.4 GPa (Table 4.9). If the pressures of multiple saturation have petrogenetic significance, the data imply that mare basalts come from sources that are distinct from those that gave rise to the picritic glasses. The mare-basalt source region would exist in the upper lunar mantle at depths from 100 to 500 km and the mantle would consist of cumulates more evolved than those parental to the picritic glasses. The Mg′ values of fine-grained mare basalts, for example, are typically much lower than comparable picritic glasses (Delano 1986; Shearer and Papike 1993). However, before the implications of these ideas to lunar evolution are considered, we must be certain that the mare basalts pass several important tests; not only must the mare basalt compositions represent quenched liquid but these compositions must represent primary or at least, very primitive liquids.

The major-element compositional diversity of mare basalts is, in part, the result of fractionation of olivine and other near-liquidus phases (Papike et al. 1976, 1997, 1998 and references therein). It follows that only the most primitive basalt of any batch of magmas related by crystal fractionation qualifies as a potential primitive or possibly primary magma. Only a few of the mare basalts survive this filter. It is useful to illustrate how rare such primitive basaltic magmas are in the terrestrial database. The example below considers only terrestrial ocean-floor basalts because the lithospheric barriers to eruption are minimized in this tectonic environment.

The following criteria are used to identify primary ocean-floor basalts (Hess 1992). A primary basalt must have only olivine on the 1-bar liquidus, an Mg′ value of 68 or greater, and an MgO content of 9 wt% or greater (Hess and Head 1990). Out of a total of more than 900 chemical analyses of ocean-floor vitrophyres, where glass assures that the basalt represents a liquid, only 21 samples had the appropriate MgO and Mg′ values, and then only marginally. At best, less than 2.5% of the liquids erupted on the ocean floor can be considered a primitive liquid. These primitive basalts are probably not primary because a primary liquid derived from appropriate depths in the terrestrial mantle should have MgO values 10-30 wt% higher than recorded in these samples. Given that mare basalts must pass through a thick crust and lithosphere, it is unlikely that most crystalline mare basalts represent even primitive compositions. If such exist, they are found only in the most radically quenched compositions, particularly those compositions that are equivalent to liquids.

The identification of quenched-liquid compositions for crystalline mare samples is not a trivial task. Fine-grained to vitrophyric basalts produced by rapid cooling will most likely have preserved the original liquid composition. Aphyric samples are relatively rare, however most contain microphenocrysts and/or complex textural domains (BVSP 1981). Walker et al. (1976) found a strong positive correlation between grain size and normative olivine content in Apollo 12 olivine basalts. They argued convincingly that most samples were differentiated by simple olivine settling perhaps in the basal portion of a cooling unit. It follows that there are sets of complementary basalts that have differentiated by losing olivine.

Experimental petrology also provides a number of tests to help identify phenocryst-enriched compositions, such as the “phenocryst-matching” test. This test compares the compositions of the earliest phenocrysts in fine-grained basalt to the composition of the liquidus phase determined by experiment at 1 bar. Provided that the containers used in the experiments do not change the bulk composition of the basalt (the use of Fe, Pt or even Mo, containers can alter the FeO content of basaltic liquids), the composition of the phenocryst and liquidus phase must be identical if the basalt in question represents a liquid composition. Green et al. (1971) used this technique to examine the petrogenesis of Apollo 12 olivine basalts.
High-pressure multiple-saturation experiments on ilmenite basalt 12022 have olivine and low-Ca pyroxene on the liquidus at pressures of about 15 kb (Green et al. 1971). Olivine of composition Mg′ = 77 is the sole liquidus phase at 1 bar pressure and 1290 °C. Pigeonite does not appear until about 1150 °C just above the temperatures of the incoming of ilmenite and plagioclase. The natural rock contains olivine phenocrysts with core compositions of Mg′ = 69, a composition attained only at about 1170 °C in the experiments. These disparate results are most readily interpreted by assuming that the basalt sample 12022 is a partial cumulate of olivine of Mg′ = 69. Because olivine is much richer in MgO than the coexisting liquid, concentrating olivine in the magma increases both the normative olivine and Mg′ of the rock so that when it is totally melted, both its liquidus temperature and Mg′ exceed the values of the true liquid parent. Because the liquid is enriched in normative olivine, the high pressure multiple-saturation point will exist at pressures that exceed the multiple-saturation point of the corresponding liquid without the added olivine. Clearly, only basalts that can be equated to liquids have phase equilibria with direct petrogenetic significance, but difficulties in interpretation exist even for phase equilibria performed on liquids.

The depths to the mare source regions are difficult to constrain uniquely notwithstanding the consensus among most lunar petrologists that the source regions must lie within the deep lunar mantle. One of the most important constraints is provided by the high-pressure-liquidus phase relations. The liquids corresponding in bulk composition to the VLT and high-Ti picritic glasses have both olivine and orthopyroxene on their liquidi at pressures between 1.6-2.5 GPa (Table 4.9) (Longhi 1992; Wagner and Grove 1997; Elkins-Tanton et al. 2000) but have only olivine and orthopyroxene at lower and higher pressures respectively (Fig. 4.45). Neither ilmenite, nor high-Ca clinopyroxene, nor plagioclase have stability fields that are within 100 °C of the liquidi at high pressure.

The traditional interpretation of these phase equilibria is that these melts were derived from a source in which olivine and orthopyroxene were restite phases at depths indicated by the pressures of the point of multiple saturation. A number of criteria and assumptions both stated and unstated must be satisfied before such broad conclusions are warranted. One of the most basic assumptions is that the mare-basalt sample represents a liquid composition. Even if the sample represents a melt composition, the melt must represent a primary or at least a very primitive basalt. To be “primary,” the liquid composition must not have been substantially modified since the liquid was equilibrated and segregated from its source and then erupted onto the lunar surface. This requirement is a tall order on any planet. Recall that primitive ocean-floor basalts on the Earth are rare even in a tectonic environment where the lithospheric is thin to absent. Mare basalts, in contrast, are dense magmas that must pass through up to 60 km or more of anorthositic and mafic crust plus a substantial thickness of mantle lithosphere to erupt onto the lunar surface. Significant modifications to the original magma composition may result through crystallization of olivine and other near liquid phases and through assimilation of crustal materials. Even if mare liquids survive this transit intact, there remains the relevant question whether true primary magmas really exist in nature.

Longhi (1992) models the green glass VLT composition as the product of polybaric fractional melting. The final composition of melt erupted onto the lunar surface has not equilibrated with a well defined mantle residue at a specific pressure and temperature, but rather is a composite of melts generated over a range of pressures, temperatures, and mantle sources. Under the most favorable circumstances, the final melt composition reflects only an average depth of melting from some weight-averaged mantle. However, if the mantle is heterogeneous on various scales, this composite melt cannot simply be inverted to reproduce even the average conditions of melting. In any case, the concept of a primary melt loses its meaning and significance. Indeed, one of the most serious challenges facing lunar petrologists is to understand and model the physics of the generation, segregation, and transport of melt within the lunar mantle.
These caveats notwithstanding, high-pressure phase relations, particularly those that record multiple-saturation points of olivine and orthopyroxene, provide important constraints for models of petrogenesis. The significance of an olivine-orthopyroxene multiple-saturation point for terrestrial basalts has long been appreciated (BVSP 1981) because melting in the terrestrial mantle proceeds by first eliminating high-Ca clinopyroxene and the aluminous accessory phases from the mantle restite (e.g., Hess 1992). After about 25% melting, only olivine and orthopyroxene are left in the mantle residue and much higher degrees of melting are required to consume orthopyroxene. It follows that, at a minimum, primitive or near primary terrestrial basalts must have both olivine and orthopyroxene on their liquidus at their point of separation and isolation from the mantle.

No such arguments are relevant to the lunar mantle. The mare sources are believed to be cumulates composed of dunite, harzburgite, or olivine pyroxenite hybridized to various degrees by additions of high-level cumulates with small, but variable quantities of intercumulus liquid. Small but variable amounts of intercumulus liquids may also derive from the cumulates themselves and through the process of migration upwards in the pile (Snyder et al. 1992). It follows that additional arguments must be developed to argue that primitive mare basalts and picritic glasses indeed are multisaturated with respect to olivine and orthopyroxene.

One curious feature of the picritic glass compositions is their limited range of Mg'. At 1 bar these melts are in equilibrium with olivine with compositions between Mg' = 77-86 (Delano 1990). High-Ti picritic glasses have Mg' that range from 48 to 54 and are not correlated with their TiO_2 contents (Delano 1986; Shearer and Papike 1993). The VLT glasses are more primitive but are confined to Mg' numbers from 57-64 with one outlier at 68. As expected, the VLT glasses coexist with the most magnesian olivines with Mg' between 81-86. If hybridized cumulates are the source regions to mare basalts what physical processes constrain the Mg' of the mantle to only 9 Mg' units?

Depending on the Mg' number of the bulk Moon, cumulates derived from crystallization of the first half of the LMO would have Mg' greater than that observed for the source regions. In fractionating LMOs with initial Mg' = 88-79, olivine cumulates at the 50% solidification point have Mg' ranging from 89-86 respectively (Hess 1994). One mechanism to remove such mafic cumulates from participating in mare magma genesis is through overturn of the cumulate pile and then sequestering the most magnesian cumulates in a growing lithosphere (Ryder 1991; Hess and Parmentier 1995). At the time of mare basalt volcanism, the elastic lithosphere as given by the 800 °C isotherm is about 100-150 km thick and thickens at a rate of 100 km/Ga (Hess and Parmentier 2001). Isotherms of 1100-1200 °C reach deeper into the cumulate mantle, perhaps to 200-300 km depths depending on model parameters. Since the liquidus temperatures of picrite and basalt magmas exceed 1300 °C at these depths, it follows that cumulates at 200-300 km depths are not thermally conditioned to be the mare source region. A 300 km mantle section on the Moon includes roughly 50% of the volume of the Moon. If the LMO encompassed the whole Moon, then 50% of the cumulates would be contained in this cold trap and would be lost to magma genesis. With mantle overturn, these ultramafic cumulates would sequester phases with Mg' ≥ 86, explaining the high-end limits for mare sources. But how do we rationalize the absence of more Fe-rich mare sources?

The simplest explanation is that the more Fe-rich members of mare basalts are products of this iron-rich mantle. This conclusion is not warranted, however, because many, if not all, mare basalts are derivative melts produced by various amounts of fractionation and, thereby, have lowered Mg' values. In fact, most of the high-Ti crystalline mare basalts have little to no modal olivine and cannot be samples of the primary melts from an olivine-bearing cumulate mantle. These basalts cannot be used to constrain source depths as argued below. If the most evolved basalts are eliminated the composition of the mantle sources could possibly be extended by
5-10 Mg' units to olivine Mg' values of 65 (Delano 1990) but not much lower. Where then are the really Fe-rich cumulates?

In a fractionating LMO, the Mg' of the resulting cumulates vary dramatically only near the terminus of crystallization. In the model of Snyder et al. (1992), the Mg' of the mantle cumulates produced in the last 10% of fractional crystallization range from the mid 80’s to 0 (Fig. 4.46). For a whole Moon LMO, these cumulates represent a 100 km thick layer in the upper mantle but cumulate overturn would sequester such cumulates in the lower 600 km of the lunar mantle. Perhaps the small volume of iron-rich cumulates is sufficient reason not to expect large volumes of mare basalts to come from such sources. Such Fe-rich cumulates are also very dense and would resist mobilization by thermal convection. Pressure-release melting would be ineffective. Lastly, melts derived from such sources may never make it to the surface in their pristine state. These magmas, instead, might act as metasomatizing agents, which hybridized more mafic and more voluminous mare mantle sources.

These arguments suggest that the cumulate sources for mare magmas are fractionation products of the LMO between 50 and about 90% solidification. If these estimates are correct, then the cumulate sources for mare basalts are likely to be olivine-bearing orthopyroxenites (see also discussion in Hess 2000), but for arguments sake let’s assume that the cumulates are dunites rather then orthopyroxenites. For a chondritic bulk Moon and LMO, the lower half of the LMO cumulate pile would consist of dunites (Fig. 4.46) (e.g., Snyder et al. 1992). Such cumulates would be magnesium - an LMO with an initial Mg' of 82 would produce dunite cumulates with Mg' no lower than 88; a more magnesian Moon would produce even more magnesium dunites. Since the Mg' of olivine that would be in equilibrium with the picritic glasses is between 77-86, it follows that picritic glasses do not originate in the dunite layers of the Moon unless their parent liquids suffered significant olivine fractionation. Arguments developed by Delano (1986) and Shearer and Papike (1993) suggest that such olivine fractionation was limited only to a few percent. Dunites are, therefore, very unlikely to be the sources of the primitive mare basalts and glasses.

Evidence for orthopyroxene in the mare source region is given by the Sm/Nd isotopic systematics of mare basalts. For an assumed value for the initial $\varepsilon_{Nd}(=0 \ CHUR)$ and the age of the Moon (4.55 Ga), the initial $\varepsilon_{Nd}$ and the time of volcanism for a particular mare basalt gives the Sm/Nd ratio of the source (e.g., Nyquist et al. 1979; Beard et al. 1998). The present day $\varepsilon_{Nd}$ or the Sm/Nd of the mare basalt determines whether Sm was fractionated from Nd during the melting of the source mantle. Because for small amounts of melting, pyroxene (but

---

**Figure 4.46.** Composition of cumulates to a model magma ocean (modified after Snyder et al. 1992 and Shearer and Papike 1999).
not olivine) is capable of fractionating Sm from Nd, orthopyroxene is indicated in the mantle residue whenever (Sm/Nd) mare basalt < (Sm/Nd) source. Beard et al. (1998) and others have used these criteria to argue that orthopyroxene is required in the source mantle of low- and high-Ti mare basalts. Sources containing olivine and orthopyroxene are also indicating by the bulk compositions of the picritic glasses.

The compositions of the picritic mare glasses projected onto the MgO-SiO$_2$-TiO$_2$ phase diagram at 2 GPa (MacGregor 1969) define and follow the olivine-orthopyroxene boundary curve (Fig. 4.47). The scatter is not surprising given the approximations required to project the picritic glass compositions into the MgO-SiO$_2$-TiO$_2$ system. The fact that the pressures of multiple saturation of high-Ti and low-Ti picritic glasses vary from 1.6 to 2.4 GPa indicate that the picritic glasses were probably generated over a range of pressures (Longhi 1992). It is noteworthy, nevertheless, that the distribution of projected points certainly reproduces the expected trend of the cotectic. These results suggest that the picritic glasses were generated by melting (polybaric?) processes buffered by the olivine-orthopyroxene cotectic surface. We conclude therefore that the pressures of multiple saturation are petrologically meaningful and constrain the depths of melting to regions that straddle the apparent depths of multiple saturation. Mare picritic glasses were derived by melting of olivine-orthopyroxene cumulates initiated at average depths of 400 km or more.

The petrogenetic significance of high-pressure liquidus phase relations and, in particular, the saturation pressures and temperatures of non-primary crystalline mare basalts are more questionable. The main variable that determines such pressures is the normative olivine/low CaO pyroxene ratio, although other compositional parameters such as normative feldspar and ilmenite contents are also important. The most primitive mare basalts have olivine on the 1-bar liquidus, so the principal effects of near-surface differentiation is the subtraction of olivine from differentiated liquids and the addition of olivine to partial cumulates from such liquids. The former has the effect of lowering the depths of multi-saturation, the latter of increasing such depths. Subtracting roughly 5 mol% olivine from an Apollo 15 green glass composition lowers the pressure of multiple saturation from about 2.5 GPa to 1.5 GPa; the MgO melt composition, however, is reduced by only 1 wt% (from 16 to 15, Longhi personal communication). Increasing the MgO contents of high-Ti mare basalt 70215 by 4 wt%, the equivalent of adding 10 wt% olivine, increases the depth of multiple saturation by 200 km. It is clear from these examples that the effects of even modest degrees of low-pressure olivine fractionation can have dramatic effects on the apparent depths of multiple saturation. Longhi (1992) noted that most mare basalts have apparent depths of multiple saturation between 0.5 and 2.0 GPa. These values could be achieved by as little as 10 wt% olivine fractionation from more primitive parent liquids or the accumulation of olivine into more evolved basalts.

5.8. Remelting of cumulates

Source regions for mare basalts are generally modeled as cumulates from the
differentiation of the LMO. Partial melting of these cumulates, some of which are hybridized to various degrees by incompatible-element-rich components, generates the complex chemical signatures that characterize the mare-basalt suite. Of particular interest is the partitioning of TiO$_2$ and other incompatible or compatible trace elements during the initial cumulate-forming process and then in the subsequent partial-melting event that generated the basalts.

For simplicity, assume that the trace elements exist in chondritic proportions in the LMO. First, consider cumulates formed without intercumulus liquid. The trace-element content in the cumulates is given by the distribution coefficients for each element multiplied by the concentrations of the element in the liquid. The first partial melt from this cumulate will contain the same relative and absolute concentrations of these elements. Additional melting will dilute the incompatible elements and enrich the compatible elements in the liquid. As an example, the following describes the evolution of the TiO$_2$ content during partial melting of a system that is undersaturated with respect to ilmenite. TiO$_2$ is moderately incompatible in pyroxene and highly incompatible in olivine. Thus, the maximum amount of TiO$_2$ that can be generated in mare basalts is equal to the TiO$_2$ content of the LMO parental to the cumulate source. If we assume that the LMO has chondritic TiO$_2$ (about 0.2 wt%; Taylor 1980), the cumulates formed by 95% fractional crystallization coexist with a melt with about 4 wt% TiO$_2$. Depending on $P,T$ and bulk composition of the magma, ilmenite crystallization will terminate the TiO$_2$ enrichment in the residual liquids. We therefore consider only those cumulates formed prior to ilmenite saturation. Partial melting of these cumulates cannot produce melts with more than about 4 wt% TiO$_2$. Indeed, if the degree of melting is comparable to the bulk distribution coefficient, say $F = 0.1$, then the TiO$_2$ content in the resulting basalts will be less than 2 wt%. Therefore, it is doubtful that low-Ti mare basalts containing 2-6 wt% TiO$_2$ can be produced from only olivine-orthopyroxene cumulates of the LMO. Certainly, low-Ti basalts at the high end of this range require a TiO$_2$-rich phase in the source.

Possible TiO$_2$-rich phases include intercumulus liquid, cumulate ilmenite, or late stage high-Ca clinopyroxene. Although cumulate with a small amount of intercumulus liquid will produce partial melts with somewhat elevated TiO$_2$ content, they cannot produce TiO$_2$-rich mare basalts unless the intercumulus liquid makes up a significant fraction of the rock. Isotopic and trace-element considerations (Nyquist and Shih 1992; Snyder et al. 1994, 1997), however, limit the amount of acceptable intercumulus component to no more than a few percent. Because mare volcanism occurred more than hundreds of millions of years after the cumulates were created, the intercumulus component should have been fully crystallized and the compatible and incompatible elements would be partitioned among the existing cumulate phases, particularly if the cumulate was pyroxene-rich. The net effect will be to increase the contents of the highly incompatible elements but hardly affect the contents of the more compatible elements. Returning to the example, the TiO$_2$ content of a pyroxene cumulate would be increased by only a few percent of the amount present; a partial melt from this cumulate would contain little more TiO$_2$ than the melt from the barren cumulate. In an olivine cumulate, however, the crystallized intercumulus liquid may contain trace amounts of ilmenite depending on the composition and abundance of liquid. The first melts would then be ilmenite-saturated and contain appropriately elevated content of TiO$_2$. Ilmenite saturated melts contain more than 20 wt% TiO$_2$ at high $T,P$ (Beck and Hess 2001).

The required ingredients to effectively hybridize the cumulate source are TiO$_2$-rich components, but without the large amounts of trace-element-rich components that compromise the isotopic systematics. The late-stage ilmenite-rich cumulates have some of the desired properties. These cumulates contain high-Ca clinopyroxene, ilmenite, orthopyroxene, olivine, intercumulus melt and whatever plagioclase is not lost to the growing anorthositic crust. For simplicity, we consider only the clinopyroxene and ilmenite cumulates; these cumulates should exist in roughly a 5/1 ratio (Van Orman and Grove 2000). The cumulates would contain about 7-8% TiO$_2$. 
Ilmenite cumulates added to the olivine-rich adcumulate produce a source that contains excess high-Ca clinopyroxene and ilmenite. Small degrees of melting of this source to produce TiO$_2$-rich mare basalts must leave high-Ca clinopyroxene in the residual melt because the high-Ca clinopyroxene/ilmenite ratio is much greater than one. Such TiO$_2$-rich mare basalts must then be saturated with high-Ca clinopyroxene, contrary to what is observed in the most primitive mare basalts and picritic liquids, which are far from saturated in CaO-pyroxene. An olivine-rich cumulate with added high-level clinopyroxene-ilmenite cumulates is not an appropriate source for high-Ti mare basalts.

A different, more acceptable result follows if the cumulates are rich in orthopyroxene because orthopyroxene is a potential reservoir for both TiO$_2$ and the wollastonite (CaSiO$_3$) components. For simplicity, assume that the orthopyroxene is sufficiently primitive and contains very little normative FeTiO$_3$ or CaSiO$_3$. From the experiments of Delano (1980), orthopyroxene saturated with ilmenite contains about 2-3 wt% TiO$_2$. Since the high-level ilmenite cumulates contain 7-8 wt% TiO$_2$, it follows that significant amounts of this cumulate can be incorporated into orthopyroxene. Suppose that the hybridized mixture contains ilmenite cumulates and orthopyroxene in a 1/9 ratio. The remainder of the cumulate is formed from olivine. The orthopyroxene is strongly undersaturated with respect to high-Ca clinopyroxene and ilmenite, thus the hybridized orthopyroxene would contain about 1 wt% TiO$_2$ and 3 wt% CaSiO$_3$. Small degrees of melting of this hybridized orthopyroxene would produce TiO$_2$-enriched mare basalts, which are not saturated with respect to either ilmenite or Ca-pyroxene as required by the phase-equilibria constraints.

An alternative method to hybridize the orthopyroxenite cumulates follows a more indirect path. Hess and Parmentier (1995) suggested that only a small fraction of the dense ilmenite cumulates would be convectively mixed into the underlying cumulates in various proportions. Most of the ilmenite cumulates would ultimately collect in or around the lunar core. This model was guided by the constraint that none of the known mare basalts were saturated with respect to ilmenite and high-Ca clinopyroxene yet all appear to have been produced by relatively small degrees of melting. To achieve both these ends, the source for mare basalts could not have attained large additions of ilmenite cumulates. This delicate balancing act might be acceptable on Earth were continuous convective mixing of the mantle would dilute mantle inhomogeneities an appropriate amount. However, on the Moon convective mixing is not an efficient mechanism to homogenize the mantle.

It is more likely that the lunar cumulate mantle remains inhomogeneous on a number of scales. The downwelling ilmenite cumulates would form discrete pods of mantle strongly enriched in ilmenite and high-Ca clinopyroxene (Ringwood and Kesson 1976). Such pods or stringers are analogous to the pyroxenite or eclogite layers in the Earth’s mantle (Hirshman et al. 1996). These ilmenite cumulates are less refractory than the pyroxene cumulate mantle and have solidus temperatures a few hundreds of degrees lower than an olivine-orthopyroxene assemblage (Wyatt 1997). The ilmenite-high-Ca clinopyroxene cumulates would melt first and create melts initially saturated with high-Ca clinopyroxene and ilmenite. Such melts would contain more than 11 wt% TiO$_2$ (Van Orman and Grove 2000). Buoyancy and perhaps surface energy forces would drive these melts into neighboring, more refractory cumulates where they would hybridize orthopyroxene cumulates. Partial melting of these variously hybridized zones would create the spectrum of mare basalts that exist on the Moon.

It is worth emphasizing that the partial melts of the deep-seated ilmenite cumulates cannot give rise to observed mare basalts. First, such melts would be saturated with high-Ca clinopyroxene and perhaps even ilmenite. Second, near solidus temperatures are several hundred degrees lower than the liquidus temperatures of mare basalts. Even the equilibrium liquidus of the ilmenite cumulates is at a lower temperature than the corresponding mare basalts at the same depths. Finally, some of the high-Ti mare basalts and picritic glasses have
TiO$_2$ contents that are greater than the composition of the minimum partial melts derived from these cumulates (Van Orman and Grove 2000). We conclude that the ilmenite cumulates alone at depth are not the appropriate source for high-Ti mare basalts.

5.9. Volatile contents of mare magmas

Picritic mare glasses are generally believed to be the products of fire-fountaining (Delano 1979 and references therein; see also Chapter 3, Section 3.1.4). The forces driving the fountaining process appear to be related to the exsolution of volatiles, a record of which now coats the glass beads. Although volatile-driven eruption is certainly a viable and likely hypothesis, questions remain as to the origin of these volatiles, the nature of the forces driving fire fountains, and the extent of degassing and volatile loss. Fogel and Rutherford (1995) and Sato (1979) before them gave a detailed analysis of the most likely eruptive mechanisms. They concluded that graphite oxidation alone is capable of generating sufficient CO/CO$_2$ gas to account for the pyroclastic effect. The oxidation of graphite can proceed by a number of steps perhaps via the reduction of various oxides of Cr$_2$O$_3$, TiO$_2$ or even FeO. The exsolution of the CO gas induces other volatiles such as S, Cl etc. to partition into the gas and be lost to the surroundings. Indeed, glassy inclusions of orange glass trapped in olivine phenocrysts contain 600 ppm S compared to 200 ppm S in the glass bead, indicating at least 400 ppm loss of S from the degassing melt sphere (Weitz et al. 1999). Delano et al. (1994) and Shearer et al. (1998) determined the sulfur contents in the interior of 14 varieties of picritic glasses varying in composition from VLT to high-Ti affinities. The mean S values were from 170 ppm to 540 ppm S. In comparison, bulk sulfur contents of the glasses which included the volatile-rich coatings were more than 4× greater. The S contents of the interiors of picritic glasses are also significantly lower than those in crystalline mare basalts. If both crystalline mare basalts and picritic glasses initially had the same S content, then up to 90% of the original S was lost from these picritic melts.

There are two most likely sources for the gases. One is primitive, undegassed lunar interior. The other is the late-stage dregs of the LMO. If most of the Moon was processed by the LMO, then the volatile elements, which in many cases are also incompatible, would have been concentrated into the KREEPy remains of the LMO. Some of these materials would have been swept back into and hybridized the mantle. Partial melting of the hybridized mantle concentrates the volatiles into the melt. The volatile contents of mare magmas would then be controlled by the volatile contents of the source region and the degree of partial melting.

5.10. Buoyancy of mare basalts

Magma compressibility studies have shown that mare basalts, especially the intrinsically dense high-Ti mare basalts, eventually become more dense than the lunar mantle residue in which these melts are generated (Delano 1990; Circone and Agee 1996). The depths at which such magmas are neutrally buoyant depend on the TiO$_2$ content of the melt and the olivine/orthopyroxene ratio of the lunar mantle. The density of molten Apollo 14 black glass (TiO$_2$ = 16.4 wt%), for example, is the same as the density of olivine (Fo = 80) at 2.0 GPa but equals the density of orthopyroxene (En = 80) at 0.66 GPa. It follows that the Apollo 14 black glass is neutrally buoyant at depths from 120 km to 400 km for orthopyroxene and dunite mantles respectively. Densities of high-Ti mare basalts of more modest TiO$_2$ contents, as for example, the Apollo 17 orange glass (TiO$_2$ = 9.3 wt%), have depths of neutral buoyancy that range from 400 km (relative to orthopyroxene) to more than 800 km (relative to olivine) (Delano 1990). Several authors have cited these as significant if not fatal barriers to deriving mare basalts from deep with the interior of the Moon. Indeed, melts that are denser than their local surroundings will sink into the interior until neutral buoyancy is achieved rather than rise to the surface (e.g., Delano 1990; Wagner and Grove 1997).

Hess (1991) has argued that even negatively buoyant melts will rise to the lunar surface within ascending lunar diapirs, provided that the ascent velocities of the diapir exceed the veloc-
ity of the downward percolating melt. In this case the melt will either be carried past the depths of neutral buoyancy and/or be trapped above the lower thermal boundary of the diapir. If the rate of melt segregation initially exceeds the rate of diapiric ascent, then the melt will descend not only relative to the diapir but also relative to the surrounding mantle. A melt that progresses downward along the mantle adiabat to higher pressures must crystallize and therefore halt its descent. The crystallized products would then be carried upward and then remelted. The net effect is that the mean velocity of downward segregation is much less than the value calculated for the melt for which no crystallization is allowed. It thus appears that the depths of neutral buoyancy are not insurmountable barriers to deriving dense magmas from the lunar interior.

The negative buoyancy of high-Ti magmas is a key physical asset in the models of Elkins-Tanton et al. (2003). According to their models, the simple sinking of solidified high-Ti cumulates is unlikely because at the temperatures of solidification at around 1100 °C the viscosity of pyroxene-rich mantle is high, too high to allow these cumulates to sink. Instead, Ti-rich hybridized mantle is created by partial melting of these cumulates. At the "eutectic" of about 1125 °C at 100 km depth, the composition of the melt is in a ratio of two parts clinopyroxene and 1 part ilmenite (Van Orman and Grove 2000). These melts are Al$_2$O$_3$-poor (near solidus melts contain only 4-5 wt% Al$_2$O$_3$), and the densities of these melts at about 3.2 gm/cm$^3$ are slightly denser than the pyroxene-rich cumulates below the ilmenite cumulate layer. If the mantle is permeable to the Ti-rich melts, these negatively buoyant liquids would percolate into the underlying mantle, producing a fertile and hybridized mantle. These and other models are critically reviewed in the “Petrogenesis” section (4.5.12).

5.11. Insights from remotely sensed data

While lunar samples provide valuable insights into the character of the lunar mantle, absolute ages of mare basalt flows, and magmatic processes, remotely sensed data provide a planetary-scale context for the sample data. Remote-sensing observations have provided insights into the distribution of mare basalts, the distribution of mare-basalt compositions, and the relative ages of basalt flows. These observations and resulting conclusions are summarized in Chapters 1-3. The remote sensing observations also provide further interpretation of the mare-basalt sample suite. The sample-return data suggest that crystalline mare basalts are bimodally distributed between high-Ti mare basalts (TiO$_2$ ≥ 9 wt%) and low-Ti basalts (TiO$_2$ ≤ 6 wt%) with peaks at about 2.5 wt% TiO$_2$ and 12-13 wt% TiO$_2$ (see Chapter 2, Section 10.5.3). These data are of limited value, however, because they are not representative of mare volcanism given the statistically small sample. These data, moreover, are not mass weighted and simply represent the total number of basalt samples that were analyzed.

Remotely sensed TiO$_2$ abundances for all near-side lunar mare are area-weighted and thus more statistically representative (Giguere et al. 2000; Chapter 2, Section 10.5.3). The remote sensing technique is reasonably well calibrated in the 0.5 to 10.0 wt% TiO$_2$ range but still includes a generous ± 1.1 wt% standard deviation. The TiO$_2$ contents are not for mare basalts but rather for lunar mare soils. A comparison of TiO$_2$ values of rocks and regolith soils suggests that the TiO$_2$ values are diluted by the addition of low-TiO$_2$ highland materials (Gillis et al. 2003). The true TiO$_2$ values are between 10-30% higher than the observed values. The remotely sensed TiO$_2$ data, nevertheless, quite clearly demonstrate that there is a continuous distribution of TiO$_2$ values with no hint of the bimodality recorded in the returned sample set. Some of the continuity may be ascribed to mixing of soils under the influence of impacts. Nevertheless, high-Ti mare samples represent only a small fraction of the area-weighted lunar surface. Low-Ti basalts in the 2-3 wt% range are the dominant mare lithology. Even VLT mare basalts far outweigh the high-Ti mare basalt compositions. Far-side data are more sparse, but they too are dominated by VLT mare samples, where TiO$_2$ <1.0 wt%. Basaltic lithologies found among the lunar meteorites support the remote-sensing result that low-TiO$_2$ basalts are the dominant mare volcanic product on the Moon.
This fact can be rationalized from several perspectives. High-Ti mare basalt magmas are 5-10% more dense than low-Ti counterparts (Delano 1990). Both magmas are denser than either the anorthositic crust or the lower mafic crust. Thus, buoyancy considerations alone favor the eruption low-Ti mare magmas. High-Ti mare basalts require a source enriched with variable amounts of normative ilmenite. The amount of normative ilmenite in a chondritic Moon is less than 0.5 wt% (Taylor 1980). The paucity of high-Ti mare basalts may simply reflect the low TiO$_2$ content of the bulk Moon.

5.12. Petrogenesis

Models for mare petrogenesis must integrate the vast amount of geochemical information, the phase equilibria, the physical settings of mare volcanism and their tectonic environment, the timing, volume and distribution of mare volcanism, the prolonged existence of mascons, and the evolution with time of the cumulate mantle and elastic lithosphere. This task is formidable; not only are some of the input parameters model dependent, but they are limited by the quality and abundance of data. This section concludes with an overview of what critical data and criteria most effectively place limits on petrogenesis. First, we give a brief critical overview of existing petrogenetic models.

Early and still influential models for mare petrogenesis called for the production of low-Ti mare basalts from deep within the lower mantle whereas high-Ti mare basalts were to have formed by melting of upper-level (shallow mantle) ilmenite-rich cumulates (Taylor and Jakes 1974). The hypothesis that high-Ti mare liquids are the direct melt products of pyroxene-ilmenite cumulates residing below the anorthositic crust still finds some support (Manga and Arkani-Hamed 1991). These authors suggested that partial melting of the ilmenite-cumulates beneath the near-side basins occurred because of the thermal insulating effect of ejecta blankets from giant impacts. Unfortunately, the model faces some considerable difficulties. Primitive high-Ti mare basalts and their ultramafic parents have liquidus temperatures more than 1300 °C at subcrustal pressures (Delano 1980; Wagner and Grove 1997). Reheating the subcrustal mantle to these temperatures in the 3.8-3.9 Ga time period conflicts with the strong and cool upper mantle needed to support lunar mascons (Hess and Parmentier 2001). A more direct argument is that primitive high-Ti mare liquids are not saturated with respect to high-Ca clinopyroxene and ilmenite at any pressures. The shallow cumulates contain high-Ca clinopyroxene and ilmenite in roughly a 5:1 ratio. Near-eutectic melts with more than 13 wt% TiO$_2$ are saturated with both ilmenite and high-Ca clinopyroxene (Van Orman and Grove 2000). Melts formed at higher temperatures appropriate to the liquidus temperatures of primitive mare liquids become undersaturated with respect to ilmenite, develop lower TiO$_2$ contents, and retain high-Ca clinopyroxene at the liquidus. These characteristics are unlike any known lunar mare basalt. These results effectively rule out a high-Ti magma source composed mainly of clinopyroxene and ilmenite.

The extended period of mare volcanism also precludes a shallow subcrustal source for high-Ti mare basalts. Radiometric ages place high-Ti mare volcanism in the early Imbrian Period (3.3-3.8 Ga) (Nyquist and Shih 1992; Snyder et al. 2000), but remote sensing observations extend the periods of volcanism to roughly 1.1 Ga ago (Shultz and Spudis 1993; Hiesinger et al. 2000, 2003). Some of the youngest basalts have 6-12 wt% TiO$_2$. The latter are surely high-Ti mare basalts. Given that the elastic lithosphere was thickening at about 100 km/Ga after the crystallization of the LMO (Parmentier and Hess 1998; Hess 2000; Hess and Parmentier 2001), it follows that the upper few hundred kilometers of the cumulate mantle was relatively cool, $T \leq 800$ °C, and would not be a source of volcanism in a young Moon, particularly high-Ti volcanics that have liquidus temperatures of 1200 °C or higher.

One set of models for the origin of high-Ti mare basalts involves the assimilation of shallow ilmenite-high-Ca clinopyroxene cumulates located below the lunar crust by low or very low-Ti ultramafic liquids similar in composition to the corresponding picritic glasses.
Shearer et al. (Hubbard and Minear 1975; Wagner and Grove 1997). In these models the cumulate pile is gravitationally stable and only the Ti-poor mare basalts are derived by melting of the cumulate pile at great depths. One difficulty with the assimilation model is that to make high-Ti mare basalts from their low-Ti counterparts, the proportions of ilmenite and clinopyroxene dissolved must be in a 3:1 ratio by weight. In contrast, the abundance of ilmenite relative to clinopyroxene in the late stage cumulates is roughly 1:5 by weight (Snyder et al. 1992; Hess and Parmentier 1995). Wagner and Grove (1997) suggested that faster dissolution rates of ilmenite relative to high-Ca clinopyroxene would produce the needed assimilant. Their experiments appeared to confirm that ilmenite dissolution rates were about 3× faster than diopside; however, subsequent experiments by Van Orman and Grove (2000) found that the dissolution rate of diopside in synthetic analogues in VLT mare basalts (not in alkali basalts as in Wagner and Grove 1997) were actually faster than the dissolution rates of ilmenite at comparable temperatures. They concluded that non-equilibrium and selective dissolution of the ilmenite and high-Ca clinopyroxene layer by VLT picritic magmas could not produce high-Ti mare basalts.

The strong version of the assimilation hypothesis seems to be discredited; nevertheless it is still worthwhile to discuss the consequences of assimilation of shallow cumulates since the layer may later be disrupted during cumulate overturn and inhomogeneous deposits of this layer might be scattered throughout the cumulate pile. These deposits might be selectively assimilated by migrating VLT picritic melts to produce low- and even high-Ti mare basalts and picritic glasses.

The constraints provided by Wagner and Grove (1997) are still relevant to this discussion; ilmenite and high-Ca pyroxene must enter the VLT picritic magmas in roughly a 3:1 ratio to produce high-Ti mare basalts. This assimilant contains about 42 wt% TiO₂. In order to produce a high-Ti mare basalt with 13 wt% TiO₂, about 30% of the assimilant must be added to the VLT basalt parent minus whatever olivine crystallized from the assimilating magma to provide some of the heat of dissolution. The geochemical consequences of this process should impart a distinctive geochemical signature on the high-Ti mare magma.

Some of the geochemical criteria for mare basalt petrogenesis were discussed in Hess (2000), but are repeated here for completeness. Additional observations relevant to the origin of mare basalts included:

1) Shallow-mantle ilmenite of the ilmenite cumulate layer should contain significant concentrations of incompatible elements such as Hf, Zr, Nb, and Ta because these elements are only slightly incompatible to compatible in Ti-rich phases (McCallum and Charette 1977; Jones 1995; Horng and Hess 2000). The assimilant should have Nb/Ce≈10⁴ and Nb/Zr≈50 as well as high concentrations of these elements (Shearer et al. 1996; Shearer and Papike 1999). The ratios of these elements in high-Ti mare basalts should far exceed those in VLT picritic glasses. This effect is not observed. High-Ti glasses have only slightly higher Nb/Ce and Nb/Zr ratios than low-Ti magmas (Shearer et al. 1996). Moreover, Zr/Ce ratios in VLT picritic glasses and in high-Ti picritic glasses are roughly the same, contrary to expectation.

2) In the same vein, variations in the initial ε_Hf and ε_Nd show that the high- and low-Ti mare basalts formed separate, long-lived mantle source regions (Beard et al. 1998). The different ε_Hf and ε_Nd trends require that the Lu/Hf and Sm/Nd compositions of the source regions be distinctly different, inconsistent with the hypothesis that these magmas were produced by assimilation of various amounts of ilmenite-rich cumulates.

3) All mare basalts have positive initial ε_Nd and are derived from variably depleted mantle sources. The initial ε_Nd vs. age systematics are not consistent with the assimilation of incompatible-rich, KREEP-like masses (Snyder et al. 2000). Moreover, the general increase in initial ε_Nd with Ti (A15 green glass → A12 olivine basalts, A12 pigeonite basalts → A12 ilmenite basalts, A17 high-Ti basalts) further rules out the assimilation of KREEP to produce
these magmas and indicates an association between the ilmenite-clinopyroxene cumulates and extreme long-lived Sm/Nd source depletion. Furthermore, although the ilmenite-clinopyroxene cumulates and KREEP were produced after extensive LMO crystallization, they generally remained isolated from one another.

4) If assimilation of ilmenite cumulates by VLT picrite magmas was a common process, why was the KREEP layer spared? Conversely, why did the parent magmas to the pre-mare Mg-suite not assimilate ilmenite-rich cumulates? The lack of assimilation of ilmenite cumulates is explained by an overturn model that brings the dunite cumulates to the upper mantle and the shallow ilmenite-rich cumulates to the deeper mantle (Hess and Parmentier 1995).

5) Chromium is strongly partitioned into ilmenite under lunar conditions; the partition coefficient at 1 bar and 1180-1200 °C is >6 and the ilmenite Cr$_2$O$_3$ contents exceed 2 wt% in high-Ti mare basalts (Delano 1980). The Cr$_2$O$_3$ contents of VLT picritic glasses are in the range 0.39-0.64 wt% whereas the range for high-Ti picritic glasses is generally between 0.54-0.82 but with three values 0.93-1.77 for melts with 16.0 wt% or more TiO$_2$ (Shearer and Papike 1992). Only the highest Cr$_2$O$_3$ values are consistent with the assimilation of ilmenite into VLT picritic glasses.

These geochemical arguments support Van Orman and Grove’s (2000) rejection of the assimilation hypothesis for the origin of high-Ti mare basalts; VLT mare basalts are not the parent magmas of high-Ti mare basalts. This conclusion, however, does not exclude the role of minor assimilation or hybridization at depths, since some of these petrogenetic processes probably played roles of varying significance in the generation of mare volcanics. Elkins-Tanton and Grove (1999) argued that generation of low-Ti green glass occurred by melting of “reshuffled” cumulates, followed by fractionation and assimilation of small amounts of high-Ti cumulates. Assimilation in this model would occur at depth and reflects the movement of melts through a heterogeneous cumulate mantle.

A second set of models for the evolution of the mare basalt source, particularly the source for high-Ti mare basalts, involve the overturn of the gravitationally unstable cumulate pile of the LMO (Ringwood and Kesson 1976; Herbert 1980; Hess 1991; Ryder 1991; Shearer et al. 1991; Spera 1992; Hess and Parmentier 1995). The overturn creates a hybrid mantle composed of mixtures of various amounts of high-Ca clinopyroxene-ilmenite cumulates and olivine + orthopyroxene cumulates. Heat-producing elements from the intercumulus liquid carried down with the shallow cumulates, moreover, provide the energy to subsequently partially melt the hybridized cumulates and produce the wide range of mare basalt magmas observed on the lunar surface. These models have several attractive features that account for key apsoects of mare-basalt petrogenesis.

The emerging view of mare volcanism is that it was long lived. Mare volcanism, peaked in the 3.6-3.8 Ga period in several Nectarian and pre-Nectarian basins and then went into a smooth decline, interrupted by short periods of vigorous activity, until volcanism devolved into more or less sporadic activity in periods younger than 2.6 Ga (Hiesinger et al. 2000, 2003). Rare outbursts of volcanic activity apparently continued to 1.1 Ga (Hiesinger et al. 2001, 2003). If we now include the earliest, pre-basin mare volcanics as recorded in the cryptomare and the high-Al basalts from Apollo 14, then mare volcanism was active for more than 3.2 b.y. or 70% of lunar history. Aluminous mare basalts show that mare volcanism initiated at least 300 m.y. earlier and the lunar mantle produced basaltic magmas that were emplaced in the lunar crust as early as 4.45 Ga. These observations suggest that the energy source for basaltic magmatism was hardly transitory (i.e., impact driven) or geographically limited.

Given the longevity of mare volcanism, it is significant that the sum of mare volcanism represents only 1% of the volume of the lunar crust and only 0.1% of the volume of the Moon (Head and Wilson 1992). It is also noteworthy that the distribution of mare volcanism is largely
concentrated in the near side lunar basins. South Pole-Aitken basin, the largest and likely the most ancient lunar basin, contains only a small volume of mare eruptions. This observation means that the existence of major basins is not sufficient by itself to induce eruption of mare volcanism (Hess and Parmentier 1999). The asymmetry of mare volcanism must therefore reflect the deep-seated distribution not only of the heat-producing elements but also of TiO$_2$-rich source materials. The varied thickness of lunar crust is not the sole or even principal determinant for controlling mare volcanism.

The features described above are consistent with the overturn of the cumulate stratigraphy, where dense ilmenite cumulates and incompatible element-rich intercumulus materials sink into the lunar interior (Hess and Parmentier 1995). This model has several important implications. (1) It sequesters heat producing elements into the lunar interior providing for a long-lived energy source. Depending on their concentrations, an energy source with $3\times$ (or perhaps less) lunar abundances of the heat producing elements would not only persist throughout most, if not all, of lunar history but might actually generate and maintain a liquid core. (2) The mare source region, particularly for the high-Ti mare volcanics, is provided by the mixing of these late-stage cumulates with a refractory olivine-orthopyroxene cumulate. The volume of mare volcanism is therefore limited by the fertility of this hybrid cumulate. (3) The cumulate stratigraphy produced after the “overturn” is gravitationally stable and produces a heterogeneous and relatively stable mantle (Parmentier and Hess 1998). Large and localized heating of the lower mantle must be invoked to induce the compositionally dense lower mantle to rise diapirically (thermal buoyancy) and generate mare volcanism by pressure-release melting (Hess and Parmentier 1995). This stable stratigraphy is consistent with the isotopic studies, including the most recent Hf-W results (Lee and Halliday 1997). The cumulate stratigraphy was established within the first 50 Ma or so of lunar history and has remained relatively unaltered throughout the magma-generating stage. This stable stratigraphy is a formidable barrier to magma production and offers one explanation for the paucity of mare volcanism.

Conductive cooling through the lunar crust and heating from below from a convecting mantle controls the rate of thickening of the elastic lithosphere (Parmentier and Hess 1996, 1997). In all models, the lithosphere is 100-150 km thick at the time of formation of the near-side lunar basins and subsequent filling of these basins with mare basalts. Since the lithosphere extends well below the lunar crust, plumes undergoing pressure-release melting are trapped at a rheological boundary well below the lunar crust (Hess 2000). This constraint on mare volcanism has several important consequences. (1) Large dikes must exist to carry mare volcanics to the lunar surface from depths of 100-200 km or more. (2) The low-density anorthositic crust probably is not a critical density filter for such magmatism. (3) The rheological boundary at the bottom of the lithosphere provides another physical barrier for mare eruptions. (4) The existence of lunar basins, even large ones, cannot strongly influence the eruption of mare basalts from such deep reservoirs.

The gravitational-overturn model rationalizes many of these geochemical and physical features of mare petrogenesis. The model, however, is not without problems. Van Orman and Grove (2000) argued that the shallow cumulates may have been physically too strong to allow the overturn to occur in a timely manner. Using a viscosity estimated from the dry flow laws of olivine and a temperature corresponding to the solidus of the LMO, diapirs 40 km in diameter are calculated to sink only 20 km in 400 million years. Only diapirs larger than 180 km in diameter would descend several hundreds of kilometers in the mantle.

Alternative approaches are needed to overcome these difficulties because there is little doubt that shallow ilmenite cumulates could not be partially melted to produce high-Ti magmas. Ti-rich materials must have sunk deep into the lunar mantle to produce suitably hybridized sources. Van Orman and Grove (2000) suggested that mixing of olivine-rich cumulates with the Ti-rich cumulates could lower the viscosity sufficiently to promote gravitational instabilities.
Parmentier and Hess (1999) showed that small ilmenite-rich diapirs would sink small distances into the subjacent cumulates but then the mixed layer would itself become gravitationally unstable. The diapirs resulting from these mixed sources would be large enough to sink deep into the Moon. Elkins-Tanton et al. (2003) suggested that evolved liquids of the LMO assimilated 10-20% of magnesian olivine from overturned lower cumulates thereby increasing the liquidus by more than 100 °C and bringing Ti-rich spinel to the liquidus. It is possible that such cumulates, being hotter and less viscous, could sink. Elkins-Tanton et al. (2003) also investigated the process of remelting the ilmenite-clinopyroxene-cumulate, which produces a high-Ti, negatively buoyant liquid that would percolate into and hybridize at least the upper mantle. However, these liquids would crystallize as they descend adiabatically. Moreover, the energy source for remelting these cumulates is not clear. If it is assumed that energy from radioactive decay in intercumulus liquids was sufficient to remelt the cumulates at 3.9 Ga, then it is hard to understand how the liquid would have crystallized to form the cumulates in the first place (Hess and Parmentier 2001). Impact-induced melting is also unlikely because it must have been a global event so as to produce widely separated (Procellarum vs. Tranquillitatis) high-Ti mare sources.

Arkani-Hamed and Pentecost (2001) argued that the SPA impact generated vigorous mantle circulation that stripped away the KREEP layer from beneath the lunar far-side crust. This circulation would have distributed radioactive elements throughout the lunar mantle, and generally cooled the mantle beneath the SPA basin, perhaps explaining the lack of substantial volcanism there. The SPA impact might have softened the upper mantle sufficiently to destabilize the ilmenite-rich layer leading to the needed gravitational instability. The gravitational instability would not be Moon-wide and would be limited to the far-side hemisphere, a result which could explain the asymmetry of basaltic volcanism and heat-producing elements on the Moon.

Certainly, a number of alternative models exist that are designed to carry products from shallow-mantle late-stage cumulates to the lunar interior. The geochemical characteristics and phase equilibria of low- and high-Ti mare magmas seem to demand such transfers. Now, there are additional geophysical data to support the convective overturn hypothesis. New information regarding the densities of the lunar interior was recently obtained from the inversion of lunar free oscillations (Khan and Mosegaard 2001). In their density model of the lunar mantle, the density increases from about 3.1 ± 0.2 gm/cm³ near the lunar crust to values of 3.7 gm/cm³ below 950 km. The increase of density with depth is consistent with the profile of a lunar cumulate mantle established after a gravitationally induced convective overturn. Moreover, the density in the central region of the Moon assumes a value of 4.7 gm/cm³, which is not consistent with a large Fe core. These densities are more appropriate to a sulfide-rich core and/or a core formed from Fe-rich high-Ca clinopyroxene-ilmenite cumulates.

5.13. Synthesis and questions

Geochemical features that best constrain petrogenetic models for mare magmatism are (1) strong, negative Eu anomalies, (2) positive initial \( \varepsilon_{\text{Nd}} \) values, (3) very large range of TiO$_2$ contents, particularly the extreme values for high-Ti picritic glasses, (4) very low CaO and Al$_2$O$_3$ contents of picritic magmas relative to values observed in terrestrial basalts, and (5) positive initial \( \varepsilon_{\text{Sr}} \) and \( \varepsilon_{\text{O}}^\text{initial} \), which place the evolution of most mare source regions in the earliest 100 m.y. of lunar history. Less constraining, yet important, are phase equilibria data that require the average mantle source to be undersaturated with respect to ilmenite, high-Ca pyroxene and plagioclase. The heterogeneity of the mare sources is evident in the range of bulk compositions, the varied contents not only of incompatible trace elements, but also in the complex isotopic signatures that characterize mare basalts. Compatible elements, however, appear to place the mantle sources to a limited stratigraphic section of an olivine-orthopyroxene cumulate pile; cumulates with very high or very low Mg’ values appear to be excluded from the mare-source mantle.
The mare source regions are best modeled as orthopyroxene-rich, olivine-orthopyroxene cumulates crystallized from an evolved plagioclase saturated LMO. Ryder (1991) pointed out that the mare basalt sources cannot be simply cumulates from plagioclase saturated LMO because these cumulates are more magnesian than the FANs and they contain too much Ni and clinopyroxene. This may be a product of cumulate mixing (Ringwood and Kesson 1976; Ryder 1991; Shearer et al. 1991; Shearer and Papike 1993). In addition, these cumulates have been fertilized to varying degrees by TiO$_2$-rich and incompatible element-rich phases. The incompatible element-rich components are constrained to less than about 1% of the cumulate. The characteristic trace-element composition of mare basaltic magmas is, therefore, achieved by small degrees of melting of a hybridized source, but the extent of melting must be sufficient to deplete the source of all minor phases leaving only olivine and orthopyroxene in the residue. This requirement suggests that the TiO$_2$-rich components must also be added to the source in relatively small amounts. The role of garnet in the source appears to be required by the Lu/Hf data (Beard et al. 1997) and trace-element ratios, such as Sm/Yb and Zr/Y exceeding those in KREEP basalts (Neal 2001). The imprint of garnet implies that melting was polybaric (Longhi 1992) and that garnet-bearing lithologies were distributed throughout the cumulate mantle, albeit in relatively minor amounts. Garnet-pyroxenite lenses may originate from plagioclase-bearing mafic cumulates that were carried deep into the mantle through convective overturn. Partial melts from these lenses would then metasomatize the mantle and create not only the garnet signatures, but possibly the requisite TiO$_2$ enrichments for the low-Ti and high-Ti mare magmas. If the LMO only encompassed the outer 500 km of the Moon, garnet may not be stable in the cumulate pile, only in the unmelted portion of the Moon (Neal 2001, and references therein). More data are needed to model the implications for mare-basalt magma petrogenesis in the case of the latter scenario.

This discussion briefly encapsulates the broad framework of petrogenesis advocated by most lunar researchers. A number of major questions, however, remain outstanding. First, what are the relationships between crystalline mare basalts and the picritic glasses? The geochemical data reviewed here suggest that the mantle sources of the basalts and the picritic glasses are not the same and that their petrogeneses are not directly related. Alternatively, these differences between picritic glasses and the crystalline basalts may reflect either sampling or differences in eruptive processes. Other important questions relate to the lunar asymmetry in mare volcanism, the mechanisms by which mare basalts erupt from deep within the Moon, the nature of melting and melt segregation, the role of assimilation and possible zone refining processes during magma transport, the evolution of the thermal lithosphere and the tectonic processes by which shallow-level cumulates were distributed throughout the mantle. The answers to these questions will have profound implications for models that deal with the size and evolution of the LMO and the composition and dynamics of the lunar mantle.

6. PHYSICAL PROCESSES GOVERNING THE THERMAL, CHEMICAL, AND MAGMATIC EVOLUTION SUBSEQUENT TO INITIAL PLANETARY DIFFERENTIATION

6.1. Introduction

The focus of this section is to incorporate new observations to better define the physical processes that influence lunar evolution subsequent to initial lunar differentiation (discussed in Section 4.3). In many cases, the latter evolution of the Moon must be placed within the context of accretion and primary lunar differentiation (i.e., LMO and core formation) because these primary processes have an intrinsic role in controlling later events. Fluid dynamics govern many of these processes, but in a wide variety of regimes and scales (length scales, time scales, viscosities, etc.). For example, the physics of solid-state convection is different than that of
magma migration, thus these processes are discussed separately. This section is divided into five parts: (1) Basic physics used in modeling lunar thermal and magmatic history in order to facilitate cross-disciplinary communication and so that new workers can be made aware of the assumptions and uncertainties involved. (2) Observational constraints that must be met by lunar thermal evolution models. (3) A synthesis of the thermal history models that have been calculated for the Moon. (4) Mechanisms involved in the migration of mare basalts. (5) Effects of very large impacts on the thermal and chemical state of the lunar mantle.

6.2. Physical processes governing thermal and magmatic evolution

By definition, thermal history models track temperature as a function of space and time within a planetary body by considering the various sources of heat, methods of heat loss (e.g., thermal conduction, convection, magmatism) and conservation of energy. A detailed tutorial on the physics of this process is given in Chapter 9 of BVSP (1981). Here, developments over the past 20 years are emphasized.

6.2.1. What heat to lose? The Moon is a small body, but it still has significant amounts of primordial and radioactive heat. Current models of lunar formation have the Moon accreting within 1-10 years of a giant impact between the proto-Earth and a Mars-size impactor (Section 4.2). This rapid accretion following an energetic giant impact leads to a Moon that begins hot. However, as outlined in Section 4.2, there are uncertainties about the timescale between the giant impact and lunar formation, such that lunar material might have significantly cooled before accreting. No matter how cold the proto-lunar material started, the process of accretion itself caused heating within the proto-Moon. The nature of the accreting material (whether small or large impacts dominated) constrains how much heat is stored deep within the lunar interior and how much remains near the surface (Melosh 1990). Given uncertainties about the initial temperatures of proto-lunatesimals, their size distribution, and the timescale for accretion, it is difficult to constrain the accretional heat within the Moon. One generalization that can be made is that impact energy increased as the planet grew in size and mass (e.g., Kaula 1979; Ransford and Kaula 1980) such that if accretion took less than 100 years, melting of the outer several 100 km of the Moon as a magma ocean would have been an unavoidable result (Pritchard and Stevenson 2000).

The late heavy bombardment had important thermal consequences for the outer layers of the Moon and may account for some of the properties of mare basalts, as well as later eruptions to produce the volcanic glasses (discussed in Section 6.6). Large impacts produce melt through three stages: surface shock melting, instantaneous decompression melting, and pressure-release melting in response to later mantle convection under the site of impact (Section 6.6). Though movement of mantle materials in the Moon is very slow, convection and a hot selenotherm continued under the large craters for 300 to 500 Ma, a scenario that could explain the longevity of mare basalt eruptions (O’Keefe and Ahrens 1993, 1999; Cintala and Grieve 1998).

Following accretion and large impacts, the long lived radioactive isotopes (235U, 238U, 232Th, and 40K) are the most important heat source. Models normally only cite a present-day bulk U content (estimates vary between 20-46 ppb; Hood 1986) and determine concentrations of the other elements using the ratios K/U = 2000, and Th/U = 3.6 (e.g., Toksöz and Solomon 1973; Korotev 1998). The vertical and lateral distribution of these elements have important implications for thermal history. For example, the radioactive elements are concentrated strongly in the crust, and within the crust, they are further concentrated in the Procellarum KREEP Terrane (Haskin 1998; Korotev 2000; Lawrence et al. 1998, 2000; Jolliff et al. 2000). This concentration would increases the amount of melting in this region (Wieczorek and Phillips 2000; Hess and Parmentier 2001). How radioactive elements were concentrated in this region is discussed in Section 7.
6.3. Observational constraints

The chemical and isotopic composition of the Apollo and Luna igneous samples provide critical constraints for thermal models. The samples provide four different types of constraints on thermal models listed below.

6.3.1. Age of mare basalts. The majority of sampled mare basalts were extruded between about 3.9-3.1 Ga, although some volcanism clearly predates the mare, and might have continued to 1 Ga or later (Schultz and Spudis 1983). A fundamental question is whether the time span of mare extrusion was mainly governed by the duration of partial melting in the lunar interior or the ability of melts to reach the surface. For example, some workers have considered whether the apparent delay between lunar formation at 4.5 Ga and the onset of mare formation at 3.9 Ga provides important information on mantle heat sources (e.g., Hess and Parmentier 1995). There may also have been extensive volcanism between 4.5 and 3.9 Ga, but the late, heavy bombardment or younger lava flows obscured or destroyed the record. Others have suggested that the end of extensive mare volcanism at 3.1 Ga might be linked to global thermal stresses from lunar cooling and contraction (e.g., Solomon and Head 1980). Other explanations for the termination of mare volcanism are possible, such as an increase in depth of melting with time so that the path length of melt to the surface became prohibitive, or diminishment of heat sources with time so that less partial melting occurred. Owing to uncertainties in the duration and volume of mare magmatism, it is best to consider the mare ages a weak constraint on thermal models. Models should have melting and basalt production peaking between 3.9 and 3.1 Ga, and models that generate melt before 3.9 Ga or after 3.1 Ga should not be rejected. Most models that have a magma-ocean initial thermal condition in the outer layer of the Moon have no problem generating melt between 3.9 and 3.1 Ga (Cassen et al. 1979).

6.3.2. Temporal variations observed in mare basalts. Most lunar thermal models involve cooling from above and all predict that the melting zone should deepen with time, except under special circumstances (e.g., Zhong et al. 2000). If the composition and/or source depth of the mare basalts changed with time, that would constrain the location of melting in thermal models. There are, however, few observational constraints on the relationship between time and depth of mantle source. For example, there is no firm evidence that the source depth determined from multiple-saturation points increased as a function of time. Early models sought to relate age of mare basalts, their Ti content, and depth of melting; however, remote sensing studies indicate that some young mare basalts also have high Ti. The oldest examples of lunar magmatism (Mg-suite, high-Al basalts) are relatively low in Ti content. Recent studies of the ages of lunar mare meteorites, which are all low-Ti or VLT, however, have crystallization ages at both the high end [3.8-3.9 Ga (Asuka 881757, Misawa et al. 1993; Yamato 793169, Torigoye-Kita et al. 1993)] and low end [2.8-3.0 Ga (NWA 032, NWA 773, Fernandes et al. 2003; Borg et al. 2004)] of the main range of mare-basalt sample ages. These observations indicate significant mantle heterogeneity and confound any simple relationship between Ti content, age, and depth of melting.

6.3.3. Isotopic systematics observed in mare basalts. Several isotopic systems have been used to determine the lunar differentiation interval and can constrain processes related to LMO crystallization and subsequent lunar differentiation. For example, the U-Pb isotopic system for the mare basalt source appears to have closed between 4.42 Ga and mare basalt extrusion hundreds of My later (e.g., Tera and Wasserburg 1974, 1975, 1976; Papanastassiou et al. 1977; BVSP 1981) and has important consequences for thermal models, although others have questioned whether such isotopic closure occurred (Unruh and Tatsumoto 1977). Brett (1977) used U-Pb isotopic closure and arguments about atomic diffusion to constrain temperatures within the mare source region and initial mantle temperatures. Conversely, Pritchard and Stevenson (2004) argued that because diffusion is inefficient, isotopic closure instead places limits on the amount of melting in the mare source between 4.42 and mare-basalt extrusion.
and says little about the initial thermal state of the Moon. However, the ability to use isotopic closure to constrain the extent of melting is limited by inadequate knowledge of the partition coefficients for U and Pb under appropriate lunar conditions and by the limited data used to define isotopic closure. Further work is needed on the U-Pb system and applying results from other isotopes like Sm-Nd (e.g., Nyquist et al. 1995) to quantitative thermal-history models.

6.3.4. Mare basalts compositional heterogeneity and lunar asymmetry. The mare basalts (crystalline mare basalts, pyroclastic glasses) that erupted between 3.8 and 3.1 Ga exhibit both compositional and isotopic variation suggesting that the near-side lunar mantle is laterally and vertically heterogeneous in both temperature and composition (e.g., BVSP 1981; Nyquist and Shih 1992). Remotely sensed compositional data, such as Ti content inferred from Clementine spectral reflectance data, also indicate significant lateral variability in basalts (Gillis et al. 2003). Although it is difficult to constrain the size of the mantle heterogeneities (Spera 1992), overturn models proposed by Parmentier and co-workers have predicted the potential size range of high-Ti cumulate pods (see Section 3.3). The heterogeneous lunar mantle indicates that it was not well mixed and that mantle convection was not vigorous enough to completely homogenize the cumulate remnants of the LMO (Turcotte and Kellogg 1986).

The majority of lunar thermal-history models have been one dimensional, representing a global average of temperature and melting as a function of time. However, new global datasets confirm older results that the distribution of mare basalts, KREEP and crustal thickness are not globally uniform. Mare basalts cover about 30% of the near side but only about 1% of the far side, KREEP is concentrated in the Procellarum region, and the crust is generally thicker on the far side but spatially variable on the near side (see Chapter 3). Any thermal history should explain the distribution of mare basalts and KREEP, but the crustal dichotomy is not understood. In order to explain lunar asymmetries, two- and three-dimensional models (e.g., Wieczorek and Phillips 2000; Zhong et al. 2000; Spohn et al. 2001) incorporating lateral variations in the distribution of radioactive elements and melting are needed. For example, the concentration of KREEP in the Procellarum region could have enhanced melting in that region, and the near side in general, partially explaining the hemispheric dichotomy of mare basalts (e.g., Wieczorek and Phillips 2000). However, extremely large amounts of melt would form an impenetrable barrier to the eruption of basalts (Hess and Parmentier 2001). Hess and Parmentier (2001) also suggested that the associated high crustal temperatures might pose problems with the long-term support of the mascons. Zhong and Zuber (2000) concluded that on a planetary body the size of the Moon the existence of mascons provided no evidence for the thickness of the lunar lithosphere. More generally, two- and three-dimensional models of thermal evolution are required to understand the importance of large impact basins and their spatial distribution on thermal evolution. Recent work on the global effects of impacts on the thermal evolution on Mars (Reese et al. 2000) and regional effects upon the Moon (Arkani-Hamed and Pentecost 2001; Elkins-Tanton et al. 2003) explored the possible effects of impacts on planetary asymmetries.

6.3.5. Lunar magma ocean. Thermal histories require an initial condition. Most models begin immediately following crystallization of the LMO and consider only the thermal effects of that event—meaning that temperatures in the region of the magma ocean start at the local solidus. There are weak bounds on the thickness of the LMO with the range of 300-1000 km favored (Warren 1985; Longhi 1992). Besides thermal effects, the LMO has important effects on chemical layering. Among the possible effects are: stable stratification that would inhibit mantle convection, or unstable stratification that promotes mantle overturn with concomitant movement of radioactive elements to deeper layers and pressure-release melting (Hess and Parmentier 1995; Alley and Parmentier 1998). More detailed models of LMO crystallization are needed to better understand the thermochemical parameters used in thermal models.

6.3.6. Surface geology. The lack of a global system of thrust faults on the lunar surface has been used to constrain the amount of volume change within the Moon and the initial temperature
profile (e.g., Solomon and Chaiken 1976). As mentioned in Section 4.2, it is more difficult to relate surface strain to internal processes than previously thought so the lack of thrust faults does not strongly constrain lunar volume change. Another way that surface faults have been related to volume change is through the formation of rilles (interpreted to be grabens) around the periphery of mascons. In Chapter 3, a discussion is given of rilles formed through flexural stresses related to the mass concentration in the mascon. Solomon and Head (1979, 1980) also attributed the formation of the rilles to the global stresses related to internal thermal evolution. In particular they noted that the apparent global cessation in rille formation at about 3.6 Ga (Lucchitta and Watkins 1978) required the global stress pattern to change from extension to contraction. This constraint on the timing of the change in global stress reduces the number of viable thermal models. However, given the uncertainties in predicting faulting around mascons (Chapter 3) and in dating the time of formation of rilles, it is not clear whether the timing of this change in global stress can be so precisely dated, is necessary to explain the global cessation of rilles, or if such a global cessation really occurred (since rilles are not globally distributed).

6.3.7. Current temperature profile. Hood (1986) reviewed the geophysical evidence for the current thermal state of the Moon from heat flow, seismic data and electromagnetic sounding. The evidence indicates that the outer layers of the Moon are subsolidus, but that temperatures below about 1000 km are likely warmer and possibly at the melting point (see Chapter 3). The fact that brittle failure in the form of moonquakes occurs to about 1000 km means that the lunar mantle is subsolidus to at least that point. However, because we cannot directly measure temperature within the Moon, many assumptions must be made to relate measurements to interior temperatures (BSVP 1981; Hood 1986). For example, the two heat flow measurements have been used to infer the bulk U content of the Moon (very important for thermal models), but depending on the assumptions made, inferred values range from 20-46 ppb (Langseth et al. 1976; Hood 1986; Warren and Rasmussen 1987).

6.3.8. Gravity field. Two aspects of the lunar gravity field bear upon temperatures within the Moon. (1) Mascons (discussed in Chapter 3) are mass concentrations associated with large impact craters that are clearly not in isostatic equilibrium. This means that the strength of the lithosphere has been able to support the excess mass of these structures since they were formed at ~4 Ga. However, the ability of a planet to support long-wavelength loads increases as planetary radius decreases. In other words, because of the small size of the Moon, the persistence of mascons does not strongly constrain the thickness of the lithosphere or temperatures in the upper mantle (Zhong and Zuber 2000). (2) Many workers have claimed that the low-order terms of the spherical harmonic expansion of lunar topography and the gravity field are anomalous and require an explanation separate from the other terms (Hood and Zuber 2000; Stevenson 2001 and references therein). The most common special explanation is that the low-order gravity and topography terms are a relic of faster rotation in the past, typically when the radius of the Moon’s orbit was 13-16 Earth radii. The fast evolution of the Moon away from the Earth early in its history (e.g., Goldreich 1966), requires that the Moon would have to cool quickly in order for the rotational bulge from this early period to be preserved. This rapid cooling would have important implications for the timing of cooling/heating of the LMO and the initial thermal state of the Moon following differentiation (Hood and Zuber 2000). However, other explanations for the low order gravity and topography exist (Stevenson 2001), and so the thermal constraints are not unique.

6.3.9. Magnetic field. As outlined in Chapter 3, several lines of evidence support the possibility of a small, Fe-rich, conducting, and liquid lunar core. Maintaining a liquid core for the duration of lunar history provides a weak constraint on interior temperatures since the eutectic temperature of an Fe-S core under lunar conditions is only 1250 K. If a lunar dynamo existed early in lunar history, as suggested by some paleomagnetic studies (see Chapter 3), it constrains temperatures within the lunar core during that time interval. There are theoretical
difficulties with a convecting lunar dynamo generating the field strengths recorded in the samples (Chapter 3), and so perhaps some other mechanism for stirring the lunar core is needed (e.g., tidal effects; Williams et al. 2001). If true, constraining the role of tidal effects would be important for understanding early lunar thermal/orbital evolution. The small size of the lunar core results in the process of core formation not having large thermal consequences for the mantle, and might even be able to coexist with a cold deep mantle (Stevenson 1980).

6.3.10. Mantle layering. The four seismometers deployed by Apollo were used to study mantle structure on the near side, and indicated a possible discontinuity at 500 km (see Chapter 3). Khan and Mosegaard (2001) used lunar free oscillations to determine that the lower mantle is denser than the upper mantle, meaning that whole-mantle convection would be inhibited by compositional stratification. If such a layer is truly global in nature it suggests that mantle convection (if it occurred) could not penetrate the 500 km discontinuity and was therefore layered, which decreases its vigor. However, because of the limited distribution of lunar seismometers and the difficulties in interpreting lunar seismograms, the sharpness, importance, and extent of the 500 km discontinuity is questionable.

6.3.11. Summary. Thermal models must agree with the available observational constraints, although in the discussion above we have highlighted the vagueness and non-uniqueness involved in each of these. Most observational constraints should not be used to rigidly divide which models are acceptable and unacceptable, although some requirements are more secure—there was an LMO and a high proportion of post-LMO mantle melting occurred between ~4.45 and 3.1 Ga. In fact, as noted above, the first requirement practically guarantees the second in most models, and so there really are few unique constraints. However, the situation is not hopeless and with future work, at least two observational constraints and two computational approaches could become more robust: (1) Additional seismometers on the lunar surface, especially on the far side, could determine if the lateral extent of the 500 km discontinuity is real and global. (2) New sample studies of lunar paleomagnetism can help to determine if there is a lunar dynamo and when it operated. (3) A mostly unexplored area is directly linking geochemical sample studies (isotopes, trace and major elements) with quantitative thermal models. A much better understanding of melt migration within the Moon and how samples were contaminated in the journey from source to surface will be required to make this link (4) Coupled thermal and dynamic models of early lunar evolution could provide better understanding of the low-order gravity and topography.

6.4. Synthesis of thermal history models

Over the past 50 years, roughly 50 thermal-history models have been published in peer-reviewed journals (Tables 4.10 and 4.11). Chapter 9 in the BSVP (1981) provides an excellent overview of the history and physics of thermal models as well as a summary of lunar thermal histories calculated before 1980. More than 80% of all thermal histories were calculated before that date. However, the lack of papers published since 1980 should not be interpreted as a lack of questions to be answered or new developments. As outlined below, there have been many developments that show how previous models were incomplete and too simplistic, and there are many fundamental observations about the Moon that are poorly understood (e.g., the asymmetric distribution of mare and KREEP). Instead, most thermal modelers have followed the data—as the number of lunar missions declined, the modelers applied their craft to explain new data from the other terrestrial planets and the galilean satellites. With the recent return to the Moon and anticipation of further new datasets, this is a good place to consider the state of this field. In this section is a summary of the types of models that have been used, simplifications and limitations of each type of model, and then synthesize the important results.

6.4.1. Types of models. Many thermal models prior to the return of Apollo samples were purely conductive and did not include the effects of melting. For example, Urey’s models (see
Shearer et al.  Thermal & Magmatic Evolution of the Moon

Table 4.10. Summary of primary types of thermal models that have been used and the deficiencies with each. See text for more discussion.

<table>
<thead>
<tr>
<th>Model</th>
<th>Deficiencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purely conductive or purely convective</td>
<td>Has to include melting.</td>
</tr>
<tr>
<td>Conductive – test for instability to convection</td>
<td>Old models didn’t include compositional buoyancy which can be as important as thermal buoyancy.</td>
</tr>
<tr>
<td>Conductive or convective with melting but no melt migration</td>
<td>Melt migration, particularly segregation of radioactive elements into melt is important. Parameterization of melting is not well constrained.</td>
</tr>
<tr>
<td>Conductive with melting and melt migration</td>
<td>Parameterizations of melt migration are crude-- New models must include physics of melt migration.</td>
</tr>
<tr>
<td>Conductive/convective including chemical effects</td>
<td>Again, must include physics of melt migration. Must further explore plausibility of initial conditions.</td>
</tr>
</tbody>
</table>

Table 4.11. A summary of the thermal models from the past 50 years. The criteria for putting papers in this list is that they had to be published in a peer-reviewed journal, and that the model had to quantitatively track temperatures in the lunar interior as a function of time.

<table>
<thead>
<tr>
<th>Type of Model</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purely Convective</td>
<td>Turcotte and Oxburgh (1969); Tozer (1972,1974); Schubert et al. (1977,1979,1980), Cassen and Young (1975)</td>
</tr>
<tr>
<td>Conductive, and test for convection</td>
<td>Reynolds et al. (1972); Cassen and Reynolds (1973,1974); Toksöz et al. (1978); Cassen et al. (1979)</td>
</tr>
<tr>
<td>Conductive with melting – No melt migration</td>
<td>Levin (1962,1966); Levin and Majeva (1960); Majeva (1964); Solomon and Chaiken (1976); Solomon (1977); Solomon and Head (1979); Wieczorek and Phillips (2000)</td>
</tr>
<tr>
<td>Convective with melting – No melt migration</td>
<td>Turcotte et al. (1972,1979); Chacko and De Bremaecker (1982); Konrad and Spohn (1997); Spohn et al. (2001)</td>
</tr>
<tr>
<td>Conductive with melting and melt migration</td>
<td>Fricker et al. (1967); Lee (1968); McConnell and Gast (1971); Wood (1972); Solomon and Toksöz (1973); Toksöz and Solomon (1973); Toksöz and Johnston (1974,1977); Strangway and Sharpe (1975); Binder and Lange (1977, 1980); Ormatskaya et al. (1977); Kirk and Stevenson (1989)</td>
</tr>
<tr>
<td>Conductive/convective including chemical effects</td>
<td>Hess and Parmentier (1995); Alley and Parmentier (1998); Zhong et al. (2000)</td>
</tr>
</tbody>
</table>

Table 4.11 for references) favored a cold Moon by relying upon three arguments later found to be flawed. (1) Urey and others assumed that in order to maintain structures in the lunar crust and the non-hydrostatic lunar shape, there could be little to no partial melting during lunar history. Thus, many early workers used models that were purely conductive and did not use latent heat to calculate the extent of partial melting (Table 4.10). To minimize any partial melting, these models required the Moon to start extremely cold, nearly 273 K at all lunar
radii, a result that is inconsistent with the LMO (e.g., Wood 1972). Recent work has shown that only the outermost lunar layers must remain cold to maintain long-wavelength lunar shape, and that interior temperatures are not constrained (Zhong and Zuber 2000). (2) Many early workers assumed a chondritic abundance of radioactive elements, whereas sample analysis has revealed that the lunar samples are depleted in K and enriched in U relative to chondrites (e.g., Toksöz et al. 1972; BVSP 1981). (3) Urey said that partial melting might not be required to form the mare and that they could be explained by impact melting. Sample analysis indicates that the mare were formed by partial melting in the lunar interior.

Even before analysis of lunar samples conclusively demonstrated partial melting in the lunar interior, many conductive models attempted to quantify melting in the lunar interior (e.g., Levin 1962) and the effects of moving melt to the surface (e.g., Fricker et al. 1967; Lee 1968). The physics of melt migration are complicated and there are many uncertainties about how the processes operated in the Moon (Section 6.5). Therefore, models crudely parameterize complex properties by instantaneously moving all of the melt from the source region to the near surface. Some models move melt to the top of the partial melting region (e.g., Wood 1972), while others distribute it exponentially with depth (e.g., Toksöz et al. 1972; Kirk and Stevenson 1989). The most important consequence of the melt migration is to remove the radioactive elements (which are incompatible and preferentially go into the melt) from the melting region to shallower, cooler layers. Different parameterizations of exactly how the radioactive elements segregate into the melt have been presented, but they all have the effect of limiting the total amount of melting. As an example, Figure 4.48 shows the total amount of melting as a function of depth and time within a one dimensional conductive thermal history, both with and without melt migration. The principal results of these calculations is that melt migration limits the total amount of melting that occurs at any given depth, and expands the region of melting at a given time to slightly shallower depths. Thus, melt migration is able

![Figure 4.48](image)

**Figure 4.48.** The total percent melted within the lunar interior as a function of depth and time for a one dimensional conductive model both with and without melt migration. The model uses the method and parameters of Kirk and Stevenson (1989). Although the radioactive elements are not initially distributed uniformly throughout the interior (they are concentrated in the center and in the crust), this does not affect the result that including melt migration reduces the total amount melted at a given depth.
to reduce possibly unrealistic amounts of melting that exist in some thermal models (e.g., Wieczorek and Phillips 2000).

Some thermal modelers have considered the effects of mantle convection within the Moon, first suggested to be important by Runcorn (1962). There is no observational evidence indicating convection has occurred or is occurring within the Moon (Cassen et al. 1978; Phillips and Ivins 1979; BVSP 1981). In fact, there are three observational arguments that indicate that mantle convection has been weak or nonexistent. (1) An analysis of the Moon’s free oscillations by Khan and Mosegaard (2001) found that while the density of the upper ~500 km of the mantle was approximately constant, a sharp increase in density occurred below this depth. While the identification of lunar free oscillations remains controversial (see Chapter 3), the existence of a 500-km seismic discontinuity in the Apollo seismic data is consistent with a change in composition (and possibly density) below this depth (e.g., Hood and Jones 1987; Mueller et al. 1988). If these results are globally representative of the mantle (all seismometers were located in a small region on the lunar near side), then a 500-km density discontinuity would presently act to inhibit whole-mantle convection. While the time of formation and the origin of the 500-km discontinuity are presently being debated (see Section 2.3), if it is similarly ancient, then convection (if it occurred) may have only been confined to the region above or below the 500-km discontinuity. If so, this significantly reduces the length scale and vigor of convection. Such a density stratification within the lunar mantle could have been achieved as a result of the overturn of a gravitationally-unstable magma-ocean cumulate pile (e.g., Herbert 1980; Spera 1992; Hess and Parmentier 1995). (2) The distribution of heat sources in the Moon may also have acted to inhibit thermal convection within the lunar mantle. Using Lunar Prospector Th abundances at the surface (e.g., Lawrence et al. 1998, 2000), Jolliff et al. (2000) estimated the quantity of heat-producing elements in the mantle and crust. They found that about 75% of the Moon’s incompatible elements were located in the crust, and furthermore, that about 40% of the crustal incompatible elements were located within the regions of Oceanus Procellarum and Mare Imbrium. Such a strong crustal enhancement of heat-producing elements could act to heat the mantle from the top down, and hence lessen the thermal buoyancy contrast between the upper and lower mantle. Estimates of the abundance of Th in the mantle are as low as 5-25 ppb based on the composition of the picritic glasses (e.g., Warren and Wasson 1979; Jolliff et al. 2000), and the low end of this range is considerable less than is usually assumed for most thermal-convection models (e.g., 20-29 ppb in the models of Spohn et al. 2001). A low concentration of heat-producing elements in the mantle would be a natural expectation if the mantle is composed of cumulates from a near-global LMO and would further act to limit the vigor of mantle convection. (3) The existence of compositional heterogeneity in the mare-basalt source region (Section 4.5) indicates that the lunar mantle was not well mixed, and that vigorous mantle convection did not occur (Turcotte and Kellogg 1986).

Consistent with the observations cited above, theoretical calculations that consider realistic viscosities also indicate that lunar-mantle convection might have been weak or non-existent. Many workers have tested various thermal models for convective instabilities that form when thermally induced positive buoyancy overcomes viscous resistance and the Rayleigh number overcomes its critical value (see Chapter 3 and Section 6.2 of this chapter). Early models that assumed a constant viscosity within the lunar interior found that the Moon and other terrestrial planets would be convectively unstable (e.g., Schubert et al. 1969). Models that started with a hot exterior (LMO) and cold interior, did not become convectively unstable until the unfavorable initial temperature profile could be reversed by the interior heating up and the exterior cooling (e.g., Cassen and Reynolds 1973). However, viscosity is a very sensitive function of temperature such that viscosity is not constant as a function of depth within the lunar mantle. Including the effects of temperature-dependent viscosity, Cassen and Reynolds (1974) calculated that the lunar mantle is currently unstable to convection. Recent progress has been made in understanding when and how convection occurs on planets with
immobile lithospheres, such as the Moon (see Section 6.2). A primary result is that convection can only occur within a region where the viscosity changes by about an order of magnitude. This severely limits the amount of positive buoyancy that can form convective instabilities, and makes convection more unlikely and less vigorous on planets with stagnant lids than was previously thought. Considering the uncertainties in lunar mantle viscosity (Section 6.2), several thermal evolution scenarios are possible, including no convective instabilities or weak convection (Pritchard and Stevenson 1999).

As the Moon could be in the regime where convection is only marginally viable, compositional effects on buoyancy must be considered as well as thermal effects. Both stable and unstable compositional layering must be considered. During LMO crystallization, dense Fe-rich cumulates are placed over less dense Mg-rich cumulates, forming an unstable arrangement that causes overturn to a more stable layering that inhibits thermal convection (Hess and Parmentier 1995; Alley and Parmentier 1998). As most of the Fe- and Mg-rich cumulates crystallized in the LMO are less dense than the deeper lunar interior, overturn of the cumulates is likely to be limited to above the bottom of the LMO (Hess and Parmentier 1995). Conversely, late-stage cumulates that include ilmenite would be more dense than all underlying material and under certain conditions, these cumulates and other late-stage materials rich in radioactive elements could sink deep into the lunar interior as Rayleigh-Taylor instabilities, perhaps all the way to the center of the Moon (Ringwood and Kesson 1976; Hess and Parmentier 1995; Zhong et al. 2000).

Alternatively, the denser material could sink into the lunar interior via percolative flow instead of as diapirs. Elkins-Tanton et al. (2001) argue that the cumulates crystallize at low temperature (about 1150 °C), and the viscosity under these conditions is very high and augmented by the viscosity law of pyroxene, which is stiffer than olivine. Following numerical modeling with a realistically cooled Moon and a solid high-Ti layer, they find the process of sinking instabilities unlikely due to viscosity constraints. If the ilmenite cumulate layer could not fall as large instabilities into the deep Moon, there is still the need to move high-Ti material deeper into the Moon in a heterogeneous manner, to create the source regions required by the picritic glasses.

Van Orman and Grove (2000) also recognized that the high viscosity of the ilmenite + pyroxene cumulate, could inhibit subsolidus flow, and suggested that mixing olivine-rich cumulate with the high-Ti cumulate could lower its viscosity enough to allow the formation of instabilities. Another possible mechanism for redistributing high-Ti material into underlying cumulates is melting by a large impact, and then sinking high-Ti liquids into underlying cumulates. As high-Ca clinopyroxene and ilmenite melt in a ratio of ~2:1 (Wyatt 1977; Van Orman and Grove 2000), remelted high-Ti cumulates would be negatively buoyant at depths as shallow as 20-80 km, and would sink as liquids, percolating downward through the underlying mantle and beginning to recrystallize ilmenite at 200 km depth, making a hybrid, heterogeneous mantle (Elkins-Tanton et al. 2001).

Subsequent to crystallization of the LMO, another process could form a stable compositional gradient that would inhibit or weaken convection. Nearly all thermal models include a deepening of the melting region as a function of time and lessening of the degree of melting with depth related to cooling from above. As the residual material left over after melting is less dense than the original material, the variable extraction of melt as a function of depth forms a stable compositional gradient that can be large enough to offset thermal buoyancy (Pritchard and Stevenson 1999). However, there is great uncertainty in the amount of melting as a function of depth and time in the mantle, and there are certainly lunar histories (especially if there is heating from below by an ilmenite/radioactive enriched core) when the variable extraction of melt with depth is not important.
6.4.2. Synthesis of results.

1.) Conduction vs. convection. As we emphasized above, there is no conclusive observational or theoretical requirement for lunar-mantle convection, and several reasons to believe that convection might have been weak or nonexistent. The primary difference between conductive and the incorrect constant-viscosity convective thermal histories is that the latter cool too quickly and efficiently, causing a more rapid disappearance of partial melt, thinner lithosphere, and lower mantle temperatures (BSVP 1981). Models that more appropriately account for the effects of variable-viscosity convection (e.g., Spohn et al. 2001) are more similar to conduction models in terms of mantle temperatures, lithosphere thickness, and duration of partial melt.

2.) Initial condition. When a thermal model begins with an accretionary thermal profile (hot exterior and cold interior, see Section 6.3), it is stable to convection. Therefore, it takes hundreds of millions to billions of years to heat up the interior before it can become thermally unstable to convection (e.g., Cassen and Young 1975; Turcotte et al. 1979; Spohn et al. 1999), although because chemical buoyancy is also important, this does not mean convection will occur. As 4.5 Ga is long enough to allow an initially cold interior to heat up, but not long enough for a hot interior to cool down, the present thermal profile does not well record the initial thermal profile (Cassen et al. 1979).

3.) Timing of partial melting. It is not difficult to account for partial melting in the lunar mantle during the primary era of mare volcanism (3.9-3.1 Ga) when the outer layers of the Moon begin at the solidus as expected for conditions following a LMO (Cassen et al. 1979). Most models also generate melt before and after this era. The record before 3.9 Ga is largely obscured so it is difficult to say much about that time period, but an unanswered question is why the majority of the visible mare seemed to have erupted between about 3.9 and 3 Ga, with minor additions occurring up to ~1 Ga. Proposed explanations include an inability of the melt generated to reach the surface because of global compressive stresses or an increase in the depth of melting, or a decrease in the amount of melt generated because the source regions were depleted in heat sources (due to cooling or loss of radioactive elements) or an increase in the solidus (due to prior melting or loss of volatiles; BVSP 1981).

4.) Chemical versus thermal buoyancy. As discussed in Section 7.5 (below), compositional gradients formed during LMO crystallization or from the variable removal of melt as a function of depth have two effects. First, stable compositional gradients can inhibit the development of convection. Second, unstable compositional gradients might overcome thermal buoyancy effects and cause sinking of dense ilmenite cumulates, if certain rheological conditions are met. Therefore, it is important to consider both chemical and thermal buoyancy when evaluating what will happen in the lunar interior.

5.) Melt migration matters. Melt migration in thermal history models is either ignored or parameterized in a simplistic manner, as discussed above. However, the extraction of melt with its preferential removal of volatiles and radioactive elements has important consequences for the possibility and extent of further melting (Section 4.6.5).

6.) Lateral variations. The distribution of heat sources, whether they are radioactive elements or large impacts, affects when and where melting occurs. While most thermal models have been one-dimensional, two- and three-dimensional models of lunar thermal evolution are needed to account for the laterally and vertically variable distributions of heat sources.

6.5. Processes governing migration of lunar basalts

6.5.1. Dynamic models of magma transportation. How the amount of melting is calculated from the conductive or convective temperature profiles was discussed in the previous sections. The processes by which melt could be transported to the surface are discussed in this section. Much progress has been made over the past two decades in modeling the dynamics of magma
transportation; however, significant gaps remain in understanding this process. A full treatment of this subject would include the effects of melt segregating from its matrix, porous flow, formation of dikes, and thermodynamics of melt interacting with the surrounding material. In this section we briefly review two aspects of magma transport that may help to decipher the complex geochemistry of the lunar basalts; porous flow and dike propagation. The reader is referred to the review by Rubin (1995), the papers in the book Magmatic Systems (Ryan 1994a) and the papers by Lister and Kerr (1991) and Sleep (1988) for more detailed accounts.

6.5.2. Porous flow. Porous flow of magma may be an important process deep in the lunar interior where melt would be generated. The migration of deep magma depends on the buoyancy of the melt with respect to the matrix, the permeability of the solid, viscosity of the melt, and deformation and compaction of the matrix (e.g., McKenzie 1984). As compaction should only be important for scales on the order of 100 m within the lunar mantle (e.g., Spera 1992), melt migration can be adequately modeled by Darcy’s law:

\[ v = \frac{k (1 - \phi) \Delta \rho g}{\eta \phi}; \quad k = k_0 \phi^n \]  

where \( v \) and \( v_s \) are the velocity of the melt and solid, \( k \) is the permeability, \( \phi \) is the porosity of the matrix (i.e., melt fraction), \( \Delta \rho \) is the density difference between the melt and matrix, \( g \) is the gravitational acceleration, and \( \eta \) is the melt viscosity. Inserting parameters that might be appropriate for the lunar mantle (\( \phi = 0.05 \), \( g = 1.6 \) m s\(^{-2} \), \( \Delta \rho = 300 \) kg m\(^{-3} \), \( \eta = 1 \) Pa s and \( k_0 = 10^{-8} \) m\(^2 \), \( n = 2.5 \) (e.g., Sparks and Parmentier 1994; Spera 1992)) the vertical velocity of the melt should be on the order of 1 m/year. Given uncertain knowledge of such parameters as the melt fraction and permeability, this estimate could be more than an order of magnitude too small or large. Indeed, the formation of high-porosity channels could significantly increase these estimated velocities (e.g., Dick 1999; Kelemen et al. 1999).

If the mare source was located about 500 km below the surface, it would take on the order of a million years for this melt to reach the surface via porous flow. Because of this long duration it might be expected that the melt would always be in chemical equilibrium with its surrounding matrix. Furthermore, because of the slow ascent velocities, if the temperature of the matrix ever dropped below the melt’s solidus, the magma would freeze, inhibiting further vertical melt migration. Subsequent to initial freezing, melt would begin to accumulate below this low-permeability barrier, possibly leading to the formation of a magma chamber and/or to the initiation of magma transport via dike propagation.

6.5.3. Static dikes. Numerous papers have been written on the subject of determining the shape of a fluid-filled dike in an elastic medium (e.g., Sneddon 1946; Weertman 1971; Pollard and Holzhausen 1979; Pollard 1987; Rubin and Pollard 1987). In these models the shape of the dike is determined by the elastic stresses in the host rock, the hydrostatic pressure in the dike, and the resistance to fracture at the dike tips. Exact solutions exist for disk-shaped as well as 2-D blade-like fractures. These studies show that in order for a blade-like dike to be completely static, with no tendency to propagate, the vertical length of the crack must be less than

\[ l_{\text{max}} = 2 \left( \frac{K_c}{\Delta \rho g} \right)^{2/3} \]  

and have a width less than

\[ w_{\text{max}} = \frac{1 - \nu}{\mu} \left( \frac{K_c^4}{\Delta \rho g} \right)^{1/3} \]  

where \( K_c \) is the fracture toughness, \( \mu \) is the shear modulus, and \( \nu \) is Poisson’s ratio of the rock.
The fracture toughness is a material property that describes the distribution of stress near a dike tip that is necessary for a rock to fracture. Using typical values that might be appropriate for the mantle ($K_c = 1-5$ MPa m$^{1/2}$, $\mu = 60$ GPa, $\nu = 0.25$), the maximum length of a static dike is ~300 m to 1 km, and its maximum width is ~0.1 mm to 1 mm.

If the length of a static dike were to exceed $l_{\text{max}}$, then the elastic stresses would act to open the upper tip of the dike and to close the lower tip. This quasi-static result has prompted many studies to assume that an individual dike will propagate upwards once this marginally unstable height has been achieved (e.g., Head and Wilson 1992; Wilson et al. 2001). However, since this formulation does not take into account any hydrodynamic effects (namely, the pressure gradients that are required for the magma to flow) it is inaccurate to use this model to describe a propagating dike. Furthermore, this static model cannot estimate how fast such a dike would propagate. Two other fundamental problems exist using such a dike propagation model (Lister and Kerr 1991). First, owing to the extreme narrowness of the dikes, they should freeze rapidly as they traverse the lithosphere. Secondly, the bottom portion of the dike will not be able to completely close because an infinite pressure gradient is required to squeeze viscous magma out of a narrowing gap. Thus, a dike is likely to propagate only if magma is continually supplied.

### 6.5.4. Vertical dike propagation.

In modeling the propagation of a fluid-filled dike in an elastic medium, several sources of pressure within the dike must be considered. These include the elastic stresses set up by deforming the surrounding medium, the stresses required to fracture the rock at the dike tips, hydrostatic pressures within the fluid (i.e., buoyancy), the pressure gradients required to drive fluid flow, external tectonic stresses, and internal overpressurization. Lister and Kerr (1991) considered the relative magnitudes of these stresses for a vertically propagating dike and determined that the balance between buoyancy forces and pressure gradients due to flow in a viscous fluid were the most important. Elastic stresses were found to play a moderate role, and the stresses required to fracture the medium were found to be negligible compared to the other sources of stress.

Two models of a vertically-propagating fluid-filled dike were developed by Lister and Kerr (1991) that might be applicable to the transport of magma within the Moon. In one model, magma was continuously supplied from a point source, and the equilibrium shape of the dike was determined. Neglecting the fracture strength of the rock, the breadth, $b$, and central width, $w$, of the dike, were calculated as:

$$b(z) = 5.25 \left( \frac{Q \eta \mu^3}{g^3 \Delta \rho^4 (1 - \nu)^3} \right)^{1/10}$$

$$w(z) = 1.808 \left( \frac{Q \eta^3 (1 - \nu)}{\mu g^2 \Delta \rho^2 z} \right)^{1/10}$$

where $Q$ is the volumetric magma production rate, and $z$ is the height above the point source. This model of magma transport might be valid for determining the width and length of a dike that erupts at the surface from a deep mantle source. Assuming that such a dike initiates 500 km below the surface and has a low magma production rate of 10 m$^3$/s, the length and width of the dike at the lunar surface would be ~50 km and ~2 cm, respectively. For a high magma production rate of 10$^6$ m$^3$/s, this model predicts a dike length of ~170 km and a width of ~0.7 m. As discussed by Lister and Kerr (1991), freezing of magma near the dike margins would tend to decrease the breadth of such a dike and increase its width.

In another model, Lister and Kerr (1991) obtained an analytical solution for a vertically propagating dike being fed by linear source at depth. It was found that the fracture strength of the rock only slightly influenced the shape of the dike near its tip. For a magma production rate $q$ per unit length, the dike width at depth, $w_\infty$, and the speed of dike propagation, $c$, were shown to be given by
The linear magma production rate can be approximated from the first of these models. If a total magma production of $10^6$ m$^3$/s is assumed to operate over a dike length of 170 km, then $q$ is $\sim 6$ m$^2$/s. This gives a dike width at depth of $\sim 0.3$ m and a propagation speed of 10 m/s. If the mare source were located 500 km below the surface, then it would only take about half of an Earth day for this dike to reach the surface. Alternatively, assuming a total magma production of 10 m$^3$/s operating over a length of 50 km, $q$ is $\sim 2 \times 10^{-4}$ m$^2$/s, $w_\infty \sim 1$ cm, and $c$ is $\sim 0.01$ m/s. In this case the dike would be able to traverse 500 km in just over a year. In modeling the transport of primary basaltic magmas (represented by the VLT pyroclastic glasses) from the deep lunar mantle, Spera (1992) calculated the requirements needed to move that magma a distance of 400 km with no or very limited degrees of fractional crystallization. He concluded that such a melt must travel at 10 m/s through a crack of 40 m in diameter.

These qualitative results show that the transportation of magma via dikes is much faster than by porous flow. These short ascent times require that a dike should be able to traverse the cool lithosphere without completely crystallizing. Furthermore, the magma is unlikely to come into chemical equilibrium with the surrounding rock, thus potentially preserving isotopic and compositional information from the mare source. In that some of the above requirements may be unlikely (400 km $\times$ 40 m dike) for the Moon, alternative models have been described to transport these magmas to the lunar surface.

Wilson and Head (2003) used the dike propagation analysis of Rubin (1993) to derive a modified dike transportation model for these near-primary magmas. They proposed a model in which the dike propagates rapidly from the magma source to the surface through gas build-up in a low-pressure micro-environment near the tip the magma-filled crack. The gas-rich region consists of a free gas cavity overlying a basaltic foam extending vertically for $\sim 20$ km. It is certain that the lunar mantle is volatile-depleted. In such a volatile-poor environment, it has been proposed that the driving mechanism for lunar fire-fountaining is the oxidation of C to produce CO (Fogel and Rutherford 1995). The depth at which this reaction occurs (Fogel and Rutherford 1995) appears to be shallower than that proposed by the Wilson and Head (2003) model. In this model, the gas-producing reaction must occur in the deep lunar mantle. Also, rapid transport will aid the preservation of the primary composition of these magmas. However, to what degree volatile exsolution and heat-loss to the adjacent wall rock affect the primary nature of these melts is unknown.

Longhi (1992) and Hess (1991) acknowledged the importance of transport of these magmas through vertical dikes from the base of the lunar lithosphere (for specifics see Section 6.6). However, they suggested that these magmas were transported to the base of the lithosphere from the deep lunar mantle within ascending lunar diapirs. In this model, pools of melt are generated over a range of pressures. Therefore, the multiple saturation pressures calculated for these near-primary magmas are an average and melting was initiated at greater depths ($\sim 1000$ km). In these models, porous flow would have an important role in pooling magmas in a deforming and compacting crystalline matrix in the diapir and might be able to account for the transport of magmas that have a negative buoyancy if the diapir velocity was significantly large (see Section 5.9).

6.6. Processes governing the distribution of basaltic magmas on the Moon

It has long been apparent that the lunar maria are distributed in a highly non-uniform manner. On the Moon’s Earth-facing hemisphere, basaltic flows comprising the maria are commonly located within the confines of large impact basins. The only major exception is the vast expanse of volcanic flows that make up Oceanus Procellarum. In contrast to the near side,
Shearer et al.

Thermal & Magmatic Evolution of the Moon

lava flows on the far side are relatively scarce, even though large impact structures are located there (e.g., South Pole Aitken Basin). One of the fundamental questions of lunar geology is the cause of this near side/far side asymmetry of basaltic eruptions (Section 6.3.4) and of the preference for lavas to erupt within large impact structures.

There are two possible end-member processes that could lead to this distribution of mare basalts on the Moon. The first is that the eruption of basaltic magmas might be controlled by a magma-transport processes that inhibits eruptions on the far side and promotes eruptions within basins. This might include factors such as crustal thickness variations, volatile exsolution, and/or magma buoyancy. Alternatively, it is possible that the current distribution of maria is the result of spatial variations in magma production. In this case the thickness of the maria would be related to the quantity of melt that was produced in the underlying mantle, and the scarcity of far-side lava flows would be a consequence of a low magma production within this hemisphere. Of course, both magma production and magma-transport processes probably play some role in affecting the distribution of basaltic eruptions on the Moon. This section addresses factors that may influence whether a basaltic magma would erupt at the surface or form a crustal intrusion instead.

6.6.1. Hydrostatic head models. Until recently, the distribution of mare basalts on the Moon was thought to be solely the result of some magma transport process. The high concentrations of Fe and Ti in most lunar magmas means they are generally more dense than their terrestrial counterparts (see the following section for more details). In fact many (though not all) basalts are more dense than the Moon’s anorthositic crust (e.g., Solomon 1975). These dense lunar magmas would not have been able to erupt merely on their inherent buoyancy—some other process must have “pushed” them dense magmas to the surface.

The collection of altimetry data from the Apollo and Clementine missions led to the discovery of the Moon’s ~2 km center-of-mass/center-of-figure offset (e.g., Kaula et al. 1972, 1973, 1974; Zuber et al. 1994; Smith et al. 1997). The higher elevations of the far side have commonly been interpreted as a result of this hemisphere possessing a thicker crust, and the topographic depressions associated with the large impact basins have been attributed to the crust there being thinner than usual. The coincidence of regions of inferred thinned crust with the distribution of lava flows immediately suggested that these two phenomena were somehow genetically related. Based on the Apollo data, Kaula et al. (1973), Runcorn (1974), and Solomon (1975) all suggested that the distribution of maria could be interpreted in terms of the concept of “hydrostatic head.” In this model a dike is envisioned to extend all the way from the surface to the mare source region in the mantle. Even though the magma is denser than the crust, it is less dense than the underlying mantle. If the walls of this dike were completely rigid, a simple force balance shows that a basaltic eruption could occur at any place on the Moon’s surface if the mare source was deeper than ~250 km below the surface (e.g., Wieczorek et al. 2001). If the mare source was located above this level, then eruptions would only be able to occur below a critical elevation.

A modified version of the hydrostatic-head model was proposed by Head and Wilson (1992). They recognized that a basaltic magma would buoyantly rise through the mantle until it reached a level in which it was neutrally buoyant. Since they assumed all lunar basaltic magmas were denser than the crust, this level was located about 60 km below the surface at the crust-mantle interface. The hydrostatic pressure of a magma chamber at this level is insufficient for a dike to reach the surface so they postulated that an eruption would only occur if the magma chamber became over-pressurized. For a given amount of over-pressurization, they showed that basaltic eruptions would preferentially occur where the crust is thin. (Note that if the crust is isostatically compensated by an Airy mechanism, that crustal thickness is linearly related to elevation.)
The above hydrostatic-head models appeared to be in concordance with the limited Apollo data and hence were (and still are) widely accepted. Near-global altimetry data obtained from the Clementine mission, however, have since shown that there are substantial problems with hydrostatic-head being the sole factor in determining whether an eruption will or will not occur. In particular, the full depth and extent of the giant far-side South Pole-Aitken basin was delineated by the Clementine data. This basin’s floor was found to have the lowest elevations of the entire Moon, and yet only a small number of lava ponds are found in this topographic depression (e.g., Yingst and Head 1997). While the maria are generally located at low to moderate elevations, the paucity of lava flows in the SPA basin is clearly in conflict with the predictions of the hydrostatic-head models (e.g., Lucey et al. 1994; Smith et al. 1997).

The thickness of the lunar crust also does not appear to be the sole factor controlling whether a basaltic eruption will or will not occur. Wieczorek et al. (2001) correlated the spatial distribution of maria with a geophysically derived crustal-thickness model. In seeming agreement with this model, they found that there was an abundance of lava flows where the crust was extremely thin (<30 km). However, they noted a paucity of flows within the SPA basin where the crust is about 40 km thick, and a relative abundance of flows in Oceanus Procellarum where the crust is thicker (~50 km).

Besides these observational inconsistencies with the hydrostatic-head models, some of the assumptions with these models have been called into question by more sophisticated fluid-dynamic models of dike propagation. In particular, the walls of a dike are unlikely to be completely rigid, and hence a dike should not by able to overshoot the neutral buoyancy level by any significant amount. An experimental model of dike propagation performed by Lister and Kerr (1991) showed that a dike would propagate vertically though an elastic medium whenever the magma is positively buoyant. When this dike encountered a level of neutral buoyancy, though, vertical propagation was found to nearly cease and the dike instead propagated primarily horizontally along the interface. This behavior results from the fact that vertical propagation beyond this level has to work against gravity, whereas horizontal propagation does little gravitational work. Nonetheless, if this horizontal bladed dike was continuously fed by magma, it would slowly overshoot the neutral buoyancy level, but at a rate much lower than that of the horizontal propagation.

Using the Lister and Kerr (1991) theoretical model of a growing bladed dike Wieczorek et al. (2001) tested whether a dike stalled at the crust-mantle interface on the Moon would be able to overshoot this level by ~60 km and erupt at the surface. Even by using extreme parameters, they found that an eruption would be unlikely to occur. In order for such a dike to reach the surface, the length of the dike would have to exceed ~1000 km and more reasonable parameters gave rise to dike lengths that exceeded the circumference of the Moon.

6.6.2. The role of magma buoyancy. From the study of magmatic processes on Earth, it is apparent that a magma’s buoyancy plays a dominant role in its storage and transportation. Specifically, terrestrial magma chambers are located at the level in which they are neutral buoyant, and bladed dikes propagate along this level as well (e.g., Rubin and Pollard 1987; Ryan 1987, 1994; Walker 1989). As the near-surface layer of the Earth is highly fractured, this neutral buoyancy horizon is typically located only a few kilometers below the surface. Eruptions from shallow depth can then occur sporadically as a result of a bladed dike overshooting its neutral buoyancy horizon, from the exsolution of volatiles, or by temporal changes in the pressure of the magma body. Thus, the crust acts as a density filter, allowing only those magmas that are less dense to erupt. It is thus reasonable to suspect that magma buoyancy would play a dominant role in the storage and transport of magma on the Moon as well.

The density of a magma can be calculated for a given composition and temperature. The volume of a melt is well represented by the following equation:
where \( X_i \) is the number of moles of oxide component \( i \), \( V_i(T) \) is the partial molar volume of oxide \( i \) at temperature \( T \), and the last term takes into account a non-linear interaction between \( \text{Na}_2\text{O} \) and \( \text{TiO}_2 \). Partial molar volumes of the common oxides can be found in Lange and Carmichael (1987), and the density of a given magma is obtained by dividing the molar volumes by the total mass of the oxides. The liquidus temperature for a given composition can be approximated by using an empirical correlation between temperature and the partitioning behavior of MgO between olivine and melt (Delano 1990) or empirical relations defined by experimental studies (i.e., MELTS; e.g., Ghiorso and Sack 1995).

Using this approach, Wieczorek et al. (2001) calculated the liquidus densities of typical lunar basaltic magmas (BVSP 1981; Delano 1986). Their results are plotted in Figure 4.49 as a function of composition, and it is seen that those magmas possessing the highest concentration of Ti are generally the most dense (Delano 1990; Circone and Agee 1996), whereas those magmas possessing high concentrations of Al are the least dense. A large number of the mare-basalt magmas are more dense than typical anorthositic crustal materials (\( \rho \approx 2.86 \text{ g/cm}^3 \)) and should not have been able to erupt owing solely to buoyancy. Nonetheless, some of the VLT, high-Al, and low-Ti basalts are seen to be less dense than the anorthositic crust and could have erupted based on their inherent buoyancy.

As is summarized in Chapter 3, a number of pieces of evidence suggest that the Moon’s crust becomes increasingly mafic (and dense) with increasing depth below the surface. These include (1) the possible existence of a seismic discontinuity ~20 km below the surface that might be compositional in origin (e.g., Toksöz et al. 1974), (2) the observation that some central peaks of complex craters are highly noritic (e.g., Tompkins and Pieters 1999), (3) the observation that the ejecta of large impact basins is generally more mafic than the surrounding highlands (e.g., Reid et al. 1977; Ryder and Wood 1977; Spudis and Davis 1986; Bussey and Spudis 2000), and (4) the inference that the floor of the South Pole-Aitken basin is likely composed of noritic lower-crustal materials (Lucey et al. 1995; Pieters et al. 1997; Wieczorek and Phillips 1998). Assuming that the Moon’s crust is stratified into upper anorthositic and

![Figure 4.49](image-url)
lower noritic layers, Wieczorek and Phillips (1998) constructed a dual-layered geophysically-derived crustal thickness model of the Moon. Correlating this model with the composition of central peaks (Tompkins and Pieters 1999), (Wieczorek and Zuber 2001) constrained the density of the lower crust to be \( \approx 3.04 \, \text{g/cm}^3 \).

All of the mare basaltic magmas are predicted to be less dense than the Moon’s lower noritic crust (Fig. 4.50). Thus, if the upper anorthositic crust was completely excavated during an impact event, any basaltic magma could have later erupted within this basin based solely on buoyancy considerations. One of the more interesting aspects of the Wieczorek and Phillips (1998) dual-layered crustal thickness model (updated and slightly modified in Chapter 3) is that the upper crust is predicted to be completely absent beneath many of the large impact basins. These basins include Imbrium, Serenitatis, Crisium, Synthii, Humboldtianum, Orientale, Humorum, Nectaris, and South Pole-Aitken. Additionally, the upper crust is predicted to be less than 5-km thick beneath Mendel-Rydberg, Moscovienne, Freundlich-Sharonov, and a portion of eastern Mare Frigoris. All of these regions are seen to possess some mare basalts, and these basaltic eruptions are hence consistent with the hypothesis that density alone determines whether a basaltic magma will erupt or form a crustal intrusion instead.

The basalts that erupted wherever the upper anorthositic crust is predicted to be present must be analyzed separately. Gillis and Spudis (1998) showed that those basalts that erupted within the far side highlands, the Australe basin, and the eastern limb have lower than typical abundances of Fe as inferred from Clementine remote-sensing data (Lucey et al. 1998). Using an empirical correlation among a magma’s liquidus density, Fe and Ti content, Wieczorek et al. (2001) showed that these magmas were likely to have been less dense than the upper anorthositic crust, and hence are consistent with erupting based solely on buoyancy considerations.

The only other mare basalts that need to be reconciled with the magma buoyancy hypothesis are those that erupted within Oceanus Procellarum. The dual-layered crustal thickness model of Wieczorek and Phillips (1998) predicts that this region possesses an anorthositic crust that is \( \approx 20 \, \text{km} \) thick, yet some of the highest Ti basalts erupted in this region. There are two possible ways out of this apparent conundrum. First, it is possible that the crust in this region of the Moon might be denser than is predicted by the geophysical model.

![Figure 4.50](image_url)

**Figure 4.50.** Plot of the density of mare basaltic magmas as a function of temperature. The dot at the left end of each line corresponds to the liquidus density of the magma and the horizontal line corresponds to the approximate density of the Moon’s upper anorthositic crust (modified after Wieczorek et al. 2001).
Since Oceanus Procellarum lies within an anomalous KREEP-rich geochemical province, the crust there may in fact be quite different in composition from more typical highland regions (see Chapter 3). Voluminous intrusions of basaltic magma over time in this region may have substantially altered the bulk composition of the upper crust as well. Secondly, it is possible that the magmas in this region erupted with superliquidus temperatures. Figure 4.49 shows the density of the basaltic magmas from Figure 4.49 as a function of temperature. A few hundred degrees of superheating would make many of these basaltic magmas less dense than the upper anorthositic crust. Thermal models that take into account the high abundance of incompatible elements within and/or beneath the crust of Oceanus Procellarum, in fact, do predict extremely high temperatures that might be sufficient to make these magmas positively buoyant in the crust (Parmentier et al. 2000; Wieczorek and Phillips 2000; Zhong et al. 2000). Analysis of the textures in samples from this area may someday be used to test the superliquidus hypothesis.

Finally, we note that the near-surface layer of the Moon has low density as a result of impact brecciation. Thus, most basaltic magmas would probably become neutral buoyant a few kilometers below the surface and accumulate at this level. Though this low-density layer should inhibit an eruption from occurring, a number of mechanisms could lead to an eruption onto the surface. First, as mentioned above, a growing bladed dike can easily overshoot its neutral buoyancy horizon by a few kilometers. Secondly, the exsolution of gas from a magma at low pressures would help propel an eruption. Carbon monoxide gas, in particular, is predicted to exsolve from lunar magmas less than about 4 km below the surface (e.g., Wilson and Head 1981). Other possibilities are S and Cl. Because of this near-surface low-density layer, it seems likely that the quantity of intrusive mare basaltic magmas is probably much more voluminous than those that are visible on the surface.

### 6.6.3. Relationships between large basin impacts and mare volcanism

Before artificial satellites ever imaged the far side of the Moon, it was noted that the largest near side impact basins were filled by the relatively smooth “maria.” The hypothesis was born that the two phenomena were genetically related, and furthermore that the basin-forming process was responsible for the generation of the mare basalts (e.g., Ronca 1966). Many pieces of evidence have since cast doubt upon this hypothesis. In particular, images of the far side demonstrate that many large impact basins are not filled with mare basalts (or at least much less so than the near side basins), and radiometrically determined ages of mare samples at the Apollo 15 and 17 sites were found to be considerably younger than the basins in which they reside. Even if basins and basaltic eruptions appear to be somewhat correlated in places, it has recently been suggested that this might solely be related to large impacts removing the low density upper crust, allowing basaltic magmas to rise buoyantly through the more dense lower crust and erupt (see Wieczorek et al. 2001 and Section 6.6.2).

Based on the indistinguishable radiometric ages of the Imbrium basin and Apollo 15 KREEP basalts, as well as the inference that KREEP basalts are stratigraphically younger than the Imbrium basin (Spudis 1978), Ryder (1994) argued that the Imbrium impact event induced KREEP basaltic volcanism within this region of the Moon. A variety of pieces of evidence (e.g., Jolliff et al. 2000; Korotev 2000; Wieczorek and Phillips 2000) suggest that KREEP basalt is confined to the Procellarum KREEP Terrane (which encompasses the Imbrium basin), and thermal modeling (Wieczorek and Phillips 2000) has shown that this material could have remained molten up until at least the time of the Imbrium impact. As such, this isolated case of impact-induced volcanism might only be a result of this impact event having denuded an already molten KREEP basaltic magma chamber (Ryder 1994; Wieczorek and Phillips 2000).

The case for a genetic relationship between mare volcanism and large impact basins has recently been revived by Elkins-Tanton et al. (2004). They distinguished between two processes that could give rise to large volumes of magma. First, the excavation of large quantities of crustal material might give rise to near-instantaneous decompression melting in
the underlying mantle. Second, as a result of lateral temperature gradients beneath an impact crater, small-scale convection could have led to further melting over a much longer period of time. The total amount of melt generated in their model was found to be about $10^4$ times greater than the volume of the extrusive mare basalts. Although subsequent convection beneath the crater only accounted for 0-2% of the total melt volume, this later process was found to operate for up to 350 m.y. following the impact event (see also Arkani-Hamed and Pentecost 2001).

The most debatable aspect of the Elkins-Tanton et al. (2004) model concerns the process of instantaneous decompression melting beneath the impact basin. Ivanov and Melosh (2003) showed that impact-generated volcanism by decompression melting on the Earth is improbable. In the Elkins-Tanton et al. model, the magnitude of the pressure drop beneath the crater was calculated by assuming that the entire mantle column between the crater floor and core-mantle boundary was uplifted by an amount equal to the depth of excavation of the crater. While this might approximate the pressure conditions just beneath the crater, this assumption probably overestimates the pressure drop in the deeper mantle. Following the excavation stage of the impact event, mantle rebound and crater collapse does not occur by strictly vertical displacements; inward lateral flow of mantle materials beneath the central portion of the basin partially offset the mass deficit caused by the excavation of near-surface materials (e.g., Melosh 1989; Ivanov and Melosh 2003). As illustrated in Ivanov and Melosh (2003), the amount of central uplift beneath a crater decreases substantially with increasing depth.

Although realistic considerations of the magnitude of the pressure drop beneath impact basins suggest that the quantity of decompression melt would be much less than that found by Elkins-Tanton et al. (2004), this question deserves further investigation. The study of Ivanov and Melosh (2003) did not explicitly model impact craters with sizes typical of the largest lunar basins, nor did they attempt to employ mantle temperature profiles that might be representative of the Moon at about 4 Ga. In addition, as discussed in Section 7, the Moon’s thermal evolution appears to be highly asymmetric, and the possibility exists that decompression melting might have occurred beneath some craters and not others. Hydrodynamic simulations of large impact events on the Moon for a variety of mantle temperature profiles should be able to delimit which conditions could plausibly give rise to decompression melting. Subsequent impact-generated convection and melting has only been investigated by a few papers and warrants further investigation as well.

### 7. ORIGIN OF CRUSTAL ASYMMETRIES

#### 7.1. Introduction

Prior to the Clementine and Lunar Prospector missions, all published thermal-evolution models possessed one commonality: they all assumed that the distribution of heat-producing elements depended solely upon radius. As such, the amount of partial melt generated beneath the surface should have been independent of location. As it is well known that the majority of the lunar basalts erupted on the near side, the scarcity of visible far-side basalts was commonly attributed to magma transport processes. The most popular explanation was based on the assumption that the mare basaltic magmas were more dense than the lunar crust, and relied upon magma-chamber over-pressurization for these magmas to reach the surface (e.g., Head and Wilson 1992). Altimetry of the South Pole-Aitken basin obtained during the Clementine mission, however, is inconsistent with this hypothesis and it now appears that magma buoyancy may have been the primary factor determining whether a basaltic eruption would or would not occur (for a more in-depth discussion, see Section 6 and Wieczorek et al. 2001).

In addition to differences in crustal structure affecting eruptive flux as discussed in Section 4.6.6, the scarcity of mare basaltic eruptions on the lunar far side (particularly within the SPA basin) is also a result of a difference in magma production rates between the near- and
far-side hemispheres. It was not until after the *Lunar Prospector* mission in 1998, however, that such asymmetric thermal-evolution models began to be quantitatively investigated. The rationale for these models is clearly evident in the global Th abundance maps generated by the LP gamma-ray spectrometer (Lawrence et al. 1998, 2000), which demonstrated that Th (and by inference KREEP) is highly concentrated within a small region of the Moon encompassing Mare Imbrium and Oceanus Procellarum. This possibility had been previously suggested on the basis of the equatorial Apollo gamma-ray data (e.g., Warren and Wasson 1979, 1980; Ryder 1994; Haskin 1998; Haskin et al. 1998), but was not widely accepted or fully appreciated.

On the basis of these and other data, Jolliff et al. (2000) divided the lunar crust and underlying mantle into distinct terranes that possess unique geochemical, geophysical, and geological histories. The two most extensive terranes are the Procellarum KREEP Terrane (PKT) and the Feldspathic Highlands Terrane (FHT), and the rationale for these divisions, as well as their properties and evolutionary histories are discussed in detail in Chapter 3 (see also Haskin et al. 2000; Korotev 2000; Wieczorek and Phillips 2000). Below, we list only their most important characteristics that are necessary for deciphering the large-scale asymmetric magmatic evolution of the Moon.

**The Feldspathic Highlands Terrane:**

1. This terrane encompasses ~60% of the Moon’s surface area and is composed primarily of ancient ferroan-anorthositic rocks that crystallized from an LMO between about 4.5 and 4.4 Ga.
2. Crustal Th concentrations range from less than 1 ppm for near surface materials, to possibly 2-3 ppm thorium for deep crustal rocks.
3. Less than 10% of this terrane by area has been resurfaced by mare basaltic lava flows, with the vast majority being emplaced almost exclusively during the Imbrian period. Remote-sensing data of central peaks in this terrane further suggest that basaltic intrusions within the underlying crust are volumetrically minor (~5% by volume).
4. Basaltic flows in this terrane generally contain low to moderate concentrations of TiO$_2$ (<6 wt%) as well as low concentrations of Th (~1-2 ppm).

**The Procellarum KREEP Terrane:**

1. This terrane encompasses the regions of Oceanus Procellarum and Mare Imbrium, and is principally characterized by high Th concentrations (3-12 ppm), as well as the abundance of mare basaltic lava flows.
2. Although this terrane only comprises ~16% of the Moon’s surface area, >60% of the Moon’s extrusive basalts by area occur there. Only those regions of high-standing topography (such as the rim of the Imbrium basin) have not been volcanically resurfaced.
3. Volcanism in this terrane appears to have extended from at least 4.2 Ga to about 1 Ga.
4. The majority of Moon’s high-Ti mare lava flows occur within this terrane.
5. Judging from the composition of Imbrium ejecta, ~50% of the crust may have a composition similar to that of Apollo 15 KREEP basalt (~12 ppm Th).
6. Global Th mass-balance calculations suggest that ~30% of the Moon’s heat-producing elements may be sequestered within the Procellarum KREEP Terrane. Furthermore, about 40% of the total amount of Th in the lunar crust resides within the PKT.
7. The heat flow at the Apollo 15 site (which lies within the PKT) is ~50% greater than that measured at the Apollo 17 site (which lies just outside of this province). In addition, only those basins that formed within, or on the edge of, this province appear to have been modified by enhanced rates of viscous relaxation.
8. On the basis that most magnesian- and alkali-suite rocks are defined by their high incompatible element contents (i.e., KREEP signature), it is unlikely that these types of rocks formed outside the PKT.

9. From orbital gamma-ray data, surface Th concentrations of the mare-basalt-dominated regoliths in this region range from ~3 to 7 ppm. It remains unclear, however, whether these Th concentrations are representative of the basalts themselves, or if they result from vertical mixing with an underlying Th-rich substrate.

The high abundances of heat-producing elements within the Procellarum KREEP Terrane are in all likelihood related to the magmatic productivity and volcanic longevity associated with this region. In order to have a complete understanding of the Moon's thermal evolution two important questions need to be addressed: (1) How did KREEP become concentrated in a single geologic province? (2) How did this enhancement in heat-production within the PKT affect its magmatic evolution? Of course, these two questions are highly coupled, but it may nonetheless be useful to initially separate these two problems. In the remainder of this section we discuss the different models that have been advocated to date in order to explain these and other characteristics. These models are somewhat preliminary; more detailed calculations are needed to fully explore their feasibility and implications.

7.2. Effects of a thicker KREEP layer within the Procellarum KREEP Terrane

Prior to the Lunar Prospector mission, most researchers assumed that the fractional crystallization of an LMO would have given rise to a KREEP layer between the crust and mantle that was nearly global in extent. Estimates for the thickness of this layer depended upon many factors, including the initial depth of the LMO, but was commonly assumed to be approximately two kilometers (assuming an LMO that was 360 km deep; e.g., Warren and Wasson 1979). If instead KREEP were to reside mainly within the Procellarum KREEP Terrane, as appears to be the case, then how much thicker would this layer be? Assuming this terrane encompasses ~16% of the Moon’s surface area, its equivalent thickness would be ~10 km. If the LMO was deeper than originally assumed (e.g., Pritchard and Stevenson 2000), then the KREEP layer within the Procellarum KREEP Terrane would be even thicker.

Setting aside the question of how KREEP came to be concentrated within a single lunar province, Wieczorek and Phillips (2000) quantitatively modeled the thermal consequences of such a localized enhancement in heat-producing elements. On the basis of geochemical and geophysical data, they argued that up to 50% of the crust within the PKT (~30 km) might be composed of a material similar in composition to KREEP basalt. This rock possesses REE concentrations ~300 times chondritic (12.4 ppm Th; Korotev 2000), which is just less than that of Warren and Wasson’s (1979) urKREEP (18 ppm Th). In order to approximate the thermal state of the mantle just after the bulk of the LMO crystallized, its initial temperature profile was assumed to be adiabatic with the top of the mantle set to its solidus. The KREEP basalt layer was approximated as a spherical cap having an angular radius of 40° and was initially placed between the crust and mantle. Mantle heat-producing abundances were estimated from the mare basalts and appropriate distribution coefficients (e.g., 25 ppb Th; Warren and Wasson 1979), whereas crustal abundances were taken from orbital gamma-ray data (0.53 ppm Th). The bulk U concentration of this model Moon is just larger than that of the Earth’s primitive mantle, and is about twice that of ordinary chondrites.

Whereas many lunar thermal evolution models have assumed that internal heat was partly transported by convection, a purely conductive model was employed in this study. As is discussed in Section 6, this assumption is at least plausible given that (1) the crystallization of an LMO would likely result in a mantle that was severely depleted in heat-producing elements, (2) the crystallization of an LMO could have resulted in a density-stratified mantle that was stable against convection, and (3) the mantle beneath the PKT would have been heated...
primarily from the top down, counteracting any tendency for large scale mantle convection. While melting was accounted for in this model using approximate latent heats and phase diagrams, melt transport was not modeled (i.e., all melt remained in the mantle).

The principal results of this model are illustrated in Figure 4.51 which shows the spatial extent of melting beneath the Procellarum KREEP Terrane as a function of time. Melting within the mantle is predicted to occur only directly below this terrane, with the depth of melting increasing with time. Owing to the lateral transport of heat near the edges of the terrane, the lateral extent of partial melting is further predicted to diminish with time. For the particular initial conditions of this model, melting should have extended ~500 km below the surface and volcanism would have spanned the entire geologic history of the Moon (though with exponentially decreasing melt-production rates with time).

In Figure 4.52 the predicted present-day heat flux is plotted as a function of distance from the center of the PKT. The heat flux decreases across the boundary between the PKT and FHT. Unfortunately, even though the measured heat-flow values at the Apollo 15 and 17 sites are consistent with this model, these landing sites are located very close to this boundary where the heat flow is predicted to vary significantly. This leaves open the possibility that these measurements are not representative of either the PKT or FHT. Knowledge of the heat flux within the heart of the Procellarum region, as well as far from it, would thus dramatically help to constrain these types of models.

The first-order result of melting within the mantle directly beneath the PKT appears to be an inevitable consequence if the equivalent of 10 km of KREEP basalt resides there. For
instance, if all the heat-producing elements were removed from the mantle, then the maximum depth of partial melting beneath the PKT would still be ~340 km and span ~4 Ga of lunar history. Conversely, if the entire complement of heat-producing elements within a 10 km KREEP basalt layer were uniformly distributed over the 60 km thick crust in this province, then melting would have reached a maximum depth of ~180 km and would have lasted for ~1.5 Ga. The first-order results of this model were confirmed by Hess and Parmentier (2001) using slightly different parameters, though they ultimately considered these model predictions to be inconsistent with petrological and geophysical constraints.

7.2.1. Model implications. Although the thermal model of Wieczorek and Phillips (2000) is admittedly simple, it nevertheless offers a coherent framework in which to understand a large variety of previously unrelated aspects of the Moon’s geologic and thermal evolution. Below we list those as advocated by Wieczorek and Phillips (2000). Most of these still remain to be more thoroughly explored.

1. The fact that the maria are primarily located on the lunar near side is a natural consequence of the high abundance of heat-producing elements within the crust of the Procellarum KREEP Terrane.

2. The fact that the measured heat flow is larger at the Apollo 15 site than at the Apollo 17 site is a direct consequence of this site being closer to the center of the PKT (e.g., Langseth et al. 1976; Warren and Rasmussen 1987).

3. The range of crystallization ages of the Mg- and alkali-suite rocks (Table 3.3 in Chapter 3) is the result of the slow crystallization of discreet intracrustal KREEP-basalt magma chambers (e.g., Snyder et al. 1995a, b). This extended range of KREEP-related magmatism is also indicated by the crystallization age of Apollo 15 KREEP basalts and their coincidence with the Imbrium formation event (Ryder 1994).

4. The reason that an igneous protolith to the Low-K Fra Mauro (“LKFM”) mafic impact-melt breccias (commonly interpreted as representing Imbrium ejecta) has not been identified is that this impact melt is a mixture, of which one component was initially molten (Spudis et al. 1991).

5. One of the interpretations of the 500-km seismic discontinuity observed beneath the near side crust is that it is the base of the LMO cumulate pile. Based on this interpretation, mare magmatism (a product of cumulate melting) represents melts produced above this discontinuity.

7.2.2. Model shortcomings. Even though this model appears to explain the major first-order features of the Moon’s post-LMO magmatic evolution, there are a few issues that still need to be critically addressed. Perhaps the most worrisome (Wieczorek and Phillips 2000; Hess and Parmentier 2001) is that the KREEP layer in these types of models generally remains partially molten for more than a billion years. Thus, even though partial melting in the mantle underlying the Procellarum KREEP Terrane might give rise to mare basaltic magmas, in order for such magmas to erupt at the surface, they must have first passed through this molten layer. Even if this were possible, one might have expected these basaltic melts to have become strongly contaminated by the KREEP component in the process, something that has not been observed in the Apollo sample collection.

There are two possible ways to resolve this conundrum. First, it is possible that the KREEP layer beneath the PKT was not a continuous giant sill, but rather that this magma resided in numerous bladed dikes and/or laterally restrictive sills. Basaltic eruptions would then be possible wherever there was a continuous solid pathway from the mare source to surface. Admittedly, this explanation is difficult to test in the absence of global heat-flow data. Alternatively, it is possible that the KREEP layer cooled more quickly than predicted in this model, solidifying around 3.9 Ga when most of the visible maria erupted. Recent reanalysis of the Apollo seismic
data have shown that the crust is probably much thinner than previously assumed (~38 km (Khan and Mosegaard 2002) as opposed to ~60 km (Toksöz et al. 1974)), and this would indeed favor more rapid cooling in this province (e.g., Hess and Parmentier 2001). It remains to be tested, though, as to how this would affect the conclusions of Wieczorek and Phillips (2000).

Another potential problem noted by Wieczorek and Phillips (2000) and Hess and Parmentier (2001) is that if the crust were indeed much hotter than typical beneath the PKT, then the lithosphere might not have been able to support the excess load associated with the Imbrium “mascon.” While this question deserves further scrutiny, it is probably not a fatal flaw to the hypothesis that the PKT is the host to a KREEP layer that is on the order of 10 km thick. By modeling the deformation associated with a load on a visco-elastic self-gravitating sphere, Zhong and Zuber (2000) have demonstrated that the Moon can support considerably greater stresses in comparison to that of the Earth as a result of its smaller planetary radius. In particular, for the case where the lunar lithosphere was approximated as having a temperature profile corresponding to 100 Ma terrestrial oceanic lithosphere, the degree of compensation of the surface load was found to be less than 0.3 for all wavelengths.

Another criticism mentioned by Hess and Parmentier (2001) concerns the stability of the anorthositic crust overlying a KREEP basaltic magma. In particular, molten KREEP basalt has a density of ~2.7 g/cm³, which would be approximately 4-7% less dense than that of the overlying anorthositic crust (2.8-2.9 g/cm³ neglecting fractures). Thus, either this magma would have buoyantly risen through the crust with some portion of it erupting at the surface, or the crust would have catastrophically foundered. While pristine KREEP basalts are rare in the Apollo sample collection in apparent contradiction with these scenarios, the much younger visible mare basalts would have certainly buried any such ancient KREEP basaltic flows. Even presuming that one of these two scenarios did indeed occur, Wieczorek and Phillips (2000) have shown that if the heat production of a 10 km KREEP basalt layer was spread over the entire crust in this province that volcanism would still be predicted to occur there for ~1.5 Ga.

Finally, we note that the volume of melt produced in this model is predicted to be somewhere between 5× and 56× that of the visible basalts, depending upon the model assumptions. The lower end of this range might be consistent with the crust possessing voluminous basaltic intrusions, but the upper bound would imply that the entire crust is composed of mare basalts. While these large volumes of basaltic melt are problematic, it is likely that this is only an artifact of the simplified model that was employed. For instance, if melt transport were taken into account then the amount of melt generated would have been reduced. The amount of melt depends upon bulk composition and the phase relationships for the mantle source, which are rarely taken into account in thermal evolution models. While Wieczorek and Phillips (2000) used the bulk-Moon liquidus and solidus of Ringwood (1976), these are expected to only crudely approximate reality. Indeed, the effects of mantle overturn (e.g., Hess and Parmentier 1995) might have given rise to a refractory and Mg-rich upper mantle that would drastically reduce the amount of melt generated in these models.

7.3. Degree-1 upwelling of ilmenite-rich cumulates

One consequence of the Moon possessing an LMO is that ilmenite-bearing mineral assemblages should precipitate near its terminal stages of crystallization (e.g., Snyder et al. 1992). As a result of their relatively high-Ti and Fe contents, these cumulates should be relatively dense (~3.7 g/cm³) and hence might have sunk through the lunar mantle (e.g., Ringwood and Kesson 1976; Herbert 1980; Spera 1992; Hess and Parmentier 1995) (Section 3.3). Remelting of these cumulates in the deep lunar interior is widely thought to be implicated in the origin of the high-Ti mare basalts (Section 5).

Alternative explanations for the hemispheric asymmetry in the distribution of mare basalts are related to the fate of these ilmenite-rich cumulates. Two models have been advocated that
rely upon the development of a long-wavelength (degree-1) convective instability in order to concentrate these materials within one lunar hemisphere and to thus give rise to a higher than typical magma production rate. The starting premise of the model proposed by Zhong et al. (2000) is that late-stage LMO ilmenite-rich cumulates were somehow able to sink through the entire mantle, forming a global layer above the Moon’s iron core. It was assumed that these cumulates were mixed in some proportion with more typical mantle materials (decreasing their density in the process), and that this material further had an enhanced abundance of heat-producing elements owing to the presence of some proportion of trapped intragranular KREEP-rich liquid. While these cumulates would have initially been gravitationally stable at the base of the mantle, radioactive heating within this layer would reduce its buoyancy with time, and at some point might have caused it to have become buoyantly unstable. The question is thus to determine under which conditions this layer would rise predominantly within a single hemisphere of the Moon, or more precisely, under which conditions a degree-1 instability would have the largest growth rate.

This question was addressed from the perspective of both a Rayleigh-Taylor linear stability analysis, as well as through finite-element modeling of thermo-chemical convection. From the Rayleigh-Taylor stability analysis, it was shown that the wavelength dependent growth rate depended primarily upon both the radius of the core and of that of the mixed ilmenite-rich cumulate layer, with the viscosity contrast between the mantle and ilmenite cumulates playing only a secondary role (Fig. 4.53). In this plot, a degree-1 instability would be the most favored when the normalized growth rate has a value of unity. The different curves within Figure 4.53 correspond to different assumed radii of the Moon’s core. The most important result is that a degree-1 instability would only have the largest growth rate if the radius of the core was less than 250 km.

These results were confirmed and expanded upon by axis-symmetric thermo-chemical convection modeling. As one example, Figure 4.54 shows how both the temperature and compositional fields evolve with time. In this particular simulation, a degree-1 upwelling structure is seen to develop after ~400 Ma. While the onset time for this upwelling to occur depends upon the assumed model parameters, it should lie between ~250 and 650 Ma for reasonable choices. Partial melting could occur within this upwelling plume either as a result of the higher temperatures found there, or as a result of decompression melting.

**7.3.1. Model implications and shortcomings.** The main successes of the Zhong et al. (2000) model are the following:

1. Melting within a single ilmenite-rich “plume” could have given rise to a hemispheric asymmetry in the distribution of mare basaltic lava flows.
2. Some portion of these basaltic magmas should have been enriched in Ti, consistent with the abundance of high-Ti lava flows on the lunar near side.

3. The onset time for the ilmenite-rich materials to become buoyant and rise is on the order of 500 Ma, which is in general agreement with the observation that the majority of the visible mare were emplaced after ~3.9 Ga.

4. The removal of an ilmenite-rich “thermal blanket” just above the core around 4 Ga could have possibly given rise to a geodynamo, potentially explaining the paleomagnetic signatures observed in some lunar rocks (Stegman et al. 2003) (See also Chapter 3, Section 4.6).

Even though this type of model has some attractive attributes, there are a few critical (and perhaps fatal) concerns that need to be addressed. For example, this model requires the lunar core to have a radius less than 250 km, whereas most geophysical estimates suggest that the lunar core is in fact larger. For instance, an analysis of the Moon’s magnetic induction response suggests that it possesses a metallic core having a radius between 250 and 430 km (Hood et al. 1999). In addition, an analysis of the Moon’s rotation from three decades of lunar laser ranging data (LLR; Williams et al. 2001) strongly implies the existence of both solid-body tidal dissipation, as well as dissipation at a liquid-core/solid-mantle interface. Assuming that the core is composed of pure iron, its radius is constrained to be between 314 and 352 km. A sulfur-rich core would possess a larger radius. The possibility exists, however, that the LLR study has overestimated the core radius by neglecting the effects of dissipation at a solid-inner core interface. Nevertheless, the fact that the two most stringent core-radius estimates are currently in excess of the 250 km maximum radius required by Zhong et al. (2000) is not encouraging for this model.

An additional concern is that while this model can explain the hemispheric asymmetry of mare basaltic lava flows, it does not address why KREEP is concentrated solely within the PKT where most of these mare basalts happen to reside. Although partial melting within an ilmenite-rich plume might have given rise to Th-rich mare basaltic magmas, such samples are extremely rare, or possibly even non-existent, in the Apollo sample collection (see Section 3.5 for discussion of this issue). Even supposing that the 3-7 ppm surface Th concentrations of the basalts within the PKT were a result of such Th-rich mare basaltic lavas, it still would remain unexplained as to why the underlying crust has even higher concentrations. As such, while this model appears to be consistent with the record of mare volcanism, the spatial correlation that exists between the basaltic flows and the high-thorium underlying crust remains an unexplained coincidence.
7.4. Degree-1 downwelling of ilmenite-rich cumulates

In contrast to a degree-1 upwelling of ilmenite-rich cumulates from the base of the mantle, it has been proposed by Parmentier et al. (2002) that the ilmenite-rich cumulates could have become concentrated within a single hemisphere as they sank through the mantle in a degree-1 manner. The model’s starting premise is that ilmenite-rich cumulates crystallized from the LMO forming a globe-encircling layer between the crust and mantle. As a result of the high density of this initially thin layer, small diapirs should have descended a small distance into the underlying mantle, forming a much thicker (though still global) layer of mixed ilmenite and mantle cumulates. The problem addressed in this study was to determine under which conditions this mixed layer might descend into the mantle solely within a single hemisphere.

Like the Zhong et al. (2000) study, this problem was addressed both from the perspectives of a Rayleigh-Taylor stability analysis as well as thermo-chemical convection simulations. From the Rayleigh-Taylor stability analysis it was shown that a predominantly degree-1 downwelling would occur only if (1) the thickness of the mixed ilmenite cumulate layer was greater than about 50 km, and (2) if the underlying mantle had a viscosity more than \(10^3\) that of the mixed cumulate layer (Figs. 4.55 and 4.56). The absolute growth rates were found to be proportional to the density contrast between the cumulates and mantle, and inversely proportional to the corresponding viscosity contrast. If it was assumed that the mantle viscosity increased with depth, then short-wavelength instabilities were always found to be favored. Nevertheless, as these short wavelength instabilities act to increase the thickness of the mixed layer, a degree-1 instability would eventually represent the dominant flow pattern at later times.

Thermo-chemical convection modeling has confirmed that such a degree-1 downwelling of dense cumulates could indeed occur. The timescales associated with the descending plume depend on factors such as the mantle and layer viscosity, whereas its associated temperature depends on the concentration of incompatible elements. Although these two factors are not well constrained, parameter values can be found that give rise to a plume that descends \(\approx 500\) km, and starts to partially melt, in about 500 Ma. Although not explicitly modeled, if a liquid KREEP-rich layer still existed at this time, it would be expected to pool above this downwelling. In this manner mare volcanism as well as a crustal KREEP enhancement would be predicted to occur in the same region.

![Figure 4.55](image_url)

**Figure 4.55.** Rayleigh-Taylor instability analysis showing growth rates (normalized to the maximum growth rate) as a function of spherical harmonic degree. (A) Growth rates as a function of layer thickness for a mantle that is \(10^4\) more viscous than the cumulate layer. (B) Growth rates as a function of viscosity contrast for a layer thickness of 100 km. [Reprinted with permission of Elsevier from Parmentier et al. (2002) *Earth & Planetary Science Letters*, Vol. 201, p. 473-480.]
7.4.1. Model implications and shortcomings. The main successes of this model include:

1. A dense and global magma-ocean cumulate layer could descend through the mantle in a degree-1 pattern, concentrating KREEP and ilmenite-rich materials within one hemisphere of the Moon.

2. Melting within this downwelling would be promoted by its high concentration of heat-producing elements and could conceivably account for the commencement of the “main phase” of mare volcanism at ~4 Ga on the lunar near side.

3. Melting within this mixed ilmenite cumulate layer could generate high Ti mare basaltic magmas.

4. If a liquid KREEP-rich layer was still present while the ilmenite cumulates were sinking through the mantle, it should have concentrated above this dense downwelling.

The main shortcoming of this model concerns the requirement that the mixed cumulate layer be some $10^3$ less viscous than the underlying mantle. This is problematic because the late-stage ilmenite-rich magma-ocean cumulates would have been rich in clinopyroxene, which is about $100\times$ more viscous than olivine. (See Elkins-Tanton et al. (2002) for a discussion of some of the problems associated with generating models in which ilmenite-rich cumulates sink into the mantle). The high-viscosity of this mineral could be partially compensated by inclusion of about 5-10% interstitial liquid within the mixed layer. While this might initially be plausible, the lower density of the liquid component would act to reduce the material’s bulk density, increasing the timescales associated with the downwelling event. Additionally, compaction of the partially molten matrix might be able to expel the interstitial liquid on a timescale associated with this downwelling plume.

Another issue is whether melting within a mixed ilmenite-rich downwelling would give rise to magmas that are geochemically similar to the high- and low-Ti mare basalts. In particular, melting within this downwelling would be facilitated by high concentrations of incompatible and heat-producing elements associated with the late-stage magma-ocean cumulates, yet the sampled mare basalts for the most part possess low concentrations of these elements.

7.5. The effect of crustal thickness variations during LMO crystallization

A conceptually simple mechanism to concentrate KREEP within a single region of the Moon is based on the premise that crustal thickness variations might have existed at the time...
The LMO was crystallizing. It is likely that the Moon’s center-of-mass/center-of-figure offset is an ancient feature dating from this time. This feature is most popularly explained by the far-side crust being thicker than that of the near side. Warren and Wasson (1979) and Wasson and Warren (1980) originally noted that if the anorthositic crust was indeed thicker than usual for the far side, then the thickness of the LMO there would have been locally thinner. As a result of the reduced heat production beneath this region of thickened crust, LMO crystallization would have proceeded at a greater rate, resulting in a feedback process leading to a yet thinner LMO. As crystallization in this region progressed towards completion, the far side crust would have eventually “grounded” on the underlying mantle, effectively removing the KREEP layer from this hemisphere. This model was reiterated and slightly modified by Warren and Kallemeyn (1998) (Fig. 4.57).

The question thus remains as to why the crust, at the time of LMO crystallization, was thinnest beneath what is now the Procellarum KREEP Terrane. Although there are plausible explanations, they are not easily testable. For instance, if an ancient and giant “Procellarum” impact occurred, forming a crater larger than the South Pole Aitken basin, the crust would have been thinned, causing accumulation of molten KREEP. However, there is no unequivocal evidence that such a “Procellarum impact basin” (Neumann et al. 1996) exists, though admittedly such an ancient feature would have been heavily modified by the Moon’s subsequent impact history. An alternative hypothesis is that a large number of statistically independent intermediate-sized impact events might have caused a net redistribution of crustal materials from one side of the Moon to the other. The expected magnitude of this process, while ultimately statistical in nature, has yet to be worked out. Finally, aspects related to convection within the LMO may have somehow caused large-scale crustal-thickness variations.

7.6. The effect of large impacts and mantle convection on the distribution of KREEP

A final model of note is that advocated by Arkani-Hamed and Pentecost (2001) for creating lateral variations in the thickness of an originally global KREEP-rich layer. This model rests upon the premise that large impact basins should have created localized near-surface thermal anomalies that subsequently affected the pattern of mantle convection in their vicinity. Convection simulations showed that the KREEP layer close to the crater should have become entrained in the mantle circulation and should have been displaced both radially outward and into the underlying mantle. For the case of a South Pole-Aitken sized impact, the convection cells were

![Figure 4.57](image-url) Conceptual manner in which KREEPy magma-ocean residua might have become concentrated within a single region of the Moon. Formation of a thinned crust beneath what is now the Procellarum KREEP Terrane gives rise to a relatively thicker magma ocean in this region. This could have occurred either by a “Procellarum impact” or by regional statistical fluctuations in crustal thickness. By the time the South Pole-Aitken basin formed on the far side, the majority of KREEP had already been concentrated beneath the PKT.
shown to strip away the KREEP layer within one crater radius and to transport it into the underlying mantle. In contrast, for an Imbrium-sized impact, the convective cells were weaker and acted to concentrate all KREEP within 1 crater radius into a narrow annular ring 1 crater radius away. From the work of Manga and Arkani-Hamed (1991), it was suggested that the combination of a low-conductivity ejecta blanket, as well as a KREEP enhancement about 1-crater radii away, could produce partial melting in the upper mantle. Favorable stress conditions might then have led to lateral transport of this magma from beneath surrounding highlands into the basin.

Although it is plausible to expect large impact basins to have influenced localized convection cells, the redistribution of KREEP in this model may be an artifact of assuming that its density was the same as that of the mantle. If the progenitor of the Mg-suite rocks was an Mg-rich KREEPy magma (such as KREEP basalt) sandwiched between the crust and mantle, its lower density would have resisted entrainment in the mantle circulation. Indeed, it may be that this layer was still molten at the time of the Imbrium impact (e.g., Hubbard and Minear 1976; Solomon and Longhi 1977). It is also unlikely to expect that a low-conductivity ejecta blanket surrounding a basin would significantly influence thermal evolution because (1) the upper layer of the Moon had already been extensively brecciated by numerous impacts, and (2) thermal conductivity depends primarily upon rock porosity, which is mostly a function of lithostatic pressure (i.e., depth below the surface).

### 7.7. Summary and conclusions

Thermal models that attempt to explain the asymmetric magmatic evolution of the Moon are still in their infancy. At present, two scenarios seem most plausible for the formation of the Procellarum KREEP Terrane and the long duration of magmatic activity there. First, KREEP may have become concentrated within the PKT as a result of the crust there being thinner than typical. Subsequently, radiogenic heating within this layer would have heated the underlying mantle giving rise to a wide compositional range of basalts. Secondly, a degree-1 downwelling of ilmenite-rich cumulates may have concentrated both the residual magma-ocean KREEP layer and subsequent magmatism in one hemisphere as has been proposed by Parmentier et al. (2000, 2001). Although the degree-1 upwelling model of Zhong et al. (2000) can not be dismissed, this seems to be the least plausible model as current estimates of core size would disallow such a hemispheric upwelling. Moreover, this model does not address the origin of the KREEP enhancement found within the Imbrium and Oceanus Procellarum region.

### 8. NEW VIEWS OF THE THERMAL AND MAGMATIC EVOLUTION OF THE MOON

#### 8.1. Introduction

The pioneering science of the Apollo missions created an observational and conceptual framework on which to develop a comprehensive understanding of the Moon and to construct scientifically well-founded lunar missions. In the two decades immediately following the Apollo missions, lunar exploration was based on new analytical-computational-technological advances and a retrospective view of the science established during Apollo. Not until the 1990s did the Galileo, Clementine and Lunar Prospector missions provide additional direct observational data. This chapter presented new insights into the thermal and magmatic evolution of the Moon afforded by the integration of new approaches and observation. A summary of some of the more important new views follows.

#### 8.2. New views

The observational data derived from the Clementine and Lunar Prospector missions further resolved the asymmetrical nature of the Moon. Although it had been long recognized that the distribution of mare volcanism was asymmetrical, the observation that a substantial
enrichment of the Procellarum terrane in incompatible and heat-generating elements (KREEP) was a decisive advance in understanding lunar evolution and early planetary differentiation. Although our understanding of the mechanism behind the distribution of KREEP is in its infancy, the potential ramifications are profound. The asymmetry of the thermal character of the lunar crust and mantle appears to affect not only the ability of the mantle to generate basalts but it also affects distinct evolutionary pathways for the lunar mantle.

New data on short-lived isotopic systems (Nd-Sm, Hf-W) and computational modeling of both large impacts and accretional processes has set limits on the timetable for the formation and differentiation of not only the Moon but all the terrestrial planets. Building on the violent encounter of the Earth with a Mars-size body, rapid accretion of the Moon is predicted by models and is consistent with isotopic systematics. An important consequence of rapid accretion is that substantial heat is stored in the early Moon. This provides a solution for the problem of not having enough energy to generate a magma ocean. Although early thermal modeling of the LMO suggests that the majority of the magma ocean crystallized in about 150 to 250 m.y., interpretation of Hf-W isotopic data and the old ages of Mg-suite plutonic rocks suggests that most of the LMO event (99%) may have been very short lived (10s of millions of years). Rapid, early differentiation of the Moon is also consistent with evidence for other planetary bodies in the solar system (i.e., Mars).

Since the realization that the Moon and potentially other terrestrial planets may have experienced rapid differentiation via a magma ocean, the concept of the LMO has become increasing both more important and complex. The increasingly complex modeling of the dynamics of a global magma ocean have yielded considerable insights and a conceptual framework on which to attach petrologic and geochemical observations. Yet to be accomplished is integrating these complex theoretical models with more simplistic petrologic perceptions of LMO crystallization.

The recognition that the pyroclastic glasses represent the best approximations of primary lunar basalts has lead to insights such as the depth of the LMO, dynamics of the lunar mantle, composition of the lunar mantle, and the depth of origin of the mare basalts. The prevalent opinion soon after the Apollo missions was that most mare basalts were derived from the shallow lunar mantle (< 400 km) and that only the outer portion of the Moon participated in the LMO. In contrast, studies of the pyroclastic glasses strongly suggest that some aspects of mare magmatism reflect melting initiated in the deep mantle. The conclusion reached from experimental studies of the pyroclastic glasses that both the high-Ti and low-Ti magmas were generated from the deep lunar mantle added credence to the Ringwood and Kesson models that were proposed in the mid-1970s that the LMO cumulate pile was gravitationally unstable. More sophisticated modeling of cumulate instability has provided a whole new perspective on both mare magmatism and earlier periods of lunar magmatism (i.e., Mg-suite).

The integration of global, remotely sensed spectral data with lunar samples has corrected early misconceptions concerning mare magmatism. Early sample analysis indicated that mare basalts were bimodal in composition with regard to Ti, that there was a temporal relationship between Ti content and eruption, and that the duration of basaltic magmatism was fairly short-lived (4.4-3.0 Ga). In contrast, current integrated data indicate that (1) low-Ti basalts are the dominate basalt type, (2) there appears to be no or limited relationship between Ti content and age (although there appears to be a relationship between age and KREEP component in lunar basalts), (3) there appears to be a relationship between mare basalt composition and its regional distribution, and (4) basaltic magmatism extended from approximately 4.5 to 1.0 Ga. These new observations have dramatic implications for the thermal history of the Moon, the thermal character of the lunar mantle, and the generation of lunar basalts.

Lunar exploration has many lessons for exploring the thermal and magmatic evolution of a planet. Exploration of the Moon has provided a paradigm for exploring a planetary body.
Achieving the scientific goals for planetary exploration requires a coordinated and integrative program involving orbital global mapping, in situ generation of geophysical, mineralogical, and geochemical data, and sample return and analysis. Sample return is not the last step in an extensive planetary exploration program. Rather, it is critical that data obtained from sample return missions be integrated into the analytical and scientific goals of subsequent global mapping and in situ missions to provide a broad planetary perspective.

8.3. Goals for future science initiatives

Lunar scientific exploration has demonstrated that for a “simple” planetary body, the early stages of planetary evolution so well preserved on the Moon are very complex. Planetary-scale remotely sensed observations and the lunar meteorites indicate that materials collected by the Apollo and Luna missions, however, sampled only a restricted subset of the Moon’s complexity.

First and foremost, an understanding of the Moon’s internal structure is needed to significantly advance our understanding of its thermal and magmatic evolution. New missions designed to collect new geophysical data are necessary to gain further insight into the size and composition of the lunar core, the depth of the LMO, and the existence and nature of the lower lunar mantle.

Key to understanding the early differentiation of the Moon and other planetary bodies is gaining a rigorous theoretical understanding of the workings of a magma ocean. Models have become much more sophisticated since the Apollo missions, but they are not yet well integrated with observational data. Integrating complex modeling within a strict petrologic framework is essential. The Moon provides the best preserved remnant of early planetary evolution and therefore provides a natural laboratory to test and constrain complex models. Unfortunately, much of our perspective of the Moon comes from biased sampling. Sampling of other regions of the Moon will surely provide additional constraints and undoubtedly many surprises.

9. ACKNOWLEDGMENTS

The authors of this chapter would like to thank the numerous science programs in NASA that funded both the original research summarized in this chapter and the organization and preparation of this manuscript. We would also like to thank the Lunar and Planetary Institute for their assistance in organizing many of the New Views of the Moon workshops and special LPSC sessions and contributing to the preparation of this manuscript. This manuscript was vastly improved by reviews provided by Marc Norman, Jim Papike, and Brad Jolliff. The team leader (CKS) is very appreciative for all the unselfish contributions made by team members, their valuable insights expressed in this manuscript and internal reviews that they provided. CKS would also like to thank members of the Institute of Meteoritics for the invaluable input they provided during the preparation of this manuscript.

10. REFERENCES

Cameron AGW (1997) The origin of the Moon and the single impact hypothesis V. Icarus 126:126-137
Elkins-Tanton LT, Van Orman JA, Hager BH, Grove TL (2002) Reexamination of the lunar magma ocean cumulative overturn hypothesis: Melting or mixing is required. Earth Planet Sci Lett 196:249-259
Shearer et al.


Green DH, Blundy JD, Adam J, Yaxley GM (2000) SIMS determination of trace element partition coefficients between garnet, clinoxyroxene and hydrous basaltic liquids at 2-7.5 GPa and 1080-1200 °C. Lithos 53:165-187


Halliday AN, Poccelli D (2001) In search of lost planets – the paleocosmochemistry of the inner solar system. Earth Planet Sci Lett 192(4):545-559


Hartmann WK (1965) Terrestrial and lunar flux of large meteorites in the last two billion years. Icarus 4:157-165


Hartmann WK, Davis DR (1975) Satellite-sized planetesimals and lunar origin. Icarus 24:504-515


Helmke PA, Haskin LA, Korotev RL, Ziege KE (1972) Rare earths and other trace elements in Apollo 14 samples. Proc Lunar Sci Conf 3:1275-1292


James OB, Flohr MK (1983) Subdivision of the Mg suite noritic rocks into Mg-gabbronorites and Mg-norites. J Geophys Res 88:A603-A614
Kaula WM (1979) Thermal evolution of Earth and Moon growing by planetesimal impacts. J Geophys Res 84:999-1008


Shearer et al.


Oldenburg CM, Spera FJ (1990) Numerical experiments of magmatic solidification and convection. Eos Transactions, American Geophysical Union 71:1660


Papanastassiou DA, Wasserburg GJ (1971) Rb-Sr ages of igneous rocks from the Apollo 14 mission and the age of the Fra Mauro formation. Earth Planet Sci Lett 12:36-48


Papike JJ (1996) Pyroxene as a recorder of cumulative formational processes in asteroids, Moon, Mars, Earth: Reading the record with the ion microprobe. Am Mineral 81:525-544


Papike JJ, Fowler GW, Shearer CK (1994) Orthopyroxene as a recorder of lunar crust evolution: An ion microprobe investigation of Mg suite norites. Am Mineral 79:790-800


Papapastasiou DA, Wasserburg GJ (1971) Rb-Sr ages of igneous rocks from the Apollo 14 mission and the age of the Fra Mauro formation. Earth Planet Sci Lett 12:36-48

Pieters CM, Englert P (eds), Cambridge University Press, p 309-336


Presnall DC, Hoover JD (1987) High pressure phase equilibrium constraints on the origin of mid-ocean ridge basalts.

Pianetti DE, Mysen BO (eds), Geochemical Society, p 75-89

Pieters CM, Englert P (eds), Cambridge University Press, p 309-336

Shearer CK, Righter K (2001) Hafnium and tungsten partitioning in silicates. A key to understanding the early evolution of both the Moon and Mars. Lunar Planet Sci XXXII:1620


Solomon SC, Toksoz MN (1973) Internal constitution and evolution of the Moon. Phys Earth Planet Inter 7:15-38


Tera F, Papanastassiou DA, Wasserburg GJ (1973) A lunar cataclysm at ~3.95 AE and the structure of the lunar crust. Lunar Planet Sci IV:723-725
Urey HC (1952) The planets: Their origin and development. Yale Univ. Press
Urey HC (1959) Primary and secondary objects. J Geophys Res 64:1721-1737


Weitz CM, Rutherford MJ, Head JW, McKay DS (1999) Ascent and eruption of a lunar high-titanium magma as inferred from the petrology of the 74001/2 drill core. Meteorit Planet Sci 34:527-540


Shearer et al.