地质专业英语教材

(研究生用)

山东科技大学

2008.9

Contents

I. Early Earth

John W. Valley, Elements 2(4) (2006) 201-204.

The earliest Earth was a strange inhospitable world, yet transitions to a more familiar planet occurred within the first billion years. In spite of sparse preservation of an ambiguous rock record, recent studies refine the nature and timing of key events. This issue reviews current knowledge of the age of the Earth, massive early meteorite impacts, the early atmosphere and hydrosphere, the rock record, and the first life.

The surface of the Earth was cruelly inhospitable at the time of its birth about 4.5 billion years ago. Red-hot oceans of magma, massive meteorite strikes, and a dense atmosphere elicit the name "Hadean" for this earliest time period. Clearly such extreme conditions had to subside before continents, oceans, and life could survive on Earth (see cover). So when did the Hadean end? Evidence from rocks has been used to place this boundary at 3.8 Ga (billion years before present), the age of the oldest known water-laid sediments. The fossil record of life begins even later at 3.5 Ga, although low carbon isotope ratios (13C/12C) in carbonaceous matter from Isua, Greenland, suggest that life got its start before 3.8 Ga. Recently, isolated detrital zircons as old as 4.4 Ga from localities in Western Australia have provided evidence that the Hadean was shorter than previously thought. While extreme opinions exist on both sides, a moderate position is that small continent-like land masses existed by 4.4 Ga and oceans hospitable to life existed at 4.2 Ga. In this interpretation, the Hadean was over by 4.2 Ga and possibly earlier. Life may have emerged quickly in the post-Hadean oceans—there is no direct evidence. If life did exist before 3.8 Ga, it was subjected to intense meteorite bombardment and possibly extinction.

The age of the Earth has been debated for centuries. Accurate estimates based on the radioactive decay of uranium to lead were first made with improved mass spectrometers following the Second World War. In the second article of this issue, "The Origin of the Earth – What's New?," Alex Halliday describes the precise chronology of events that has been established from the many isotope systems now available for determining the age of rocks. A traditional type of geochronometer uses parent–daughter isotope pairs with long half-lives for radioactive decay (i.e. $^{238}U \rightarrow ^{206}Pb$, half life of 4468 million years, Myr); this method has established the geological time scale in years before the present. A more recently exploited technique uses isotope pairs with short half-lives (i.e. 26Al→26Mg, half-life of 0.73 Myr), in which the parent isotope is now extinct; ages are measured relative to the formation of the solar system. Thus, we can count forward or backward. This situation leads to the irony that the occurrence of some events is well established during the first 50 million years, but the next 500 Myr period from which no known rocks have been preserved (sometimes called the geological "Dark Ages") is poorly understood, and the following 4000 Myr of Earth history are increasingly well documented. No single date can be assigned for the formation of the Earth, which accreted over a period of time from

swirling dust, rocks, and planetesimals (~1 km diameter bodies) in the solar nebula. However, the age of the solar system is known with extreme precision from studies of meteorites to be 4.567 Ga. The Earth grew rapidly after this start, and calculations show that it attained most of its mass within 10 million years. Other events within the first 50 Myr include the gravitational settling of iron to form the metal core of the Earth and the formation of the Moon when a planet struck the Earth—a terrestrial impact so violent that both bodies melted and partly vaporized.

The Earth continues to grow today, with 106–107 kg of meteorites added each year, including thousands that are fist-size or larger. This is a trace contribution to the mass of the Earth and largely goes unnoticed. In "Impact Processes on the Early Earth," Christian Koeberl portrays a far more violent early history of meteorite impacts and cratering. During accretion, the mass flux of material added to the Earth was extreme, but it dropped five to ten orders of magnitude during the first 100 Myr and has continued to decline ever since, except during a resurgence called the Late Heavy Bombardment at ~3.85 Ga. While no evidence of impacts during the first billion years is preserved on Earth, largely due to tectonic reworking, nearly 200 younger impacts are known, the oldest being the Vredefort Dome (2023 Ma) in South Africa. Only the ancient surfaces of the Moon and Mars have provided evidence of the earlier ravages.

The early atmosphere was thick, hot, and poisonous. Carbonic "greenhouse" gases prevailed, and before 2.3 Ga, levels of oxygen were too low to sustain aerobic life. In "Earth's Earliest Atmosphere," Kevin Zahnle traces the evolution from a Hadean atmosphere to more clement conditions that could nourish life. Early on, not only were the oceans vaporized by energetic impacts, so was the silicate surface of the Earth, followed by magmatic "rain" and "hailstones" of rock! However, the faint young Sun was 30% weaker than today, and cooling of Earth's surface was surprisingly rapid. Calculations suggest that post-Hadean temperatures subsided enough to precipitate steam as ocean water and, depending on poorly known levels of atmospheric insulation, to freeze water. Zahnle argues that the cool early Earth was actually a snowball with pools of water localized around geothermal vents or at impact sites. These changing conditions could have spurred the emergence of life.

Rocks at least 2.5 billion years old are found on every continent and are relatively common. However, most of the more ancient rocks have been destroyed or reworked beyond recognition by the Earth's tectonic processes. Only a handful of small localities are known to be older than 3.6 Ga. In "Antiquity of the Oceans and Continents," Allen Nutman reviews studies of these key areas, with emphasis on Isua, Greenland, the most diverse early terrane. In spite of moderate metamorphism, ~3.8 Ga rocks from Isua preserve evidence of plate tectonics, oceans, and perhaps life.

The only known older rocks are gneisses from Acasta, Canada, at 4.0 Ga. The only record of rocks older than this comes from isolated crystals of the mineral zircon from Western Australia that is as old as 4.4 Ga.

Zircons are exceptionally robust and retentive. The most ancient grains were removed from unknown parent rocks, transported as wind-blown dust and river mud, and deposited as detrital grains in sedimentary rocks. They carry chemical clues to their origin in the form of mineral inclusions, trace elements, growth zoning, and isotopes of

uranium, lead, oxygen, and hafnium. The study of these tiny time capsules requires the use of advanced instruments, including large secondary-ion mass spectrometers called ion microprobes, which measure the age (U–Pb) and oxygen isotope ratio in 10–20 μm domains. These characteristics can be correlated with other features in the 100–300 μm crystals. The existence of zircons as old as 4.4 Ga indicates that small amounts of granitic (sensu lato) proto-continent existed at that time. Without such buoyant crust, the zircon-bearing rocks would have sunk into the mantle and been destroyed. High oxygen isotope ratios in some early zircons suggest that liquid water oceans existed at 4.2 Ga. Evidence from Ti and Hf has been interpreted as indicating even earlier granites and that plate tectonics was operating, but the uniqueness of these interpretations is debated (Harrison et al. 2006; Valley et al. 2006).

The oldest-known fossil evidence for life, along with appropriate geochemical signatures, is found in ~3.5 Ga rocks, but this does not necessarily date the emergence of life. If the first life was at 3.5 Ga, that would suggest that it was delayed by 700 million years after the end of the Hadean. In "The First Billion Years: When Did Life Emerge?" William Schopf discusses the evidence for life found in rocks from the oldest localities. It appears that one-celled organisms were diversified, flourishing, and relatively common by ~3.5 Ga, indicating that the first life came much earlier. If the carbon isotope record in metamorphic rocks is correct, then the emergence of life was before 3.8 Ga. In fact, all of the essential ingredients for life were assembled on Earth as soon as near-surface waters cooled enough for DNA to be stable (~4.2 Ga). If life emerged on Earth or was delivered from space at this time, its main challenge would have been possible annihilation during the Late Heavy Bombardment (~3.85 Ga). Survival chances would have been enhanced if primitive microbes, like archaea, were capable of subsisting underground in the absence of sunlight. Alternatively, the earliest life on Earth may have evolved and become extinct many times before 3.85 Ga, and the present inhabitants of Earth may be descended from the first continuously successful life, but not the first life.

Taken together, these articles summarize a new understanding of the first one billion years of history of our planet and portray its evolution from highly energetic to clement. This is a rapidly emerging field of study that will aid interpretation of other planets and the origin of life, inside the solar system and possibly beyond.

New words and expressions

Hadean 冥古代　　detrital zircon 岩屑的锆石　　chronology 年代学
Archaea 太古代的　　geochronometer 地质年代计　　isotope 同位素
secondary-ion mass spectrometers 二次离子质谱仪　　uranium 铀
ion microprobe 离子探针　　lead 铅　　hafnium 铪
Ma 百万年(地质年代单位)　Ga =10 亿年(地质年代单位)
planetesimals 星子的, 星子组成的, 星子假说 小行星体
nebula 星云, 云翳　　annihilation 歼灭, 灭绝

II. Water, Melting, and the Deep Earth H_2O Cycle

Marc M. Hirschmann

INTRODUCTION

Cycling of H_2O between the mantle and the exosphere and between different reservoirs in the mantle is one of the critical processes governing Earth's geodynamical and geochemical evolution. However, the distribution and inventory of H_2O in the mantle remains uncertain and the mechanisms of transport between principal mantle reservoirs are not fully known. The total H_2O stored in Earth's mantle is poorly constrained, with estimates from about a quarter the mass of H_2O in the world's oceans to ~ 4 ocean masses (Ringwood 1975, Ahrens 1989, Jambon & Zimmermann 1990, Bolfan-Casanova et al. 2003). The distribution of H_2O in the mantle may be heterogeneous on a large scale, but apart from the upper mantle, there is considerable ambiguity regarding concentrations of H_2O in Earth's interior. Narrowing uncertainties in the inventory of H_2O in the different portions of the mantle remains among the chief challenges to understanding the global deep H_2O cycle. In the mantle H_2O may be stored in solid minerals or present as hydrous fluids or melts. Such hydrous fluids are critical agents of mass transfer and mantle differentiation. Cycling of elements between the mantle and the crust is influenced by the loci of melting in basalt source regions (**Figure 1**), which in turn are strongly affected by H_2O (Hirth & Kohlstedt 1996, Ito et al. 1999). Further, deep hydrous melting may play a principal role in redistributing elements between different large-scale reservoirs in the mantle (Bercovici & Karato 2003). The strong effect of H_2O on mantle creep strength (Hirth & Kohlstedt 1995, Mei & Kohlstedt 2000) means that the distribution of H_2O within a planet governs the style of convective flow in its interior, and this may be the chief factor influencing the distinct tectonic behavior of Earth and Venus (Mackwell et al. 1998, Richards et al. 2002). This distribution is a product of both the capacity of mantle phases to store H_2O and the fluxes of H_2O between regions of the mantle and between the mantle, the crust, and the exosphere.

Here we consider several topics related to the deep water cycle. We begin by reviewing constraints from basalt geochemistry on H_2O storage in the mantle. We then discuss recent developments from mineral physics and experimental petrology that constrainH_2O storage capacities in the upper mantle, transition zone, and lower mantle and discuss the consequences for the loci of melting and cycling of H_2O between mantle reservoirs. The review emphasizes the role of melting in modulating fluxes of H_2Oin the mantle, with particular attention to melting in the source regions of oceanic basalts, above the 410 km discontinuity, and in the lower mantle.

The key role for H_2O melting of the mantle was first recognized for subducting oceanic lithosphere and processes associated with dynamics of the mantle beneath arc volcanoes (e.g., Green & Ringwood 1967). This topic has been investigated intensively for nearly 40 years and has been the subject of recent reviews (Poli & Schmidt 2002, Eiler 2003, van Keken&King 2005). Consequently, the present chapter focuses chiefly

on the role of H_2O in other regions of the mantle.

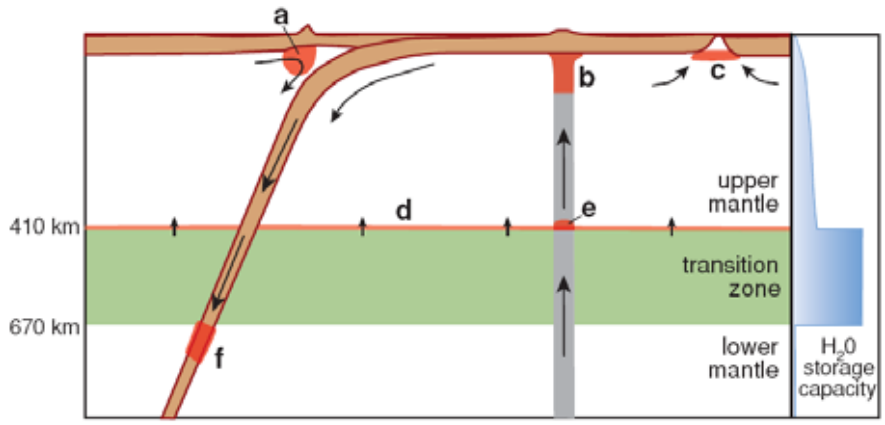

Figure 1

Possible regions of hydrous melting in the mantle. Shaded regions show possible loci of hydrous melting in the mantle discussed in the text (*a*) melting in mantle wedges above subduction zones, (*b*) melting in the deeper portions of OIB source regions, (*c*) melting in the deeper portions of MORB source regions, (*d*) pervasive melting above the 410 km discontinuity, (*e*) melting in localized upwellings above the 410 km discontinuity, and (*f*) melting in downwellings below the 670 km discontinuity. Residues from melting in regions (*a*), (*b*), and (*d*) have been suggested as possible principal sources of the upper mantle (see text). At right is a schematic view of the H_2O storage capacity as a function of depth.

PRELUDE: H_2O STORAGE IN THE MANTLE VIEWED BASALT GEOCHEMISTRY

A primary consideration regarding operation of the deep Earth H_2O cycle is the inventory of H_2Ostorage in the mantle. Geochemical arguments based onK_2O/H_2O ratios of basalts suggest that the bulk silicate Earth (BSE) contains 500–1900 ppm H_2O(Jambon&Zimmermann 1990), corresponding to 1.5–5.5 oceans in the mantle. More detailed studies of basalt geochemistry reveal the H_2O concentration may vary significantly between mantle reservoirs (e.g., Dixon et al. 2002). H_2O behaves incompatibly during mantle melting and subsequent shallow differentiation processes (e.g., Michael 1995). Effects of degassing can be avoided by studying submarine eruptions of sufficient depth or melt inclusions. Comparison of H_2O to other incompatible elements shows that its behavior during melting is similar to La, Ce, or K, with perhaps the greatest similarity to Ce (Michael 1988, Jambon & Zimmermann 1990, Michael 1995). Consequently, establishment of H_2O/Ce ratios of undegassed basalts, combined with estimates of Ce in mantle source regions (e.g., Salters & Stracke 2004,Workman & Hart 2005) provides the best estimates of H_2O in basalt source regions. Combined with independent determinations of the behavior of Ce during partial melting of peridotite (i.e., Salters & Longhi 1999), similar behavior of H_2O and Ce indicates that the H_2O partition coefficient during mantle melting H_2O, Dperidotite/melt H_2O , is ~ 0.007–0.009. Studies of H_2O and H_2O/Ce in normal mid-ocean ridge basalts (MORBs) indicate the mantle beneath normal ridges has between 50 and 200 ppm H_2O (Michael 1988, 1995; Danyushevsky et al. 2000; Dixon et al. 2002; Saal et al. 2002; Simons et al. 2002). The source of normal MORB is a global layer in the shallow oceanic mantle and is characterized by distinctive depletions

in certain trace elements and radiogenic isotopes (e.g., Zindler & Hart 1986) and so this 50–200 ppm is likely to be a good estimate of the concentration of H_2O in the upper mantle, apart from mantle wedges above subduction zones and those regions that have been affected by plumes. Geochemical studies of oceanic island basalts (OIBs) and oceanic plateaux (Dixon et al. 1997, 2002; Jamtveit et al. 2001; Hauri 2002; Nichols et al. 2002;Wallace et al. 2002; Seaman et al. 2004) and from MORB that incorporate a component from a plume-associated source (Dixon et al. 2002) suggest H_2O concentrations in plume sources ranging from 300–1000 ppm. These studies suggest that the mantle has one or more H_2O-rich reservoirs that feed plume-related magmatism. These reservoirs are conventionally assumed to reside in the lower mantle or the core-mantle boundary region.

Importantly, the MORB source is the driest of geochemically identifiable mantle reservoirs; all others are more enriched in H_2O (Dixon et al. 2002). Estimates for the volume of this relatively dry, depleted source of MORB range from 25% to 90% (Jacobsen &Wasserburg 1979, Hofmann 1997). The OIB source region, consisting of a range of reservoirs with distinct isotopic characteristics reflecting contributions from recycled and/or primitive materials (Hofmann 1997), could constitute much of the rest of the mantle. It is possible that parts of the mantle are not sampled by any oceanic basalts, but these are generally thought to represent a relatively small volume restricted to the core-mantle boundary (Blichert-Toft & Albarede 1997, Rudnick et al. 2000, Tolstikhin & Hofmann 2005).

If all major reservoirs in the mantle are sampled by oceanic basalts, the total mass of H_2O in the mantle depends on the relative proportions of MORB and OIB sources, with larger OIB reservoirs corresponding to more total H_2O storage. For 25%MORBsource with 50–200 ppm H_2O and 75% OIB source with 300–1000 ppm H_2O, the total H_2O in the mantle amounts to 1.5 ± 0.8 oceans. For an intermediate case with 75% MORB source and 25% OIB source (Kellogg et al. 1999), total H_2O storage equals 0.7 ± 0.4 oceans, and for an extreme dry case, with 90% MORB source and 10% OIB source, storage equals 0.5 ± 0.3 oceans.

A critical question is which, if any, of these of possibilities is consistent with mineral physics and petrologic constraints on H_2O storage capacity as a function of depth in the mantle. H_2O storage capacity is the maximum H_2O that a mineral or an assemblage of minerals can accommodate in their structures without saturating in an H_2O-rich fluid or hydrous melt (Hirschmann et al. 2005). In the following sections, we discuss some of the constraints on H_2O storage capacities in the principal layers of the mantle and their relationship to melting, H_2O fluxes between the layers, and between the mantle and the exosphere.

THE UPPER MANTLE

The H_2O storage capacity of olivine, the principal phase in the upper mantle, is perhaps

better characterized than that of any other mantle phase, although important uncertainties remain. The pioneering work of Kohlstedt & coworkers (Bai & Kohlstedt 1993, Kohlstedt et al. 1996) showed that olivine storage capacity increases greatly with H_2O fugacity and consequently with depth, from ~ 25 ppm at 10 km depth to ~ 1300 ppm at 410 km. Recent reevaluation of analytical techniques for H_2O analysis in olivine have lead to an upward revision of these values by a factor of 3 to 3.5 (Bell et al. 2003, Koga et al. 2003), meaning that the storage capacity of olivine at 410 km approaches 0.4 wt.%. This accords with more recent detailed infrared and secondary ion mass spectrometry (SIMS) analyses (Chen et al. 2002, Smyth et al. 2005). Storage capacities for pyroxenes and garnet are less well-known than for olivine, but at low pressure the capacity of pyroxene to store H_2O greatly exceeds that of olivine, meaning that it is the principal host of H_2O in the upper mantle (Aubaud et al. 2004). This may or may not be the case at greater depth, but if it is, the storage capacity at the base of the upper mantle could approach 1 wt.% (Hirschmann et al. 2005). Importantly, these storage capacities derive from relatively low temperatures (1000◦C–1250◦C), and may not be directly applicable to the actual mantle geotherm.We return to this subject below.

Where the upper mantle H_2O content exceeds the local storage capacity, hydrous melting results. This occurs in basalt source regions of the shallow mantle (Hirth & Kohlstedt 1996, Asimow et al. 2004, Aubaud et al. 2004) and possibly at the bottom of the upper mantle, if H_2O-rich material is advected from the transition zone (Young et al. 1993, Kawamoto et al. 1996, Litasov&Ohtani 2002, Bercovici&Karato 2003). Melting at the top of the upper mantle governs the flux of H_2O from the mantle to the exosphere and likely forms dry stiff lithospheric plates (Hirth & Kohlstedt 1996). Melting at the bottom of the upper mantle may govern the flux of H_2O to the upper mantle (Bercovici & Karato 2003), thereby modulating its geochemical and rheological properties.

EFFECT OF H_2O ON MELTING IN BASALT SOURCE REGIONS

The now classic work of Hirth & Kohlstedt (1996) first alerted the Earth science community to the profound influence of small amounts of H_2Oon the melting regime beneath oceanic ridges. Subsequent studies have considered the relevance of this critical process to lateral variations in mantle viscosity (Dixon et al. 2004), the origin of the low velocity zone (Karato&Jung 1998), the pattern of flow and melting beneath ridge-centered hotspots (Ito et al. 1999), the major and trace element compositions of MORB (Asimow & Langmuir 2003, Asimow et al. 2004), and the survival of continental and oceanic lithosphere in convecting mantle (Lee et al. 2005). However, as Hirth & Kohlstedt (1996) made clear, their parameterization of the effect of small amounts of H_2Oon mantle melting was an estimate based on incomplete experimental constraints.

A major difficulty is that direct experimental determinations of the effect of small amounts of H_2O on mantle melting are infeasible. The tiny amounts of melt (<0.1 wt.%) generated by 50–200 ppm H_2O are difficult or impossible to detect in the laboratory.

Instead, we must rely on theoretical or empirical models to understand this critical process. These in turn depend on experimental measurements of the behavior of H_2O in partial melts and residual minerals found in peridotite. One key constraint is partitioning of H_2O between mantle minerals and partial melts, $D^{mineral/melt}_{H2O}$. For example, estimates of $D^{mineral/melt}_{H2O}$ provide constraints for several proposed dehydration melting models (Hirth & Kohlstedt 1996, Katz et al. 2003, Aubaud et al. 2004).

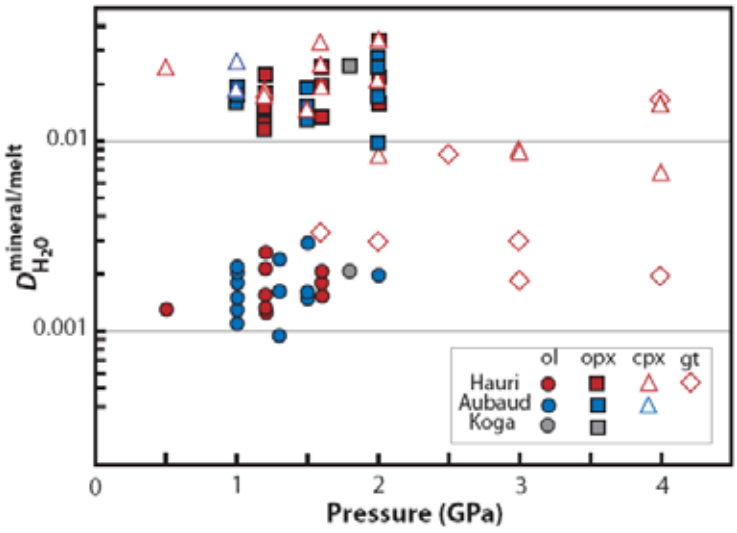

Figure 2

Experimental partition coefficients of H_2O for nominally anhydrous upper mantle minerals (olivine, pyroxenes, and garnet) coexisting with basaltic melt determined by SIMS as a function of pressure. Data from Koga et al. (2003), Aubaud et al. (2004), Hauri et al. (2004), and C. Aubaud (unpublished data).

A widely applied technique for H_2O determinations in nominally anhydrous minerals is Fourier transform infrared spectroscopy (FTIR), but FTIR requires doubly polished thin sections of minerals, and in most cases is best performed with crystallographically oriented grain mounts. Thus, it is poorly suited for experimental determinations of $D^{mineral/melt}_{H2O}$ between nominally anhydrous minerals (NAMs) and silicate melts, as such experiments generally yield small (<100 μm) crystals embedded in glass. New developments of low blank SIMS techniques that allow accurate analyses of H_2O in NAMs down to 5–10 ppm by weight (Koga et al. 2003) have made experimental determinations more tractable. SIMS determinations of $D^{mineral/melt}_{H2O}$ for mantle minerals (Koga et al. 2003, Aubaud et al. 2004, , Hauri et al. 2004; C. Aubaud, unpublished data) are restricted to a relatively narrow pressure range (chiefly 1–2 GPa), and are summarized in Figure 2. They show $D^{olivine/melt}_{H2O}$ partition coefficients in the range of 0.001–0.003 and $D^{pyroxene/melt}_{H2O}$ partition coefficients in the range of 0.007–0.03. The small amount of data available for garnet shows values of $D^{garnet/melt}_{H2O}$ ranging from 0.002–0.015. Together, these yield $D^{peridotite/melt}_{H2O}$ of ~ 0.009, a value similar to that inferred from H_2O/Ce ratios of basalts (see above). Available data do not suggest a strong pressure dependence to the mineral partition coefficients (Figure 2), although they do indicate an important role for pyroxene composition (Aubaud et al. 2004). However, it is important to emphasize that hydrous melting of peridotite beneath oceanic ridges and oceanic islands likely commences at 50–200 km, corresponding to pressures of 2–7 GPa, and so, experimental determinations at higher pressure are needed.

A thermodynamic approach to the effect of H_2O on mantle melting begins with examining H_2O storage capacity in olivine and H_2O solubility in silicate melts. Bai & Kohlstedt (1993) and Kohlstedt et al. (1996) showed that the concentration of H_2O in olivine is proportional to the fugacity of H_2O, $C^{olivine/melt}_{H2O} \sim f^{n}_{H2O}$, where the exponent, n, is close to unity. On the other hand, at low concentrations, H_2O dissolves into melts as hydroxyl according to the reactions similar to O^{2-} (melt) + H_2O (fluid) = $2OH^{-}$(melt), and the concentration of hydroxyl in silicate liquids is given by $C^{melt}_{OH^-} \sim f^{0.5}_{H2O}$.With increasing concentration of H_2O in silicate melts, molecular H_2O species become important, and the effective value of n may be greater than 0.5. The combined relations for hydrous incorporation in olivine and melt yield an expression for the olivine/melt partition coefficient: $C^{olivine/melt}_{H2O} \sim f^{0.5}_{H2O}$. This relationship predicts that the mineral/melt partition coefficient of H_2O should diminish severely with increasing H_2O fugacity, which in turn implies that $D^{olivine/melt}_{H2O}$ should increase with increasing pressure and with increasing H_2O concentration (Asimow et al. 2004). However, such pressure or concentration dependences are not evident in the limited available experimental partition coefficient data (Figure 2). They may emerge as the number of partition coefficient determinations increase.

A thermodynamic approach to the effect of H_2O on mantle melting begins with examining H_2O storage capacity in olivine and H_2O solubility in silicate melts. Bai & Kohlstedt (1993) and Kohlstedt et al. (1996) showed that the concentration of H_2O in olivine is proportional to the fugacity of H_2O, $C^{olivine/melt}_{H2O} \sim f^{n}_{H2O}$, where the exponent, n, is close to unity. On the other hand, at low concentrations, H_2O dissolves into melts as hydroxyl according to the reactions similar to O^{2-} (melt) + H_2O (fluid) = $2OH^{-}$(melt), and the concentration of hydroxyl in silicate liquids is given by $C^{melt}_{OH-} \sim f^{0.5}_{H2O}$.With increasing concentration of H_2O in silicate melts, molecular H_2O species become important, and the effective value of n may be greater than 0.5. The combined relations for hydrous incorporation in olivine and melt yield an expression for the olivine/melt partition coefficient: $C^{olivine/melt}_{H2O} \sim f^{0.5}_{H2O}$. This relationship predicts that the mineral/melt partition coefficient of H_2O should diminish severely with increasing H_2O fugacity, which in turn implies that $D^{olivine/melt}_{H2O}$ should increase with increasing pressure and with increasing H_2O concentration (Asimow et al. 2004). However, such pressure or concentration dependences are not evident in the limited available experimental partition coefficient data (Figure 2). They may emerge as the number of partition coefficient determinations increase.

The thermodynamic properties of H_2O in silicate melts become highly uncertain above~ 1 GPa owing to the lack of available experimental constraints. Constraints are derived chiefly from solubility experiments at pressures<1 GPa. Based on equilibrium between water-rich fluid and hydrous silicate melt, H_2O (melt) ⇔ H_2O (fluid), the chemical potential of H_2O in the melt is equated with the known chemical potential of pure H_2Oat a given temperature and pressure (e.g., Pitzer&Sterner 1994, Zhang&Duan 2005) and related to the concentration of H_2O (or its dissociated species) measured in quenched glass. However, with increasing pressure, increasing mutual solubility of

H_2O and silicate make such measurements intractable. Concentrations of H_2O cannot be measured accurately in the silicate liquids, which do not quench to glasses. At the same time, there are few constraints on the thermodynamic properties of soluterich H_2O-rich fluids. Thus, the data and models needed to extend thermodynamic treatment of the behavior of H_2O during mantle melting to higher pressures are largely unavailable. Significant breakthroughs in experimental measurements and thermodynamic models are needed.

One rather crude way to estimate the effect of H_2O on mantle melting is the cryoscopic calculation of Hirschmann et al. (1999), whereby the freezing point depression caused by ideal dilutent (in this case H_2O) can be approximated as

$$T = \frac{1}{\frac{1}{T^{\text{fusion}}_{\text{peridotite}}} - \frac{R}{\Delta H^{\text{fusion}}_{\text{peridotite}}} \ln\left(1 - X^{\text{melt}}_{OH^-}\right)}, \quad (1)$$

where $T^{\text{fusion}}_{\text{peridotite}}$ and $\Delta H^{\text{fusion}}_{\text{peridotite}}$ are the melting temperature and enthalpy of fusion of dry peridotite at the pressure of interest, R is the gas constant, and $X^{\text{melt}}_{OH^-}$ is the mole fraction of OH^- in the melt. For a given concentration of H_2O in a bulk rock and a given mineral/melt partition coefficient, the concentration of H_2O in the melt $C^{\text{melt}}_{H_2O}$ can be estimated at the solidus ($F = 0$) from the batch melting relation:

$$C^{\text{melt}}_{H_2O} = \frac{C^{\text{bulk}}_{H_2O}}{D^{\text{peridotite/melt}}_{H_2O}}. \quad (2)$$

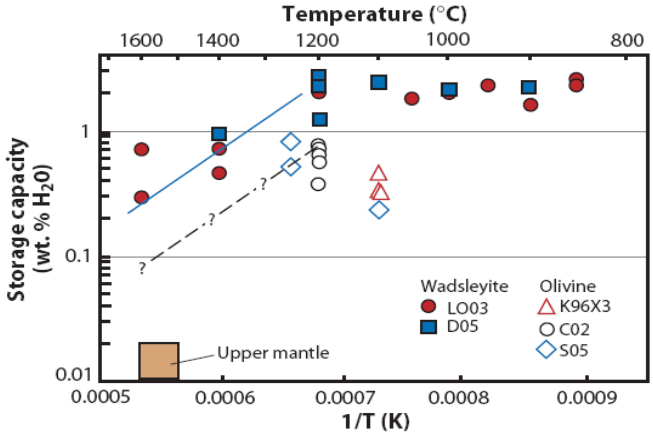

Figure 3

Storage capacities of Mg_2SiO_4 wadsleyite (LO03: Litasov & Ohtani 2003, D05: Demouchy et al. 2005) and $(Mg, Fe)_2SiO_4$ olivine at the base of the upper mantle (12–13.5 GPa) (K96X3: Kohlstedt et al. 1996, with the revisions of Bell et al. 2003; C02: Chen et al. 2002; S05: Smyth et al. 2005) as a function of reciprocal temperature. The storage capacity of Mg_2SiO_4 wadsleyite decreases at high temperature (*solid line*) owing to increased dissolved silicate component in the fluid and consequent diminished H_2O fugacity. An analogous effect, although not yet observed experimentally, is expected for olivine (*dashed line*). However, the expected olivine trend still predicts storage capacities at the base of the upper mantle that are much higher than the 50–200 ppm characteristic of the upper mantle (see text for references), making it unlikely that upper mantle H_2O can be the residue of partial melting at 410 km.

Although this approach has the virtue of simplicity, it also has a number of limitations.

First, it neglects possible nonideal mixing between H_2O and silicate. Second, it requires an estimate for the enthalpy of fusion under the conditions of interest. Third, because it employs a simple partition coefficient, it is limited to melt H_2O concentrations where Henry's Law applies. Fourth, it requires a method for converting the mass fractions given by the batch melting equation to mole fractions of the hydrous melt component. This requires assumptions (or independent constraints) about hydrogen speciation (hydroxyl in the case of Equation 1) as well as estimates of the molecular weights of the silicate components. Following Burnham (1975), Aubaud et al. (2004) calculated these by assuming an eight-oxygen molecular unit.

Aubaud et al. (2004) applied the cryoscopic approach to estimation of the effect of H_2O on melting of peridotite. The results (Figure 3) indicate that 50–200 ppm H_2O incite incipient melting just 5–20 km below the dry solidus. This result contrasts with the calculation of Asimow et al. (2004) and Hirth&Kohlstedt (1996), who surmised 20 and 55 km solidus depression for 50 and 150 ppm H_2O, respectively.

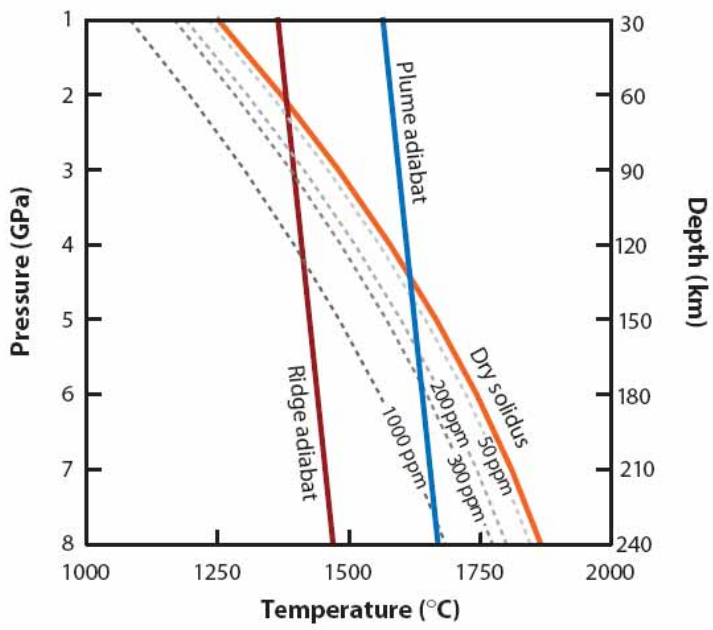

Figure 4

Calculated effect of H_2O on the solidus of peridotite based on the cryoscopic approximation (after Aubaud et al. 2004). Isopleths of mantle H_2O content show the depth of initial melting along ridge or plume adiabats. Compared with melting of dry mantle beneath ridges and plumes, 50–200 ppm H_2O in the MORB source incites melting 5–20 km deeper, and 300–1000 ppm H_2O in the OIB source incites melting 40–100 km deeper.

The difference lies partly in upward revisions in the storage capacity of olivine (e.g., Bell et al. 2003), but also may be related to differences in assumed enthalpy of fusion or to differences in assumed structure and speciation of silicate liquids. At this time, it is difficult to be sure which calculation is more accurate, although other things being equal, the upwardly revised storage capacity in olivine would certainly diminish the estimates of Hirth & Kohlstedt (1996) and Asimow & Langmuir (2003).

When applied to melting beneath oceanic islands, the cryoscopic calculation predicts initial melting depths (170–230 km) that are significantly deeper than the intersection with the dry peridotite solidus (Figure 4). The much larger effect of H_2O

beneath OIB source regions compared to MORB sources is partly due to the greater H_2O concentrations (300–1000 ppm) and partly due to the steepening P-T slope of the solidus with increasing pressure.

One way to test the strengths and weaknesses of the cryoscopic estimation of solidus depression is to compare the calculation to well-established melting relationships in simple systems. In Figure 5, we apply the cryoscopic model to the effect of melting on pure forsterite at 1 GPa, as established from experiments (Hodges 1974). The model predicts the location of the solidus to within ±40◦C for 0–10 wt.%H_2O, but diverges at greater H_2O contents. This divergence is the natural consequence of the breaking down of Henry's Law. If similar limits apply to partial melting of peridotite, then the calculation should yield reasonable results for bulk compositions with up to 1000 ppm H_2O (i.e., for a bulk partition coefficient of 0.009, up to 11 wt.% H_2O in the near-solidus partial melt).

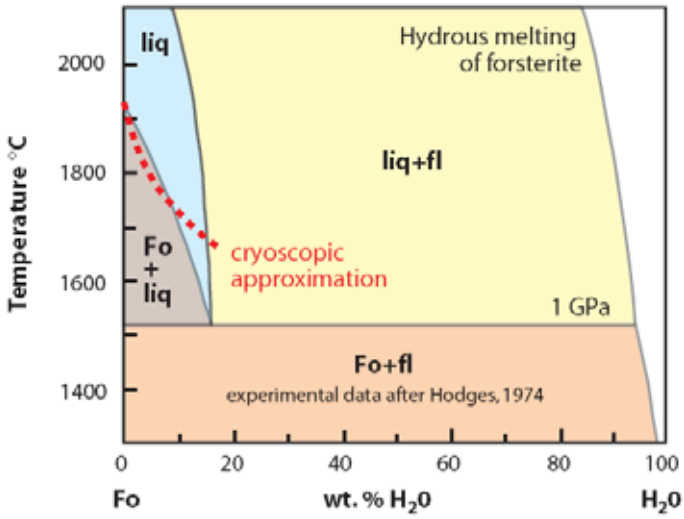

Figure 5

Calculated effect of H_2O on melting of forsterite at 1 GPa compared with experimental results (Hodges 1974). Calculation using Equation 1 follows the methodology of Hirschmann et al. (1999) and Aubaud et al. (2004), using $\Delta H^{fusion}_{Mg_2SiO_4}$ estimated from statistical mechanics (3nR, where n is the number atoms per molecule and R is the gas constant), Mg_2SiO_4 liquid mole fractions calculated assuming an eight-oxygen molecular unit, and complete dissociation of H_2O to OH^-. The reasonable match between the observed and calculated freezing point depression suggests that the cryoscopic approach may give plausible estimates for the effect of H_2O on melting temperatures of peridotite (**Figure 4**), so long as the concentration of H_2O in the melt is not large.

H_2O IN THE UPPER MANTLE: A RESIDUE OF MELTING

In previous sections, it was noted that the upper mantle has 50–200 ppm H_2O, but what is the origin of that H_2O? It is well established that the MORB source has been geochemically depleted by ancient melting events (e.g., Salters & Stracke 2004, Workman & Hart 2005). Is the H_2O left behind the residue of a melting event? Or does the H_2O content of the upper mantle reflect mixing between dry, highly depleted residues and more hydrous primitive or recycled materials? Are the observed H_2O contents clues to the geodynamic history of the upper mantle?

Recently, three intriguing hypotheses have been advanced, all arguing that MORB-source upper mantle is a residue of partial melting, but each postulating that the melting occurs in distinct geodynamic environments (Figure 1). Bercovici & Karato (2003) suggested that partial melting of H_2O-rich material advected through the 410 km discontinuity leaves a residue with an H_2O content similar to the average upper mantle. Hirth & Kohlstedt (2003) proposed that typical upper mantle H_2O inventories may be the residues of hydrous melting in mantle wedges beneath volcanic arcs. Finally, Phipps Morgan & Morgan (1999) considered formation of the depleted MORB-source mantle as the residue of partial melting of plumes beneath oceanic islands. Phipps Morgan & Morgan (1999) did not focus specifically on H_2O, but their hypothesis also implies that the H_2O content of the MORB-source mantle is a residue of partial melting.

Origin of the upper mantle by partial melting at 410 km. Several have suggested that the large contrast in storage capacity between the transition zone and overlying upper mantle could lead to partial melting of material advecting across the 410 km discontinuity (Young et al. 1993, Kawamoto et al. 1996, Litasov & Ohtani 2002, Bercovici & Karato 2003). Irrespective of the fate of the melts, the residues of such melting are delivered to the upper mantle with the H_2O retained by upper mantle minerals (olivine, pyroxene, garnet) residual from the melting process. These residues are therefore a potential source of upper mantle H_2O. Although this process could be restricted to localized H_2O-rich upwellings, Bercovici & Karato (2003) propose that it may be a global feature, with melts eventually returning to an H_2O-rich transition zone and the residual solids advected to upper mantle, constituting a principal source of the depleted upper mantle, and hence, of its H_2O.

A critical prediction of this hypothesis is that the H_2O concentration in the upper mantle should be equal to that left in the upper mantle residue following melting. Hirschmann et al. (2005) argued that the storage capacity of the upper mantle assemblage at 400 km may be >0.4 wt.%. The reasoning is based partly on revised estimates of the storage capacity of olivine at these pressures, which is in the range of 0.3–0.8 wt.% at 1100◦C–1250◦C (Kohlstedt et al. 1996, with the revisions of Bell et al. 2003; Chen et al. 2002; Smyth et al. 2005) and partly on observations of high storage capacity in pyroxene relative to that of olivine (Aubaud et al. 2004, Hauri et al. 2004). However, it is important to note that the upper mantle above the 410 km discontinuity is ~ 1500◦C–1600◦C (Ita & Stixrude 1993), and the storage capacities of upper mantle minerals are not well-known at these temperatures.

At high pressure, the fluid coexisting with silicate minerals becomes increasingly melt-like with increasing temperature (Stalder et al. 2001, Mibe et al. 2002, Kessel et al. 2005). The consequent reduction in H_2O concentration reduces the H_2O fugacity, which in turn diminishes the storage capacity of the mineral. This effect has now been well documented for Mg_2SiO_4 wadsleyite (Litasov & Ohtani 2003, Demouchy et al. 2005), for which storage capacity declines significantly above 1200◦C (Figure 5). (Note however, that a similar trend is not evident from existing studies of $(Mg, Fe)_2SiO_4$

wadsleyite, Hirschmann et al. 2005). An analogous drop should occur in olivine, and if a similar temperature dependence applies, then the storage capacity of the upper mantle above 410 km could be as low as 1000 ppm. This should be considered a minimum, as it is likely that the pyroxene portion of the upper mantle assemblage sequesters considerably more than 1000 ppm (Inoue et al. 1995, Stalder & Skogby 2002, Yamada et al. 2004). But the principal point is that it is unlikely that the decrease in storage capacity with temperature owing to changing fluid composition is much more pronounced for olivine than for wadsleyite, as would be required if the 50–200 ppm H_2O in the upper mantle were a residue of partial melting at 410 km.

Do hydrous melts percolate up or down? A key consideration regarding the fate of hydrous melts in the deep mantle and their impact on mantle geochemistry and dynamics is whether the partial melts are negatively or positively buoyant relative to their matrix. For example, the transition zone water filter model of Bercovici & Karato (2003) postulates that hydrous melts formed above 410 km are negatively buoyant and consequently drain back into the transition zone. In addition, as discussed below, hydrous melts forming from regions downwelling in the lower mantle may conceivably percolate upward and get deposited in the transition zone or percolate downward and collect in the core-mantle boundary region.

Experimental investigations show that there are several circumstances in which liquidus minerals float in the liquids from which they precipitate, and these may be applicable to formation of perched layers during magma ocean solidification (e.g., Ohtani et al. 1995, Agee 1998). However, the buoyancy of small-degree hydrous melts relative to their matrices is a different question. Olivine floatation occurs near the liquidus of peridotite at pressures (~ 12–13 GPa) corresponding to the bottom of the upper mantle. Other conditions being equal, hydrous silicate melts should be less dense than anhydrous equivalents owing to the relatively high partial molar volume of H_2O (Ochs & Lange 1999). But melts have greater thermal expansion than minerals, so the increase in volume associated with H_2O may be offset by the much lower temperatures relevant to near-solidus conditions compared with near-liquidus conditions (Bercovici & Karato 2003). However, negative melt buoyancy near the solidus requires that the hydrous melt be denser than peridotite, rather than simply olivine. Peridotite at 400 km consists of ~ 28% garnet and 12% clinopyroxene (Irifune & Isshiki 1998), and both of these phases are denser than olivine. Also, near-solidus minerals are richer in Fe and therefore denser than those precipitating at the liquidus. This may be partially offset if the near-solidus melts are also Fe-rich, but melt density may be influenced more by the low-density hydrous component. Thus, at the base of the upper mantle, olivine with 95%–97% forsterite content and density of ~ 3300 kg/m3 can float in liquid peridotite (e.g., Agee 1998), but it seems much less likely that small-degree hydrous partial melts will be negatively buoyant in solid peridotite with a density of ~ 3600 kg/m3 (Ita & Stixrude 1993).

Origin of the upper mantle in mantle wedges. The volume of mantle processed

through mantle wedges through geologic time is poorly constrained. It depends on the dynamics of the wedges, which in turn depends on the rheological effects of temperature and fluid on peridotite as well as the amount of viscous coupling between the wedges and the underlying subducting slabs, all topics of current research (Hirth & Kohlstedt 2003, Kelemen et al. 2003, van Keken 2003). However, if mantle wedges are dragged down to the deeper mantle at rates comparable to those of their underlying subducting slabs, and if the region dragged down has a thickness in the range of 10–100 km, then fossil mantle wedges comprise a volume within an order of magnitude of subducted lithosphere and are therefore a sizable fraction of the mantle.

Hydrous magmas characteristic of arc volcanism typically have 2–6 wt.% H_2O (Roggensack et al. 1997, Cervantes & Wallace 2002, Grove et al. 2002). If the bulk peridotite/melt partition coefficient (0.009) determined by Aubaud et al. (2004) applies, then the peridotite residua in equilibrium with these melts have between 180 and 540ppm H_2O.Thelower of these values lies at the high end of estimates (50–200 ppm) for the H_2O concentrations in the MORB source, so if relatively dry (2 wt.%) magmas are volumetrically predominant in arcs, then wedge residues could be plausible sources of upper mantle. Alternatively, the 2% H_2O may be representative of the last magmas to separate from mantle wedges during a dynamic melting process.

Although the residues of mantle wedges may potentially have H_2O contents similar to the typical upper mantle, it seems unlikely that the geochemical features of mantle wedges are appropriate for the principal sources of MORB. Arc lavas have characteristic enrichments in fluid-mobile elements such as Sr and Pb, as well as depletions in high field-strength elements, which are not present in MORB, and at least some of these features are likely to be present in mantle wedge residues. In particular, compared with typical upper mantle or other oceanic basalt source regions that have H_2O/Ce in the range of 100–250 (e.g., Dixon et al. 2002), H_2O/Ce ratios of arc and back-arc lavas are extraordinarily high (350–10,000) (Sobolev&Danyushevsky 1994, Danyushevsky et al. 1995, Cervantes &Wallace 2002,Walker et al. 2003), reflecting addition of anH_2O-rich fluid from the slab. Melting residua in mantle wedges should not have H_2O/Ce ratios much smaller than their associated arc magmas, as it is very unlikely that $D^{peridotite/melt}_{Ce} \approx D^{peridotite/melt}_{H2O}$ in arc lava source regions. Thus, residual mantle wedges are not likely candidates for the principal source of the MORB-source mantle.

Upper mantle as the residue of melting of plumes. The two-stage melting hypothesis of Phipps Morgan & Morgan (1999) posits that OIB represent the low-melting-point fraction of heterogeneous upwelling plumes, and that MORB represent further melting of the residua of these sources. In this model, the average amount of melting beneath oceanic islands is 1%–4%, although the more fertile portions of the plume melt more and the more refractory portions melt less. For a typical plume source composition of 300–750 ppm (Dixon et al. 1997, Dixon & Clague 2001, Jamtveit et al. 2001, Hauri 2002), which produces residues with 50–400 ppm

H_2O, assuming batch melting and $D^{perid/melt}_{H2O} = 0.009$ (Aubaud et al. 2004, Hauri et al. 2004), although residue concentrations could be far less if melt were removed dynamically during melting. Under some circumstances, this model may be capable of producing a residual upper mantle with appropriate H_2O concentrations. On the other hand, this model may have certain shortcomings. For example, it requires that the flux of plume material into the upper mantle be similar to the flux of upper mantle material participating in lithosphere formation at ridges. The buoyancy flux of plumes is generally thought to be an order of magnitude less than the flux of mantle beneath ridges (Sleep 1990). Also, the Phipps Morgan&Morgan hypothesis may not be consistent with either the absolute or relative abundances in noble gases in MORB (Porcelli & Ballentine 2002).

In summary, it is unlikely that the H_2O in the upper mantle represents the residue of melting at 410 km because residues should have much more H_2O than observed in the upper mantle. Partial melting in mantle wedges above subduction zones may yield residues with H_2O concentrations appropriate for the average upper mantle, but other geochemical characteristics of mantle wedges are distinct. The residue of small-degree melting of plumes may have H_2O concentrations that match the average upper mantle, but other problems with this model remain to be resolved.

If none of the melting residue models for the origin of the upper mantle are wholly satisfactory, then it seems likely that the upper mantle and its characteristic H_2O concentration derives from a blend of sources, some of which may include contributions from localized melting at 410 km, fossil mantle wedges, or recently depleted plumes. Models based on trace element and isotope geochemistry generally consider the upper mantle to be a mixture of depleted and enriched components, where the former are chiefly the residue from earlier crust formation events and the latter derive either from recycled crust or more primitive sources (Hofmann 1997, Helffrich &Wood 2001, Campbell 2002).

THE TRANSITION ZONE: A POTENTIAL SPONGE

Subduction may deliver considerable H_2O to great depths, potentially allowing formation of a H_2O-enriched shell in the transition zone, where minerals have very large storage capacities. Although, subducted oceanic crust is effectively dehydrated before it reaches the transition zone (Schmidt & Poli 1998), subducted peridotite may potentially carry H_2O deeper than 410 km, particularly along cold geotherms that might apply to fast subduction of older oceanic lithosphere (Kawamoto et al. 1996, Schmidt & Poli 1998, Ohtani et al. 2004). The amount of H_2O delivered to such depths is unknown, in part because the hydration state of subducting peridotite is poorly constrained. Penetration of H_2O to considerable depths in the slab may occur along extensional faults at the so-called outer-rise, which results from plate flexure outboard of oceanic trenches (Peacock 2001, R¨ upke et al. 2004). Intermediate depth earthquakes in the mantle section of downgoing slabs may signal partial dehydration of hydrous phases

(Peacock 2001), but phase equilibria studies suggest that the downgoing peridotite remains partly hydrated so long as the geotherm remains colder than the so-called choke point that marks the intersection of the highpressure limit of serpentine stability with the low-pressure limit of Phase A, a dense hydrous magnesium silicate (DHMS) (Schmidt & Poli 1998). Although many slabs subduct along geotherms that are hotter than this, some proportion of subduction is likely to have occurred along cold geotherms through much or all of the past few billion years. Even in the absence of subduction of H_2O to great depth, the transition zone may have gained significant H_2O by upward percolation of hydrous melts from the lower mantle. This possibility is considered in greater detail in The Lower Mantle: A Desert in the Mantle, below.

Water delivered to the transition zone can be stored in high-pressure polymorphs of olivine, wadsleyite, and ringwoodite, which have enormous H_2O storage capacities (Kohlstedt et al. 1996, Smyth et al. 1997). It has been speculated that the H_2O sequestered in the transition zone rivals that of the oceans (Smyth 1994, Bercovici & Karato 2003, Ohtani et al. 2004). On the other hand, the upper boundary of the transition zone at 410 km is not known to be a barrier to convection, so advection across the 410 km discontinuity, either from focused plume-like upwellings or from pervasive passive upwellings required by formation and subduction of oceanic lithosphere, may remix subducted H_2O into the upper mantle, preventing enhanced concentration of H_2O in the transition zone (Richard et al. 2002). But such mixing on a large scale does not preclude significant heterogeneities on a smaller scale, and subduction of H_2O may produce local H_2O-rich regions in the transition zone.

Because large H_2O contents have considerable effects on physical properties, H_2O enrichments in the transition zone may be detectable from observable geophysical properties, such as electrical conductivity or the character of seismic discontinuities (Wood 1995). Huang et al. (2005) presented electrical conductivity measurements of wadsleyite and ringwoodite for a range of H_2O contents and compared those with conductivity profiles for the mantle in the north Pacific. They found closest matches for wadsleyite and ringwoodite with ~ 0.1–0.2 wt.% H_2O, which supports the view of a modestly hydrous transition zone. However, conductivity of these phases depends critically on oxygen fugacity. Huang et al. (2005) posit that the conditions in the transition zone are quite oxidizing and near the Ni-NiO buffer. Because their experiments were reduced owing to the presence of metallic Mo, they corrected their experimental measurements accordingly. However, contemporary estimates of transition zone oxidation state indicate fugacities similar to or more reduced than Mo-MoO_2 (O'Neill et al. 1993, McCammon 2005). Consequently, the experiments of Huang et al. (2005) may be applied to the transition zone without correction for oxygen fugacity, in which case the H_2O concentrations required to match the conductivity measurements would not be markedly different from typical upper mantle values.

THE LOWER MANTLE: A DESERT IN THE MANTLE

The H_2O storage capacity of the lower mantle is uncertain and a source of recent controversy. Murakami et al. (2002) and Litasov et al. (2003) reported storage capacities in the principal lower mantle phases, aluminous Mg-perovskite, Ca-perovskite, and ferropericlase, of up to 1500–2000, 4000, and 2000 ppm H_2O, respectively, at 1400◦C–1650◦C and 25.5 GPa. If accurate, these results suggest that the lower mantle can store more than three oceans of H_2O. In contrast, Bolfan-Casanova et al. (2000, 2002, 2003) found <10 ppm H_2O in aluminous Mg-perovskite at 1400◦C and 24 GPa, and <20 ppm in ferropericlase at 1200◦C and 25 GPa. If these low storage capacities are correct, the lower mantle is much drier than the upper mantle, and the total storage of H_2O in the lower mantle is limited to less than 3% of Earth's oceans. These large discrepancies highlight the great challenges in performing hydrous experiments at extreme pressures and temperatures and in analyzing small concentrations of H_2O in the tiny crystals that result from these experiments. Erroneously high H_2O concentrations may be inferred if microbeam or spectroscopic analyses inadvertently include contributions from H_2O-rich fluids, grain boundaries, or hydrous mineral inclusions. Indeed, Bolfan-Casanova et al. (2003) observed inclusions of superhydrous B [$Mg_{10}Si_3O_{14}(OH)_4$] in their aluminous perovskite crystals and suggested that similar inclusions may have lead Murakami et al. (2002) and Litasov et al. (2003) to infer exaggerated H_2O concentrations for perovskite. K. Litasov (personal communication, December, 2004) affirms that hydrous inclusions likely biased the analyses of Litasov et al. (2003). These experiments were conducted up to 1600◦C, more than 200◦C hotter than the likely stability limit of superhydrous B (see below), but the inclusions could have persisted metastably.

Experimental difficulties may also lead to spurious low storage capacities if fluid saturation was not maintained during the experiment. Bolfan-Casanova et al. (2000, 2003) do not report that perovskite in their experiments coexisted with an H_2O-rich fluid. On the other hand, they do report fluid present at the end of their ferropericlase storage capacity experiments (Bolfan-Casanova et al. 2002). This suggests that the storage capacity for perovskite is also small, as experiments with coexisting ferropericlase and (Mg, Fe)SiO_3 perovskite show that H_2Opartitions preferentially into the former (Bolfan-Casanova et al. 2000). Unfortunately similar data for coexisting ferropericlase and aluminous perovskite are not available. Also, the observation of superhydrous B inclusions in nearly dry aluminous perovskite in the experiments of Bolfan-Casanova et al. (2003) suggests high H_2O fugacities. More experiments are required to resolve the discrepancies in this crucial problem, but for the present time, it seems more likely that the storage capacities of the principal lower mantle minerals, aluminous magnesioperovskite and ferropericlase, are very low.

Much of what is known about H_2O storage capacity in lower mantle minerals (e.g., Bolfan-Casanova et al. 2000, 2002, 2003; Murakami et al. 2002; Litasov et al. 2003) is restricted to pressures corresponding to the uppermost~ 100 km of the lower mantle. Thus, it has been speculated that storage capacities properties could increase at much higher pressures (Navrotsky 1999). On the other hand, defect concentrations in

aluminous, Fe^{3+}-bearing Mg-perovskite in equilibrium with ferropericlase diminish with increasing pressure (Walter et al. 2004), which may actually reduce the tolerance of this predominant lower mantle phase for H substitution.

If the principal minerals in the lower mantle have very low H_2O storage capacities, must the lower mantle be nearly bone-dry? One possibility is that significant H_2O is stored in minor hydrous phases, such as DHMS. Both phase superhydrous B and phase D [$MgSi_2O_4(OH)_2$] are stable under lower mantle conditions. PhaseDdecomposes to anhydrous phases and fluid above 44 GPa (Shieh et al. 1998). Superhydrous B decomposes above ~ 30 GPa (Shieh et al. 1998, Ohtani et al. 2003). However, the stability of these phases (in natural peridotite as well as in simple synthetic systems) appears to be restricted to temperatures below 1300◦C–1400◦C (Frost 1999, Litasov & Ohtani 2003, Kawamoto et al. 2004). This is ~ 300◦C cooler than the mantle adiabat, meaning that these DHMS are likely restricted to colder geotherms characteristic of recent subduction.

It has been suggested that DHMS may be stabilized to high temperature by fluorine substitution (Williams & Hemley 2001, Larsen et al. 2003). However, the abundance of fluorine in the mantle is quite low—the depleted upper mantle contains 11±5 ppm (Salters & Stracke 2004). Therefore, to accommodate a significant fraction of the 50–200 ppm H_2O of typical depleted mantle, DHMS with F/H_2O ratios <0.2 would have to be stable along the convective geotherm. Less depleted regions of the mantle may have more F; i.e., similar to the 25 ppm F in the bulk silicate Earth (McDonough & Sun 1995), but these are presumably much richer in H_2O (Dixon et al. 2002).

Melting in Lower Mantle Downwellings

Possibly very low H_2O storage capacities in lower mantle assemblages suggest that rocks descending from the transition zone may release partial melt. Material brought downward through the 670 km discontinuity in subducted slabs may sequester H_2O in superhydrous B or phase D, perhaps to depths of 1250 km (44 GPa) (Shieh et al. 1998, Ohtani et al. 2003), but if no other host of H_2O is stabilized at those depths, then melting must result.

Importantly, hydrous melting in lower mantle downwellings may not require transport of H_2O-rich phases in subducting slabs. The apparently very low storage capacity of nominally anhydrous lower mantle minerals indicates that they are incapable of storing even normal upper mantle concentrations of H_2O (50–200 ppm). Stability of DHMS along subduction geotherms through the upper mantle and transition zone may enhance delivery of H_2O to the lower mantle and thereby enhance melting. Stability of DHMS below the 670 km discontinuity may delay this melting to significant depths, but the likelihood of such melting is not predicated on the transport of H_2O-rich material through the 670 km discontinuity. If such melts are positively buoyant, they percolate upward through the 670 km discontinuity and solidify in the

transition zone (Bercovici & Karato 2003). If such melts are negatively buoyant, they may percolate downward, concentrating H_2O in the core-mantle boundary (D") region. Such H_2O could be effectively trapped in D", as plumes ascending from D" through the lower mantle may not be able to carry H_2O either in minerals (which could have low storage capacity) or in melts (which may be negatively buoyant and percolate back down to D"). This rather speculative combination of circumstances could imply that a large portion of Earth's H_2O remains sequestered at D", perhaps accounting for the widespread partially molten regions there (Revenaugh & Meyer 1997).

The possibility of very low H_2O storage capacity in the lower mantle raises some critical questions about geochemical reservoirs in the mantle. First, where do the H_2O-rich sources of OIB reside? A dry lower mantle may imply that plumes are chiefly shallow features, originating from the upper mantle or the transition zone. Alternatively, they may originate deep and entrain H_2O-rich components from more shallow depths, but this would raise the question as to why the dry lower-mantle component of plumes is not evident in OIB geochemistry.

Additionally, if the lower mantle really is drier than the MORB source, H_2O in the upper mantle must come from shallow sources, such as subducting slabs, mantle wedges, or the transition zone, and theMORBsource would be a mixture made in the upper mantle, but potentially incorporating lower mantle components. This may be difficult to reconcile with geochemical models that require the MORB geochemical reservoir to constitute the majority of the mantle.

CONCLUDING REMARKS: FREEBOARD AND THE MANTLE H_2O BUDGET

One of the key features of Earth's surface geology is the near constancy of continental freeboard through geologic time (Wise 1974). Although the geologic record provides extensive evidence for eustatic sea level changes, these changes over the past 3 Ga [and perhaps 4 Ga (Eriksson 1999)] are small relative to the total volume of the ocean basins; i.e., a few hundred meters compared to an average ocean depth of 4 km. This reflects a steady budget of surface H_2O relative to H_2O stored in the mantle, which is particularly interesting considering that the H_2O in the oceans could be a modest fraction of the total terrestrial H_2O budget. Evidently, the proportions of H_2O stored in Earth's interior and present in the exosphere were established relatively early in Earth history and have changed little since. On the other hand, the rate of subduction of H_2O is sufficient to desiccate the oceans in 1–2 Ga (Ito et al. 1983). Thus, the fluxes of H_2O into the deep mantle must be closely in balance with degassing at ridges and regassing at subduction zones (McGovern & Schubert 1989, R¨ upke et al. 2004).

The near-steady-state balance between exosphere and interior H_2O reservoirs means that it is unlikely that there have been large changes in the hydration state of the various subreservoirs of the mantle. For example, the rate of H_2O degassing at ridges is

proportional to the average H_2O concentration in the upper mantle, which therefore must not have experienced significant secular increases or decreases over the past 3–4 Ga. If, as suggested above, the upper mantle is a heavily processed reservoir derived from a complex blend of sources, then its recipe must have been established early in Earth history. Also, if one posits that the transition zone contains a significant fraction of Earth's H_2O, then either this must have been true through much of Earth history or the hydration of the transition zone has occurred at the expense of other mantle reservoirs and not from the oceans. In this context, models in which the transition zone becomes hydrous by subduction of H_2O (Kawamoto et al. 1996, Bercovici & Karato 2003, Ohtani et al. 2004) may be problematic. Finally, if there is significant H_2O storage in the D" boundary region, it was either established early in Earth history or has occurred at the expense of other mantle reservoirs. Thus, if negatively buoyant hydrous melts in the lower mantle deliverH_2O from subducted slabs to D", either the total flux over the past 3–4 Ga has been negligible or upward fluxes in plumes recycle that H_2O back to the shallower mantle.

New words and expressions

BSE: bulk silicate Earth,

MORB: mid-ocean ridge basalts

OIB: oceanic island basalt

SIMS: secondary ion mass spectrometry

FTIR: Fourier transform infrared spectroscopy

NAM: nominally anhydrous mineral

DHMS: dense hydrous magnesium silicate

III. Sedimentation and Sedimentary rock

Burial and Accumulation of Sediment

A bar in a river may exist for only a few months, from the high-water stage that formed it to the next one, which destroys it. Of all the sediments that are made day by day, only a small fraction is preserved to record geologic time. Sedimentary environments differ in the probability of preservation, and the differences are related, in part, to the rapidity of sedimentation and burial. In the deep sea, deposits accumulate at very low rates, only a few millimeters every thousand years. Once deposited, many of these deep-water sediments are unlikely to be eroded and redeposited because bottom currents strong enough for the task are unusual. That is why deep-sea pelagic sediments offer the most complete historical records of organic evolution, of temperature changes during ice ages, and of other geologic patterns.

In contrast, the shallower waters of the continental shelf deposit sediments at much higher rates—many centimeters every thousand years, on the average. These sediments are buried faster, but waves and currents are much more likely to scour the bottom and rework it mechanically. Furthermore, the higher the sedimentary accumulation, the more likely the waves are to rework and redistribute it; so that there is a level of sedimentation above which accumulation stops. The importance of **subsidence** of the crust is that it allows sedimentation to continue and the accumulation to build up. This process is one of slight positive feedback, in the sense that tectonic subsidence is reinforced by the sedimentary load. The further sedimentation proceeds, the greater the weight on the crust, because each layer of sediment is much denser than the sea water it replaces. The effect of the weight is to push down the crust (see the discussion of isostasy in Chapter 18), allowing still more sedimentation to proceed. This process does not continue endlessly, nor can it account for all of the subsidence shown by most continental shelves. The reason for this is that, as subsidence continues and sediments are laid down, more and more of the crust at that point consists of sediments; because these are much less dense than average crust, the effect of additional sediment tends to become negligible. Subsidence initiated by tectonic mechanisms and enhanced by sediment weight is evidenced by great thickness of sediments.

Geosynclines

The first to theorize about the origin of sediments of great thicknesses was the State Geologist of New York in the middle of the nineteenth century, James Hall. As a result of extensive field mapping of sediments of the Appalachian Mountains, correlating them stratigraphically, and measuring their thicknesses, Hall concluded that there were many tens of thousands of feet of sediment that had accumulated in a long, relatively narrow trough bordering the continent. He envisioned it as having been something like a large downfold in the surface that steadily accumulated sediment as it subsided. Later, the name geosyncline was applied to the sediment-filled trough.

For the past century the term has undergone a long and complex series of redefinitions tied to changing notions of the origins of the supposed troughs, mechanisms of subsidence in relation to sedimentation, the evolution of igneous and metamorphic rocks in the geosynclines, and of the nature of the structural deformation

of these thick sedimentary sections. In the 1930s geosynclines were subdivided into an inner belt closer to the continental platform, the **miogeosyncline**, and an outer belt closer to an ocean basin, the **eugeosyncline**. The miogeosyn cline was characterized by lesser thicknesses of limestones, shales, and alluvial sandstones, though other kinds of sediment could also be found. The greater thicknesses of the sediments characteristic of the eugeosyncline included turbidite sandstones and shales, pelagic limestones and shales mixed with volcanic rocks, submarine basalts, and ash falls and flows.

An explosion of information on continental margins has accompanied the rise of plate-tectonic theory and its amplification in the last decade. This knowledge has led to new definitions, solidly based on geophysical and geological theory and observation of both oceans and continents. We now see geosynclines—there is still much of value in keeping the name—as thick, linear piles of sediments deposited along continental margins at present or former plate boundaries. The sandstones, shales, and carbonates of the continental shelves of passive continental margins accumulate to thicknesses of several kilometers as the edge of a rifted continent moves away from a mid-ocean ridge, cools, and subsides. These sediments correspond roughly to the miogeosyncline. The volcanics, turbidites, and forearc sediments of active continental margins are mixed with scraped-off pelagic sediments at subduction zones. This is the thicker sequence of the eugeosyncline. The two are brought together by a change in plate motions as a formerly passive continental margin converges with an active continental margin. When the two continents collide, their margins crumple first. Continental-rise turbidites of the one continent are mashed and faulted against the turbidites, volcanics, and pelagic sediments of the other. A collisional mountain chain is created of the thick sedimentary accumulations of the geosynclines.

Structurally deformed, metamorphosed, and injected with igneous intrusions, geosynclines were the forerunners of the major mountain chains of the world: the Alps, the Himalayas, the Appalachians, and the Western Cordilleran belt of the Americas. Only in the youngest mountain belts are the youngest geosynclinal beds preserved. As the mountains wear away, older and older sedimentary rocks are stripped off by erosion, in many places uncovering igneous and metamorphic rocks below. As geologic time goes on, the less likely it is that sediments uplifted into mountain ranges will survive erosion. In general, less and less sediment is found preserved as one moves back through the history of the Earth.

Continental platforms

Wide areas of the North American continent between the Appalachians and the Rocky Mountains have been relatively stable at least since the Cambrian Period began. Sedimentary accumulations are much thinner than in the geosynclines along the continental margins to the east and west, though there are a number of sedimentary basins, such as the Illinois Basin, that have received as much as 3000 meters (10,000 feet) of sediment. Deep-water formations are conspicuously absent, and many of the sediments are shallow carbonate-platform sediments and alluvial (or near-shore marine) sandstones and shales. Individual formations more or less continuous with those in geosynclines are much thinner on the platform. Subsidence has been much less on the platform, and no great mountain-making episodes have affected it. Conversely, erosion

has not eaten so deeply into the surface sedimentary rocks, so that Precambrian igneous and metamorphic rocks, the basement, are exposed in only a few places, such as in the Ozark Mountains or the Black Hills. The key feature of platform sedimentation is not so much in sediment types and the environments in which they formed, for there is almost complete overlap with geosynclines in that regard, but in the thinness of the accumulation, a response to only mild and intermittent subsidence. The origin of this mild subsidence in areas far from plate boundaries remains in doubt.

Subsidence, which influences environment and thickness of accumulation, is only one aspect of the tectonic control of sedimentation. Just as important is the indirect effect that tectonics has on the mineral composition of detrital sediments.

Tectonics and sediment composition

Contrasts in the mineral composition of clastic sediments can be startling. For example, compare two alluvial sandstones, and arkose and a quartz arenite. The arkose may contain as much as 30 to 40 percent feldspar, the remainder being quartz and other mineral and rock fragments. The grains are coarse and angular, showing little sign of abrasion. The quartz sandstone may be more than 95 percent quartz with no feldspar. The grains are well sorted and well rounded. The contrast is not in the environment, for both are alluvial, but in the source of the erosional debris. The source of the arkose is likely a granitic igneous rock; the rocks eroded to make the quartz sand might well have been another sandstone.

As shown in Chapter 5, feldspar decomposes chemically to clay minerals, given enough time for reaction with rain and soil water. In comparison, quartz is essentially unaffected by most chemical weathering. The survival of feldspar (and thus the quartz-feldspar ratio) in a weathering terrain depends upon the rate of chemical decomposition in relation to the rate of mechanical erosion. That ratio is related to the topography of the terrain- the more elevated and rugged the mountains, the more rigorous the mechanical erosion. In relating the detrital quartz-feldspar ratio to mountainous topography, we have tied it to the product of tectonic activity. Tectonically quiet low-lying areas not subject to severe mechanical erosion produce detritus low in feldspar; much of the feldspar decomposes to clay before it can be transported.

This tectonic background to sediment composition is a major basis for the reconstruction of ancient tectonic events, of high mountain chains that once shed abundant detritus and are now gone. For example, the Devonian sandstones of the Catskill Mountains of New York are a very thick accumulation of alluvial and deltaic sands that were transported from a great mountain range to the east, roughly parallel to the present low hills of the Taconic Mountains east of the Hudson River. The mineralogy of the sandstones, including the types of feldspar, mica, other silicates and rock fragments, indicates a mixture of granitic igneous and high-grade metamorphic source rocks that are typical products of the mountain-building that results from continent-continent plate collision. Continental drift reconstruction shows this mountain range to have been formed by the junction of eastern and western continents at the beginning of the assembly of Pangaea. A similar analysis has been applied to the Tertiary formations of the San Joaquin Valley of California. Their mineralogy indicates volcanic island arc source rocks. The stratigraphic and structural relationships, coupled

to plate tectonic reconstructions of the sea floor off California, show this sequence to have been formed in the fore-arc region of a subduction zone formed by convergence of a former oceanic plate, now entirely gone, and the North American plate. In Canada geologists have reconstructed fragmentary patterns of vanished mountain ranges, some more than three billion years old, from Precambrian sediments. In these ways, geologists have learned to use sediments as the hieroglyphics of geologic history.

Sediment into rock: diagenesis

Once a sediment is deposited and buried by other sediments, it is not immune to change. One has only to compare freshly deposited muds and sands with ancient shales and sandstones to see the obvious differences in hardness, cohesion, and porosity. The many processes that produce the changes in a rock's composition and texture after deposition are lumped together in the term **diagenesis**. Generally, they operate to harden the soft sediment into rock—that is, to **lithify** it. Diagenesis may also alter the mineral composition by dissolving some of the original minerals and precipitating new ones. The nature of oil, gas, and coal is almost completely the result of diagenesis of original sedimentary organic matter.

The major physical diagenetic change is **compaction**, a decrease in porosity caused by mineral grains being squeezed closer to each other by the weight of overlying sediment. Sands are fairly well packed as they are deposited, so they compact little, even when buried deeply. Newly deposited muds, however, are highly porous. They have a high water content when deposited, often well over 60 percent, and their layers thin drastically by compaction. The porosity equilibrium mixture of diverse minerals that have been brought together as detritus in the sediment (and, perhaps, mixed with some chemical precipitate from the environment, such as calcite). Thus, sedimentation may mix minerals from two very different kinds of igneous rocks, minerals that would have been incompatible under the original conditions of formation of the igneous rocks, such as a sodium-rich plagioclase feldspar from a granite and a calcium-rich one from a basalt. Diagenetic chemical changes tend to dissolve the calcium-rich feldspars and precipitate sodium-rich feldspars, thus moving the rock toward chemical equilibrium—namely, a more homogeneous plagioclase composition. A different example of lack of equilibrium is a grain of aragonite in a carbonate sediment which given enough time, tends to change to calcite, the form of calcium carbonate stable at low pressures and temperatures. This tendency to equilibrium results in many other chemical reactions between incompatible or unstable minerals, which result in the formation of new minerals and, thus, bring the rock to a new composition in equilibrium with its surroundings.

The second tendency is for a sediment to be buried more or less deeply in the crust. As a sediment is buried, it is subjected to increasingly high temperatures--on the average 1°C for each 30 meters (100 feet) of depth on the continents (see Chapter 13) - and high pressures - on the average about 1 atmosphere for each 4.4 meters of depth (1 pound per square inch for each foot of depth). As minerals and the surrounding groundwater in pore spaces are heated and put under greater pressure, they tend to react chemically to form new minerals. This process, when carried far enough, becomes metamorphism, in which the entire character of the rock alters. The boundary between

diagenesis and metamorphism is somewhat arbitrary, usually drawn at a temperature of about 300℃.

The list of specific diagenetic changes is great. Few of these are universal; which ones occur depends on the geologic situation. A few important examples include the change of unstable opaline silica shells of diatoms to quartz, the stable form; transformation of the unstable swelling clay, smectite, to a stable mica type of clay, illite; the precipitation of calcite and quartz as pore-filling cement in sandstones; and the alteration of volcanic glass to clay minerals and zeolite minerals. All of these changes must be understood before we can deduce the nature of the original sediment that was laid down, as we must do in order to interpret geologic history.

Diagenesis is of practical value, because recovery of oil, gas or groundwater from formations depends on the porosity of the rock, which may have been largely determined by diagenetic cementation. The formation of oil itself is a diagenetic process by which organic remains of many kinds of organisms are gradually altered to liquid petroleum or natural gas, or both. The transformation of peat through bituminous coal to anthracite is another diagenetic process. Understanding all of these processes better enables us to search more intelligently and successfully for new mineral and energy resources.

New words and expressions

sediment 沉积物
deposit 沉积（物），矿床
burial 埋葬，埋藏
trough 槽，海槽，地堑
miogeosyncline 冒地槽，副地槽
turbidite 浊积岩，浊流沉积
alluvial 冲积的
isostasy 地壳均衡
passive continental margin 被动大陆边缘
collid (v), collision (n.) 碰撞，冲击
arkose 长石砂岩
abrasion 磨蚀（作用）
mica 云母
hardness 硬度
porosity 孔隙度
lithify (v.) 石化，使…成石
aragonite 霰石（文石）
opal 蛋白石
smectite 蒙脱石
illite 伊利石
bituminous 沥青质，沥青的
deltaic 三角洲的
sedimentation 沉积作用
geosyncline 地向斜,地槽,沉降带
subsidence 下沉，沉降
downfold 向斜槽
eugeosyncline 优地槽，主地槽
topography 地形，地形学
pelagic 远洋的，深海的
continental shelf 大陆架
forearc 前弧
platform 台地，地台，平台
arenite 砂碎屑岩，砂岩
weathing 风化作用
hieroglyphics 象形文字
cohesion 粘性，凝结力
diagenesis 成岩作用
compaction 压实，致密
cementation 胶结作用
montmorillonite 蒙脱石
bituminous coal 烟煤
peat 泥炭
anthractite 无烟煤

IV. Plate Tectonics: a Review and Summary

Plate tectonics is the conceptual framework of this book, and we have already introduced the basic ideas in earlier chapters. In this chapter we draw together and review the diverse lines of evidence that support the theory of plate tectonics, using primarily illustrations you have already seen in earlier chapters. We will begin a discussion of rock associations and orogeny within the framework of plate tectonics. The fragmentation of Pangaea since the Jurassic will be reviewed together with some speculation about continental drift and extinct plates in the pre-Jurassic. The chapter will close with some brief remarks on the driving mechanism of plate tectonics, but the reader shouldn't expect more than vague speculations, for the subject is just beginning to receive serious study.

The mosaic of plates

According to the theory of plate tectonics, the rigid plates whose outlines are shown in the illustration inside the back cover slide over a partially molten, plastic asthenosphere in the general directions shown. According to the relative motions of adjacent plates, we can define three kinds of plate boundaries: (1) zones of divergence or spreading, typically ocean ridges; (2) fracture zones, or transform faults; and (3) zones of convergence.

Zones of divergence are boundaries along which plates separate. In the process of plate separation, partially molten mantle material upwells along linear ocean ridges, and new lithosphere is created along the trailing edges of the diverging plates. Such boundaries are characterized by active basaltic volcanism, shallow-focus earthquakes caused by tensile (stretching) stresses, and high rates of heat flow. The outpouring of magma along ocean ridges and the building of the oceanic lithosphere are volumetrically the most significant form of volcanism. Figures 1-13, 1-17, 13-8, 13-11, 15-19, and 15-37 emphasize the different aspects of divergence zones.

Typically transform faults are boundaries along which plates slide past one another, with neither creation nor destruction of lithosphere. Sometimes marked by scarps, transform faults are characterized by shallow-focus earthquakes with horizontal slips. Occasionally there occur "leaky" transforms, in which some volcanism and slight plate separation accompanies the transform. Examples are in Figures 1-17,17-13, and 18-17.

Zones of convergence are boundaries along which the leading edge of one plate overrides another, the overridden plate being subducted, or thrust into the mantle, where lithosphere is resorbed. The thrusting mechanisms that operate along these collision boundaries tend to produce deep-sea trenches, shallow- and deep- focus earthquakes, adjacent mountain ranges of folded rocks, and both basaltic and andesitic volcanism. Convergence zones are illustrated in Figures 1-16, 13-8, 13-11, and 15-37.

Each plate is bounded by some combination of these three kinds of zones, as can be seen on the inside of the back cover. For example, the Nazca Plate in the Pacific is bounded on three sides by zones of divergence, along which new lithosphere forms, and

on one side by the Peru-Chili trench, where lithosphere is consumed. Continental margins may or may not coincide with plate boundaries. If they do, the continents tend to remain “afloat” because continental plates are thicker and too buoyant to be readily subducted. Where two plates with continents at their leading edges converge, the crust thickens to form great mountain ranges like the Himalayas.

The global sum of plate creation and consumption is approximately zero The Earth would otherwise change size in order to accommodate the new sea floor, and this doesn’s seem to be happening. Instead, the plates form and disappear and change in size and shape as they evolve.

The structure and evolution of plates

Figure 19-1 depicts some of the structural details of a rigid lithospheric slab, a plate, from its region of generation at a ridge axis to its region of subduction, where it is resorbed. Both oceanic and continental crusts cap the plate; the continent, embedded in the moving plate, is carried along passively by it. Thus, in a real sense, continental drift is simply a consequence of plate movements. Underneath is the plastic, partially molten asthenosphere-source of the raw materials that build new lithosphere. Once heated and partially melted, subducted lithosphere becomes a source of magma, which rises to feed the overlying volcanic chain. A generalized heat-flow profile. Shows a large amount of heat emerging along the ridge axis, less from the older, cooled slab, and more from the volcanic chain of the subduction zone and the marginal basin behind it, where a small region of secondary spreading occurs.

Geophysicists have made theoretical studies and computer models of the evolution of a plate, from its creation out of hot rising matter at ocean ridges, through its spreading and cooling phase to its subduction, with reheating, melting and final resorption in the underlying mantle. The models help explain some important geological and geophysical observations: the major features of the ocean floors, the variation of heat flow from the sea floor, the occurrence of volcanism at plate margins, and the location and mechanism of earthquakes in the subducted slab.

Ocean depths increase with age, t, of the sea floor in a remarkably simple manner. For the first 80 million years the data fit a curve in which ocean depth increases as $\sqrt{t}$. This is precisely the relationship predicted if a plate cools and contracts as it spreads. Beyond 80 million years ocean depths tend to flatten out, compared to the theoretical cooling curve, as would be expected if a small amount of heat is flowing into the plate from the underlying hot asthenosphere.

When a cold plate is subducted, it remains cooler than the surrounding hot mantle for about 12 million years, only gradually warming as it penetrates more deeply. Slow-moving plates heat up and are assimilated at shallower depths, perhaps 400 kilometers (250 miles), than fast-moving plates, which can penetrate to about 700 kilometers before heating to the point of assimilation. The process of subduction involves very large forces, and in a general way these forces must be responsible for the deep-focus earthquakes that occur only in downgoing plates. The sudden failures associated with earthquakes take place until the plates become so warm that stress is relieved by slow plastic deformation rather than by faulting. This seems to be the likely explanation for the fact that no earthquakes occur below 700 kilometers.

Rates of plate motion

The velocities of moving plates are measured by dating ocean-floor magnetic anomalies (using the time scale of magnetic stratigraphy) and dividing the age of each anomaly into the distance between it and the ridge axis. The procedure was outlined graphically in the preceding chapter.

The worldwide pattern of sea-floor spreading is being worked out by using a combination of magnetic, seismic, and bathymetric data. The charts used earlier (Fig. 18-21 and inside back cover) map the world's zones of spreading, subduction, and fracture; their geographic locations were obtained from the positions of ocean ridges, deep-sea trenches, earthquake epicenters, and other indications of activity. On the basis of spreading rates determined from magnetic data, isochrones (contours that connect points of the same age) were drawn to show the age of the sea floor in millions of years. The distance from a ridge axis to a 50-million-year isochron, for example, indicates the extent of new ocean floor created in that period. In Figure 18-21, note the closer spacing of the isochrones in the Atlantic than in the Pacific, where the spreading rate is higher. Because the fracture zones offset the isochrones, the age of the sea floor changes abruptly as one crosses a fault. A summary of the rates and directions of plate motions, measured in centimeters per year relative velocity, is given in Figure 19-4. The largest spreading velocity, 18.3 centimeters per year, occurs between the Pacific and Nazca plates.

The first results announced by the Deep Sea Drilling Project represented a great triumph for the magneticians who worked out spreading rates. The goal of this joint project of major oceanographic institutions and the National Science Foundation was to drill through the sediments of the sea floor at many places in the world's oceans. Studying the sedimentary cores makes it possible to work out the history of the ocean basin directly, in contrast to the indirect methods of magnetic anomalies. Since sedimentation begins as soon as an ocean forms, the age of the oldest sediments in the core, those closest to the basaltic bedrock, dates the ocean floor at that spot. The age is obtained from the fossils found in the cores. No sediments older than about 150 million years have been found, attesting to the "youth" of the sea floor. The sediments become older with increasing distance from mid-ocean ridges, confirming the prediction of the sea-floor-spreading hypothesis. Figure 19-5 is a plot of the ages determined from drill cores from the Atlantic and Pacific Oceans against ages predicted from magnetic data. It is remarkable how closely the experimental points approach the straight line with slope of one, which represents perfect agreement. This agreement clinches the concept of magnetic stratigraphy and the hypothesis of sea-floor spreading.

As an interesting aside, we have included a photograph of the drilling vessel Glomar Challenger. It is 400 feet long, and amidships it carries a drilling derrick 140 feet high. The only ship of its kind in the world, it can lower drill pipe several kilometers to the sea floor and drill thousands of meters into the sediments and underlying volcanic rock. For the ship to accomplish such a feat required a technological breakthrough. A means had to be devised to hold the ship stationary, regardless of current, wind, or waves, during drilling. Otherwise, the drill pipe would break off. The problem was solved by developing a positioning device that uses sound

waves from acoustic beacons planted on the sea floor. Any change in the ship's position is sensed by a computer that monitors the time of arrival of the sound pulses. The same computer controls bow and stern side thrusters and the ship's main propulsion to keep the vessel on station. The Glomar Challenger was the answer to those who said when lunar exploration started, "Better to explore the ocean's bottom than the backside of the Moon." We ended up doing both.

Geometry of plate motion

If the individual plates behave as rigid bodies, as seems a reasonable first assumption, several interesting and useful geometric consequences follow. By "rigid" we simply mean that the distances among three points on the same plate-say, New York, Miami, and Bermuda-do not change, no matter how the plate moves. But the distance between New York and Lisbon, of course, increases because the two cities are on different plates that are being separated along a narrow zone of spreading on the mid-Atlantic ridge. Listed here are some geometric principles, mostly self-evident, that govern the sliding of plates on a planet.

1. Along transform-fault boundaries, no overlap, buckling, or separation occurs; the two plates merely slide past each other without changing the surface area. Look for a transform fault if you want to deduce the direction of plate motions because the orientation of the fault is the direction of relative sliding of two plates, as Figures 1-17 and 17-13 show. Surface area obviously changes at zones of convergence or divergence where plates are subducted or created. The plates can move perpendicularly or obliquely to the trend of convergent boundaries, which are therefore not as reliable indicators of directions of movement as transform faults or divergence zones.

2. Magnetic anomaly stripes and isochrones are roughly parallel and are symmetrical with respect to the ridge axis along which they were created. Look at Figure 18-17 to see why this must be so. Since each magnetic strip or isochron marks the edge of an earlier plate margin, isochrones that are of the same age but on opposite sides of an ocean ridge can be brought together to show the positions of the plates and the configuration of the continents as they were in that earlier time. By this means we can reconstruct, for example, the opening of the Atlantic Ocean, as shown in Figure 19-7[1].

3. The point at which three plates meet is called **a triple junction.** Figure 19-8 shows an example of a point at which a spreading zone, a subduction zone, and a transform fault meet. If the relative motion between two pairs of plates is known, we can solve for the third by using a simple equation (see Box 19-1).

The point where the Pacific, Cocos, and Nazca plates meet (see inside back cover) is an actual triple junction. Three spreading zones meet at this junction, as shown in the enlarged view in Figure 19-9. The unknown motion, found by vector addition, was that between the Nazca and Pacific plates, the motions between the Pacific-Cocos and Cocos-Nazca plates having been worked out from transform faults and magnetic-anomaly stripes. The arrows show the resultant plate movements. Note also

[1] The Great Pyramid of Egypt is aimed slightly east of true north. Did the ancient Egyptian astronomers make a mistake in orienting the pyramid 40 centuries ago? Probably not. Over this period of time Africa drifted enough to rotate the pyramid out of alignment with true north.

how the isochron bend to become parallel to the spreading centers, where they originated, and how they are offset by the transform faults. The spacing of the isochron reflects the spreading rates, which are largest for the Pacific-Nazca plates and least for the Cocos-Nazca plates.

Up to this point we have considered plates sliding on a plane. Although much can be learned about plate motions by making this simplification, plates actually move on the Earth's spherical surface. Box 19-2 explains how plate movements on a sphere can be described. With the application of these geometric principles to find spreading directions and magnetic anomalies to deduce spreading rates, the relative motions of the lithospheric plates are being worked out worldwide. Some results have already been pictured in Figures 18-21 and 19-4. However, geophysicists are searching for ways to measure the absolute motions of individual plates rather than their motions relative to each other. If the hot spots discussed in Chapter 15 turn out to be fixed in the mantle below plates, then the string of extinct volcanoes trailing from the hot spot would record the movement of individual plates as they glide over the mantle. This is currently a subject of active research.

Sea-floor spreading and continental drift: rethinking Earth history

One of us (F. P.) once helped write a paper dealing with the permanence of ocean basins. If he were allowed to expunge from the scientific record the one contribution he regrets the most, this would be it. The notion of the stability of global geographic features was not only a main tenet of the old geology but seems to be firmly rooted in the human psyche. We now know that on the geological time scale the sea floor is far from permanent. The present ocean basins are being created by spreading and recycled by subduction on a time scale of about 200 million years, which is about 4 percent of the age of the Earth. The likelihood of finding extensive older remnants of sea floor is slight. Continents. On the other hand, are mobile but permanent features. They are too buoyant to be subducted. They may be fragmented, moved, reassembled, deformed, and eroded at their surfaces, but their bulk does not seem to be much diminished. Old terrains with ages of around 3.5 to 3.7 billion years can still be found. Continents grow with time by the gradual accumulation of materials along their margins. New continental strips can therefore be added on in different places at different times, depending on the history of fragmentation, movement, and reassembly.

With the emergence of these revolutionary ideas, geologists are rethinking Earth history. Most of the evidence for plate tectonics comes from the sea floor, a relatively simple place compared to the enormously complicated continents. Just how plate tectonics explains continental geology is now receiving much attention. New developments reported in nearly every issue of the geological journals show that the subject has definitely been revitalized. Rock associations, volcanism, metamorphism, the evolution of mountain chains—all are being reexamined in the framework of plate tectonics. Some of the new interpretations that we describe in this chapter may not stand the test of time. In this connection, future editions of this book may show some changes, not so much in the big picture of plate tectonics as in the details of fitting regional geology into the overall framework. The student (as well as the authors of this book) should be cautioned against calling on plate tectonics for easy explanations of

everything geological. It is not clear, for example, how or whether the origin of such structures as the Ozarks, the Black Hills, the Colorado Plateau, or such intracontinental, sediment-filled depressions as the Michigan Basin are related to plate movement.

Rock assemblages and plate tectonics

The only record we have of past geologic events is the incomplete one found in the rocks that have survived erosion or subduction. Since only sea floor younger than 200 million years (the last 4 percent of Earth history) has survived subduction, we must focus on the continents to find the evidence for most of Earth history. Some of the methods of reading the rock record have been described in earlier chapters. Here we explore the nature of the rock assemblages that characterize different plate-tectonic regimes as a first step in unraveling the history of past plate motions. Our aim is to reconstruct the process of continent fragmentation and ocean development, to locate the sites of vanished oceans, and to recognize the sutures that mark ancient plate collisions.

Of the three kinds of plate boundaries, we might expect distinct suites (assemblages) of rocks to be associated with plate divergence and convergence. At transform faults no distinct or characteristic rock assemblages are to be expected. Discontinuities across the fault are found, however, since rock formations formed and altered elsewhere have slipped past one another, and once-continuous formations or structural features are displaced.

Think of all that happens at a zone of divergence, where plate accretion and spreading occur, and you can predict the kinds of rock that would characterize the place and the process. Because there is extensive undersea volcanism, one would expect to find submarine basaltic lava, perhaps pillow lavas, the volcanic rock formed when hot lava is quenched by cold sea water (Chapter 15). Suboceanic crust and mantle are created here; dredge hauls and geophysical data show these layers to consist of mafic rocks, such as gabbro and peridotite, often showing evidence of alteration in a water environment (hydrous metamorphism). A carpet of deep-sea sediments would cover all of this. From Chapters 10 and 11 we remember that these deposits are recognized by thin layers of shale, limestone, and the siliceous rock chert, often with thin, discontinuous turbidites between them. Some or all of these layers may contain fossil remains of open-ocean marine organisms. A combination of deep-sea sediments, submarine basaltic lavas, and mafic igneous instrusions like that shown in idealized section in Figure 19-10, is called an **ophiolite suite**. The presence of narrow ophiolite zones in convergence features like the Alpine-Himalayan belt and the Ural and Appalachian belts may indicate that slices of oceanic crust and mantle originally produced at accreting plate margins were thrust onto land when an ancient ocean finally disappeared as two continents converged. It is generally believed that the Appalachians, for example, mark the site at which the ancestral Atlantic Ocean (called Iapetus for one of the Greek gods) closed when North America and Africa converged about 375 million years ago. The Atlantic reopened a few hundred kilometers east of this old suture, about 200 million years ago, in a spreading episode that is still underway.

Continental-shelf deposits are sedimentary rock assemblages that are laid down in an orderly sequence under tectonically quiet conditions in a geosyncline at a receding continental margin. Figures 19-11 and 10-29 show the orderly sequence of deposits in

the geosyncline that is still forming off the Atlantic coast of the United States. The continental margin there was formed when the American plate separated from the European plate about 200 million years ago. Resting on the offshore shelf is a wedge-shaped deposit of sediments eroded from the continent and carried into shallow water. Because the trailing edge of the continent slowly subsides as the spreading lithosphere cools and contracts, the geosyncline continues to receive sediments for a long time. The load of the growing mass of sediment further depresses the crust isostatically, so that the geosyncline can receive still more material from land. For every three meters of sediments received, the crust sinks two meters. The result of these two effects is that the geosynclinal deposits can accumulate in an orderly fashion to thicknesses of 10 kilometers or more. At the same time, the supply of sediments is sufficient to maintain the shallow-water environment of the geosyncline, or miogeosyncline, as we called it in Chapter 11.

The deposits show all of the characteristics of shallow-water conditions (Chapter 11). At the bottom of the entire sequence are rift valleys containing basaltic lavas and nonmarine deposits formed during the early stages of continental fissuring. In the early stages of shelf deposition, sandy materials started to fill the depression. Much was dropped on the continental slope, only to be moved later to the continental rise by turbidity currents. In deep water, very thick deposits can be built up in this way. As the shelf miogeosyncline builds up, deposition may become dominated by shales and carbonate platform deposits—indicators of a decrease in the supply of detritus from the continent.

Think what might happen to these geosynclines if the orderly, sequentially layered, gently dipping sediments were to become the leading edge of a plate in collision. In the following sections we describe some of the many possibilities.

Just as the events that take place in a convergence zone are different from divergence-zone phenomena, so do the rock assemblages have different characteristics. The main features of ocean-ocean or ocean-continent collision are shown in transverse section in Figure 19-12. Thick marine sediments, mostly turbidites, eroded from the continent or the island arc, rapidly fill the long marginal depressions. In descending, the cold oceanic slab stuffs the region below the inner wall of the trench with these sediments and with deep-sea materials brought with the incoming plate. Regions of this sort are enormously complex and highly variable, as they included turbidities and ophiolitic shreds scraped off the downgoing slab by the edge of the overriding plate—all highly folded, intricately sliced and metamorphosed. They are difficult to map in detail but recognizable by their distinctive mix of materials and structural features. Such a chaotic mess has been called a **mélange**. The metamorphism is the kind characteristic of high pressure and low temperature because the material may be carried relatively rapidly to depths as great as 30 kilometers, where recrystallization occurs in the environment of the cold slab. Somehow, perhaps by buoyancy and mountain-building, the material rises back to the surface much later. Find a mélange and you can't be too far from the place of downturn of an ancient plate, long since consumed, but leaving this relic of its existence.

Refer again to Figure 19-12. Parallel to the mélange is a magmatic belt that makes

up the arcuate system of volcanoes, intrusions, and metamorphic rocks formed on the edge of the overriding plate. Here the conditions are dominated by the rise of magma from the descending plate. At the interface, where the descending plate slides past the overriding one, perhaps friction is great enough to melt the upper part of the downturned slab, including the subducted wet sediments and ocean crust. The liquids rise buoyantly from depths of 100 to 200 kilometers to erupt and build the volcanic chains on the leading edges of plates. The characteristic igneous rocks produced are andesitic lavas and granitic intrusives. Island arcs, built up from the sea floor, may contain larger amounts of basalt; continental margins typically erupt rhyolitic ignimbrite and are intruded by granitic batholiths below. In contrast to that in a mélange, the metamorphism in the magmatic belts is typically the result of recrystallization under conditions of high temperatures and low pressures. This is because the hot fluids rise close to the surface, delivering much heat to a low-pressure environment.

Paired belts of mélange and magmatism are the signatures of subduction. The essential elements of these features of collision have been found in many places in the geologic record. One can see mélange in the Franciscan Formation of the California Coast Ranges and magmatism in the parallel belt of the Sierra Nevada to the east. This paired belt marks the Mesozoic boundary between the colliding Pacific and American plates. It even shows the polarity of the convergence by the location of mélange on the west and magmatism on the east; the Pacific plate was the subducted one. Other paired belts—for example, in Japan—can be found along the continental margins framing the Pacific basin. The central Alps, a European example, were produced by the convergence of a Mediterranean plate with the European continent.

Seismic reflection profiles are beginning to provide “x-ray” views of layers deep within the crust. Figure 19-14, a remarkable example of this new technique, shows the Australian plate being subducted under the Eurasian plate at the Jave trench.

Orogeny and plate tectonics

Orogeny means mountain-making, particularly by folding and thrusting of rock layers. In the framework of plate tectonics, orogeny occurs primarily at the boundaries of colliding plates, where marginal sedimentary deposits are crumpled and magmatism and volcanism are initiated.

Consider first some scenarios of plate convergence. In Figure 19-15a, a plate with a continent at the leading edge collides with another plate carrying a continent. In the early stage, during which the convergence is between continent and subducted oceanic lithosphere, a magmatic belt, folded mountains, and mélange deposits may be features of the overriding continental boundary. An example exists today along the Pacific coast of South America, where the American and Nazca plates are colliding. Look at the illustration inside the back cover to see the setting of the plates. The Andes, from which the name of the volcanic rock andesite is derived, lie in the magmatic belt; subduction is taking place under the Peru-Chile trench.

In a later stage, continent may meet continent, as shown in Figure 19-15b. Since continental crust is too light for much of it to be carried down, the plate motions could be slowed or halted. Another possibility, the one depicted in the figure, is that the plate motions continue, with subduction ceasing at the continent-continent suture but starting

up anew elsewhere. Cold and dense as the descending slab is, chunks of it may break off, fall freely into the mantle, and be resorbed. As Figure 19-15c shows, the suture is marked by a mountain range made up of either folded or thrust rocks, or both, coincident with or adjacent to the magmatic belt, and by a much-thickened continental crust. A prime example of continent-continent collision is the Himalayas, which began forming some 25 million years ago when a plate carrying India ran into the Asiatic plate (the collision and uplift are still going on). This may be how the root underlying the Himalayas originated (see Chapter 18). The plate-tectonic cycle of the closing of an ocean basin, a continent-continent collision, and the formation of an intracontinental mountain belt has been called the **Wilson cycle**, after the Canadian geologist J. Tuzo Wilson, who first suggested the idea that an ancient ocean closed to form the Appalachian mountain belt and then reopened to form the present day Atlantic Ocean.

Displaced terrains

Geologists have come across blocks within continents whose rock sequences, fossils, and paleomagnetism are alien to their surroundings. The rock assemblages and the fossils indicate different environments and ages than the surrounding terrain, and the paleomagnetic poles imply that the block originated in a different latitude. These are now believed to be fragments of other continents or of ocean crust that were swept up and plastered onto a continent in the process of plate collisions and separations. Coastal New England and Newfoundland may be slices of Europe; parts of Alaska, British Columbia, and Nevada may have been scraped off Asia; and central Florida may be a fragment of Africa. Displaced terrains have also been found in Japan, Southeast Asia, China, and Siberia, but their original locations have yet to be worked out.

The grand reconstruction

At the close of the Paleozoic, some 250 million years ago, there was a single supercontinent Pangaea, stretching from pole to pole. The fragmentation of Pangaea as a result of plate tectonics and continental drift over Mesozoic and Cenozoic time to form the modern continents and oceans is documented in the well-preserved record of magnetic reversal stripes on the ocean floor. But what of the pre-Pangaean distribution of continents? What were their shapes and where were they located? There is growing evidence that Pangaea was formed by the collision of continental blocks——not the same continents we know today but continents that existed earlier in the Paleozoic. The ocean-floor record for this period has been destroyed by subduction, so we must rely on the older evidence preserved on continents to identify and chart the movements of these paleocontinents. Old mountain belts like the Appalachians and the Urals mark the collision boundaries of the paleocontinents. Rock assemblages there reveal ancient episodes of rifting and subduction. Rock types and fossils also indicate the distribution of shallow seas, glaciers, lowlands, mountains, and climatic conditions. Paleomagnetic data can be used to find the latitude and the north-south orientation of the paleocontinents. Latitudes can also be checked by paleoclimatic data. Although it is not possible to assign longitudinal position to the paleocontinents, the relative sequence of continents around the globe can be pieced together from the fossil record. One of the first efforts to depict the per-Pangaean configuration of continents using these methods is shown in Figure 19-16. The ability of modern science to recover the geography of

this strange world of hundreds of millions of years ago is truly impressive. Geologists may be able to continue to sort out more details of this complex jigsaw puzzle, whose individual pieces change shape over geologic time.

Figure 19-17 reconstructs the most recent breakup of Pangaea as we now understand it. Figure 19-17a shows the world as it looked in Permian times, a little more than 200 million years ago. Pangaea was an irregularly shaped land mass surrounded by a universal ocean called Panthalassa, the ancestral Pacific. The Tethys Sea, between Africa and Eurasia, was the ancestor of part of the Mediterranean. The fit of North and South America with Europe and Africa is very good in detail when taken at the outer edge of the continental shelves, instead of at the present shorelines, which are some distance from the original rift. It is the fit for which we have the firmest evidence. The positions of Central America, India, Australia, and Antarctica are less certain.

The breakup of Pangaea was signaled by the opening of rifts from which basalt poured. Relics of this great event can be found today in the Triassic basalt flows all over New England. Radioactive dating of these flows provides the estimate of about 200 million years for the beginning of drift.

The geography of the world after 20 million years of drift—at the end of the Triassic some 180 million years ago—is sketched in Figure 19-17b. The Atlantic has opened, the Tethys has contracted, and the northen continents (Laurasia) have all but split away from the southern continents (Gondwana). New ocean floor has also separated Antarctica-Australia from Africa-South America. India is off on a trip to the north.

By the end of the Jurassic period, 135 million years ago, drift had been underway for 65 million years. The big event at this time is the splitting of South America from Africa, which signals the birth of the South Atlantic. The North Atlantic and Indian oceans are enlarged, but the Tethys Sea continues to close. India continues its northward journey.

The close of the Cretaceous period 65 million years ago sees a widened South Atlantic, the splitting of Madagascar from Africa, and the close of the Tethys to form an inland sea, the Mediterranean. After 135 million years of drift, the modern configuration of continents becomes discernible.

The modern world, produced over the past 65 million years, is shown in Figure 19-17e. India has collided with Asia, bringing its trip to an end. Australia has separated from Antarctica. Nearly half of the present-day ocean floor was created in this period. Figure 19-18 shows several schematic sections that summarize modern plate, ocean, continent, and island-arc relationships for the American, African, Eurasian, and Indian plates.

Most of the modern Pacific Ocean basin consists of the Pacific plate side of the East Pacific rise spreading zone, as can be seen inside the back cover and in Figure 19-18b. This implies that an area equal to most of the Pacific Ocean has disappeared by subduction under the Americas in the past 130 million years. As much as 7000 kilometers (4300 miles) of Pacific sea floor may have been thrust under North America!

Not one branch of geology, except perhaps crystallography, remains untouched by

this grand reconstruction of the continents. Economic geologists are using the fit of the continents to find mineral and oil deposits by correlating the formations in which they occur on one continent with their predrift continuations on another continent. Paleontologists are rethinking some aspects of evolution in the light of continental drift. For example, during most of the age of reptiles, the continents we know today were grouped together in two supercontinents, Laurasia and Gondwanaland. There continents were fragmented during most of the age of mammals, with faunas developing on the daughter continents isolated from one another. Is this why mammals diversified into so many more orders than the reptiles did, and in a much shorter period of time? Structural geologists and petrologists are extending their sights from regional mapping to the world picture, for the concept of plate tectonics provides the means of interpreting such geological processes as sedimentation and orogeny in global terms. For example, the Caledonian mountain belt that runs along the northwest margin of Europe is the predrift continuation of the Appalachian belt, and the trend of the Andes may be followed into Antarctica and Australia, as Figure 19-19 shows.

Oceanographers are reconstructing currents as they might have existed in the ancestral oceans, to understand better the modern circulation and to account for the variations in deep-sea sediments. Paleoclimatologists are "forecasting" backward in time to describe temperature, winds, the extent of continental glaciers, and the level of the sea as they were in predrift times. What better testimony to the triumph of this once-outrageous hypothesis than its ability to revitalize and shed light on so many diverse topics!

The driving mechanism of plate tectonics

Up to this point everything we have discussed might be categorized as descriptive plate tectonics. The geometry and rates of plate motions, the consequences of plate separation and collision have been described. But what drives it all? We will not fully understand plate tectonics until we can answer this question. The International Geodynamics Project enlisted the efforts of thousands of scientists in seeking the underlying cause of plate motions.

It is generally accepted that most of the mantle is a hot solid, capable of flowing like a liquid at a speed of about a centimeter per year, about the rate your fingernails grow. The lithosphere is broken into rigid plates, somehow responsive in their motions to the flow in the underlying mantle.

As is generally the case when there is an abundance of data in search of a theory, many hypotheses have been advanced. Some would have plates pushed by the weight of the ridges at the zones of spreading or pulled by the heavy down going slab at subduction zones. Others hold that the plates are dragged along by currents in the underlying asthenosphere. Figure 19-20 shows some of these ideas. In line with the discussion in Chapter 13, we agree with those who view the process not in piecemeal but as a highly complex convective flow, involving rising, hot, partially molten materials and sinking, cool, solid materials, under a variety of conditions ranging from melting to solidification and remelting. A significant part of the mantle must be involved, for slasbs are known to penetrate to depths of some 700 kilometers before being completely resorbed. Figure 19-20c, shows one of the first computer models of

the process-one that neglects some of the effects just mentioned, but that nevertheless accounts for many observations. A rising plume of hot material, heated from below, reaches the surface at a center of spreading. It moves away from the center, cools near the surface, and the cooled boundary becomes solid, strong lithosphere. Finally becoming heavier after it has cooled, the lithospheric slab sinks back into the mantle in a subduction zone, where it is reassimilated, to be heated and to rise again in the future. Another theory proposes that hot, narrow jet-like plumes rise from the bottom of the mantle, feed the growing plate, and drive it laterally away from spreading centers where the plumes mostly occur. These same plumes are evidenced at the surface by hot spots. Among the problems left to the next generation of Earth scientists is the incorporation of such important details as the shapes of plates, the history of their movements, and the formation and growth of continents into an explanation of the distribution of convective currents in time and space.

Summary

1. According to the theory of plate tectonics, the lithosphere is broken into about a dozen rigid, moving plates. Three types of plate boundaries are defined by the relative motion between plates: boundaries of divergence, boundaries of convergence, and transform faults.

2. In addition to earthquake belts, many large-scale geological features are associated with plate boundaries, such as narrow mountain belts and chains of volcanoes. Boundaries of convergence are recognized by deep-sea trenches, inclined earthquake belts, mountains and volcanoes, and paired belts of mélange and magmatism. The Andes Mountains and the trenches of the west coast of South America are modern examples. Divergent boundaries (for example, the mid-Atlantic ridge) typically show as seismic, volcanic, mid-ocean ridges. A characteristic deposit of this environment is the ophiolite suite. Transform faults, along which plates slide past one another, can be recognized by their topography, seismicity, and offsets in magnetic anomaly bands. Ancient convergences may show as old mountain belts, such as the Appalachians.

3. The age of the sea floor can be measured by means of magnetic-anomaly bands and the stratigraphy of magnetic reversals worked out on land. The procedure has been verified and extended by deep-sea drilling. Isochrons can now be drawn for most of the Atlantic and for large sections of the Pacific, enabling geologists to reconstruct the history of opening and closing of these oceans. Based on this method and on geological and paleomagnetic data, the fragmentation of Pangaea over the last 200 million years can be sketched.

4. Although plate motions can now be described in some detail, the driving mechanism is still a puzzle. An attractive hypothesis proposes that the upper mantle is in a state of convection with hot material rising under divergence zones and cool material sinking in subduction zones, The plates, according to this model, would be the cooled, upper boundary region of the convection cell.

New words and expressions

mosaic of plates 板块拼接
transform fault 转换断层
convergence zone 聚合带（俯冲带）
divergence zone 背散带（扩散带）
lithosphere，岩石圈
leading edge 主动端
thrust 逆断层，(v.) 冲，插
geophysicist 地球物理学家
magnetic anomaly 磁异常
seismic 地震的
bathymetric 等深的
acoustic beacon 声纳
accretion 侧向加积
ophiolite suite 蛇绿岩套
ignimbrite 熔结凝灰岩(熔灰岩）
Tethys 特提斯（古地中海）

asthenosphere 软流圈

scarp(s) 悬崖

buoyant 有浮力的

magnetic_stratigraphy 磁性地层学
positioning device 定位装置
thruster（螺旋） 推进器
pillow lava 枕状熔岩
mélange 混杂岩 (法语来源)

V. Geochemistry

Geochemistry of the Continental Mantle-Basaltic Rocks

Introduction

Bowen (1928) noted that'at 37-60km [one] encounter[s] material of a peridotitic character… and if this is true the only source of basaltic magma… lies in the peridotite zone, from which it must be produced by selective fusion. Moreover, he was one of the first to suggest that extraction of basalt from peridotite led to the peridotite being more barren'. Although there was a very acceptable idea towards the end of the twentieth century, these ideas were advanced at a time when some of his colleagues believed in basaltic layers deep in the Earth as a source for basaltic magmas (Daly 1926,1933).

Experimental petrology greatly advanced our understanding of the origin of basalts. Kuno(1960) linked the genesis of different magmas to depth and some of the earliest basalt-peridotite experiments (Yoder and Tilley 1962) studied the formation of melts to depths of 90km. O'Hara (1965, 1968) noted that melts in equilibrium with peridotite were different from so-called 'primary magmas' formed at depth and the need for fractionation of olivine became apparent. Green and Ringwood (1967) presented a comprehensive scheme for the production of tholeiitic and alkaline basaltic magmas in terms of polybaric partial melting of undepleted mantle. Soon the importance of CO_2 in mantle processes was realized and the stability of mantle carbonate was linked to the genesis of kimberlites and carbonatities (Eggler 1975, 1976, 1978; Wyllie 1987; Wallace and Green 1988; Green and Wallace 1988). In contrast, the compositions of basaltic magmas produced by melting of dry or volatile-free spinel lherzolite were perhaps best estimated from the experimental work of Takahashi and Kushiro (1983) and Fujii and Scarfe (1985). Throughout the 1970s and early 1980s most models of basalt genesis were based on the assumption that large degrees of melting were necessary to produce basaltic magmas from a peridotitic source and that small degree melts were unlikely to the easily mobilized. However, recent theoretical and experimental considerations indicate that small volume melts are mobile at mantle pressures and temperatures and, earlier models may need to be revised (McKenzie 1984, 1985; Watson and Brenan 1987; Brenan and Watson 1988; McKenzie and Bickle 1988; Hunter and McKenzie 1989; McKenzie 1989；Watson *et al.,* Chapter 6, this 1968) and Schilling (1971) pointed out the complementary geochemical relationship between oceanic tholeiites (depleted in large ion lithophile (LIL) and light rare earth (LRE) elements) and alkaline basalts (enriched in LIL and LRE elements). Gast (1968) proposed that the source of mid-ocean ridge basalt (MORB) had undergone a partial melting episode prior to extraction of MORB and be related the origin of alkaline (i.e. ocean island) and tholeiitic (i.e. mid-ocean ridge) basalts to variable degrees of melting of a heterogeneous mantle source (cf. McMenzie and Bickle 1988). Basalts within ocean basins that were not produced at ocean ridges were believed to be produced by hot-spot magmatism (Wilson 1965; Morgan 1971) and it was soon realized that ridge basalts were derived from shallow depleted asthenosphere and ocean island basalts were supplied by deeper mantle plumes (Schilling 1973; Hofmann and Hart 1975; Sun and Hanson 1975).

Continental alkaline volcanic rocks and mantle domains

Lewis (1888,1897) studied the matrix of diamond (blue ground) and concluded that it was a porphyritic peridotite of igneous eruptive character. Bonney (1899) prefaced an article on the parent rock of the diamond by stating that to much had been written on the occurrence of diamond by 1899 that brevity was the best introduction (Bonney 1899 and references therein). Despite that much of the next 50 years saw a rapid increase in the much of the database concerning the geology and mineralogy of kimberlites (Wagner 1914; Williams 1932). Wagner (1914) believed diamonds to have formed from kimberlitic magmas and noted that some pipes were barren of diamonds. He suggested that this might be due to a lack of carbon or carbon compounds in the deep-seated peridotite zone which he believed to the ultimate source of kimberlite. Wagner was also the first to note that diamondiferous kimberlites were erupted in stable geological terrains, a point that was re-investigated over 50 years later (Frantsesson 1968; frantsesson and Prokopchuk 1968; Dawson 1970). They noted that most diamondiferous kimberlites in southern Africa and Siberia were confined to the ancient cratons and that their abundance decreased into the surrounding mobile belts.

Probably because of the association of kimberlites with old crust enriched in radiogenic strontium the high $^{87}Sr/^{86}Sr$ ratios in kimberlites were believed to be due to: (1) incorporation of micro-xenoliths of crust (Powell 1966); (2) contamination with Sr derived from overlying limestones and shales (Brookins 1967); or (3) postemplacement alteration (Berg and Allsopp 1972). The underlying assumption at this time was that all mantlederived rocks had a source similar to oceanic basalts and as such should be isotopically identical. If their isotopic composition was different, then this was taken as evidence of crustal contamination and must constitute ‘high level’ interaction with the crust. This was accepted despite the fact that many ‘altered’ kimberlites had apparently lower $^{87}Sr/^{86}Sr$ ratios than fresh kimberlites had apparently lower $^{87}Sr/^{86}Sr$ rations than fresh kimberlites (Paul 1979; Demaiffe and Fieremans 1981). Whilst a source similar to oceanic basalts for basaltic kimberlites is consistent with much of the modern isotopic data (Kramers 1977; Basu and Tatsumoto 1980; Kramers *et al*. 1981)，not all kimberlites are genetically the same. In a landmark contribution, Smith (1983) demonstrated that basaltic or Group I kimberlites had a mantle source with the Sr, Nd, and Pb isotopic characteristics of ocean island basalts (OIB) and were therefore interpreted as asthenospheric melts. In contrast, micaceous or Group II kimberlites had Sr, Nd, and Pb isotopic characteristics similar to basalt-borne and kimberlite-borne xenoliths and were there fore different from oceanic basalts. These data were interpreted to mean that micaceous kimberlites had a significant source contribution from the subcontinental lithospheric mantle. The Group I and II terminology introduced by Smith (1983) is now more commonplace than the ‘basaltic-micaceous kimberlite’ terminology of Wagner (1914). The model of depleted asthenosphere (OIB)-enriched lithosphere interaction proposed by Smith (1983) was re-interpreted by McCulloch *et al*. (1983) who believed that the data was better explained by interaction of asthenospheric (MORB) melts and enriched lithospheric mantle. The retraction of micaceous diamondiferous kimberlites to the craton the dominant contribution from lithospheric mantle in their magmagenesis supports the contention that an enriched

Archaean keel exists beneath the cratons.

Geochemical data for alkaline and tholeiitic basalts provide us with important constraints on the global distributions of mantle reservoirs. Many continental alkaline basalts (e.g. basanites, nephelinites, and alkali olivine basalts), like basaltic kimberlites, are chemically indistinguishable from ocean island basalts (Allegre *et al*. 1982), a fact that requires parts of the upper mantle to have been a common reservoir beneath both ocean basins and continents. Moreover, this reservoir must have been isolated for billions of years(Sun and Hanson 1975; Tatsumoto 1978; Sun 1980; Chase 1981). Whilst most investigators agreed that an enriched reservoir was required for alkaline ocean island and continental basalts and a depleted reservoir for midocean ridge basalts the exact geometry of these reservoirs remained a hot debated topic. Gast (1968) and Schilling (1971) proposed that the source for alkali basalts was deep and the source of MORB was shallow, Tatsumoto (1978) proposed the opposite with a shallow enriched reservoir and a deep depleted reservoir. Anderson (1979 *a, b*) queried a more fundamental assumption—that the upper mantle was necessarily peridotite. He believed that, as a result of early differentiation within the Earth, a layer of eclogite lay beneath a layer of peridotite. MORB was derived from this eclogite and the source was replenished by subduction.

To this day the exact geometry of MORB and OIB sources, the latter feeding continental and alkaline intraplate volcanoes, remains the source of some debate (Allegre *et al*. 1981,1982; Thompson *et al*. 1984; Zindler and Hart 1986). It is generally accepted that much of the asthenosphere is MORB-like and that OIB melts may be generated as small volume melts in the lithosphere, the asthenosphere, or in the deeper mantle. Detailed studies of the geochemistry of oceanic basalts (Zindler and Hart 1986) have produced evidence for several chemically distinct reservoirs: DMM or depleted MORB mantle, PREMA or prevalent mantle for ocean islands (e.g. Hawaii), HIMU or mantle with high μ(i.e.U/Pb), EM2 or enriched mantle 2, and EM1 or enriched mantle 1. An estimate of the composition of the bulk silicate earth from the work of Zindler and Hart (1986) compares favorably with that of Wanke (1981).

Within the continental regions the correlation between crustal age and kimberlite type helped define chemical provinces in the mantle beneath the Archaean Kaapvaal craton and the surrounding Proterozoic mobile belt in South Africa. Similarly, compilations of isotopic data for continental alkaline volcanic rocks from the western USA reveal isotopic provinciality (Leeman 1982). Menzies (1989) used xenolith-bearing volcanic rocks as a probe of the geometry and chemistry of lithospheric mantle domains and was able to assign a domain identity to subcrustal regions. In summary, the isotopic characteristics of the lithosphere beneath the mobile belts is similar to that beneath the ocean basins (Hart *et al*. 1989) and the lithosphere beneath stable cratons is to some degree isotopically unique. A close correlation exists between areas of high heat flow (>2 HFU), thin crust (<30km), and the eruption of continental volcanic rocks, whose source is in the asthenosphere, whereas continental volcanic rocks with a significant source contribution from the lithospheric mantle tend to be restricted to regions of low regional heat flow (<2 HFU) and thick crust (>40km). Chemical provinciality has also been recorded in the lithosphere beneath Scotland

(Thirlwall 1982) and has similarly been interpreted as a result of lateral and vertical variability in the lithospheric mantle (Menzies and Halliday 1988). It should be noted that a shallow lithospheric source for continental alkaline basalts (Leeman 1982; Perry *et al*. 1987) from the western USA and Canada contrasts with the classic demonstration of an asthenospheric source for continental alkaline volcanic rocks from West Africa and the Gulf of Guinea (Fitton and Dunlop 1985). Indeed these data seem to indicate that under different circumstances alkaline melts can be generated in the convecting asthenosphere (or deeper) and in the lithosphere. If the chemical provinciality observed in continental volcanic rocks indicates the involvement of lithospheric mantle domains in their genesis. then continental volcanic rocks can be used to literally 'map' mantle domains (Fitton *et al*. 1988; Leat *et al*. 1988; Ormerod *et al*. 1988; Menzies 1989; Menzies and kyle, Chapter 8, this volume). Similarly, magma distribution and the variability in composition can be used to monitor the vertical movement of the asthenosphere-lithosphere boundary (McKenzie and Bickle 1988). It should be stressed that not all continental alkaline and tholeiitic volcanic rocks are suitable candidates for 'mapping' of the subcrustai lithosphere. Ewart *et al*. (1988) demonstrated that high-level processes involving continental crust may produce similar chemical provinciality in erupted basalts.

Continental flood basalts

The petrology of continental flood basalts was dealt with in many of the early texts (Washington 1922; Kuno 1969) and, more recently, Thompson (1977) deduced that certain flood basalt provinces were extruded at rates equivalent to mid-ocean ridge basalts. To account for these large volumes he proposed that continental basalts had an origin similar to oceanic basalts and that any modification of their trace element or isotopic ratio was due to digestion of variable quantities of crust during upwelling or accumulation of the magmas in the crust (e.g. Cox 1980). Whilst this is undoubtedly the case for several continental flood basalt provinces, including the British Tertiary, many others retain evidence of mantle heterogeneity. Leeman (1975,1977) and brooks and Hart (1978) proposed that the continental regions were underlain by lithospheric mantle that was very different from that beneath the ocean basins. Leeman (1977) believed in the existence of old enriched subcontinental mantle characterized by high Rb/Sr rtios and high $^{87}Sr/^{86}Sr$ rations and Brooks and Hart (1978) suggested that flood basalts inherited age information from their lithospheric mantle source, thus explaining the linear arrays of isotopic and trace element data. Hawkesworth *et al*. (1983,1986,1988) have more recently invoked the participation of Proterozoic lithospheric mantle in the production of the Karoo and the Parana flood basalt provinces. This was substantiated by recent Sm-Nd isotopic studies (Ellam and Cox 1989) which point to a Proterozoic age in the source region of some of the Karoo volcanic rocks. Pb isotope data for continental tholeiites from the western USA (Leeman 1975, 1977) also point to a Proterozoic rather than Archaean ages are forthcoming from these continental flood basalt provinces is of importance. The major, minor, and trace element chemistry of archaean lithoshpheric mantle is more refractory than post-Archaean lithospheric mantle which is closer to pyrolite in composition. It is therefore likely that such lherzolitic material could melt to produce vast quantities of tholeiitic magma than the

refractory harzburgites found beneath the Archaean cratons. Whether continental flood basalts are derived from lithospheric mantle or asthenosphere with high-level contamination with crust remains unresolved in many cases (e.g. Hawkesworth and Norry 1983 and references therein; Moorbath *et al*. 1984 and references therein; MacDougall 1989 and references therences therein; Leeman and Fitton 1989 and references therein; Menzies and Kyle, Chapter 8, this volume).

Concluding statement

Volcanic rocks and xenoliths provide a unique insight into the petrology and geochemistry of the asthenosohere and lithosphere but it must be remembered that lherzolites and metasomatized xenoliths constitute a small proportion of the lithosphere. The continental lithospheric mantle may be largely comprised of a depleted protolith that remains in a 'pristine' state at some distance from magma conduits within the mechanical boundary layer and without the thermal boundary layer where the ingress of small volume melts from the asthenosphere produces a 'secondary' assemblage. The geochemistry of this depleted protolith holds the key to the evolution of the Archaean lithosphere, perhaps the only part of the upper mantle whose major element chemistry can be explained by extraction of crustal material (e.g. komatiites). Similarly, the precise significance of the lithosphere beneath Proterozoic and Phanerozoic terrains has not been thoroughly investigated. It can be demonstrated that Proterozoic and Phanerozoic lherzolites are a suitable source for tholeiitic melts with the isotopic characteristics of mid-ocean ridge basalts. The existence of such material beneath post-Archaean crust requires some explanation in that MORB mantle was depleted in the first billion years of Earth's history. Moreover, the recent recognition of important correlations between seismic tomography and the geoid and tectonic features like subducted slabs may help focus our attention on active margin processes and the petrology and geochemistry of eclogites. Can active margin processes produce highly magnesian peridotites like those found beneath the Archaean crust? Finally, the importance of small volume melts and metasomatic processes are only now being realized with the study of wetting angles. The hitherto unspecified nature of metasomatic melts can now be more accurately defined (i.e. carbonate metls) but the exact relationship of such chemical transfer mechanisms to the origin of alkaline magmas is unknown as small volume melts can be generated within the asthenosphere and the lithosphere.

New words and expressions

peridotitic 橄榄岩的	tholeiite(ic) 拉斑玄武岩（的）	Alkaline 碱性的
polybaric 多压的	undepleted 无亏损的	pyrolite 地幔岩
kimberlite 金伯利岩	metasomatic 交代的	lherzolite 辉橄
matrix 母岩，基质	protolith 母岩	craton 克拉通,
strontium 锶	micaceous 云母的,含云母的	isotop 同位素
eclogite 榴辉岩	lithophile 亲岩的，亲石的	upwelling 上涌
basanite 碧玄岩，试金石	harzburgite (saxonite) 方辉橄榄岩	geoid 地球体
proterozoic 元古	tomography 层面 X 线照相机	nephelinite 霞岩
komatiite 科马提岩	conduit 管道、排水渠	spinel 尖晶石

VI. Remote Sensing

Integrated use of satellite images, DEMs, soil and substrate data in studying mountainous lands

Fabio Giannetti et al.

Introduction

Any analytical approach to mountainous areas, aimed at evaluating their resources and promoting their sustainable' development requires objective knowledge of the territory-based on maps and linked geographical databases. Information extracted from satellite image, DEMs and lithological data from the northwestern Alps was integrated into a GIS to meet such goals and to define eco-pedological units at a 1:250,000 scale. Operational earth observation satellites such as landsat TM and SPOT are very useful for studying mountainous areas at this scale. In particular the synoptic view of the terrain and the geometric resolution of 20-30m makes it possible to map the main environmental characteristics and monitor dynamic aspects by multi-temporal sequence analysis (Allan, 1980).

This paper presents a method for defining eco-pedological cartographic units in mountainous areas. The work described was carried out by IPLA and the European Soil Bureau at lspra as a preliminary step in the project "Ecopedological map of Italy".

Study area

The study area includes a sector of the northwestern Alps in the Piemonte region, from the Pellice valley in the north to the Maira valley in the south. The area is located southwest of Torino, the main town of the region.

This area has undergone considerable deformation and metamorphosis during the different phases of alpine orogenesis and it is characterized by a great variety of lithologies massif of the Dora-Maira, the Piedmontese zone and the Brianconnaise zone are the tectonic zones encountered when moving from the inner part of the chain towards the meso-european foreland (Societa Geologica Italiana, 1992). Acid rocks such as mica schists and gneiss are the main rock types of the Dora-Maira massif. The mica schist rocks have undergone considerable weathering in the Quaternary, resulting therefore in a smooth morphology and deep soils. The gneiss rocks are less weathered and mainly covered by shallow soils.

The Piedmontese zone is composed of series of lime schists and some large ophiolitic sheets. The lime schists are particularly weatherable and, because of the long period of snow cover, often give rise to neutral soils. Ophiolitic rocks, less weathered and more erosion-resistant than lime schists, often produce subalkaline soils. Quartztitic, calcareous and dolomitic rocks cover the limited outcrop area of the brianconnaise zone in this sector of the northwestern Alps.

The Quaternary deposits in the study area are represented mainly by terraced fans in the lower valleys and by some moraines. These deposits cover only a limited portion of the area.

The Italian boundary of the Alps in this sector is very steep with no more than 30-35 km between the basin ridges, elevations of around 2700-3000m a.s.l., and river debouchments onto the Po plain at altitudes ranging from 250m to 400m.

Materials

The following materials were used:(1) SPOT 4 images of 05/07/1998 and 18/09/98; (2) LANDSAT TM images of 15/09/92; (3) a DEM with a grid of 250m×250m converted into a raster layer; (4)a DEM with a grid of 50m×50m on grassland area; and (5) geolithological, geomorphological and soil-landscape maps of the study area at 1:250,000 scale.

The images were first georeferenced (nearest neighbour algorithm) by means of ortho-rectification using about 40 ground control points with an averaged bias of about 10m in their *x* and *y* coordinates. All image processing was done on the SPOT 4 image acquired in July 1998 as this was the most suitable period. The other two images were used to confirm the result obtained through interpretation and statistical processing of SPOT 4 July image. The work was carried out using "Cartha for Windows", Italian software that integrates image processing and GIS functions.

Cartographic unit definition - systems

In this paper we have defined systems as eco-pedological units characterized by a homogeneous pattern of lithological, landscape and morphological characteristics (Finke *et al*. 1998) and we propose a method to delineate them at 1:250,000 scale. Preliminary systems delineation originates from a GIS integration to two layers: the first concerning geological substrates; and the second the result of processing and interpretation of satellite images.

The work was conducted in several steps. The first step concerns integration in a GIS of a geo-lithological map of 1:250,000 scale. This kind of data is essential to define the distribution of substrates of pedogenic characteristics. The pedogenic value of the substrates is derived mainly from the lithological composition of each geological formation. It is therefore necessary to obtain a preliminary re-interpretation of the geological legend to join the different geological formations in three groups: basic, acid and neutral substrates.

Each group can be further divided in two subgroups if there are differences in weatherability. In this case, differences in intensity of pedogenic processes enable the formation of deeper or shallower soils.

For the second step, the main landscape components, considered as one of the factors related to soil forming processes in mountainous areas, were extracted from satellite image and grouped into six classes: (1) rock outcrop, fresh debris fan and landslides (mineral zones almost absent of soil and vegetation); (2) areas of low-density vegetation in which vegetation of scattered and poor development, indicates the presence of discontinuous shallow soils (such as lithosoils); (3) deciduous forest (broadleaf); (4) evergreen forest (coniferous); (5) herbaceous cover zones (alpine and subalpine grassland). (6) arable zones and grasslands of the valley bottoms.

The first two classes can be best derived from computations of the NDVI vegetation index, which is a normalized difference of near infrared and red bands. As this index is very sensitive to the presence of photosynthetic biological matter

(Schowengerdt, 1997), it is possible to discriminate, by means of thresholds, areas with no vegetation cover from those with poor photosynthetic activity.

Some attempts were made in order to distinguish the other classes through utilization of NDVI and other vegetation indices, such as infrared index (Normalised difference between near infrared and medium infrared bands), which is particularly sensitive to the water content of vegetation. Forest and herbaceous vegetation are easily separable using the lnfrared index to identify their different moisture status. However, discrimination between broadleaf and conifer trees is not possible with this index.

The best solution seems to be a supervised classification (maximum likelihood algorithm) based on comprehensive large sample areas of all the different morphological locations in which each class can be retrieved. Classification was carried out using synthetic bands like arithmetic indices and statistical processing, such as Principal Component Analysis, to minimize the influence of illumination difference (Richards, 1994). Comparison between the classification layer obtained and other landuse or land cover data that was available for some parts of the study area confirmed the reliability of our results.

Some problems were faced in correctly differentiating arable lands that in narrow and steep-profile valleys are often so small that they cannot form a significant, clearly distinguishable training sample of the class for running the classification algorithm. In such cases we decided not to include them in the classification and, if allowed by cartographic scale, to delineate these zones by a simple visual interpretation based on the geometric shapes of the fields and a spectral characterization of the crops.

After a first delineation of the systems that integrate main landscape components and pedogenic substrates, these units were refined and characterized by the following techniques: (1) slope and aspect computing from DEMs; (2) video interpretation of satellite image RGB compositions useful in differentiating distinctive patterns of relief features or landforms; and (3) comparison and integration with a soil-landscape map realized by IPLA at the same scale by means of stereoscopic interpretation of aerial photographs.

DEMs and interpretation of RGB composite images led to further subdivisions of the previous units based on slope, aspect and/or general physiographic aspects (Calzolari *et al.*, 1996; Dubuq *et al*. 1991). The third refining step was useful for validating the resulting units and for verifying the consistency with the map scale.

Analysis of grassland and area 1 at 1:50,000 scale

A sub-alpine grassland area located on the Po valley head was chosen to verify the possible use of instruments such as satellite images and DEMs for a more detailed analysis of grassland resources at 1:50,000 scale. The area of interest was extracted from the SPOT 4 July image and techniques developed for the characterization of the main landscape systems at 1:250,000 scale were applied at a finer scale (1:50,000). The following results were obtained: (1) Principal Component Analysis revealed its utility in stretching differences among the main elements of the image (different kinds of herbaceous cover, shrub cover, alders colonization, rocks, debris fans, etc.); (2) NDVI and infrared index are useful to show differences in density of vegetation cover and water content; and (3) application of the redness index can help in identifying areas

with erosion in progress.

Integration of the satellite-derived information with a more detailed DEM (50m×50m grid) gave a better understanding of the relation between slope, aspect and soil distribution in the area examined. Soil profile data, flora inventory, measurement of ligneous-herbaceous cover percent, etc., in some sample areas should then the be collected for better knowledge of the area. Once acquired, these data can be spatially organized by raster information, leading to a pastoral resource map.

Conclusions

The potential of integrated use of satellite images, DEMs and soil and substrate data within a geographic information system was shown. This approach contributed significantly to the elaboration of a common, uniform 1:250,000 scale eco-pedological map of the Italian northwestern Alps.

Output of this nature will not create only a cartographic document, useful to delineate main landscape elements related to soil distribution in an alpine environment, but it also represents an open database that can be updated by satellite images, allowing dynamic processes such as soil erosion and the expansion or contraction of wooded areas to be monitored and compared with the extension of pastoral land.

New words and expressions

Remote sensing 遥感
Pedology 土壤学
pedologic (adj.) 土壤学的
Cartography 制图学、制图法
Alpine orogenesis 阿尔卑斯式的造山运动
massif 地块
piedmont 山麓
foreland 前陆
mica schists 云母片岩
gneiss 片麻岩
subalkaline 亚碱性的
terracea 成阶地的
moraine 冰碛
debouchment 河口
deciduous forest (broad leaves) 落叶森林（阔叶）
evergreen forest 常绿森林（针叶）
herbaceous 草本的
vegetation 植被
arable 可耕的
pastoral 畜牧的

VII. Orogenic gold deposits: A proposed classification in the context of their crustal distribution and relationship to other gold deposit types

D.I. Groves et al., Ore Geology Reviews 13 1998 7–27

The so-called 'mesothermal' gold deposits are associated with regionally metamorphosed terranes of all ages. Ores were formed during compressional to transpressional deformation processes at convergent plate margins in accretionary and collisional orogens. In both types of orogen, hydrated marine sedimentary and volcanic rocks have been added to continental margins during tens to some 100 million years of collision. Subduction-related thermal events, episodically raising geothermal gradients within the hydrated accretionary sequences, initiate and drive long-distance hydrothermal fluid migration. The resulting gold-bearing quartz veins are emplaced over a unique depth range for hydrothermal ore deposits, with gold deposition from 15–20 km to the near surface environment. On the basis of this broad depth range of formation, the term 'mesothermal' is not applicable to this deposit type as a whole. Instead, the unique temporal and spatial association of this deposit type with orogeny means that the vein systems are best termed orogenic gold deposits. Most ores are post-orogenic with respect to tectonism of their immediate host rocks, but are simultaneously syn-orogenic with respect to ongoing deep-crustal, subduction-related thermal processes and the prefix orogenic satisfies both these conditions. On the basis of their depth of formation, the orogenic deposits are best subdivided into epizonal -6 km , mesozonal 6–12 km and hypozonal)12 km classes.

1. Introduction

This thematic issue of Ore Geology ReĺIews includes a wide variety of papers on a single type of quartz–carbonate lode-gold deposit. The deposit type in this issue alone is referred to as synorogenic, turbidite-hosted, mesothermal and Archaean lodegold. This reflects the proliferation of such terms throughout the economic geology literature during the last ten years and a subsequent increase in confusion for the readers. For example, is a synorogenic Mother-lode type gold deposit different from an Archaean gold-only type or from a mesothermal greenstone–gold type? Many researchers working on such deposits would recognize these as essentially a variety of subtypes of a single deposit type, i.e. epigenetic, structurally-hosted lode-gold vein sys-tems in metamorphic terranes Kerrich, 1993 . However, the consistent usage of a single and widely accepted classification term for this deposit type as a whole is clearly warranted. 'Mesothermal' is such a term that has been widely adopted during the last ten years, but is a term that, as originally defined by. Lindgren 1933 for deposits formed at about 1.2–3.6 km, is more applicable to sedimentary rock-hosted ' Carlin-type' deposits and the gold porphyry skarn. environment (Poulsen, 1996).

A principal aim of this introductory paper is to present and justify a unifying

classification for these lode-gold deposits. An attempt is made to place these so-called ‘mesothermal’ deposits into a broader class that emphasizes their tectonic setting and time of formation relative to other gold deposit types. A second aim is to review briefly their more significant defining features in the light of current inconsistent terminology and the recognition that this deposit group may form over a wider range of crustal depths and temperatures than commonly recognized (Groves, 1993; Hagemann and Ridley, 1993; Gebre-. Mariam et al., 1995). The term orogenic is introduced and justified as a term to replace ‘mesothermal’ and other descriptors for this deposit type. It is also suggested that the terms epizonal, mesozonal and hypozonal be used to reflect crustal depth of gold deposition within the orogenic group of deposits.

2. Definition of so-called mesothermal gold deposits

The so-called ‘mesothermal’ gold deposits Table. 1 are a distinctive type of gold deposit which is typified by many consistent features in space and time. These have been summarized in a variety of comprehensive ore-deposit model descriptions that include Bohlke (1982), Colvine et al. (1984), Berger(1986), Groves and Foster (1991), Nesbitt (1991). Hodgson (1993) and Robert (1996). Kerrich (1993) summarizes many of the steps that led to these evolving modern-day models. A unifying tectonic theme has recently been evaluated by workers such as Wyman and Kerrich (1988), Barley et al. (1989), Hodgson and Hamilton (1989), Kerrich and Wyman (1990), Kerrich and Cassidy (1994) and Goldfarb et al.(1998-this issue).

2.1. Geological characteristics

2.1.1. Geology of host terranes

Perhaps the single most consistent characteristic of the deposits is their consistent association with deformed metamorphic terranes of all ages. Observations from throughout the world’s preserved Archaean greenstone belts and most recently-active Phanerozoic metamorphic belts indicate a strong association of gold and greenschist facies rocks. However, some significant deposits occur in higher metamorphic grade Archaean terranes e.g. McCuaig et. al.(1993) or in lower metamorphic grade domains within the metamorphic belts of a variety of geological ages. In the Archaean of Western Australia, a number of synmetamorphic deposits extend into granulite facies rocks Groves et al.(1992). Premetamorphic protoliths for the auriferous Archaean greenstone belts are predominantly volcano-plutonic terranes of oceanic back-arc basalt and felsic to mafic arc rocks. Clastic marine sedimentary rockdominant terranes that were metamorphosed to graywacke, argillite, schist and phyllite host most younger ores, and are important in some Archaean **t**erranes (e.g. Slave Province, Canada).

2.1.2. Deposit mineralogy

These deposits are typified by quartz-dominant vein systems with F3–5% sulfide minerals mainly. Fe-sulfides and F5–15% carbonate minerals. Albite, white mica or fuchsite, chlorite, scheelite and tourmaline are also common gangue phases in veins in greenschist-facies host rocks. Vein systems may be continuous along a vertical extent of 1–2 km with little change in mineralogy or gold grade; mineral zoning does occur, however, in some deposits. Gold:silver ratios range from 10 (normal) to 1 (less common), with ore in places being in the veins and elsewhere in sulfidized wallrocks. Gold grades

are relatively high, historically having been in the 5–30 g/t range; modern-day bulk mining methodology has led to exploration of lower grade targets. Sulfide mineralogy commonly reflects the lithogeochemistry of the host. Arsenopyrite is the most common sulfide mineral in metasedimentary country rocks, whereas pyrite or pyrrhotite are more typical in metamorphosed igneous rocks. In fact, the Salsigne gold deposit in Cambrian sedimentary rocks of the French Massif Central is the world's largest producer of arsenic (Guen et al., 1992). Gold-bearing veins exhibit variable enrichments in As, B, Bi, Hg, Sb, Te and W; Cu, Pb and Zn concentrations are generally only slightly elevated above regional backgrounds.

2.1.3. Hydrothermal alteration

Deposits exhibit strong lateral zonation of alteration phases from proximal to distal assemblages on scales of metres. Mineralogical assemblages within the alteration zones and the width of these zones generally vary with wallrock type and crustal level. Most commonly, carbonates include ankerite, dolomite or calcite; sulfides include pyrite, pyrrhotite or arsenopyrite; alkali metasomatism involves sericitization or, less commonly, formation of fuchsite, biotite or K-feldspar and albitization and mafic minerals are highly chloritized. Amphibole or diopside occur at progressively deeper crustal levels and carbonate minerals are less abundant. Sulfidization is extreme in BIF and Fe-rich mafic host rocks. Wallrock alteration in greenschist facies rocks involves the addition of significant amounts of CO , S, K, H_2O, SiO ±Na and LILE.

2.1.4. Ore fluids

Ores were deposited from low-salinity, near-neutral, H_2O–CO±CH fluids which transported gold as a reduced sulphur complex. Fluids associated with this gold deposit type are notable by their consistently elevated CO concentrations of >5 mol%. Typical $\delta^{18}O$ values for hydrothermal fluids are about 5–8 per ml in the Archaean greenstone belts and about 2 per ml higher in the Phanerozoic gold lodes.

2.1.5. Structure

There is strong structural control of mineralization at a variety of scales. Deposits are normally sited in second or third order structures, most commonly near large-scale often transcrustal compressional structures. Although the controlling structures are commonly ductile to brittle in nature, they are highly variable in type, ranging from: (a) brittle faults to ductile shear zones with low-angle to high-angle reverse motion to strike-slip or oblique-slip motion; (b) fracture arrays, stockwork networks or breccia zones in competent rocks; (c) foliated zones pressure solution cleavage or d fold hinges in ductile turbidite sequences. Mineralized structures have small syn- and post-mineralization displacements, but the gold deposits commonly have extensive down-plunge continuity hundreds of metres to kilometres. Extreme pressure fluctuations leading to cyclic fault-valve behavior (Sibson et al., 1988) result in flat-lying extensional veins and and mutuallycross-cutting steep fault veins that characterize many deposits (e.g. Robert and Brown, 1986).

2.2. Tectonic setting and timing of 'mesothermal'vein emplacement

The so-called 'mesothermal' gold deposits Table.1 occupy a consistent spatialrtemporal position Fig. 1 , having formed during deformational processes at convergent plate margins orogeny irrespective of whether they are hosted in Archaean or

Proterozic greenstone belts or Proterozoic and Phanerozoic sedimentary rock sequences (e.g. Barley and Groves,. 1992; Kerrich and Cassidy, 1994). The placing of these deposits in a plate tectonic setting was a logical outgrowth of the acceptance of plate tectonic theory in the early 1970's. Guild (1971) initially discussed the ''orogen-associated endogenic mineral deposits of Mesozoic and Tertiary age on the sites of Cordilleran-type continentrocean collisions''. Sawkins (1972) noted, soon after, how both these Circum-Pacific gold ores and spatially associated felsic magmas were probable products of subduction-related tectonism. Just as significant was. Sawkins 1972 observation that Archaean gold lodes in the Superior Province, Canada, may have some relationship to the southward younging of igneous ages, interpreted as being reflective of a seawardmigrating trench. It would be, however, another sixteen years cf. Wyman and Kerrich, 1988 before workers would follow-up on this important concept and begin to widely look at Archaean gold as a product of continental-margin deformational events.

The concept of a general spatial association between the gold deposits and subduction-related thermal processes in accretionary orogens oceanic-continental plate interactions became commonplace in the mid-1980's. Fyfe and Kerrich 1985 presented a model at that time to explain the massive fluid volumes required for the numerous gold-bearing vein swarms adjacent to crustal-scale thrust zones of continental margins. They hypothesized that underplated hydrated rocks contained the required water and such water was released during thermal reequilibration as subduction ceased. Subsequent models for the Mesozoic and Cenozoic gold fields of western most North America relied heavily on correlating gold vein emplacement with subduction-driven processes (Bohlke and Kistler, 1986; Goldfarb et al., 1988). Landefeld (1988) expanding on the ideas in (Fyfe and Kerrich 1985), detailed how the seaward stepping of subduction accompanying terrane accretion could have been crucial for the formation of the Sierra foothills gold districts including the Mother lode belt. With an abundance of new geochronological data from western North America, recent models of gold genesis in accretionary orogens have been able to look closely at specific processes e.g. changing plate motions, changing collisional velocities, ridge subduction, etc. occurring during accretionrsubduction that tend to be most closely associated with veining (e.g. Goldfarb et al., 1991b; Elder and Cashman, 1992; Haeussler et al., 1995). Theoretically, as a subduction zone steps seaward, a series of gold systems and plutonic bodies should develop and young towards the trench-part of a so-called Turkic-type (Sengor and Okurogullari, 1991) orogen. This type of scenario crudely characterizes Alaska, USA, a part of the North American margin almost entirely composed of accreted oceanic rock sequences (Plafker and Berg, 1994).

Collisional orogens continent–continent collision, including the Variscan, Appalachian and Alpine, also are host environments for gold deposits. In fact, collisional or internal and accretionary or peripheral orogens may represent end-members of a continuous process. Any continent–continent collision will be preceded by closure of an ocean basin, and hence is nothing more than a final stage of a peripheral orogen. The gold systems that are associated with the Phanerozoic internal orogens are actually all spatially associated with marine rocks that have been caught up within the suture. In addition, within peripheral orogens, accretion of microcontinents such as Wrangellia

along western North America （Plafker and Berg, 1994） or Avalonia along Laurentia （Keppie, 1993） may be viewed as a type of small-scale continent–continent collision. A key point in all examples is that hydrated marine sedimentary and volcanic rocks were added to continental margins and, at some time during this growth, the accreted rocks experienced relatively high geothermal gradients.

Oligocene veins in the western European Alps (Curti, 1987) are the youngest recognized, economic examples of this deposit type. They also serve to point out that more than simple plate subduction is required for vein formation. The closure of an ocean basin between Europe and Adria perhaps a part of northern Africa occurred during an 80-m.y.-long period of Early Cretaceous–early Tertiary oceanic crust subduction without any preserved evidence of gold veining or magmatism; blueschist metamorphic facies in the Alps now record the low thermal gradients. By the early Eocene, complete closure of the ocean had led to continent–continent collision and a partial subduction of the European continental margin between 55 and 45 Ma Blanckenburg and . Davies, 1995 . It was not until almost 100 m.y. subsequent to the onset of convergence, perhaps due to slab delamination resulting in the cessation of subduction at 45–40 Ma (Blanckenburg and Davies, . 1995), that magmatism and high temperature metamorphism impacted the obducted upper nappes of the western Alps near the collisional suture. Much of the Alpine gold veining occurred during the early. Oligocene peak of magmatism (Curti, 1987).

The understanding of gold-forming processes and timing in older Phanerozoic orogens may be complicated by the hundreds of millions of years of additional geological time, but certainly such Palaeozoic continental margins were favorable environments for veining. Geochronological study of the gold deposits in the Meguma terrane of Nova Scotia, Canada, indicates veining between 380 and 362 Ma (Kontak et al., 1990), during the late part of Acadian deformation of the Appalachian orogen. The Meguma was the final terrane accreted to the Atlantic margin during the poorly-understood late Palaeozoic Laurentia–Gondwanaland collision. Keppie and Dallmeyer (1995) noting that magmatism and high-temperature metamorphism were restricted to a narrow time range of about 380–370 Ma, rather than the prolonged 100 m.y. of Meguma collision, suggest a distinct episode of lower lithospheric delamination for the thermal perturbation. This brief thermal event, occurring at the same time as gold veining, is also likely to be important to the ore-forming process. Whereas little is certain about the subduction-related tectonics of the northern Appalachians, mesothermal-type gold ores such as the Hammer Down in northwestern. Newfoundland Gaboury et al. (1996) indicate that a broad belt of gold systems accompanied continental growth.

Palaeozoic gold veins of the Tasman orogenic system in eastern Australia make it clear that the ore-forming process need not require a ‘Cordilleran style’ of terrane accretion. Unlike the collage of small terranes that formed the accreted margin of western North America, eastern Australia is mainly composed of a single lithotectonic assemblage (the Lachlan ‘terrane’) that represents a 2,000-km-wide Palaeozoic turbidite fan sequence developed adjacent to the Gondwanan craton Coney(1992). Such an environment lacks deep-crustal terrane-bounding faults located between accreted material and the active margin, which, where present in the North American Cordillera,

expose a variety of crustal levels and often serve as the focus of hydrothermal fluid flow. Compression-related deformation is solely intraplate rather than concentrated along sutures between terranes. The fact that such a large percentage of gold has been concentrated in the Bendigo–Ballarat area of Victoria (Phillips and Hughes, 1996; Ramsay, 1998 - this issue) indicates some significant and still poorly-understood, local control on vein emplacement in the orogenic system. Nonetheless, similar to the North American Cordillera, the Tasman orogenic system is characterized by significant growth of the eastern Australian margin addition of the Lachlan 'terrane' and a subduction zone east of the Lachlan assemblage throughout much of the Palaeozoic (Solomon and Groves, 1994).

The abundance of geological similarities between the gold ores of the Phanerozoic orogens and those in Archaean greenstone belts began to be interpreted by the late 1980's as evidence of a similar tectonic setting for ore formation. Wyman and Kerrich 1988 hypothesized that gold mineralization in the Superior Province of Canada was ''related to convergent plate margin-style tectonics''. At roughly the same time, Barley et al. (1989) independently reached the same conclusion to explain the development of gold lodes in Western Australia. Subduction of oceanic rocks into the zone of partial melting appeared to be significant in the development of these gold ores within orogens of all ages (Hodgson and Hamilton, 1989). Major fault zones spatially associated with auriferous belts in the Archaean terranes were now recognized by several researchers as ancient terrane boundaries. Kerrich and Wyman (1990) pointed out that, as observed in present-day convergent margins, Archaean ore-forming fluids were products of deeper crustal thermotectonic events which occurred subsequent to magmatism and metamorphism in ore-hosting supracrustal rocks. Detailed geochronological studies now recognize such lower- to mid-crustal, late deformational regimes in Archaean terranes (Jackson and Cruden, 1995; Kent et al., 1996). Gold deposits in any given Archaean province may all be a part of the same supercontinent cycle (cf. Barley and Groves, 1992), but can show a wide variation in age Table 1, reflecting a variety of thermal events during many tens of millions of years of accretion and subduction.

2.3. Crustal enĺironment of 'mesothermal' gold deposition

The majority of deposits of this ore style are sited in ductile to brittle structures, have proximal alteration assemblages of Fe sulfide–carbonate–sericite ±albite (in rocks of appropriate composition to stabilise the assemblage) and were deposited at 300±50°C and 1–3 kbar, as indicated by fluid inclusion and other geothermobarometric studies (Groves and Foster, 1991; Nesbitt, 1991). They are consistently syn- to post-peak-metamorphic and were emplaced at temperatures generally within 100°C of peak metamorphic temperatures experienced by the surrounding host rocks. However, recent studies in mainly Archaean greenstone belts have extended the ranges of temperature and pressure, and hence extended the inferred crustal range of formation of the deposits into higher- and lower-grade metamorphic rocks e.g. the crustal continuum model of Groves (1993). The evidence for formation of these gold deposits over P–T ranges of about 180–700°C and <1–5 kbar (Groves, 1993); Hagemann and Brown, (1996); Ridley et al.(1996) implies vertically extensive hydrothermal systems that contrast sharply with other continental-margin gold systems that are apparently restricted

to the upper 5 km or so of crust Fig. 2 .

Studies in the early 1990's, summarized in McCuaig et al.(1993), identified higher P–T examples of these gold ores in amphibolite facies terranes of Western Australia, the Superior and Slave Provinces in Canada, India and Brazil. Most such mineralization occurred between 450–600^{o}C and 3–5 kbar. A few examples in granulite terranes formed at even higher P–T regimes (Barnicoat et al., 1991; Lapointe and Chown, 1993). The gold ores were still precipitated from the same low salinity, CO - and δ^{18}O-rich fluids, but, because of the higher temperatures and different mineral stabilities, there is a scarcity of carbonate phases and an abundance of calc-silicate minerals characterizing alteration haloes (Mikucki and Ridley, 1993). Such assemblages are similar to those typifying skarn systems (Mueller and Groves, 1991).

It is somewhat problematic as to why a similar continuum of gold deposits has not been widely recognized in higher metamorphic-grade portions of Phanerozoic orogenic belts. Was there something inherently different between the tectonics of Archaean and Phanerozoic continental growth? Or do such gold deposits occur in high-grade terrains of the Phanerozoic and they have just been classified differently? Perhaps a re-evaluation of the classification of some of the gold-bearing 'skarns' or contactmetamorphosed deposits in younger orogenic belts might help to solve this problem. Ore fluid salinity might be a key discriminator in the case of the skarns, with relatively high ore-fluid salinities being associated with typical gold skarn deposits that are more directly linked to intrusive sources (Meinert, 1993). The late Palaeozoic Muruntau deposit in Uzbekistan is apparently one example of a post-Archaean, higher metamorphic grade 'mesothermal type' deposit. The abundance of thin quartz layering, fluid inclusion data suggesting trapping temperatures in excess of 400^{o}C Berger et al., 1994 and a skarn-like, calc-silicate assemblage Marakushev and Khokhlov (1992) from deeper parts of the ore system all suggest that the deposit may represent a deeper part of the crustal continuum.

Ore formation at temperatures of 200–25^{o}8C and at crustal depths of only a few kilometers is not uncharacteristic of these ores where hydrothermal fluids have migrated to shallower crustal levels. However, a few anomalies from shallow gold systems in the Yilgarn block of Western Australia are notable. Comb, cockade, crustiform and colloform textures at the Racetrack deposit, deposited from CO_2 -poor fluids in lower greenschist facies rocks at depths </=2.5 km, are more like those developed in classic epithermal vein deposits (Gebre-Mariam et. al., 1993). Similar textures at the Wiluna gold deposits in subgreenschist facies rocks, as well as δ^{18}O measurements as light as 6–7 per ml, quartz provide some of the strongest evidence of meteoric water involvement in some of the 'mesothermal' hydrothermal systems (Hagemann et al., 1992, 1994).

Gold solubility relationships at temperatures below 200–250^{o}C best explain the observation that the continuum of this type of gold deposit does not continue into the uppermost few kilometres of the crust. The moderately-reducing and only moderately sulphur-rich conditions likely to characterize ' mesothermal' gold ore-fluids at low temperature (Mikucki, 1998 - this issue), would favor low gold solubilities at these low temperatures e.g. Shenberger and Barnes (1989). However, hydrothermal fluids that have been depositing 'mesothermal' gold along crustal-scale fault zones at depth, must still advect along these faults to the surface. Such is probably the case in the westernmost part

of North America where CO_2 -rich, isotopically-heavy fluids migrated to near-surface environments of very low P–T in the Cordilleran orogen. Cinnabar"stibnitebearing epithermal, silica–carbonate veins, which were deposited within the upper few kilometres of crust, define such flow (Nesbitt and Muehlenbachs, 1989). Examples include the Hg–Sb deposits of the Kuskokwim basin in SW Alaska, the Pinchi belt of British Columbia and the coast ranges of northern California. In fact, it has been recognized now for thirty years that many of the thermal springs within the accreted margin of western North America have a unique chemical character (White, 1967) and could be the surface expression of deeper 'mesothermal' gold deposits. $\delta^{18}O$ values for Hg-rich veins quartz emplaced in the near surface are as heavy as +30 per ml because of greater quartz–water fractionation, as temperatures of ore fluids cooled to as low 150^{o}C.

Such heavy oxygen values are very distinct from $\delta^{18}O_{quartz}$ values of other types of vein systems deposited in classical epithermal environments, such as those of the Nevada Basin and Range (Goldfarb et. al., 1990). The identification of this type of Hg–Sb epithermal system in a continental margin terrane with limited erosion may be a valuable guide to the down-dip existence of a so-called 'mesothermal' gold occurrence.

2.4. Comparisons with other lode-gold deposit types

Most deposit types that contain ore-grade gold Table 2 , whether with gold as the principal metal or together with copper, are sited along convergent plate margins (Sawkins, 1990). There are notable exceptions, such as gold-rich volcanogenic massive sulfide deposits developed along spreading ocean ridges e.g. Bousquet and other deposit styles associated with possible anorogenic hot spots (e.g. Olympic Dam). However, as a rule, many of the Phanerozoic gold-bearing epithermal vein, Carlin Carlin type sedimentary rock-hosted and porphyryrskarn deposits developed within the same active continental margins as the so-called 'mesothermal' deposits Fig. 1. Notable distinctions, however, can be made that relate to local changes in tectonism within a developing orogen and to crustal depth range a reflection of regional geothermal gradient of the auriferous hydrothermal systems.

As shown schematically in Fig. 1, a significant proportion of epithermal and porphyry deposits are distinct in that they form above subduction zones distal to continental margins or within continental margins, but during post-collisional extension. Many other gold-rich epithermal and porphyry systems develop in oceanic regimes within the top few kilometers of crust of volcano-plutonic island arcs located above intermediate- to steeply-dipping subduction zones (e.g. Sawkins, 1990; Sillitoe, 1991), with a vertical transition from porphyry-style to classic epithermal vein-style mineralization e.g. White and Hedenquist (1995). Other epithermal lodes, including some of the world-class deposits (Muller and Groves, 1997) are associated with alkalic, mantle-related rocks that reflect extensional episodes in a convergent orogen in either a near-arc region (e.g. Porgera: Richards et al., 1990) or far inland of the accretionary wedge (e.g. Cripple Creek: Kelley et al., 1996). Certainly, many of the well-studied Tertiary epithermal ores associated with volcanic rocks throughout Nevada are products of post-orogenic Basin and Range extension. Geochronological evidence is beginning to favour a similar temporal setting for Carlin-type mineralization (Hofstra, 1995; Emsboo et al., 1996).

The gold-bearing epithermal vein and porphyry systems that are, however, associated with collisional, subduction-related tectonics (Sillitoe, 1993) are typically located in different crustal regimes in the orogen than the so-called ‘mesothermal’ gold systems. Whether in an island arc, compressional orogen, or a zone of back-arc rifting, the porphyry skarn-epithermal vein continuum normally is telescoped into the upper 2–5 km of crust (Figs. 1 and 2; Poulsen, 1996). Magmatism generally I-type and high temperatures impose a very steep geothermal gradient on the upper crust, often locally far in excess of 100^{o}Crkm. An abundance of subvolcanic to volcanic rocks necessitates that much of the gold ore is hosted in lithologies of roughly equivalent age. The shallow level of the hydrothermal activity restricts much of the lode-gold emplacement to rocks that are unmetamorphosed to only slightly regionally metamorphosed.

In contrast, the so-called ‘mesothermal’ ore deposit type is deposited over a very broad range of the upper crust (Groves, 1993; Poulsen, 1996). Rather than bringing a concentrated heat source to the near surface, the fluids, granitic magmas and heat are carried to higher crustal levels along major fault zones that may have been suture zones between accreted terranes. Crustal geotherms of perhaps <30^{o}C/km are elevated, but not to the levels of the more telescoped group of ore deposit types. Where hydrothermal fluids reach the near-surface environment, their relatively low temperature hinders significant gold transport; however, bisulphide complexes still may carry significant Sb and Hg into the upper few kilometres of crust Fig. 2 . Where such fluids migrate into the realm of the typical porphyry skarn-epithermal continuum, complex overlapping of deposit styles may develop. Such a situation may characterize southwestern Alaska, where epithermal Hg–Sb ores that suggest so-called ‘mesothermal’ gold deposits at depth (Gray et al., 1997) are spatially associated with volcano-plutonic-related gold deposits (Bundtzen and Miller, 1997), or northern California where the McLaughlin gold deposit sits among a series of Hg-rich hot springs (Sherlock and Logan, 1995).

3. Problem of nomeclature

Prior to 1980, the so called ‘mesothermal’ group of Archaean through Tertiary deposits was not widely recognized as a single special type of gold ore. Most classifications scattered the deposits among the mesothermal and hypothermal regimes of Lindgren (1933). Others, such as Bateman (1950), divided these deposits into groups within a very broad ‘cavity filling’ type of epigenetic ore deposit. Hence, many Archaean lodes were classified as fissure filling type deposits, Otago was a shear zone deposit type, Bendigo was a saddle reef deposit type, Treadwell, Alaska was a stockwork type deposit, etc. The relatively low price of gold correlated with a limited research interest in gold genesis studies. In fact, in the 75th Anniversary Volume of Economic Geology 1981 , there is notably no chapter that is devoted to this economically important ore deposit type. Economic geologists had begun to notice the basic association of the Phanerozoic deposits with subduction zones and convergent margins during the growth of plate tectonic theories. However, books on tectonics and ore deposits barely mentioned these gold systems (e.g. Mitchell and Garson, 1981).

As the price of gold increased dramatically in the late 1970’s, so did interest in the understanding of these gold deposits. ‘Mesothermal’ lode-gold deposits began to receive

extensive study by ore geologists, and were subsequently described by a variety of terms during the last fifteen years as workers began recognizing them as a single mineral deposit type. The abundance of terms that define these ores reflects both the great expansion of knowledge about these systems accumulated during the 1980's (e.g. Robert et al., 1991) and the efforts by various groups to establish ore deposit model volumes that classify deposits by type (e.g. Cox and Singer, 1986). One consequence of so many terms for the same deposits is the resulting confusion for those not extremely familiar with the gold literature. Certainly, a single deposit type title would be helpful for all workers.

The paper by Nesbitt et al. (1986) on lode-gold deposits of the Canadian Cordillera seemed to initiate popularity of the phrase mesothermal. They define a group of Canadian 'mesothermal' gold deposits that formed between 200–350^{o}C within a series of accreted terranes. Prior to this paper, the broad class of 'mesothermal' gold deposits did not exist. Major gold volumes such as 'Gold '82' (Foster, 1984), 'Turbidite-hosted Gold Deposits' (Keppie et al., 1986) and 'Gold '86' (Macdonald, 1986) lacked any mention of such a deposit type. However, since the Nesbitt et al. (1986) paper, the mesothermal' terminology has become well-entrenched in the literature. This may be a response, in part, to the need to easily contrast this group of gold deposits with the generally more shallowly-deposited types of gold ores that had already been classified as' epithermal' for many years previous. Because of this widespread acceptance of the mesoth ermal label, subsequent comprehensive descriptions of these gold deposits have tended to group them under such a ' mesothermal heading' (Groves et al., 1989; Kerrich,. 1991; Hodgson, 1993).

Whereas 'mesothermal' has become the most common term used in referring to this type of deposit during the last ten years, Poulsen (1996) has recently shown how it is very inconsistent with the meaning originally proposed by Lindgren (1907, 1933). Lindgren's description of such a deposit type is for that which formed at depths of about 1,200–3,600 m and at temperatures of 200–300^{o}C. Because of the restrictive temperature range, high-temperature alteration phases, including tourmaline, biotite, hornblende, pyroxene and garnet, were stated as being absent in and surrounding mesothermal type ores. Gold districts such as those of the California foothills belt, the Meguma domain of Nova Scotia, central Victoria, and Charters Tower in Queensland were classified by Lindgren (1933) as mesothermal.

If one such Lindgren-type term was used to define the broad observed range for P–T conditions of these deposits, it probably is 'xenothermal'. The term, coined by Buddington 1935 , covers the P–T conditions from lepothermal a vague P–T regime between epithermal and mesothermal to hypothermal. As such, it would include the broad range of ore forming pressures and temperatures that is welldocumented in the Yilgarn block of Western Australia, as summarised by Groves (1993). However, other factors, such as structural control, wall rock type and fluid chemistry play a major role in the localization of a gold deposit and definition of a gold deposit type solely on P–T environment is not advisable (Bateman, 1950).

The contrasting tectonic setting between the sites of most 'epithermal' gold deposits and the sites of all so-called 'mesothermal' deposits presents another basic problem with usage of the Lindgren model. As envisioned by Lindgren (1907, 1933), the epithermal,

mesothermal and hypothermal terms were intended to define a continuum among deposits. However, as implied in Fig. 2, the term 'epithermal' is now entrenched in the literature as a specific mineral deposit type that most commonly describes high-level veining and alteration broadly associated with volcanism or subvolcanic magmatism (e.g.. Berger and Bethke, 1985). As discussed above, such epithermal gold deposits may form in oceanic arcs long before continental margin orogenesis or, as in the Basin and Range of the USA, during post-orogenic extension, as shown schematically in Fig. 1. Hence, there are typically neither consistent spatial nor temporal relations between the two gold deposit types.

Many other terms relating to host rocks, vein mineralogy or ore-fluid chemistry are equally unacceptable in the overall description of these deposits. Commonly used terms, such as 'greenstone gold', ' slate belt gold', or 'turbidite-hosted gold', disguise the fact that the deposits have many similarities despite their different hosting sequences the theme of this special Ore Geology ReĺIews issue and should be used, if at all, to describe subgroups of the major deposit type, and not the deposit type itself. The use of 'Archaean' or 'Mother lode-type' gold deposits is also unacceptable, clearly reflecting a specific temporal or spatial preference, respectively. Metamorphic gold' implies an understanding of the ore-forming process which is, however, still strongly under debate. The fact that these deposits contain only a few percent sulfide minerals, in most cases, has led to classifications referring to them as 'low sulfide' (Berger, 1986), and the fact that gold is enriched by orders of magnitudes over base metals and Au:Ag ratios are generally)1 has led to their classification as 'gold only' (Hodgson and MacGeehan, 1982; Phillips and Powell, 1993) deposits. However, many other types of gold deposits, including the sedimentary rock-hosted ores at Carlin and elsewhere in Nevada, show the same low sulfide content. Similarly, 'lode-gold' (McCuaig and Kerrich, 1994) may be interpreted to contain a variety of gold deposit types.

A critical feature of all these deposits seems to be their common tectonic setting, as described in detail above. These deposits were classified as 'pre-orogenic' by Bache (1980, 1987), who recognized their association with the world's orogenic belts. However, at the same time, the classification assumed a syngenetic exhalative origin for the auriferous lodes, an assumption clearly in conflict with modern geochronological data. Goldfarb et al. (1991a, 1998 -this issue) have often preferred the term 'synorogenic', given the clear overlap of gold-forming events in the North American Cordillera with a broad, 120-m.y.-long period of continental margin growth. The term 'post-orogenic' has been used by other workers (Gebre-Mariam et al., 1993; Groves, 1996) who emphasize that deformation and metamorphism of ore host rocks commonly predate hydrothermal vein emplacement (Groves et al., 1984; Colvine, 1989; Hodgson and Hamilton, 1989).

4. Proposed classification

These gold deposits, throughout the world's collisional orogenic belts, can actually be viewed as both syn- and post-orogenic in origin. Whereas host rocks for ore may already be undergoing uplift and cooling thus 'post-orogenic' , the ore-forming fluids may be generated or set in motion by simultaneous thermal processes at depth thus 'syn-orogenic' as described by Stuwe et al. (1993). For example, Kent et al. (1996) show

that the main episode of gold mineralization in the Yilgarn craton postdates thermal events in the ore-hosting upper crust, but temporally correlates with melting and magmatism of lower-middle Archaean crust. Because of this, it is suggested that the gold ores simply be classified as 'orogenic' lode types, as was originally suggested by Bohlke 1982 . A remaining problem is whether to classify many ' intrusion-related gold deposits' within this group of orogenic gold deposits. Sillitoe 1991 places deposits such as Muruntau and Charters Tower in such an intrusion-related deposit type. McCoy et al. (1997) distinguish 'plutonic-related mesothermal gold deposits' of interior Alaska, such as Fort Knox, as those where ore fluids are derived from evolving magmas. The Proterozoic gold lodes of northern Australia and the Mesozoic deposits of the north China craton and Korea are also commonly suggested to be genetically associated with igneous processes. Are such deposits, with ore fluid chemistries essentially identical to those of typical orogenic gold deposits, a different deposit type? Sillitoe (1991) indicated that the intrusion-related gold deposits also form in Phanerozoic convergent plate margins above zones of active subduction, although regional extension is stressed as an important characteristic and thus indicates some difference from the orogenic class defined here. Sillitoe 1991 does stress that the apparent overlap between orogenic and intrusion-related gold systems requires further attention. We would certainly agree.

A convenient terminology that both retains the prefixes 'epi', 'meso', and 'hypo' used by Lindgren (1907, 1933), and subdivides the orogenic gold deposit type, is introduced by Hagemann and Ridley (1993) and then further modified by Gebre-Mariam et al. (1995). Its continued usage is recommended. In such a scenario, epizonal deposits form within 6 km of the surface at temperatures of 150–300^{o}C, mesozonal deposits form at depths of 6–12 km and at temperatures of 300–475^{o}C and hypozonal deposits form below 12 km and at temperatures exceeding 475^{o}C. It is critical to note that this terminology has been defined solely as a subdivision for orogenic gold deposits based on many modern geothermobarometric studies. Because of this, the depth zones for these orogenic subclasses do not correspond to those in Lindgren's epithermal, mesothermal, and hypothermal regimes.

New words and expressions

orogenic gold deposit 造山型金矿床
mesothermal gold deposits 中温金矿床
compressional 压性的
transpressional 压剪性的
epizonal （6 km）浅成带
mesozonal（6–12 km）中成带
hypozonal（12 km）下部带
lodegold 矿脉
porphyry 斑岩
skarn 矽卡岩
graywacke 硬砂岩
argillite 泥岩、泥质板岩
Palaeozoic 古生代，古生代的
phyllite 千枚岩，硬绿泥石
Phanerozoic 显生宙的
arsenopyrite 毒砂, 含砷黄铁矿
Pyrrhotite 磁黄铁矿
sericitization 绢云母化(作用)
Fuchsite 铬云母
albitization 钠长石化
sulfidization 硫化
Proterozoic 元古代
Archaean 太古代的
endogenic 内成的,内生的
Appalachian 阿巴拉契亚（美国东部一地区）