General palaeontology (Palaeobiochemistry)

Characterizing habitable extrasolar planets using spectral fingerprints

Lisa Kaltenegger\textsuperscript{a,}\textsuperscript{*}, Franck Selsis\textsuperscript{b}

\textsuperscript{a} Harvard Smithsonian Center for Astrophysics, 60, Garden Street, 02138 MA Cambridge, USA
\textsuperscript{b} Laboratoire d’astrophysique de Bordeaux (CNRS, Université Bordeaux-1), BP 89, 33271 Floirac cedex, France

Received 9 January 2009; accepted after revision 3 July 2009
Available online 4 November 2009
Written on invitation of the Editorial Board

Abstract

The detection and characterization of an Earth-like planet is approaching rapidly thanks to radial velocity (RV) surveys (e.g. HARPS) and transit searches (Corot, Kepler). A rough characterization of these planets will be already achievable in 2014 with the James Webb Space Telescope, and more detailed spectral studies will be obtained by future large ground based telescopes (ELT, TNT, GMT), and dedicated space-based missions like Darwin, Terrestrial Planet Finder, New World Observer. In this article we discuss how we can read a planet’s spectrum to assess its habitability and search for the signatures of a biosphere. Identifying signs of life implies understanding how the observed atmosphere physically and chemically works, and thus gathering information on the planet in addition to observing its spectral fingerprint. To cite this article: L. Kaltenegger, F. Selsis, C. R. Palevol 8 (2009).

© 2009 Published by Elsevier Masson SAS on behalf of l’Académie des sciences.

Keywords: Habitable planets; Extrasolar planet search; Biomarkers; Planetary atmospheres

Mots clés : Planètes habitables ; Recherche des planètes extrasolaires ; Biomarqueurs ; Atmosphère planétaire

1. Introduction

The emerging field of extrasolar planet search has shown an extraordinary ability to combine research by astrophysics, chemistry, biology and geophysics into a
new and exciting interdisciplinary approach to understand our place in the universe. Space missions like CoRoT (CNES, [45]) and Kepler (NASA, [4]) will give us statistics on the number, size, period and orbital distance of planets, extending to terrestrial planets on the lower mass range end as a first step, while future space missions are designed to characterize their atmospheres. After a decade rich in giant exoplanet detection, indirect ground based observation techniques have now reached the ability to find planets of less than 10 M_{Earth} (so called Super-Earths) around small stars that may potentially be habitable [35,64]. These planets can be characterized with future space missions.

The current status of exoplanet characterization shows a surprisingly diverse set of planets. For a subset of these, some properties have been measured or inferred using observations of the host-star, a background star, or the combination of the star and planet photons (radial velocity [RV], microlensing, transits, and astrometry). These observations have yielded measurements of planetary mass, orbital elements and (for transits) the planetary radius and during the last few years, physical and chemical characteristics of the upper atmosphere of some of the transiting planets. Specifically, observations of transits, combined with RV information, have provided estimates of the mass and radius of the planet, planetary brightness temperature [6,12], planetary day-night temperature difference [20,30], and even absorption features of giant planetary upper-atmospheric constituents: sodium [7], hydrogen [66], water [58,60], methane [59], carbon monoxide and dioxide [61]. Using the Earth itself as a proxy shows the opportunities, limitations and importance of co-adding transits to detect biomarkers on Earth-analog exoplanets [23]. The first imaged exoplanet candidates or Brown Dwarfs around young stars show the improvement in direct detection techniques that are designed to resolve the planet and collect its photons. This can currently be achieved for widely separated young objects, and has already detected exoplanet candidates [2,8,18,21,31,34,43].

Future space missions are designed to image and characterize smaller planets, down to Earth-size (e.g. James Webb Space Telescope [JWST], Darwin, Terrestrial Planet Finder [TPF], New World Observer [NWO]), and to measure the color and spectra of terrestrial planets, giant planets, and zodiacal dust disks around nearby stars [3,16,21–23,32]. These missions have the explicit purpose of detecting other Earth-like worlds, analyzing their characteristics, determining the composition of their atmospheres, investigating their capability to sustain life as we know it, and searching for signs of life. They also have the capacity to investigate the physical properties and composition of a broader diversity of planets, to understand the formation of planets and interpret potential biosignatures. On Earth, some atmospheric species exhibiting noticeable spectral features in the planet’s spectrum result directly or indirectly from biological activity: the main ones are O2, O3, CH4, and N2O. Sagan et al. [46] analyzed a spectrum of the Earth taken by the Galileo probe, searching for signatures of life and concluded that the large amount of O2 and the simultaneous presence of CH4 traces are strongly suggestive of biology. CO2 and H2O are, in addition, important as greenhouse gases in a planet’s atmosphere and potential sources for the high O2 concentration from photosynthesis. In this article we describe how to characterize a habitable planet in section 2, focus on low resolution biomarkers in the spectrum of an Earth-like planet in section 3, and discuss cryptic worlds, abiogenic sources of biomarkers, spectral evolution of a habitable planet and Earth’s spectra around different host stars in section 4.

2. First steps to characterize a habitable planet

A planet is a very faint, small object close to a very bright and large object, its parent star. In the visible part of the spectrum we observe starlight, reflected off the planet; in the IR we detect the planet’s own emitted flux. The Earth-Sun intensity ratio is about 10^{-7} in the thermal infrared (∼10 μm), and about 10^{-10} in the visible (∼0.5 μm). The spectrum of the planet can contain signatures of atmospheric species, which create its spectral fingerprint. The tradeoff between contrast ratio and design is not discussed here, but lead to several different configurations for a space-based mission concept. The interferometric systems suggested for Darwin and the TPF Interferometer (TPF-I) mission operate in the mid-IR (6–20 μm) and observe the thermal emission emanating from the planet. The TPF Coronagraph (TPF-C) and the occulter of NWO look at the reflected light and operate in the visible and near infrared (0.5–1 μm). The viewing geometry results in a different flux contribution of the overall detected signal from the bright and dark side, for the reflected light, and the planet’s hot and cold regions for the emitted flux. Both spectral regions contain the signature of atmospheric gases that may indicate habitable conditions and, possibly, the presence of a biosphere: CO2, H2O, O3, CH4, and N2O in the thermal infrared, and H2O, O3, O2, CH4 and CO2 in the visible to near-infrared. The presence or absence of these spectral features (detected individually or collectively) will indicate similarities or differences with the atmospheres of terrestrial planets, and the astrobiological potential (Fig. 1).
Our search for signs of life is based on the assumption that extraterrestrial life shares fundamental characteristics with life on Earth, in that it requires liquid water as a solvent and has a carbon-based chemistry [4,5,13,39]. Life on the base of a different chemistry is not considered here because the vast possible life-forms produce signatures in their atmosphere that are so far unknown. Therefore we assume that extraterrestrial life is similar to life on Earth in its use of the same input and output gases, that it exists out of thermodynamic equilibrium, and that it has analogs to bacteria, plants, or animals on Earth [33]. The term biomarkers is used here to mean detectable species, or a set of species, whose presence at significant abundance strongly suggests a biological origin (e.g. couple $\text{CH}_4 + \text{O}_2$, or $\text{CH}_4 + \text{O}_3$, [33]). Bio-indicators are indicative of biological processes, but can also be produced abiotically. It is their quantities, and detection along with other atmospheric species, and in a certain context (for instance the properties of the star and the planet) that points toward a biological origin.

With arbitrarily high signal-to-noise and spatial and spectral resolution, it is relatively straightforward to remotely ascertain that Earth is a habitable planet, replete with oceans, a greenhouse atmosphere, global geochemical cycles, and life. The interpretation of observations of other planets with limited signal-to-noise ratio and spectral resolution, as well as absolutely no spatial resolution, as envisioned for the first generation instruments, will be far more challenging and implies that we need to gather as much information as possible in order to understand what we will see. The following step-by-step approach can be taken to set the system in context. After detection, we will focus on the main properties of the planetary system, its orbital elements as well as the presence of an atmosphere using the light curve of the planet or/and a crude estimate of the planetary nature using very low-resolution information (3 or 4 channels) [62]. Then a higher resolution spectrum will be used to identify the compounds of the planetary atmosphere, constrain the temperature and radius of the observed exoplanet. In that context, we can then test if we have an abiotic explanation of all compounds seen in the atmosphere of such a planet. If we do not, we can work with the exciting biotic hypothesis. $\text{O}_2$, $\text{O}_3$, $\text{CH}_4$ are good biomarker candidates that can be detected by a low-resolution (Resolution < 50) spectrograph. Note that if the presence of biogenic gases such as $\text{O}_2/\text{O}_3 + \text{CH}_4$ may imply the presence of a massive and active biosphere, their absence does not imply the absence of life. Life existed on Earth before the interplay between oxygenic photosynthesis and carbon cycling produced an oxygen-rich atmosphere. The results of a first generation mission will most likely result in an amazing scope of diverse planets that will set planet formation, evolution as well as our planet in an overall context.

2.1. Temperature and radius of a planet

The search for signs of life implies that one understands how the observed atmosphere physically and chemically works. Knowledge of the temperature and planetary radius is crucial for the general
understanding of the physical and chemical processes occurring on the planet (e.g. tectonic, hydrogen loss to space). This question is far from being easy. In theory, spectroscopy can provide some detailed information on the thermal profile of a planetary atmosphere. This, however, requires a spectral resolution and a sensitivity that are well beyond the performance of a first generation spacecraft. Thus we will concentrate on the initially available observations here.

One can calculate the stellar energy of the star $F_{\text{star}}$ that is received at the measured orbital distance. The surface temperature of the planet depends on its albedo and on the greenhouse warming by atmospheric compounds. However, with a low-resolution spectrum of the thermal emission, the mean effective temperature and the radius of the planet can be obtained. The surface temperature can be estimated by adjusting thermal emission to a model. The ability to associate a brightness temperature to the spectrum relies on the existence and identification of spectral windows probing the surface or the same atmospheric levels. For an Earth-like planet there are some atmospheric windows that can be used in most of the cases, especially between 8 and 11 $\mu$m as seen in Fig. 1. This window would however become opaque at high H$_2$O partial pressure (e.g. the inner part of the Habitable Zone [HZ] where a lot of water is vaporized) and at high CO$_2$ pressure (e.g. a very young Earth [Fig. 3] or the outer part of the HZ). The accuracy of the radius and temperature determination will depend on the quality of the fit (and thus on the sensitivity and resolution of the spectrum), the precision of the Sun-star distance, the cloud coverage and also the distribution of brightness temperatures over the planetary surface. For transiting planets, for which the radius is known, the measured IR flux can directly be converted into a brightness temperature that will provide information on the temperature of the atmospheric layers responsible for the emission. If the mass of non-transiting planets can be measured (by RV and/or astrometric observations), an additional estimate of the radius can be made by assuming a bulk composition of the planet, which can then be used to convert IR fluxes into temperatures.

2.2. Biomarkers detection set in context

The orbital flux variation in the IR can distinguish planets with and without an atmosphere in the detection phase [17,56]. Strong variation of the thermal flux with phase reveals a strong difference in temperature between the day and night hemisphere of the planet, a consequence of the absence of a dense atmosphere. In such a case, estimating the radius from the thermal emission is made difficult because most of the flux received comes from the small and hot substellar area. The ability to retrieve the radius in such a case would depend on the assumption that can be made on the orbit geometry and the rotation rate of the planet. For instance, if the rotation is likely to have been tidally synchronized, and the inclination of the system is known, then the temperature variations from the substellar point to the terminator can be modeled accurately and the radius can be derived. However, in most of the other cases, degenerate solutions will exist. When the mean brightness temperature ($T_b$) is stable along the orbit, the estimated radius is more reliable. The radius can be measured at different points of the orbit and thus for different values of $T_b$, which should allow one to estimate the error made. The thermal light curve (i.e. the integrated infrared emission measured at different position on the orbit) exhibits smaller variations due to the phase (whether the observer sees mainly the day side or the night side) and to the season. Important phase-related variations are due to a high day/night temperature contrast and imply a low greenhouse effect and the absence of a stable liquid ocean. Therefore, habitable planets can be distinguished from airless or Mars-like planets by the amplitude of the observed variations of $T_b$. Note that also a Venus-like exoplanet would exhibit nearly no measurable phase-related variations of its thermal emission, due to the fast rotation of its atmosphere and its strong greenhouse effect, and can only be distinguished through spectroscopy from habitable planets. The mean value of $T_b$ estimated over an orbit can be used to estimate the albedo of the planet, $A$, through the balance between the incoming stellar radiation and the outgoing IR emission. In the visible ranges, the reflected flux allows us to measure the product $A \times R^2$, where $R$ is the planetary radius (a small but reflecting planet appears as bright as a big but dark planet). The first generation of optical instruments will be very far from the angular resolution required to directly measure an exoplanet radius. Presently, such a measurement can only be performed when the planet transits in front of its parent star, by an accurate photometric technique. If the secondary eclipse of the transiting planet can be observed (when the planet passes behind the star), then the thermal emission of the planet can be measured, allowing the retrieval of mean brightness temperature $T_b$, thanks to the knowledge of the radius from the primary transit. If a non-transiting target is observed in both visible and IR ranges, the albedo can be estimated in the visible once the radius is inferred from the IR spectrum, and compared with one derived from the thermal emission only.
3. Habitable planets

The circumstellar HZ is defined as the region around a star within which starlight is sufficiently intense to maintain liquid water at the surface of the planet, without initiating runaway greenhouse conditions vaporizing the whole water reservoir and, as a second effect, inducing the photodissociation of water vapor and the loss of hydrogen to space (Kasting et al. [27–29] and Selsis [55] for a detailed discussion). The semi-major axis in the middle of the HZ \(a_{\text{HZ}}\), is derived by scaling the Earth-Sun system using \(L_{\text{star}}/L_{\text{sun}} = (R_{\text{star}}/R_{\text{sun}})^2 (T_{\text{star}}/T_{\text{sun}})^4\), so \(a_{\text{HZ}} = 1 \text{ AU} (L_{\text{star}}/L_{\text{sun}})^{1/2} (R_{\text{star}}/R_{\text{sun}})^{1/2}\) is molecular oxygen extracted from water, allows terrestrial plants and photosynthetic bacteria (cyanobacteria) to use abundant \(H_2O\) instead of having to rely on scarce supplies of electron donor to reduce \(CO_2\), like \(H_2\) and \(H_2S\). With oxygenic photosynthesis, the production of the biomass becomes limited only by nutrients and no longer by energy (light in this case) nor by the abundance of electron donors. Oxygenic photosynthesis at a planetary scale results in the storage of large amounts of radiative energy in chemical energy, in the form of organic matter. For this reason, oxygenic photosynthesis had a tremendous impact on biogeochemical cycles on Earth and eventually resulted in the global transformation of Earth environment. Less than 1 ppm of atmospheric \(O_2\) comes from abiotic processes [67]. Cyanobacteria and plants are responsible for this production by using the solar photons to extract hydrogen from water and using it to produce organic molecules from \(CO_2\). This metabolism is called oxygenic photosynthesis. The reverse reaction, using \(O_2\) to oxidize the organics produced by photosynthesis, can occur abiotically when organics are exposed to free oxygen, or biotically by eukaryotes breathing \(O_2\) and consuming organics. Because of this balance, the net release of \(O_2\) in the atmosphere is due to the burial of organics in sediments. Each reduced carbon buried lets a free \(O_2\) molecule into the atmosphere. This net release rate is also balanced by weathering of fossilized carbon when exposed to the surface. The oxidation of reduced volcanic gasses such as \(H_2\), \(H_2S\) also accounts for a significant fraction of the oxygen losses. The atmospheric oxygen is recycled through respiration and photosynthesis in less than 10,000 yrs. In the case of a total extinction of the Earth biosphere, the atmospheric \(O_2\) would disappear in a few million years.

Reduced gases and oxygen have to be produced concurrently to be detectable in the atmosphere, as they react rapidly with each other. Thus, the chemical imbalance traced by the simultaneous signature of \(O_2\) and/or \(O_3\) and of a reduced gas like \(CH_4\) can be considered as a signature of biological activity [33]. The spectrum of the Earth has exhibited a strong infrared signature of ozone for more than 2 billion years, and a strong visible signature of \(O_2\) for an undetermined period of time between 2 and 0.8 billion years (depending on the required depth of the band for detection and also the actual evolution of the \(O_2\) level) [25]. This difference is due to the fact that a saturated ozone band appears already at very low levels of \(O_2\) (10^{-3} ppm) while the oxygen line remains unsaturated at values below one PAL [52]. In addition, the stratospheric warming decreases with the abundance

3.1. Potential biomarkers

Owen [39] suggested searching for \(O_2\) as a tracer of life. Oxygen in high abundance is a promising bi-indicator. Oxygenic photosynthesis, which by-product

is molecular oxygen extracted from water, allows terrestrial plants and photosynthetic bacteria (cyanobacteria) to use abundant \(H_2O\), instead of having to rely on scarce supplies of electron donor to reduce \(CO_2\), like \(H_2\) and \(H_2S\). With oxygenic photosynthesis, the production of the biomass becomes limited only by nutrients and no longer by energy (light in this case) nor by the abundance of electron donors. Oxygenic photosynthesis at a planetary scale results in the storage of large amounts of radiative energy in chemical energy, in the form of organic matter. For this reason, oxygenic photosynthesis had a tremendous impact on biogeochemical cycles on Earth and eventually resulted in the global transformation of Earth environment. Less than 1 ppm of atmospheric \(O_2\) comes from abiotic processes [67]. Cyanobacteria and plants are responsible for this production by using the solar photons to extract hydrogen from water and using it to produce organic molecules from \(CO_2\). This metabolism is called oxygenic photosynthesis. The reverse reaction, using \(O_2\) to oxidize the organics produced by photosynthesis, can occur abiotically when organics are exposed to free oxygen, or biotically by eukaryotes breathing \(O_2\) and consuming organics. Because of this balance, the net release of \(O_2\) in the atmosphere is due to the burial of organics in sediments. Each reduced carbon buried lets a free \(O_2\) molecule into the atmosphere. This net release rate is also balanced by weathering of fossilized carbon when exposed to the surface. The oxidation of reduced volcanic gasses such as \(H_2\), \(H_2S\) also accounts for a significant fraction of the oxygen losses. The atmospheric oxygen is recycled through respiration and photosynthesis in less than 10,000 yrs. In the case of a total extinction of the Earth biosphere, the atmospheric \(O_2\) would disappear in a few million years.

Reduced gases and oxygen have to be produced concurrently to be detectable in the atmosphere, as they react rapidly with each other. Thus, the chemical imbalance traced by the simultaneous signature of \(O_2\) and/or \(O_3\) and of a reduced gas like \(CH_4\) can be considered as a signature of biological activity [33]. The spectrum of the Earth has exhibited a strong infrared signature of ozone for more than 2 billion years, and a strong visible signature of \(O_2\) for an undetermined period of time between 2 and 0.8 billion years (depending on the required depth of the band for detection and also the actual evolution of the \(O_2\) level) [25]. This difference is due to the fact that a saturated ozone band appears already at very low levels of \(O_2\) (10^{-3} ppm) while the oxygen line remains unsaturated at values below one PAL [52]. In addition, the stratospheric warming decreases with the abundance

is molecular oxygen extracted from water, allows terrestrial plants and photosynthetic bacteria (cyanobacteria) to use abundant \(H_2O\), instead of having to rely on scarce supplies of electron donor to reduce \(CO_2\), like \(H_2\) and \(H_2S\). With oxygenic photosynthesis, the production of the biomass becomes limited only by nutrients and no longer by energy (light in this case) nor by the abundance of electron donors. Oxygenic photosynthesis at a planetary scale results in the storage of large amounts of radiative energy in chemical energy, in the form of organic matter. For this reason, oxygenic photosynthesis had a tremendous impact on biogeochemical cycles on Earth and eventually resulted in the global transformation of Earth environment. Less than 1 ppm of atmospheric \(O_2\) comes from abiotic processes [67]. Cyanobacteria and plants are responsible for this production by using the solar photons to extract hydrogen from water and using it to produce organic molecules from \(CO_2\). This metabolism is called oxygenic photosynthesis. The reverse reaction, using \(O_2\) to oxidize the organics produced by photosynthesis, can occur abiotically when organics are exposed to free oxygen, or biotically by eukaryotes breathing \(O_2\) and consuming organics. Because of this balance, the net release of \(O_2\) in the atmosphere is due to the burial of organics in sediments. Each reduced carbon buried lets a free \(O_2\) molecule into the atmosphere. This net release rate is also balanced by weathering of fossilized carbon when exposed to the surface. The oxidation of reduced volcanic gasses such as \(H_2\), \(H_2S\) also accounts for a significant fraction of the oxygen losses. The atmospheric oxygen is recycled through respiration and photosynthesis in less than 10,000 yrs. In the case of a total extinction of the Earth biosphere, the atmospheric \(O_2\) would disappear in a few million years.

Reduced gases and oxygen have to be produced concurrently to be detectable in the atmosphere, as they react rapidly with each other. Thus, the chemical imbalance traced by the simultaneous signature of \(O_2\) and/or \(O_3\) and of a reduced gas like \(CH_4\) can be considered as a signature of biological activity [33]. The spectrum of the Earth has exhibited a strong infrared signature of ozone for more than 2 billion years, and a strong visible signature of \(O_2\) for an undetermined period of time between 2 and 0.8 billion years (depending on the required depth of the band for detection and also the actual evolution of the \(O_2\) level) [25]. This difference is due to the fact that a saturated ozone band appears already at very low levels of \(O_2\) (10^{-3} ppm) while the oxygen line remains unsaturated at values below one PAL [52]. In addition, the stratospheric warming decreases with the abundance
of ozone, making the O₃ band deeper for an ozone layer less dense than in the present atmosphere. The depth of the saturated O₃ band is determined by the temperature difference between the surface-clouds continuum and the ozone layer. Note again that the non-detection of O₂ or O₃ on an exoplanet cannot be interpreted as the absence of life.

N₂O is produced in abundance by life but only in negligible amounts by abiotic processes. Nearly all of Earth’s N₂O is produced by the activities of anaerobic denitrifying bacteria. N₂O would be hard to detect in the Earth’s atmosphere with low resolution, as its abundance is low at the surface (0.3 ppmv) and falls off rapidly in the stratosphere. Spectral features of N₂O would become more apparent in atmospheres with more N₂O and/or less H₂O vapor. Segura et al. [52] have calculated the level of N₂O for different O₂ levels and found that, although N₂O is a reduced species compared to N₂, its levels decreases with O₂. This is due to the fact that a decrease in O₂ produces an increase of H₂O photolysis resulting in the production of more hydroxyl radicals (OH) responsible for the destruction of N₂O. The detection of H₂O and CO₂, not as biosignatures themselves, are important in the search for signs of life because they are raw materials for life and thus necessary for planetary habitability. The methane found in the present atmosphere of the Earth has a biological origin, except for a small fraction produced abiotically in hydrothermal systems, where hydrogen is released by the oxidation of Fe by H₂O, and reacts with CO₂. Depending on the degree of oxidation of a planet’s crust and upper mantle, such non-biological mechanisms can also produce large amounts of CH₄ under certain circumstances. Therefore, the detection of methane alone cannot be considered a sign of life, while its detection in an oxygen-rich atmosphere would be difficult to explain in the absence of a biosphere. Note that methane may have been detected on Mars [38], while the atmosphere of Mars contains 0.1% of O₂ and some ozone.

In this case, the amounts involved are extremely low and the origin of the Martian O₂ and O₃ is known to be photochemical reactions initiated by the photolysis of CO₂ and water vapor. If confirmed, the presence of methane could be explained by subsurface geochemical processes, assuming that reducing conditions exist on Mars below the highly oxidized surface. The case of NH₃ is similar to that of CH₄. They are both released into Earth’s atmosphere by the biosphere with similar rates but the atmospheric level of NH₃ is orders of magnitude lower due to its very short lifetime under UV irradiation. The detection of NH₃ in the atmosphere of a habitable planet would thus be extremely interesting, especially if found with oxidized species. There are other molecules that could, under some circumstances, act as excellent biomarkers, e.g., the manufactured chloro-fluorocarbons (CCl₂F₂ and CCl₃F) in our current atmosphere in the thermal infrared waveband, but their abundances are too low to be spectroscopically observed at low resolution.

3.1.1. Low resolution spectral information in the visible to near-infrared

In the visible to near-infrared one can see increasingly strong H₂O bands at 0.73 μm, 0.82 μm, 0.95 μm, and 1.14 μm. The strongest O₂ feature is the saturated Fraunhofer A-band at 0.76 μm. A weaker feature at 0.69 μm cannot be seen with low resolution (see Fig. 1). O₃ has a broad feature, the Chappuis band, which appears as a broad triangular dip in the middle of the visible spectrum from about 0.45 μm to 0.74 μm. The feature is very broad and shallow. Methane at present terrestrial abundance (1.65 ppm) has no significant visible absorption features, but at high abundance, it has strong visible bands at 0.88 μm, and 1.04 μm, readily detectable e.g. in early Earth models (see Fig. 3). CO₂ has negligible visible features at present abundance, but in a high CO₂-atmosphere of 10% CO₂, as in an early Earth evolution stage, the weak 1.06 μm band could be observed. In the UV O₃ shows a strong feature, not discussed here. The red edge of land plants developed about 0.44Ga. It could be observed on a cloud-less Earth or if the cloud pattern is known (see section 4).

3.1.2. Low resolution spectral information in the mid-infrared

In the mid-infrared on Earth the detectable signatures of biological activity in low resolution are the combined detection of 9.6 μm O₃ band, the 15 μm CO₂ band and the 6.3 μm H₂O band or its rotational band that extends from 12 μm out into the microwave region. The 9.6 μm O₃ band is highly saturated and is thus a poor quantitative indicator, but an excellent qualitative indicator for the existence of even traces of O₂. CH₄ is not readily identified using low resolution spectroscopy for present-day Earth, but the methane feature at 7.66 μm in the IR is easily detectable at higher abundances (e.g. 100 × on early Earth [24]), provided, of course, that the spectrum contains the whole band and a high enough SNR. Taken together with molecular oxygen, abundant CH₄ can indicate biological processes [34,52]. CH₄ and N₂O have features nearly overlapping in the 7 μm region, and, additionally, both lie in the red wing of the 6 μm water band. Although methane’s abundance is less than 1 ppm in Earth’s atmosphere, the 7.75 μm shows up in a
medium resolution (Res = 100) infrared spectrum. Three N_2O features in the thermal infrared are detectable at 7.75 μm, 8.52 μm, and 16.89 μm for levels higher than in the present atmosphere of the Earth.

4. Abiotic sources, cryptic worlds, geological evolution and host stars

4.1. Abiotic sources of biomarkers

We need to address the abiotic sources of biomarkers, so that we can identify when it might constitute a “false positive” for life. CH_4 is an abundant constituent of the cold planetary atmospheres in the outer solar system. On Earth, it is produced abiotically in hydrothermal systems where H_2 (produced from the oxidation of Fe by water) reacts with CO_2 in a certain range of pressures and temperatures. In the absence of atmospheric oxygen, abiotic methane could build up to detectable levels. Therefore, the sole detection of CH_4 cannot be attributed unambiguously to life.

O_2 also has abiotic sources: the first one is the photolysis of CO_2, followed by recombination of O atoms to form O_2 (O + O + M → O_2 + M); a second one is the photolysis of H_2O followed by escape of hydrogen to space. The first source is a steady state maintained by the stellar UV radiation, but with a constant elemental composition of the atmosphere while the second one is a net source of oxygen. In order to reach detectable levels of O_2 (in the reflected spectrum), the photolysis of CO_2 has to occur in the absence of outgassing of reduced species and in the absence of liquid water, because of the wet deposition of oxidized species. Normally, the detection of the water vapor bands simultaneously with the O_2 band can rule out this abiotic mechanism [52], although one should be careful, as the vapor pressure of H_2O over a high-albedo icy surface might be high enough to produce detectable H_2O bands. In the infrared, this process cannot produce a detectable O_3 feature [57]. The loss of hydrogen to space can result in massive oxygen leftovers: more than 200 bars of oxygen could build up after the loss of the hydrogen contained in the Earth Ocean. However, the case of Venus tells us that such oxygen leftover has a limited lifetime in the atmosphere (because of the oxidation of the crust and the loss of oxygen to space): we do not find O_2 in the Venusian atmosphere despite the massive loss of water probably experienced in the early history of the planet. Also, such evaporation-induced build up of O_2 should occur only closer to a certain distance from the Star and affect small planets with low gravity more dramatically. For small planets (< 0.5 M_{Earth}) close to inner edge of the HZ (< 0.93 AU from the present Sun), there is a risk of abiotic oxygen detection, but this risk becomes negligible for big planets further away from their star. The fact that, on the Earth, oxygen and indirectly ozone are by-products of the biological activity does not mean that life is the only process able to enrich an atmosphere with these compounds. The question of the abiotic synthesis of biomarkers is crucial, but only very few studies have been dedicated to it [32,44,54,57,62].

4.2. Surface-, vegetation-, cloud-features

While they efficiently absorb the visible light, photosynthetic plants have developed strong infrared reflection (possibly as a defense against overheating and chlorophyll degradation), resulting in a steep change in reflectivity around 700 nm, called the red-edge. The primary molecules that absorb the energy and convert it to drive photosynthesis (H_2O and CO_2 into sugars and O_2) are chlorophyll A (0.450 μm) and B (0.680 μm). The exact wavelength and strength of the spectroscopic “vegetation red edge” (VRE) depends on the plant species and environment. On Earth around 440 million years ago [42,48], an extensive land plant cover developed, generating the red chlorophyll edge in the reflection spectrum between 700 and 750 nm. Averaged over a spatially unresolved hemisphere of Earth, the additional reflectivity of this spectral feature is typically only a few percent [25,36]. Several groups [1,9,37,63,68] have measured the integrated Earth spectrum via the technique of Earthshine, using sunlight reflected from the non-illuminated, or “dark”, side of the moon. Earthshine measurements have shown that detection of the Earth’s VRE is feasible if the resolution is high and the cloud coverage is known, but is made difficult owing to its broad, essentially featureless spectrum and cloud coverage [36]. Our knowledge of the reflectivity of different surface components on Earth – such as deserts, ocean and ice – helps in assigning the VRE of the Earthshine spectrum to terrestrial vegetation. The Earth’s hemispherical integrated vegetation red-edge signature is weak, but planets with different rotation rates, obliquities, land-ocean fraction, and continental arrangement may have lower cloud-cover and higher vegetated fraction [50]. Knowing that other pigments exist on Earth and that some minerals can exhibit a similar spectral shape around 750 nm [49], the detection of the red-edge of the chlorophyll on exoplanets, despite its interest, will not be unambiguous. Assuming that similar photosynthesis would evolve on a planet around other stellar types, possible different types of spectral signature have been modeled that could be a guide to interpreting other spectral signatures. These signatures will be difficult to
verify through remote observations as being of biological origin.

Another topic that has been proposed to discover continents and seas on an exoplanet is the daily variation of the surface albedo in the visible [14,49]. On a cloud-free Earth, the diurnal flux variation in the visible caused by different surface features rotating in and out of view could be high, assuming hemispheric inhomogeneity. When the planet is only partially illuminated, a more concentrated signal from surface features could be detected as they rotate in and out of view on a cloudless planet. Earth has an average of 60% cloud coverage that prevents easy identification of the features without knowing the cloud distribution. Clouds are an important component of exoplanet spectra because their reflection is high and relatively flat with wavelength. Clouds hide the atmospheric molecular species below them, weakening the spectral lines in both the thermal infrared and visible [25,36]. In the thermal infrared, clouds emit at temperatures that are generally colder than the surface, while in the visible the clouds themselves have different spectrally-dependent albedo that further influence the overall shape of the spectrum. Clouds reduce the relative depths, full widths, and equivalent widths of spectral features. If one could record the planet’s signal with a very high time resolution (a fraction of the rotation period of the planet) and SNR, one could determine the overall contribution of clouds to the signal [11,40].

During each of these individual measurements, one has to collect enough photons for a high individual SNR per measurement to be able to correlate the measurements to the surface features, which precludes this method for first generation missions that will observe a minimum of several hours to achieve a SNR of 5 to 10. For the Earth [11,40], these measurements show a correlation to the Earth’s surface features because the individual measurements are time resolved as well as having an individual high SNR, making it a very interesting concept for future generations of missions.

4.3. Cryptic worlds

On the Earth, photosynthetic organisms are responsible for the production of nearly all of the oxygen in the atmosphere. However, in many regions of the Earth, and particularly where surface conditions are extreme, for example, in hot and cold deserts, photosynthetic organisms can be driven into and under substrates where light is still sufficient for photosynthesis. These communities exhibit no distinguishable detectable surface spectral signature. The same is true of the assemblages of photosynthetic organisms at more than a few meters depth in bodies of water. These communities are widespread and dominate local photosynthetic productivity. Fig. 2 shows known cryptic photosynthetic communities and the calculated disk-averaged spectra of such hypothetical cryptic photosynthesis worlds. Such worlds are Earth-analogs that would not exhibit a distinguishable biological surface feature like the VRE in the disk-averaged spectrum but still be inhabited, and remotely detectable [10].

4.4. Evolution of biomarkers over geological times on Earth

One crucial factor in interpreting planetary spectra is the point in the evolution of the atmosphere when its biomarkers and its habitability become detectable. Spectra of the Earth exploring temperature sensitivity (hot house and cold scenario) and different singled out stages of its evolution [41,47], as well as exploring the evolution of the expected spectra of Earth [25], produce a variety of spectral fingerprints for our own planet. These spectra will be used as part of a big grid to characterize any exoplanets found and influence the design requirements for a spectrometer to detect habitable planets.

The spectrum of the Earth has not been static throughout the past 4.5 Ga. This is due to the variations in the molecular abundances, the temperature structure, and the surface morphology over time. At about 2.3 Ga oxygen and ozone became abundant, affecting the atmospheric absorption component of the spectrum. At about 0.44 Ga, an extensive land plant cover followed, generating the red chlorophyll edge in the reflection spectrum. The composition of the surface (especially in the visible), the atmospheric composition, and temperature-pressure profile can all have a significant influence on the detectability of a signal. Fig. 3 shows theoretical visible and mid-infrared spectra of the Earth at six epochs during its geological evolution [25,28]. The epochs are chosen to represent major developmental stages of the Earth, and life on Earth. If an extrasolar planet is found with a corresponding spectrum, we can use the stages of evolution of our planet to characterizing it, in terms of habitability and the degree to which it shows signs of life. Furthermore, we can learn about the evolution of our own planet’s atmosphere and possible the emergence of life by observing exoplanets in different stages of their evolution. The Earth’s atmosphere has experienced a dramatic evolution over 4.5 billion years, and other planets may exhibit similar or greater evolution, and at different rates. It shows epochs that reflect significant changes in the chemical composition of the atmosphere. The oxygen and ozone absorption features could have been used to indicate the
Fig. 2. Two examples of land-based cryptic photosynthetic communities: (top) a cryptoendolithic lichen (arrow) inhabiting the interstices of sandstone in the Dry Valleys of the Antarctic, (bottom) endoevaporites inhabit a salt crust visible as pink pigmentation (arrow) (photo: Marli Bryant Miller) and their respective calculated clear reflection spectra. Spectra of granite and seawater are shown as two other habitats for cryptic photosynthesis communities.

presence of biological activity on Earth anytime during the past 50% of the age of the solar system. Different signatures in the atmosphere are clearly visible over Earth’s evolution and observable with low resolution.

4.5. Influence of host-stars

The range of characteristics of planets is likely to exceed by far our experience with the planets and satellites in our own Solar System. For instance, models of planets more massive than our Earth – rocky SuperEarths – need to consider the changing atmosphere structure, as well as the interior structure of the planet [51,65]. Also, Earth-like planets orbiting stars of different spectral type might evolve differently [19,23,52,53,55]. Modeling these influences will help to optimize the design of the proposed instruments used to search for Earth-like planets. The spectral resolution required for optimal detection of habitability and biosignatures has to be able to detect those features on our own planet for the dataset we have over its evolution.

Using a numerical code that simulates the photochemistry of a wide range of planetary atmospheres, several groups have simulated a replica of our planet orbiting different types of stars: an F-type star (more massive and hotter than the Sun) and a K-type star (smaller and cooler than the Sun) (see Fig. 4). The models assume same background composition of the atmosphere as well as the strength of biogenic sources.

A planet orbiting a K star has a thin O$_3$ layer, compared to Earth’s one, but still exhibits a deep O$_3$ absorption: indeed, the low UV flux is absorbed at lower
Fig. 3. The visible to near-IR (left panel, a) and mid-IR (right panel, b) spectral features on an Earth-like planet change considerably over its evolution from a CO₂ rich (epoch 0) to a CO₂/CH₄-rich atmosphere (epoch 3) to a present-day atmosphere (epoch 5). The bold lines show spectral resolution of 80 and 25 comparable to the proposed visible TPF and Darwin/TPF-I mission concept, respectively.

altitudes than on Earth which results in a less efficient warming (because of the higher heat capacity of the dense atmospheric layers). Therefore, the ozone layer is much colder than the surface and this temperature contrast produces a strong feature in the thermal emission. The process works the other way around in the case of an F-type host star. Here, the ozone layer is denser and warmer than the terrestrial one, exhibiting temperatures about as high as the surface temperature. Thus, the resulting low temperature contrast produces only a weak and barely detectable feature in the infrared spectrum.

This comparison shows that planets orbiting G (solar) and K-type stars may be better candidates for the search for the O₃ signature than planets orbiting F-type stars. This result is promising since G and K-type stars are much more numerous than F-type target stars [26], the latter being rare and affected by a short lifetime (less than 1 Gyr).

5. Summary

Any information we collect on habitability is only important in a context that allows us to interpret what we find. To search for signs of life we need to understand how the observed atmosphere works, physically and chemically. Knowledge of the temperature and planetary radius is crucial for the general understanding of the physical and chemical processes occurring on the planet. These parameters, as well as an indication of habitability, can be determined with low-resolution spectroscopy and
Fig. 4. Spectra of different host stars [53] (a), calculated IR spectrum of an Earth-analog with resolution $\lambda/\Delta\lambda = 30$ around a K star (b) and F star (c) [55].

low photon flux, as assumed for first generation space missions. The combination of spectral information in the visible (starlight reflected off the planet) as well as in the mid-IR (the planet’s thermal emission) allows a confirmation of detection of atmospheric species, a more detailed characterization of individual planets but also to explore a wide domain of planet diversity. Being able to measure the outgoing short wave and long wave radiation as well as their variations along the orbit, to determine the albedo and identify greenhouse gases, would, in combination, allow us to explore the climate system at work on the observed worlds, as well as probe planets similar to our own for habitable conditions. The results of a first generation mission will most likely show an amazing scope of diverse planets that will set planet formation, evolution as well as our planet in an overall context.

Acknowledgement

L. Kaltenegger acknowledges the support of the Harvard Origins of Life Initiative and the NASA Astrobiology Institute.

References