Building Terrestrial Planets

A. Morbidelli,¹ J.I. Lunine,² D.P. O’Brien,³ S.N. Raymond,⁴ and K.J. Walsh⁵

¹Department of Lagrange, Université de Nice-Sophia Antipolis, Observatoire de la Côte d’Azur, CNRS, 06304 Nice Cedex 4, France; email: morby@oca.eu
²Department of Astronomy, Cornell University, Ithaca, New York 14853; email: jlunine@astro.cornell.edu
³Planetary Science Institute, Tucson, Arizona 85719; email: obrien@psi.edu
⁴Université de Bordeaux, Observatoire Aquitain des Sciences de l’Univers, CNRS, UMR 5804, Laboratoire d’Astrophysique de Bordeaux, F-33271 Floirac Cedex, France; email: rayray.sean@gmail.com
⁵Department of Space Studies, Southwest Research Institute, Boulder, Colorado 80302; email: kwalsh@boulder.swri.edu

Keywords
protoplanetary disk, planetesimals, meteorites, giant impacts, giant planet migration, water

Abstract
This article reviews our current understanding of terrestrial planet formation. The focus is on computer simulations of the dynamical aspects of the accretion process. Throughout the review, we combine the results of these theoretical models with geochemical, cosmochemical, and chronological constraints to outline a comprehensive scenario of the early evolution of our solar system. Given that the giant planets formed first in the protoplanetary disk, we stress the sensitive dependence of the terrestrial planet accretion process on the orbital architecture of the giant planets and on their evolution. This suggests a great diversity among the terrestrial planet populations in extrasolar systems. Issues such as the cause for the different masses and accretion timescales between Mars and Earth and the origin of water (and other volatiles) on our planet are discussed in depth.
1. INTRODUCTION

Although we have more information about the terrestrial planets than about virtually any other celestial bodies, the processes leading to their formation have long remained elusive. Only in the past few decades has there been significant progress. On the one hand, geochemical and cosmochemical analyses performed with laboratory instruments of unprecedented precision have produced a huge amount of data on the chemical and isotopic composition of the planets and of their precursors—meteorites—as well as constraints on the chronology of their accretion and thermal evolution. On the other hand, the remarkable increase in computer performance has allowed modelers to undertake increasingly realistic simulations of the dynamical process of terrestrial planet accretion. The results achieved on each front have reached a sufficient level of reliability to be integrated in a comprehensive view of the early evolution of the inner solar system. This review focuses on the key processes as seen from an astrophysical point of view, but we make reference to geochemical, cosmochemical, and astronomical constraints. Our aim is that this review will be useful to both theorists and observers in confronting the results of the modeling with laboratory and observational constraints.

Terrestrial planet formation occurred through three distinct modes of growth that were ordered in time in the protoplanetary disk. An observer privileged to watch this process live would see a distinct distribution of gas and solids, of particle sizes, and of orbits of solid bodies in each of the three steps; they are potentially distinguishable from one another through disk observations by the next generation of powerful ground- and space-based observatories. In step I, planetesimals are formed in a disk of gas and dust. In step II, the collisional evolution of the planetesimal population leads to the growth of a new class of objects named planetary embryos, which represent an intermediate stage between planetesimals and planets. Giant planets probably form at roughly the same time as do planetary embryos. In step III, after the disappearance of gas from the protoplanetary disk, the embryos’ orbits become unstable, and their mutual collisions give birth to a small number of massive objects, the terrestrial planets. These steps are described in Sections 2–4. Section 5 details how the results of the simulations change depending on various assumptions (e.g., the mass distribution in the disk, the giant planet orbits, and the outcome of collisions). Section 6 discusses a new model that aims to link in a coherent scenario the dynamical evolution of the giant planets with the formation of the terrestrial planets, with the specific goal of explaining the small mass of Mars relative to Earth. Section 7 discusses the origin of water on Earth and other chemical implications of the terrestrial planet accretion models, and Section 8 details future observing capabilities that might test these ideas.

We dedicate this review to G.W. Wetherill, who was the first to investigate terrestrial planet formation by combining dynamical simulations and geochemical/cosmochemical constraints. He traced the path that we are continuing to follow here.

2. STEP I: FROM DUST TO PLANETESIMALS

When a molecular cloud collapses under its own gravity to form a star, the material forms a disk-like structure in orbit around the central object owing to the conservation of angular momentum. These protoplanetary disks are now routinely observed around pre-main-sequence stars (e.g., Kenyon & Hartmann 1995, McCaughrean & O’Dell 1996), whereas that for the Sun—termed the solar nebula—is understood from the physical and chemical evidence left behind in our solar system.

In protoplanetary disks, dust grains sediment into a thin layer at the midplane of the disks (Weidenschilling 1980). The sedimentation timescale and the volume density of grains near the
midplane depend on the severity of turbulence in the disk gas, as strong turbulence inhibits sedimentation. Even in laminar disks, though, the volume density of grains in the midplane could not reach arbitrarily large values because the sedimentation of an excessive amount of solids would itself generate turbulence in the disk via the so-called Kelvin-Helmholtz instability (Weidenschilling 1995).

The growth from these grains to kilometer-size planetesimals is still quite a mystery. In principle, one could expect that grains, once sufficiently concentrated near the midplane, should stick to one another to form progressively larger objects in an ordered-growth process. However, particles of centimeter size are too small for gravity to be effective in particle-particle collisions and too big to stick together through electrostatic forces. Moreover, grains are subject to gas drag, which makes them drift toward the central star (Weidenschilling 1977). The drift speed is size dependent; thus, particles of different sizes should collide with non-negligible relative velocities on the order of several meters per second. At these speeds, particles should break rather than coagulate (but see Wettlaufer 2010). Because the drift speed toward the central star is maximal for meter-size boulders, this issue is known as the meter-size barrier problem, even though it is likely that the fragmentation bottleneck for accretion starts at much smaller sizes (centimeters or decimeters).

A new alternative to this ordered-growth process is that planetesimals form as a result of the collective gravity of massive swarms of small particles, concentrated at some locations (local maxima of the gas density distribution or intervortex regions, depending on particle sizes) by the turbulence of the disk (Cuzzi et al. 2008, Johansen et al. 2007). These models, termed gravoturbulent models, can explain the formation of planetesimals of 100-km size or larger without passing through intermediate small sizes, thereby circumventing the meter-size barrier problem. The size distribution of objects in the asteroid belt and in the Kuiper belt, where most of the mass is concentrated in 100-km objects, supports this scenario (Morbidelli et al. 2009; but see Weidenschilling 2010). The existence and the properties of binary Kuiper belt objects also are best explained by the gravitational collapse of massive swarms of small particles that have angular momenta too large to form single objects (Nesvorny et al. 2010b). Although more work is needed to explore all the facets of this novel view of planetesimal formation, it seems to resolve many of the outstanding problems that have plagued other models.

In the gravoturbulent models, once enough small particles are concentrated at some location, the formation of a planetesimal is extremely rapid (Cuzzi et al. 2008, Johansen et al. 2007). However, the formation of self-gravitating clumps of small particles is sporadic (Chambers 2010, Cuzzi et al. 2010); hence, planetesimal formation can, in principle, proceed over an extended period of time. Moreover, one should not assume that planetesimal formation starts at the same time zero in every region of the protoplanetary disk, where time zero is usually identified with the formation time of the first solids, the calcium-aluminum inclusions [CAIs, dated at 4.568 billion years ago (Gya); Bouvier et al. 2007, Burkhardt et al. 2008]. Sufficient clumping of small particles to form planetesimals is possible only if the solid/gas density ratio is larger than some local threshold value (Johansen et al. 2009). This ratio, in principle, increases with time owing to the progressive removal of gas from the disk (Throop & Bally 2005). Thus, in some parts of the disk—for instance, the innermost regions—this condition might have been met early, thus leading to a first generation of planetesimals rich enough in short-lived radionuclides (principally $^{26}$Al and $^{60}$Fe; Ghosh et al. 2006) to undergo melting or metamorphism of the rocky component. Farther out in the disk, this condition might have generally been met late enough that the remaining quantities of short-lived radionuclides were too small to allow thermally driven melting or metamorphism of 100-km-scale planetesimals. The threshold for differentiation of 100-km planetesimals might have been reached inward of (Bottke et al. 2006), or even within (Ghosh et al. 2006), the asteroid belt.
In an idealized chemical model, planetesimals should accrete all elements and molecules that are in condensed form at their corresponding locations of the disk. Because at a given time the temperature in the disk decreases with distance from the Sun, a classic condensation sequence characterized by a clear radial gradient of chemical properties should be obtained (e.g., Dodson-Robinson et al. 2009). However, in graviturbulent models, planetesimals form sporadically, so that at any given location they need not have formed at the same time. This effect probably makes the radial gradient of chemical compositions in the planetesimal disk less “clean” than one would predict from a purely condensation-driven sequence.

Observational constraints are available to deduce properties and confront models of the planetesimal disk. There are three classes of chondritic meteorites: enstatite, ordinary, and carbonaceous. Their chemistry and mineralogy suggest that these three classes formed at decreasing temperatures: enstatites at the warmest, followed by ordinary and then carbonaceous. For instance, water is essentially absent on enstatite meteorites (Hutson & Ruzicka 2000) and quite abundant in (some subclasses of) carbonaceous chondrites, whereas the water content in ordinary chondrites is between the two (Robert 2003). Spectroscopic observations link these three classes of meteorites to asteroids of different taxonomic type (although not necessarily to specific parent bodies): Enstatite chondrites can be linked with E-type asteroids (Fornasier et al. 2008), which are predominant in the Hungaria region at 1.8 astronomical units (AU; 1 AU is the mean Sun–Earth distance); ordinary chondrites are linked to S-type asteroids (Binzel et al. 1996), which are predominant in the inner belt (2.1–2.8 AU); and carbonaceous chondrites are linked to C-type asteroids (Burbine et al. 2000), which are predominant in the outer belt (beyond 2.8 AU). There is, however, substantial overlap in the radial distributions of these three types of asteroids (Gradie & Tedesco 1982). This could be the result of the formation of subsequent generations of asteroids in a cooling disk, as mentioned above. However, the analysis of the ages of the individual chondrules in meteorites does not show any clear difference in accretion ages between the parent bodies of ordinary chondrites and those of carbonaceous chondrites (Villeneuve et al. 2009), both of which seem to have formed 3–4 million years (Ma) after CAI formation. This supports the alternative idea that the partial overlapping of the radial distributions of asteroids of different types is due to dynamical mixing that occurred after the formation of the asteroids (see Sections 4 and 6).

Comets are representative of the planetesimal disk that formed at larger distances than the asteroid belt, i.e., between and beyond the giant planet orbits. The classical view is that, whereas the parent bodies of carbonaceous chondrites are rich in hydrated minerals, comets are rich in water ice and lack hydrated silicates, presumably because they formed in a colder environment. However, new discoveries are making the difference between carbonaceous chondrites (or C-type asteroids) and comets less well defined. The close flyby images of comets (e.g., Comet Borrelly) show little surface ice and small active regions (Sunshine et al. 2006). The samples from Comet Wild 2 by the Stardust Mission turned out to be quite similar to meteoritic samples (Zolensky et al. 2006). Modeling work on the origin of the dust that produces the zodiacal light (Nesvorný et al. 2010a) predicts that at least 50% of the micrometeorites collected on Earth are cometary; however, we see no clear separation of micrometeorites into two categories, which could be traced to asteroidal and cometary dust (Levison et al. 2009). Water ice has been found on the C-type asteroid Themis (Campins et al. 2010, Rivkin & Emery 2010), and some C-type asteroids in the main belt show cometary activity (Hsieh & Jewitt 2006). The possibility of a continuum in physical and chemical properties between carbonaceous asteroids and comets is well described in Gounelle et al. (2008).

Putting all this information together in a coherent picture is not a simple task. At the least, however, we can say that there is evidence that a radial gradient in temperature existed in the disk at the time(s) when planetesimals formed. In particular, planetesimals in the inner disk
(in the inner asteroid belt region and presumably also in the terrestrial planet region) appear to have been dry and volatile poor. As pointed out by Albarede (2009), the gaseous solar nebula probably dissipated before the temperature decreased enough to allow volatiles to condense in the inner solar system, thus explaining the lack of close-in volatile-rich primitive objects (i.e., objects with solar composition). This consideration opens the question of how water and other volatiles have been delivered to Earth, which we address in Section 7.

3. STEP II: FROM PLANETESIMALS TO PLANETARY EMBRYOS

Once the protoplanetary disk contains a substantial population of planetesimals and the dynamics of accretion becomes dominated by the effect of the gravitational attraction between pairs of planetesimals, the second stage of planet formation can start. During this phase of accretion—runaway growth—big bodies grow faster than the small ones and hence rapidly increase their relative mass difference (Greenberg et al. 1978). This process can be formalized by the inequality

$$\frac{d}{dt}(M_1/M_2) > 0,$$

where $M_1$ and $M_2$ are, respectively, the characteristic masses of the “big” and the “small” bodies. The reasons for runaway growth can be explained as follows.

At the beginning of the runaway growth phase, the largest planetesimals, by necessity, must represent only a small fraction of the total mass. Hence, the dynamics is governed by the small bodies, in the sense that the relative velocities between bodies are on the order of the escape velocity of the small bodies, $V_{\text{esc}(2)}$. This velocity is independent of the mass $M_1$ of the large bodies and is smaller than their escape velocity, $V_{\text{esc}(1)}$. For a given body, its collisional cross section is enhanced with respect to the geometrical cross section by the so-called gravitational focusing factor $F_g$:

$$\frac{dm}{dt} \sim R^2 F_g.$$

The gravitational focusing factor is given by:

$$F_g = 1 + \frac{1}{V_{\text{esc}}^2} \frac{V_{\text{rel}}^2}{V_{\text{esc}}^2},$$

(Greenberg et al. 1978, Greenzweig & Lissauer 1992, Safronov & Zvjagina 1969), where $V_{\text{esc}}$ is the body’s escape velocity, and $V_{\text{rel}}$ is the relative velocity of the other particles in its environment. Because $V_{\text{rel}} \sim V_{\text{esc}(2)}$, the gravitational focusing factor of the small bodies (with $V_{\text{esc}} = V_{\text{esc}(2)}$) is of order unity, whereas that of the large bodies (with $V_{\text{esc}} = V_{\text{esc}(1)} > V_{\text{esc}(2)}$) is much larger, of order $V_{\text{esc}(1)}^2/V_{\text{rel}}^2$. In this situation, because both $V_{\text{esc}(1)}^2$ and the geometrical cross section are proportional to $M_1^{2/3}$, the mass growth of a big body is described by this equation (Ida & Makino 1993):

$$\frac{1}{M_1} \frac{dM_1}{dt} \sim M_1^{1/3} V_{\text{rel}}^{-2}.$$

Therefore, the relative growth rate is an increasing function of the body’s mass, which is the condition for the runaway growth (Figure 1).

Runaway growth stops when the masses of the large bodies become important (Ida & Makino 1993) and the perturbations from the large bodies begin to govern the dynamics. The condition for this to occur is

$$n_1 M_1^2 > n_2 M_2^3,$$

where $n_1$ is the number of big bodies, and $n_2$ is the number of small bodies. In this case, $V_{\text{rel}} \sim V_{\text{esc}(1)}$, so that $F_g \sim 1$; hence, $(1/M_1)(dM_1/dt) \sim M_1^{-1/3}$. The growth rate of the embryos gets slower
Figure 1

An illustration of the process of runaway growth. Each panel represents a snapshot of the system at a different time. The coordinates represent the semimajor axis and the eccentricity of orbits of the objects in a portion of the disk centered at 1 astronomical unit (AU). The sizes of the representations of embryos and protoplanets are proportional to the cubic roots of their masses. Initially, the system is made of a planetesimal population, in which two objects (dark blue circles) are two times more massive than the others. These objects accrete planetesimals (light blue loops) quickly, increasing exponentially their mass ratio relative to the individual planetesimals, until they become planetary embryos. The eccentricities of the planetary embryos remain low, whereas the eccentricities of the planetesimals are excited with time. The embryos also separate from one another as they grow. At the end, the embryos have grown by a factor of 200, whereas the mean mass of the planetesimals has grown only by a factor of 2. From Kokubo & Ida (1998).
as the bodies increase in size, and the relative differences in mass among the embryos also slowly decrease. In principle, one might expect that the small bodies continue to grow, approaching the mass of the embryos, but in reality, the now large relative velocities prevent the small bodies from accreting one another; rather, their mutual collisions become disruptive. Thus, the small bodies can participate in the growth of the embryos only. This phase is termed oligarchic growth (Chambers 2006; Kokubo & Ida 1998, 2000).

The runaway growth phase happens throughout the disk on a timescale that depends on the local dynamical time (Keplerian time), on the planetesimal size, and on the local density of available solid material. This density also determines the maximum size of the embryos when the runaway growth ends (Lissauer 1987). Assuming a reasonable surface density of solid materials, the runaway growth process forms planetary embryos of lunar to martian mass at 1 AU in $10^5$–$10^6$ years, separated from one another by a few times hundredths of an AU.

Thus, the planetary embryos are not yet the final terrestrial planets (at least Earth and Venus). They are not massive enough, they are too numerous, and they are too closely packed relative to the terrestrial planets that we know. Moreover, they form too quickly compared with the timescale of tens of millions of years suggested for Earth by radioactive chronometers (Kleine et al. 2009).

Because runaway growth is a local process, the embryos form from the planetesimals in their neighborhoods. Little radial mixing is expected at this stage. Thus, if the planetesimal disk is characterized by a radial gradient of chemical properties, such a gradient should be reflected in the embryos’ chemical compositions.

In view of their large masses and their rapid formation timescales, embryos can undergo differentiation. We stress, though, that embryo formation cannot be faster than planetesimal formation because a massive planetesimal population is needed to trigger the runaway growth of the embryos. Thus, if planetesimals in the asteroid belt accreted approximately 3–4 Ma after CAI formation (Villeneuve et al. 2009; see description of step I), the embryos in the asteroid belt cannot have formed earlier than this time. In this case, the low content of short-lived radioactive elements might have allowed them to avoid differentiation, as in the cases of Callisto (Canup & Ward 2002) or Titan (Sotin et al. 2010). This lack of differentiation could have helped the embryos formed in the outer asteroid belt to preserve the water inherited from the local carbonaceous chondrite-like planetesimals. Nevertheless, even if differentiation had occurred on water-rich embryos, water would not have necessarily lost; water ice could have formed a mantle around a rocky interior that possibly differentiated itself into a metallic core and a silicate outer layer, as on Europa and Ganymede.

The formation of the giant planets is intimately related to the runaway/oligarchic growth of embryos. Beyond the so-called snow line between 2 and 4 AU, the low temperatures allow water ice to condense, enhancing the surface density of solid material and drastically increasing the typical embryo mass to perhaps an Earth mass or more (e.g., Ciesla & Cuzzi 2006, Dodson-Robinson et al. 2009, Kokubo & Ida 2002, Stevenson & Lunine 1988). The formation of the massive cores of the giant planets (of approximately 10 Earth masses each) remains poorly understood. It has been proposed that convergent migration processes could have brought these embryos together, favoring their rapid mutual accretion (Lyra et al. 2010, Morbidelli et al. 2008, Sandor et al. 2011). Once formed, the cores started to accrete massive atmospheres of hydrogen and helium from the protoplanetary disk, thus becoming the giant planets that we know. The formation of gas giant planets is a complex subject that merits its own review (see, for example, Lissauer & Stevenson 2007). What is most important for our purposes is that giant planets form quickly, before the final step of terrestrial planet formation, and therefore influence this last phase of terrestrial accretion.
4. STEP III: FROM EMBRYOS TO TERRESTRIAL PLANETS—OVERVIEW

At the time of the disappearance of the gas from the protoplanetary disk, the solar system should have had the following structure: (a) in the inner part, a disk of planetesimals and planetary embryos with roughly equal total mass in each component; (b) in the central part, a fully formed system of giant planets; and (c) beyond the orbits of the giant planets, another disk of planetesimals. The orbits of the giant planets were inherited from their previous dynamical evolution dominated by their gravitational interactions with the gas disk, and they were likely different from the current orbits. We revisit this important issue in the following sections.

When the nebular gas is present, it has a stabilizing effect on the system of embryos and planetesimals because it continuously damps their orbital eccentricities. Thus, when the gas is removed, the eccentricities can grow rapidly, leading to intersecting orbits and collisions among embryos (Chambers & Wetherill 1998). However, numerical simulations (Chambers & Wetherill 2001; O’Brien et al. 2006; Raymond et al. 2004, 2005) show that the dynamical evolution in the terrestrial planet region differs greatly from that in the asteroid belt. In the terrestrial planet region, where the perturbations exerted by Jupiter are weak, the embryos’ eccentricities remain relatively small, and the embryos can accrete one another in low-velocity collisions. However, the asteroid belt is crossed by several powerful resonances with Jupiter that excite the eccentricities of the resonant objects. The orbits of the embryos and planetesimals are continually changing owing to mutual encounters; additionally, every time that one of them temporarily falls into a resonance, its eccentricity is rapidly enhanced. As a result, most of the original population of embryos and planetesimals eventually leaves the asteroid belt region by acquiring orbits that are eccentric enough to collide with the Sun or cross Jupiter’s orbit and be ejected on a hyperbolic orbit; a fraction of them can be accreted by the growing planets inside 2 AU, and a small number of planetesimals remain on stable orbits in the belt (see Figure 2 for an illustration of this process). The typical result of this highly chaotic phase—simulated with several numerical N-body integrations—is the elimination of all the embryos originally situated in the asteroid belt and the formation of a small number of terrestrial planets on the order of the mass of Earth. The latter are formed on stable orbits in the 0.5–2-AU region on a timescale of several tens of millions of years (e.g., Chambers & Wetherill 2001; O’Brien et al. 2006; Raymond et al. 2004, 2005).

This scenario has several strong points:

1. Typically, two to four planets are formed on well-separated and stable orbits. If the initial disk of embryos and planetesimals contains approximately 4–5 Earth masses of solid material, typically the two largest planets are approximately one Earth mass each. Moreover, in the most modern simulations, which account for the dynamical interaction between embryos and planetesimals (O’Brien et al. 2006; see also Morishima et al. 2010, Raymond et al. 2009), the final eccentricities and inclinations of the terrestrial planets that form are comparable with or even smaller than those of the real planets (an explanation of this fact is provided in Section 5.3).

2. Quasi-tangent collisions of Mars-mass embryos with the protoplanets are quite frequent (Agnor et al. 1999, Elser et al. 2011, Morishima et al. 2010). These collisions are expected to generate a disk of ejecta around the protoplanets (Canup & Asphaug 2001), from which a satellite is likely to accrete (Canup & Esposito 1996). This is the standard, generally accepted scenario for the formation of the Moon (e.g., Benz et al. 1987).

3. The accretion timescale of the terrestrial planets in the simulations is ~30–100 Ma, in general agreement with the timescale of Earth accretion deduced from radioactive chronometers.
Figure 2
The growth of terrestrial planets from a disk of planetary embryos and planetesimals. Each panel shows the semimajor axis and eccentricity of the bodies in the system at a given time, reported in the top left of each panel. The sizes of the representations of embryos and protoplanets are proportional to the cubic roots of their masses. The representation of Jupiter is not to scale with respect to the representations of the embryos. A system of three terrestrial planets, the most massive of which has approximately an Earth mass, is eventually formed inside 2 AU, whereas only a small fraction of the original planetesimal population survives within the asteroid belt boundaries. Abbreviations: AU, astronomical units; Ma, million years. From O’Brien et al. (2006).

whose estimates vary from one study to another over a comparable range; Allègre et al. 1995, Touboul et al. 2007, Yin et al. 2002).

4. A small fraction of the original planetesimals typically remain in the asteroid belt on stable orbits at the end of the terrestrial planet formation process, and all embryos are ejected from the belt in most, but not all, simulations (O’Brien et al. 2007, Petit et al. 2001). The depletion of the belt by embryos and their subsequent removal are essential to explaining the asteroid belt as we see it today, including its substantial mass deficit. The orbital eccentricities and inclinations of these surviving particles compare relatively well with those of the largest asteroids in the current belt. Moreover, because of the scattering suffered from the embryos, the surviving particles are randomly displaced in semimajor axes, relative to their original position, by approximately 0.5 AU. This can explain the partial mixing of asteroids of different taxonomic types, discussed in Section 2.
However, not all simulations are equally successful. A discussion on how the results change depending on parameters and initial conditions is reported in the next section. Moreover, a general problem of the simulations is that the synthetic planet produced at the approximate location of Mars is systematically too massive (Chambers 2001, Raymond et al. 2009, Wetherill 1991). Mars is an oddity not only for its mass, but also for its accretion timescale: In fact, it formed in a few millions of years only, like asteroids, i.e., much faster than Earth formed (Dauphas & Pourmand 2011). We discuss these issues in more detail in Section 6.

5. FROM EMBRYOS TO TERRESTRIAL PLANETS: DEPENDENCE OF THE RESULTS ON SIMULATION AND MODEL PARAMETERS

5.1. Outcome of Giant Collisions

In all simulations of the accretion of terrestrial planets (with the exception of Alexander & Agnor 1998 and Kokubo & Genda 2010; see below), it has been assumed that all collisions between embryos are perfectly accretional, i.e., every time two embryos collide, they merge. However, this is a gross simplification. Simulations of giant collisions between embryos conducted with the Smooth Particle Hydrodynamics (SPH) technique (Agnor & Asphaug 2004, Asphaug et al. 2006) show that perfect merging is rare. In some cases, most of the masses of the two embryos merge, but a fraction of the total mass is ejected into space in the form of small objects, or fragments. In other cases, there are so-called hit-and-run collisions in which the embryos do not merge: They bounce off each other, again losing part of their masses as ejecta. The consequences of such “imperfect collisions” are not fully understood, although preliminary studies have given us some clues.

Kokubo & Genda (2010) investigated the effects of bouncing collisions on the dynamics of terrestrial planet accretion. First, from a database of SPH simulations of embryo-embryo collisions, they derived accretion conditions in terms of impact velocity, angle, and masses of colliding bodies. Then, in N-body simulations of terrestrial planet accretion, they treated each collision as either a perfect merger or an inelastic bounce as determined through comparison of the impact parameters of the collision with comparable cases in the collision database. They found that, on average, half of the collisions do not lead to accretion. However, the final number, mass, orbital elements, and even growth timescale of the terrestrial planets are barely affected by this large fraction of bouncing collisions. The reason is that, if two embryos bounce off each other, they both remain in the system and tend to accrete during their next encounter. Thus, whether accretion occurs at the first collision or at every other collision has little influence on the broad evolution of the simulation.

In reality, of course, collisions are neither perfect mergers nor bounces. As SPH simulations show, part of the colliding bodies’ mass is merged and part is released to space as collisional fragments. These fragments might be quickly reaccreted by the embryo, leading to little net change in the overall planet formation history, as found by Kokubo & Genda (2010). Alternatively, the fragments could undergo collisional grinding into dust and be removed by radiation forces, meaning that a fraction of the total disk mass would be lost.

The interest in the latter possibility lies in the possible consequences that it can have on the chemistry of the terrestrial planets. If most planetary embryos are differentiated, the material that is released to space during giant collisions should come preferentially from the mantles of the embryos rather than from their cores. This would lead to a change in the bulk composition of the growing terrestrial planets relative to the original composition of their progenitors, particularly increasing the final iron/silicate ratio (O’Neill & Palme 2008). The metal content of the terrestrial planets is indeed higher than in any undifferentiated meteorite (Jarosewich 1990).
5.2. Disk Mass and Radial Profile

Kokubo et al. (2006) performed the most recent systematic exploration on how the outcomes of the terrestrial planet formation process depend on the basic properties of the protoplanetary disk (surface density, surface density profile, orbital separation of the initial protoplanet system, and bulk density of protoplanets). For computational reasons, in all simulations, the disk of solids was assumed to range from 0.5 to 1.5 AU. No giant planets were considered. For the standard disk model, with a surface density of solids equal to 10 g cm$^{-2}$ at 1 AU and decaying as $1/r^{3/2}$, typically two Earth-sized planets formed in the terrestrial planet region. The number of planets slowly decreased as the surface density of the initial protoplanets increased, whereas the masses of individual planets increased almost linearly. The basic structure of the resulting planetary systems depended only slightly on the initial distribution of protoplanets (individual masses and mutual separations in units of Hill radius) and the bulk density, as long as the total mass was fixed. For a steeper surface density profile, large planets formed, on average, closer to the star. The dependence of the results on the disk’s surface density profile had also been investigated in Raymond et al. (2005), who found that, for steeper profiles, the final planets were more numerous, formed more quickly, were more massive, and had higher iron contents and lower water contents.

5.3. Particle Size Distribution in the Disk

The first simulations of the terrestrial planet accretion process using modern integration techniques (Agnor et al. 1999; Chambers 2001; Chambers & Wetherill 1998, 2001) used as initial conditions a set of a few tens of planetary embryos and produced planets on orbits that were systematically too eccentric and inclined. A commonly used measure of the orbital excitation of the terrestrial planets is the normalized angular momentum deficit (AMD), defined as (Laskar 1997)

$$\text{AMD} = \frac{\sum m_j \left[ a_j (1 - e_j^2) \right]^{1/2} \cos i_j - m_j a_j^{1/2}}{\sum m_j a_j^{1/2}},$$

where the sum is calculated over all the planets, with mass $m_j$, semimajor axis $a_j$, eccentricity $e_j$ and inclination $i_j$. For reference, the AMD of the actual terrestrial planets is $-0.0018$, when averaged over million-year timescales, compared with a median value of $-0.0050$ in the simulations of Chambers (2001). The small orbital excitation of the real terrestrial planets, therefore, was a major property that the simulations failed to reproduce for many years.

To solve this problem, one proposal was that a small fraction of gas from the protosolar nebula was still present in the system when the terrestrial planets formed, so that the eccentricities and inclinations of embryos and protoplanets were continuously damped by the gravitational interactions with the gas. However, the simulations that accounted for this process (Kominami & Ida 2002, 2004) systematically resulted in too many terrestrial planets, forming in less than 10 Ma (too short a timescale, relative to radioactive chronometer constraints; Halliday & Klein 2006) on quasi-circular orbits that are much closer to one another than the real orbits are. Ogihara et al. (2007) alleviated the problem of the number and separation of terrestrial planets by assuming that the disk was strongly turbulent; nevertheless, the problem of the excessively short formation timescale remained (see also Thommes et al. 2008).

A different approach to solving the problem of the small orbital excitation of the terrestrial planets was to explore different mass distributions in the original disk of solid bodies. Chambers (2001) showed that starting the simulations with a larger number of smaller embryos than that used in Chambers & Wetherill (1998) resulted in a significant reduction of the final AMD, although the latter was still too large, in absolute value, relative to that of the real terrestrial planets. This
trend continued in the simulations of Raymond et al. (2004, 2005, 2006, 2007) and eventually led to AMDs consistent with that of the real system (Hansen 2009, Raymond et al. 2009).

Instead, O’Brien et al. (2006) assumed that the initial disk consisted of a population of planetary embryos, well separated from one another, embedded in a population of planetesimals. The total mass of each population was the same. In total, 25 Mars-mass embryos were considered from 0.3 to 4 AU, whereas the planetesimal population was modeled with 1,000 equal mass particles. The planetesimals were assumed to interact with the embryos, but not with one another. With these initial conditions, some simulations (those with Jupiter and Saturn initially on their current orbits; see Section 5.4) achieved, for the first time, a system of synthetic terrestrial planets with an AMD of \(-0.001\), which is even smaller than the real one (the reader should keep in mind that the terrestrial planets might have achieved the current AMD after their formation, during a phase of late orbital migration of the giant planets; Brasser et al. 2009).

The success of these simulations argues strongly that the key to explaining the small AMD of the terrestrial planets is dynamical friction (Wetherill & Stewart 1993). The latter is the result of the gravitational interaction between populations of bodies of different individual masses. As a result of equipartition of orbital excitation, more massive bodies tend to remain on less eccentric and inclined orbits, and less massive bodies tend to remain on more eccentric and inclined orbits. Thus, planetesimals acquire orbits that are more and more excited (until they are mostly removed by collisions with the Sun or ejection onto hyperbolic orbits), whereas the eccentricities and inclinations of the embryos and forming planets are progressively damped. Of course, given the prevalence of giant impacts during the late stages of growth, embryos must acquire large enough eccentricities for their orbits to cross. However, the key to maintaining low-eccentricity final orbits is the presence of planetesimals after giant impacts; collisional debris could play a role in this process, especially as they are naturally generated by the impacts themselves.

5.4. Giant Planet Architecture

There is a clear dependence of the final outcomes of the simulations on the orbital architecture assumed for the giant planets. By definition, the giant planets must be fully formed by the time the gas is removed from the disk, i.e., well before the formation of the terrestrial planets is complete. Thus, in most terrestrial planet formation simulations, the orbital architecture of the giant planet system is considered an environmental property of the inner solar system, i.e., an initial condition.

Terrestrial planet formation with orbital configurations of giant planets vastly different from our solar system was investigated by Levison & Agnor (2003). They found that the number, mass, and location of the terrestrial planets are directly related to the amount of dynamical excitation experienced by the planetary embryos near 1 AU, which, in turn, is related to the proximity, mass, and eccentricity of the giant planets. In general, if the embryos’ eccentricities are large, each crosses the orbits of a larger fraction of its cohorts, leading to a smaller number of more massive planets. In addition, embryos tend to collide with objects near their periastrons. Thus, in systems where the embryos’ eccentricities are large, planets tend to form close to the central star. Comparable results were found by Raymond (2006), who focused on the giant planet configurations that permit habitable terrestrial planets to form.

For systems of giant planets resembling our solar system (Jupiter- and Saturn-mass planets with moderate orbital eccentricities, comparable with those of their current orbits), the results are more subtle. Unexpectedly, the simulations with giant planets that have larger orbital eccentricities form terrestrial planets on a shorter timescale and on more circular final orbits than do simulations with giant planets that are more circular. Moreover, for eccentric giant planets, the terrestrial planets accrete less material from the asteroid belt, and the synthetic planet produced at the orbital location...
of Mars is smaller (Raymond et al. 2009). All these properties are interrelated. Eccentric giant planets deplete the asteroid belt more rapidly than do giant planets on more circular orbits, so that embryos and planetesimals originally in the belt are removed by collisions with the Sun or by ejection on hyperbolic orbit before they have a chance to interact with the growing terrestrial planets inside 2 AU. Thus, in the presence of eccentric giant planets, terrestrial planet formation proceeds as in a “closed” system, with little or no material from outside 1.5–2.0 AU incorporated into surviving planets. Therefore, on the one hand, accretion proceeds more quickly because the material that builds the planets is more radially confined. On the other hand, faster accretion timescales imply that more planetesimals are still in the system at the end of the terrestrial planet accretion process, so that the orbital eccentricities of the terrestrial planets can be damped more by dynamical friction. Finally, Mars accretes less mass because it is close to the edge of the radial distribution of the mass that participates in the construction of the terrestrial planets.

Given that the giant planets probably underwent planetesimal-driven orbital migration after the dissipation of the gaseous protoplanetary disk (Fernandez & Ip 1984, Malhotra 1995, Tsiganis et al. 2005), the orbits of Jupiter and Saturn at the time of terrestrial planet formation are unknown. To understand which giant planet orbital architecture was more likely at early times, one must address the dynamical evolution that the giant planets should have had in the gaseous solar nebula. It is well known that, by interacting gravitationally with the gaseous disk, the orbits of the giant planets migrate (see Ward 1997 for a review; this process is distinct from so-called planetesimal-driven migration). Eventually the planets tend to achieve a multiresonance configuration, in which the period of each object is in integer ratio with that of its neighbor (Morbidelli et al. 2007). The interaction with the disk also damps the planets’ orbital eccentricities. Thus, at the disappearance of the gas disk, the giant planets should have been closer to one another, on resonant and quasi-circular orbits (the giant planets could have achieved their current orbits at a much later time, corresponding to the so-called Late Heavy Bombardment; see Morbidelli 2010 for a review). However, simulations show that this kind of orbital configuration of the giant planets systematically leads to synthetic planets at \( \sim 1.5 \) AU that are much more massive than the real Mars (Raymond et al. 2009).

6. THE GRAND TACK SCENARIO

Hansen (2009) convincingly showed that the key parameter for obtaining a small Mars is the radial distribution of the solid material in the disk. If the outer edge of the disk of embryos and planetesimals is at approximately 1 AU, with no solid material outside this distance, even simulations with giant planets on circular orbits systematically produce a small Mars (together with a big Earth). The issue is, then, how to justify the existence of such an outer edge and how to explain its compatibility with the existence of the asteroid belt between 2 and 4 AU. The asteroid belt has today a very small total mass (approximately \( 6 \times 10^{-4} \) Earth masses; Krasinsky et al. 2002), but it must have contained at least a thousand times more solid material when the asteroids formed (Wetherill 1989).

The result by Hansen (2009) motivated Walsh et al. (2011) to look in more detail at the possible orbital history of the giant planets and their ability to sculpt the disk in the inner solar system. For the first time, the giant planets were not assumed to be on static orbits (even if different from the current ones); instead, Walsh et al. studied the coevolution of the orbits of the giant planets and of the planetesimal and embryo precursors of the terrestrial planets, during the era of the disk of gas. Walsh et al. built their model on previous hydrodynamical simulations that showed that the migration of Jupiter can be in two regimes: When Jupiter is the only giant planet in the disk, it migrates inward (Lin & Papaloizou 1986), but when Jupiter is paired with Saturn, both planets
The orbits of embryos and planetesimals at the end of the inward-then-outward migration of Jupiter as modeled in the Grand Tack scenario, when the gas is fully removed. From this state, the system evolves naturally on a timescale of tens of millions of years into two Earth-mass planets at ∼0.7 AU and 1 AU and a small Mars at 1.5 AU (see Figure 4).

typically migrate outward, locked in a 2:3 mean motion resonance (in which the orbital period of Saturn is 3/2 that of Jupiter; Masset & Snellgrove 2001, Morbidelli & Crida 2007). Thus, assuming that Saturn formed later than Jupiter, Walsh et al. envisioned the following scenario: First, Jupiter migrated inward while Saturn was still growing; then, when Saturn reached a mass close to its current one, it started to migrate inward more rapidly than did Jupiter, until it captured the latter in the 3/2 resonance (Masset & Snellgrove 2001, Pierens & Nelson 2008, Pierens & Raymond 2011); finally, the two planets migrated outward together until the complete disappearance of the disk of gas. The extent of the inward and outward phases of migration is unconstrained a priori because they depend on properties of the disk and of giant planet accretion that are unknown, such as the time lag between Jupiter’s and Saturn’s formations, the speed of inward migration (which depends on the disk’s viscosity), the speed of outward migration (which depends on the disk’s scale height), and the time lag between the capture in resonance of Jupiter and Saturn and the photoevaporation of the gas. However, the extent of the inward and outward migration phases of Jupiter can be deduced by looking at the resulting structure of the inner solar system. In particular, Walsh et al. showed that a reversal of Jupiter’s migration at 1.5 AU would provide a natural explanation for the existence of the outer edge of the inner disk of embryos and planetesimals at 1 AU, which is required to produce a small Mars (see Figures 3 and 4). Because of the prominent reversal of Jupiter’s migration that the Walsh et al. scenario assumes, it is nicknamed the Grand Tack scenario.

Several giant extrasolar planets have been discovered orbiting their stars at a distance of 1–2 AU, so the idea that Jupiter was at one time 1.5 AU from the Sun is not shocking in itself. A crucial diagnostic of this scenario, though, is the survival of the asteroid belt. Given that Jupiter should have migrated through the asteroid belt region twice—first inward, then outward—one could expect that the asteroid belt should now be totally empty. However, the numerical simulations by Walsh et al. (2011) show that the asteroid belt first is fully depleted by the passage of the giant planets, but then, while Jupiter leaves the region for the last time, it is repopulated by a small
Figure 4
The mass distribution of the synthetic terrestrial planets produced in the Walsh et al. (2011) simulations. The red squares refer to simulations starting with 80 embryos of 0.026 Earth masses (series 1), and the dark yellow triangles to simulations from 40 embryos of 0.051 Earth masses (series 2). The horizontal gray lines denote the perihelion-aphelion excursion of the planets on their eccentric final orbits. The solid blue squares show the real planets of the solar system. The large mass ratio between Earth and Mars is statistically reproduced.

fraction of the planetesimals scattered by the giant planets during their migration. In particular, the inner asteroid belt is repopulated mainly by planetesimals that were originally inside the orbit on which Jupiter formed, whereas the outer part of the asteroid belt is repopulated mainly by planetesimals originally between and beyond the orbits of the giant planets. Assuming that Jupiter initially formed at the location of the snow line, it is tempting to identify the planetesimals originally closer to the Sun with the anhydrous asteroids of E- and S-type and those originally between and beyond the orbits of the giant planets with the “primitive” C-type asteroids. With this assumption, the Grand Tack scenario can explain the coexistence of asteroids of different types in the main belt in a natural way. In fact, it is difficult to explain the differences between ordinary/enstatite chondrite (E- and S-type) and carbonaceous chondrite (C-type) parent bodies if they had both formed in the asteroid belt region, given that they are coeval (Villeneuve et al. 2009) and that the radial extent of the asteroid belt is small (only ~1 AU). Instead, if ordinary/enstatite and carbonaceous chondrite parent bodies were implanted into the asteroid belt from originally well-separated reservoirs, as in the Grand Tack model, the differences in physical properties would be easier to understand in the framework of the classical condensation sequence. The origin of C-type asteroids from the giant planet region would also explain, in a natural way, the similarities with comets that are emerging from recent observational results and sample analyses. The small mass of the asteroid belt and its eccentricity and inclination distribution are also well reproduced by the Grand Tack scenario.
This scenario also explains qualitatively why the accretion timescales of Mars and those of the asteroids are comparable (Dauphas & Pourmand 2011). In fact, the asteroids stopped accreting when they got dispersed and injected onto excited orbits of the main belt; Mars stopped accreting when the inner disk was truncated at 1 AU and the planet was pushed beyond this edge by an encounter with the proto-Earth (Hansen 2009). In the Grand Tack scenario, these two events coincide and mark the time of the passage of Jupiter through the inner solar system. All these results on the asteroid belt, together with the fact that the mass distribution of the terrestrial planets is also statistically reproduced (see Figure 4), make the Grand Tack scenario an appealing comprehensive model of terrestrial planet formation.

7. ORIGIN OF TERRESTRIAL WATER

Laboratory analyses show a clear correlation between the water content in meteorites and the heliocentric distance of the parent bodies from which said meteorites are thought to originate (see Figure 5). Although there is the theoretical possibility that water-rich planetesimals formed in the hot regions of the disk by water-vapor absorption on silicate grains (King et al. 2010, Muralidharan et al. 2008), the empirical evidence strongly suggests that the planetesimals in the terrestrial planet region were extremely dry.

In contrast, Earth has a water content that, although small, is non-negligible and definitely larger than what the aforementioned correlation would suggest for material condensed at 1 AU. In fact, the mass of the water contained in Earth’s crust (including the oceans and the atmosphere) is $2.8 \times 10^{-4}$ Earth masses; the mass of the water in the present-day mantle is uncertain. Lécuyer et al. (1998) estimated it to be in the range of $0.8–8 \times 10^{-4}$ Earth masses. More recently, Marty

![Figure 5](image_url)

CI and CM meteorites are the most rich in water; water amounts to approximately 5–10% of their total mass (Kerridge 1985, Robert & Epstein 1982). They are expected to come from C-type asteroids, predominantly in the asteroid belt and possibly accreted even farther out (Walsh et al. 2011). Water in ordinary chondrites amounts to only 0.1% of the total weight (McNaughton et al. 1981; Robert et al. 1977, 1979), or a few times as much (Jarosewich 1966); ordinary chondrites are spectroscopically linked to S-type asteroids, which are predominant between 2 and 2.5 AU. Finally, enstatite chondrites are very dry, with only 0.01% of their total mass in water (Hutson & Ruzicka 2000); they are expected to come from E-type asteroids, which dominate the Hungaria region in the very inner asteroid belt at 1.8 AU. Abbreviation: MB, main belt.
(2012) provided arguments in favor of a mantle water content of $\sim 2 \times 10^{-3}$. However, a larger quantity of water might have resided in the primitive Earth and been subsequently lost during core formation and impacts. Thus, the current Earth has water content larger than that of enstatite chondrites, and it is possible that the primitive Earth had water content comparable with or larger than that of ordinary chondrites. Where did this water come from, if the local planetesimals were dry?

An important constraint for models on the origin of Earth’s water is the DHO/H$_2$O ratio. This ratio in the present-day mantle and in the oceans is $1.49 \pm 3 \times 10^{-4}$ (Lécuyer et al. 1998), which is approximately 6 times higher than for molecular hydrogen in the protoplanetary disk (deduced from measurements in the atmosphere of Jupiter; Mahaffy et al. 1998). It is believed that the recycling of water in the deep mantle does not significantly change the D/H ratio. However, the D/H ratio could have significantly increased if Earth had had in the past a massive hydrogen atmosphere (with a molar ratio H$_2$/H$_2$O larger than 1) that experienced a slow hydrodynamical escape (Genda & Ikoma 2008).

Given the expectation that water was not present in sufficient quantities in the local planetesimals and embryos around 1 AU, there are three potential sources for the origin of water on Earth. The first model invokes a nebular origin. Ikoma & Genda (2006) assumed that at the end of Earth’s formation, there was still some nebular hydrogen in the protoplanetary disk, and Earth gravitationally captured a hydrogen-rich atmosphere of nebular gas of up to $10^{21}$ kg in mass. Then, the atmospheric hydrogen was oxidized, perhaps by FeO in the magma ocean, to produce water. However, in this model, the initial D/H ratio of the water would be solar. The D/H ratio can be subsequently increased by hydrodynamical escape of the hydrogen-rich atmosphere (Genda & Ikoma 2008). However, the factor-of-6 increase in the D/H ratio that is required to match observations yields unrealistically long timescales; i.e., the hydrodynamical escape of the primitive atmosphere should have occurred over billions of years, in contrast with constraints from the $^{129}$I/$^{129}$Xe chronometer (Wetherill 1975).

A second possibility is that the water was brought to Earth by the bombardment of comets (Delsemme 1992, 1999). For more than a decade, it was believed that a major problem with this model is that the D/H ratio in comets is approximately twice that on Earth (Balsiger et al. 1995, Bockelée-Morvan et al. 1998, Eberhardt et al. 1995, Meier et al. 1998), but there are no known terrestrial processes that could decrease the D/H ratio of the original water. However, a comet with terrestrial D/H ratio (103P/Hartley 2) has just been found (Hartogh et al. 2011). So, this problem may not be as severe as previously thought. A more severe problem, however, is that the collision probability of comets with Earth is very small. Of the planetesimals scattered by the giant planets from the protoplanetary disk, only one in a million would strike our planet (Morbidelli et al. 2000). From studies on the range of radial migration that the giant planets should have suffered after the disappearance of the disk of gas, owing to their interaction with planetesimals (Gomes et al. 2004, 2005; Hahn & Malhotra 1999; Malhotra 1995), it is expected that the total mass of the cometary disk was 35–50 Earth masses; moreover, from measurements of the ice/dust ratio in comets (Küppers et al. 2005), it is now believed that less than half of the mass of a comet is in water ice. When all these elements are put together, the mass of water delivered by comets to Earth should have been only $\sim 2.5 \times 10^{-5}$ Earth masses (neglecting impact losses), i.e., only 10% of the crustal water.

The third possibility is that Earth accreted water from primitive planetesimals and/or planetary embryos originally from the outer asteroid belt (Lunine et al. 2007; Morbidelli et al. 2000; O’Brien et al. 2006; Raymond et al. 2004, 2005, 2006, 2007). From an isotopic point of view, this makes sense because the average D/H ratio of water in carbonaceous chondrites is almost identical to that of Earth ($1.59 \times 10^{-4}$, with individual values in the range of 1.28–1.80 $\times 10^{-4}$; Dauphas
et al. 2000). To address the delivery of water by asteroids from the modeling point of view, we need to distinguish between the classical scenario, in which the outer belt is originally inhabited by primitive objects that are removed by mutual scattering and resonant interactions with Jupiter (see Section 4), and the Grand Tack scenario (see Section 6). In the first case, as we have seen above, the amount of material accreted by the terrestrial planets from the asteroid belt depends on the eccentricity of the orbit of Jupiter. If Jupiter’s orbit was almost circular, the terrestrial planets should have accreted 10–20% of their mass from beyond 2.5 AU (O’Brien et al. 2006, Raymond et al. 2006), most of which should have been of carbonaceous chondritic nature. Thus, in the classical scenario, the terrestrial planets should have originally been very water rich, possibly even as much as envisioned by Abe et al. (2000), and should have lost most of their water during impacts and/or their geophysical evolution. However, Drake & Righter (2002) have argued that the detailed geochemistry of the carbonaceous chondrites is inconsistent with more than a few percent of Earth being composed of this material. Furthermore, the amount of material accreted from the outer asteroid belt drops with the increasing eccentricity of Jupiter. If Jupiter had originally had an orbit with an eccentricity comparable with or larger than the current one, little material would have been accreted from the outer asteroid belt, and the terrestrial planets would have been almost completely dry (O’Brien et al. 2006, Raymond et al. 2009), as noted in Section 5.4.

In the Grand Tack scenario, the primitive asteroids (like the C-types) were implanted into the asteroid belt from between and beyond the orbits of the giant planets. In Walsh et al.’s (2011) simulations, for every primitive planetesimal implanted in the outer asteroid belt, 10–30 planetesimals ended up on orbits that crossed the terrestrial planet–forming region, for a total of $3-11 \times 10^{-2}$ Earth masses. O’Brien et al. (2010) showed that, in this situation, Earth could accrete approximately 2–3% of its mass from these objects, enough to supply a few times the current amount of water on our planet (assuming that the primitive planetesimals were 5–10% water by mass). Walsh et al. and O’Brien et al. did not consider primitive planetary embryos in their simulations, so, in principle, the total amount of primitive material supplied to Earth could be somewhat larger than the reported estimate. In view of its smaller yield, the Grand Tack scenario avoids the geochemical inconsistencies that arise from the delivery of a too-large amount of carbonaceous chondritic material, characteristic of the classical scenario with giant planets on quasi-circular orbits.

A common feature of the classical and Grand Tack scenarios for the asteroidal delivery of water to the terrestrial planets is that the water is accreted during the formation of the planets and not in a late veneer fashion (i.e., after the planets have already achieved their final masses). The accretion of water, though, is not uniform throughout the planet accretion history; instead, it accelerates toward the end. Typically, 50% of the water is accreted after Earth has reached 60–70% of its final mass (in some cases, even 80%).

From the geochemical point of view, Albarede (2009) made a strong case that the deficiency of volatiles (and water) in Earth relative to solar abundances cannot be explained by invoking the notion that Earth lost most of its volatiles during its geophysical evolution. In fact, if Earth had accreted volatiles in solar proportion and had subsequently lost them to space, the abundance of elements would be correlated with their molecular masses. Instead, the relative abundance of elements in our planet is grossly correlated with their condensation temperatures. This strongly suggests that the planetesimals and embryos from which Earth inherited most of its mass were volatile poor because they formed in a disk too hot for the condensation of these elements.

Mann et al. (2009) and Wood et al. (2010) found strong evidence that the accretion of moderately volatile elements into Earth occurred mostly during the late stages of accretion, but before core formation was complete. The evidence against a late veneer of volatiles is illustrated in Figure 6, which shows that the relative abundance of elements in our planet is correlated not
The abundance of elements on Earth, as a function of their condensation temperatures and chemical affinities. The open dots represent the abundances of elements in the Allende CV3 chondrite, and the filled dots represent the abundances of elements in the silicate Earth, relative to CI chondrites, plotted as a function of 50% condensation temperature \(T_{50}\). Red dots represent elements that are highly siderophile; purple dots represent moderately to weakly siderophile elements; blue dots represent lithophile elements. From Wood et al. (2010).

only with condensation temperature but also with chemical affinity. Highly siderophile volatile elements are more depleted in the mantle than are moderate siderophile elements or lithophile elements with the same condensation temperature. This implies that these volatile elements saw the formation of Earth’s core.

Wood et al. (2008) and Rubie et al. (2011) showed that the oxidation state of Earth increased progressively as Earth was growing. Schematically, 60–70% (by mass) of Earth should have accreted from highly reduced material, and the final 30–40% of accreted mass should have been more oxidized. This is consistent with the possibility that H\(_2\)O was added, together with moderately volatile elements, during the final 30–40% of accretion.

Together, all these results of geochemical models are in good qualitative agreement with the scenario of terrestrial planet accretion emerging from the astrophysical simulations. This gives confidence that, although many issues have yet to be understood, we are approaching a coherent and global view of the terrestrial planet formation history.

8. SUMMARY AND FUTURE PROSPECTS

Models of the formation of the terrestrial planets have come a long way since the pioneering models of Wetherill and his students, and they now take into account the compelling likelihood of significant giant planet migration within our own solar system. That one might consider such an apparently catastrophic model with a degree of confidence is the result of more than a decade of observations of giant planets in a rich variety of orbits around other stars, the evidence from their orbital properties of significant dynamical interactions (Jurić & Tremaine 2008), and the detection by microlensing of a sufficient number of free-floating planets (Sumi et al. 2011). These suggest that giant planet migration and interactions, even catastrophic, might be the rule rather
than the exception. Applying this line of thinking to our own solar system leads to solutions for the small size of Mars, the structure of the asteroid belt, and the origin of water on Earth. However, numerous key problems remain, including the correct treatment of accreting versus nonaccreting impacts and the geochemistry of the planetesimals from the giant planet region that seeded Earth with water.

The future observations required to test this picture are multidisciplinary, ranging from in situ studies of comets to astronomical observations of protoplanetary disks around other stars. Starting in our solar system and moving outward, the following key studies are needed:

- Observations of asteroids that better tie meteorite types to their asteroidal parent bodies: For example, the geochemistry of Ceres, to be assessed by the Dawn mission, and sample return missions from primitive asteroids will give us more confidence in our associations of meteorite classes with specific asteroid types.
- A more comprehensive inventory of D/H in water in solar system bodies: This is key to understanding the connections and differences between primitive asteroids and comets, which, in turn, are diagnostic for models of implantation of primitive objects into the asteroid belt from a cometary reservoir, such as the Grand Tack. Further observations of D/H in comets are forthcoming. Also, a better understanding of the D/H values in reservoirs of water on Mars would be helpful (Lunine et al. 2003).
- Determination of the size of heavy element cores in Jupiter and Saturn and of the oxygen abundances in each: This would constrain the composition of the planetesimals that formed those planets, which would indirectly provide indications of where Jupiter and Saturn formed and hence more tests for the Grand Tack model. The Juno mission will obtain this information at Jupiter, and the Cassini proximal orbits will help provide constraints for Saturn; however, oxygen in Saturn's atmosphere will be difficult to obtain (via water) from anything other than a microwave sounder or deep probe.
- A more complete census of the orbital properties of extrasolar giant planets and of the occurrence of free-floating giant planets (the latter from microlensing surveys): This would better constrain giant planet formation models and the extent to which dynamical interactions are common after formation.
- High-resolution observations of disks from gaseous to transitional, from next-generation facilities such as the Atacama Large Millimeter/submillimeter Array and James Webb Space Telescope: These can provide information on the timing and extent of giant planet migration.

**DISCLOSURE STATEMENT**

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

**LITERATURE CITED**


Delsemme AH. 1999. The deuterium enrichment observed in recent comets is consistent with the cometary origin of seawater. *Planet. Space Sci.* 47:125–31


Tucson: Univ. Ariz. Press


Tucson: Univ. Ariz. Press


Morbidelli A, Bottke WF, Nesvorny D, Levison HF. 2009. Asteroids were born big. *Icarus* 204:558–73


Contents

Reminiscences From a Career in Geomicrobiology
   Henry L. Ehrlich ................................................................. 1

Mixing and Transport of Isotopic Heterogeneity in the Early Solar System
   Alan P. Boss ................................................................. 23

Tracing Crustal Fluids: Applications of Natural $^{129}$I and $^{36}$Cl
   Udo Fehn ................................................................. 45

SETI@home, BOINC, and Volunteer Distributed Computing
   Eric J. Korpela ................................................................. 69

End-Permian Mass Extinction in the Oceans: An Ancient Analog for the Twenty-First Century?
   Jonathan L. Payne and Matthew E. Clapham ........................................... 89

Magma Oceans in the Inner Solar System
   Linda T. Elkins-Tanton ................................................................. 113

History of Seawater Carbonate Chemistry, Atmospheric CO$_2$, and Ocean Acidification
   Richard E. Zeebe ................................................................. 141

Biomimetic Properties of Minerals and the Search for Life in the Martian Meteorite ALH84001
   Jan Martel, David Young, Hsin-Hsin Peng, Cheng-Yeu Wu, and John D. Young .... 167

Archean Subduction: Fact or Fiction?
   Jeroen van Hunen and Jean-François Moyen ........................................... 195

Molecular Paleohydrology: Interpreting the Hydrogen-Isotopic Composition of Lipid Biomarkers from Photosynthesizing Organisms
   Dirk Sachse, Isabelle Billault, Gabriel J. Bowen, Yoshito Chikaraishi, Todd E. Dawson,
   Sarah J. Fokin, Katherine H. Freeman, Clayton R. Magill, Francesca A. McInerney,
   Marcel T.J. van der Meer, Pratigya Polissar, Richard J. Robins, Julian P. Sachs,
   Hanns-Ludwig Schmidt, Alex L. Sessions, James W.C. White, Jason B. West,
   and Ansgar Kahmen ................................................................. 221
Building Terrestrial Planets
A. Morbidelli, J.I. Lunine, D.P. O’Brien, S.N. Raymond, and K.J. Walsh ........... 251

Paleontology of Earth’s Mantle
Norman H. Sleep, Dennis K. Bird, and Emily Pope ....................................... 277

Molecular and Fossil Evidence on the Origin of Angiosperms
James A. Doyle ................................................................................................. 301

Infrasound: Connecting the Solid Earth, Oceans, and Atmosphere

Titan’s Methane Weather
Henry G. Roe ..................................................................................................... 355

Extratropical Cooling, Interhemispheric Thermal Gradients, and Tropical Climate Change
John C.H. Chiang and Andrew R. Friedman ....................................................... 383

The Role of H₂O in Subduction Zone Magmatism
Timothy L. Grove, Christy B. Till, and Michael J. Krawczynski ......................... 413

Satellite Geomagnetism
Nils Olsen and Claudia Stolle ............................................................................. 441

The Compositions of Kuiper Belt Objects
Michael E. Brown ............................................................................................... 467

Tectonics of the New Guinea Region
Suzanne L. Baldwin, Paul G. Fitzgerald, and Laura E. Webb ............................ 495

Processes on the Young Earth and the Habitats of Early Life
Nicholas T. Arndt and Euan G. Nisbet ............................................................... 521

The Deep, Dark Energy Biosphere: Intraterrestrial Life on Earth
Katrina J. Edwards, Keir Becker, and Frederick Colwell ................................... 551

Geophysics of Chemical Heterogeneity in the Mantle
Lars Stixrude and Carolina Lithgow-Bertelloni ............................................... 569

The Habitability of Our Earth and Other Earths: Astrophysical, Geochemical, Geophysical, and Biological Limits on Planet Habitability
Charles H. Lineweaver and Aditya Chopra ....................................................... 597

The Future of Arctic Sea Ice
Wiesław Maslowski, Jaclyn Clement Kinney, Matthew Higgins, and Andrew Roberts .......................................................... 625

The Mississippi Delta Region: Past, Present, and Future
Michael D. Blum and Harry H. Roberts ............................................................ 655
Climate Change Impacts on the Organic Carbon Cycle at the Land-Ocean Interface

Elizabeth A. Canuel, Sarah S. Cammer, Hadley A. McIntosh, and Christina R. Pondell .......................................................... 685

Indexes

Cumulative Index of Contributing Authors, Volumes 31–40 ........................................... 713
Cumulative Index of Chapter Titles, Volumes 31–40 ..................................................... 717

Errata

An online log of corrections to Annual Review of Earth and Planetary Sciences articles may be found at http://earth.annualreviews.org