Field Review

The Origin of the Moon Revisited

HELEN ASHCROFT

Helen is a second year DPhil student in Earth Sciences specialising in Experimental Petrology, and her research interests involve looking into the formation, differentiation and evolution of the terrestrial planets. Helen also contributes to the online and print editions of Bang! Science Magazine, and in her free time enjoys making and listening to music.

The Moon is the closest object to us in the solar system. We observe it every night in our sky and we know it controls the rotation and orientation of the Earth's axis, causing the tides and making possible a stable climate in which life can thrive. But how the Moon formed, and what it formed from, remain among the biggest mysteries in planetary science.

Any theory of how the Moon formed from the Earth has to explain the physics and dynamics of the alleged collision with and subsequent changes to the Earth and its orbit, as well as the chemical similarities and differences between the Earth and the Moon.

According to the Giant Impact Theory, now widely accepted, the Moon formed as a result of a Mars-sized impactor colliding with the early Earth and creating a debris disk around the Earth which formed the Moon. This hypothesis was first suggested at the first 'Origin of the Moon' meeting in 1984, and gained popularity and strength over the next decade as lunar meteorites and terrestrial rocks were being analysed for their geochemistry. For it was then that the chemical similarities between the Moon and the Earth were becoming apparent. Such geochemical advances also suggested constraints on the age of the Moon. The

Giant Impact Theory is supported by the fact that intense bombardment of the accreting planets by smaller planetary building blocks was a commonly occurring process early on in Solar System history.

At first, these chemical similarities seemed to support the Giant Impact Theory by suggesting that some kind of mixing of material between the Earth and the impactor occurred during the formation of the Moon. However the discovery that the Earth and the Moon have nearly identical oxygen, iron, hydrogen, silicon, magnesium, titanium, potassium, tungsten and chromium isotope compositions threw a spanner in the works. Isotopes are atoms of a chemical element that have the same number of protons, but different numbers of neutrons, and the ratios between two isotopes of one element are governed by a variety of chemical and physical processes. In fact isotopes are so sensitive to a range of factors that they provide chemical fingerprints for different planets – Earth has a very different isotope signature from Mars, for example, and this fact is used in the classification of meteorites.

Although the Earth and the Moon have similar isotope ratios, there are clear differences in their composition, for instance the Earth is comprised of 30 % iron, whereas it is believed the Moon only contains 10 % iron. The Moon also has fewer volatile elements, suggesting that they may have been lost during formation.

Isotope similarities could suggest that the Earth and the Impactor were formed from the same material, or that the Earth and the Impactor equilibrated fully. However both of these processes require additional explanation or information. For example numerical modelling of timescales of isotope equilibration and the timescale of the impactor don't match up, and that different planetary bodies show different isotope signatures suggests that two planetary bodies, however closely they form, should show some differences.

Therefore the similarity between the isotope signatures of the Earth and the Moon places constraints on the impact and the Impactor, and raises more questions than it answers. Initially the similarity of some of the isotope systems supported the Giant Impact theory, which suggested that there was some equilibration, or mixing, between the Earth and the Impactor, however the similarity of silicon and tungsten isotopes that are controlled by very different processes requires additional explanation.

Any theory of Moon formation will have to explain the dynamics of the Earth–Moon system, which have to be accounted for in terms of the angular moment and spin of the Earth. Many numerical models of the impact have been run, and the optimal situation appears to be one in which an impactor around $1/10^{th}$ of the size of the Earth grazes it and forms a debris disk predominantly of impactor material and to a lesser degree of terrestrial material. Hence the Moon is mainly formed from impactor material. More recent modelling suggests that a larger impactor may be responsible; however the probability of this occurring is much lower than the probability of impact by a smaller body.

In some ways the evolution of the Earth-Moon system will remain enigmatic, as a lot of the hypothesized events or processes are relatively improbable. This does not however mean that they cannot occur, but currently we do not have the information that would support the Giant Impact Theory.

The problem remains that all of the current models require secondary processes overprinting the primary formation information. Sometimes the simplest model is the best, but as a lot of the unknown factors are untestable we may never know.

However the future for lunar research is not bleak, as was evident at this year's Royal Society Meeting in London, which encompassed a lively range of debates and discussions. The general consensus from the meeting was that the Giant Impact Theory is the most probable mechanism for moon formation; however more information regarding the other terrestrial planets, for example Venus and certain asteroids is required to shed more light on early solar system processes.