## A step towards the molecular detection of life on Mars

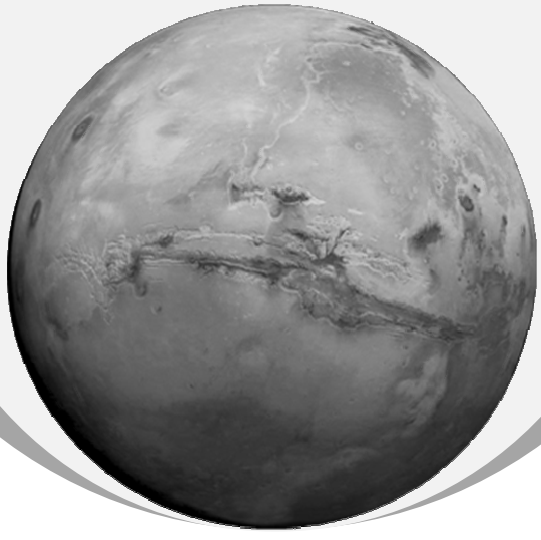
Oliveira Lebre Direito, M.S.

2012

***document version***
Publisher's PDF, also known as Version of record

# Chapter 6

General discussion



## Microbial ecology

The results described in this thesis may be relevant not just for astrobiology, but for microbial ecology in general. This relates to several methodological developments and as well as enhanced understanding on microbe-mineral interactions.

Adsorption of released nucleic acids during extraction is a major issue in the recovery of nucleic acids in many microbial ecology research projects. Low DNA recovery has a strong impact on the outcome of microbial ecology studies, in particular when addressing low biomass environments. In Chapter 2, low extractions yields were verified when spiking experiments were performed. This result appeared to be due to adsorption or degradation, revealing that DNA extraction methods needed to be optimized to improve extraction efficiencies. This issue was further addressed in Chapter 3, where first a systematic study revealed that the recovery of nucleic acids from minerals can be a major issue with losses of over 99% which could negatively influence the detection of life. Low DNA yields will also affect the detection of less abundant members of microbial communities and impact the estimation of their diversity. Therefore, an optimized extraction protocol was designed, based on an ethanolic, high phosphate buffer-based extraction at high temperature, which improved extraction yields up to 100 fold. Despite the significant improvements of nucleic acid extraction, extraction from some minerals (clays in particular) remains submaximal. It is also not yet established how the interactions between minerals and other biomarkers (e.g. lipids) occur. Further research is needed on this issue and on the development of methods to circumvent the negative impact of adsorption of these biomarkers.

In addition, spiking experiments are essential for truly quantitative detection. As developed in Chapters 2 and 3 of this thesis, spiking techniques such as the spiking of samples with a known quantity of DNA before extraction is advised in order to quantify the

recovery of DNA from any sample. Chapter 3 also addressed which type of spike should be best. An experiment revealed that genomic DNA or cells may be best as spike and that the advantage of using intact cells over genomic DNA is that most DNA in environmental samples is probably found inside cells. However, caution should be exercised so that the spike does not contain genes that one would want to characterise in the environmental samples. This introduces complications, for example if 16S rRNA gene-based analysis is planned. One possibility would be to use Archaeal cells or *petite* mutants of *Saccharomyces cerevisiae* (without mitochondrial DNA) as spike when aiming to study bacteria in the samples. More research should be performed in order to optimize spiking techniques.

In Chapter 3, the proposed optimized extraction method promoted DNA recovery from environmental samples, while maintaining reproducible DGGE profiles. The method also proved suitable for the recovery of low molecular weight DNA (<1.5 kb). This suggests that it may also be useful for studies on ancient DNA such as DNA recovered from archaeological sites and from a wide variety of ancient sediments: such DNA tends to be fragmented and degraded. This extraction method can also be adapted for other uses such as in a variety of genetic studies or studies of newly emerging diseases and in forensic sciences.

Even if low DNA yields are achieved, higher amounts of DNA for subsequent DNA-based microbial community characterisation can be obtained by random amplification (Vora et al., 2004). In Chapter 5, Whole Genome Amplification (WGA) methods were addressed. WGA methods are a good alternative to PCR (Polymerase Chain Reaction). They are useful not only for the amplification of DNA from low biomass environments (Abulencia et al., 2006) but also for life detection investigations, because they are highly sensitive (only small amounts of template DNA are necessary), provide wide genome coverage, do not require previous sequence information and require less energy consumption (no thermal cycling at high temperatures).

However, the tested WGA methods also revealed biases correlating with differences in GC content of the genomes of species, with DNA integrity and with fragment size. MDA was biased towards low GC content species while pWGA provided results that were more similar to those of the PCR that was used as control, and less biased towards low GC content species. DNA fragmentation (an issue for ancient DNA and often also occurring during extraction employing harsh cell lysis methods) interferes severely with WGA: species with fragmented DNA were less amplified with both MDA and pWGA. pWGA did not work for amplification of low molecular weight DNA (<1.5 kb) where MDA was inefficient. For short fragmented DNA, MDA may thereby be the method of choice.

Chapter 4 revealed that the mineral composition of the sample influences microbial community structure and diversity. Importantly, the effect is more quantitative than qualitative, i.e. the same species are present independent of mineral compositions, but in different proportions. This is true for the most dominant taxa (with >1% average abundance occurring on all minerals) while diversities are still different for different types of minerals. Differences in soil mineral element content will contribute to differences and heterogeneity in microbial community structure of soil and sediment samples. This is a crucial point for ecology and for microbiology, which are too often presented as qualitative sciences, in which a molecule, a species, or a mechanism is present or absent. In reality, even as proposed around 1900 by Beijerinck (Beijerinck, 1901), all species are everywhere and can be grown from almost any real world sample. It is the relative quantities of species and molecules that determine life and it is the mechanisms that determine these quantities and their activities. For cell biology this is being highlighted by systems biology (Alberghina and Westerhoff, 2005) for ecology it has been highlighted by the systems ecology of the same origin (Westerhoff et al., 2002; Roling et al., 2007).

## Implications for the search of extraterrestrial life: conclusions and future perspectives for astrobiology

In Chapter 2, a large variation in occurrence and diversity of life was observed over short distances (even over centimetre scale distances) in an extreme Mars analogue environment. By way of example, PCR was successful for all domains of life in sample P-13 (a surface sample from Dakota Sandstone), while PCRs were unsuccessful for sample P-14 (15 cm deep at the same location). Since a similar scenario may occur for Mars (because this is a cold arid desert and an environment with mineralogy and erosion processes comparable to those on Mars), these results indicate that any future sampling mission should require a high-density sampling at any location of interest. If such a high-density sampling is omitted the chances are that life will be missed. This may have occurred on previous Mars missions.

These differences over short distances may be partially explained by differences in mineral composition. In Chapter 4, the influence of minerals on microbial communities was addressed. Mineral composition is an important factor to take into consideration when choosing landing and sampling sites for life detection missions since minerals influence bacterial communities. Especially the elemental composition of the minerals had a large effect on microbial community structure. This is particularly so for the availability of Fe(III) in minerals. Biological iron-based metabolism is considered one of the earliest metabolisms on Earth, and a possible metabolism on other iron-rich planets like Mars (Weber et al., 2006). Iron-rich minerals and other minerals containing essential nutrients for life (e.g. apatite which contains phosphate) should be considered targets for space missions. In this sense our results may aid the selection of sampling locations for life detection on Mars, such as the forthcoming Exomars mission (Ehrenfreund et al., 2011).

Mineral influence was investigated for one environment only in Chapter 4. However, other environments with relevance for Mars can also be studied, for example cave environments. Possible cave entrances of volcanic origin have been spotted on Mars (Cushing, 2011) and these may provide a protected environment for life.

One should also take into consideration that even terrestrial Mars analogue sites may present differences from the original Mars environment and these differences should be accounted for. For example, on terrestrial surface analogues the mixing of soil particles may be different from that on Mars (McSween and Keil, 2000), levels of UV radiation on the Martian surface are higher than on Earth (Kuhn and Atreya, 1979; Pavlov et al., 2002).

In Chapter 3, the proposed optimized extraction method enhanced DNA recovery from clay up to hundredfold. This is of great relevance since clays are considered to be important targets for future Mars missions (Ehrenfreund et al., 2011). They indicate past aqueous activity, may protect DNA against degradation for long periods of time (Trevors, 1996), and have been implicated in the origin of life as catalysts of reactions (Ferris, 2005). Also in Chapter 3, the recovery of nucleic acids from a range of terrestrial and Mars analogue minerals was assessed. The same was not performed for other biomarkers that may be of great relevance as indicators for life (e.g. ATP, amino acids, membrane lipids). For a life discovery program not to be restricted to a single type of biomarker is important for at least two reasons. One is that evidence of a range of organic compounds is bound to be more convincing than the identification of a single type of molecular identification (Parnell et al., 2007). The second is that life on a different planet may well differ in at least some aspects from life on Earth. If the detection methods focus exclusively on one such aspect, then existing life may go undetected, notwithstanding enormous investment in time and effort.

Problems with the extraction from mineral matrices are most likely to occur with polar or charged biomarkers like DNA which should interact with positively charged minerals (e.g. Cai et al., 2006). We have shown strong evidence for this complication. Our methods for reducing this problem may still be insufficient. Therefore more research is needed on this issue and better methods to circumvent the negative impact of adsorption are required since great losses can negatively influence the detection of life. In order to assess the efficiency of extraction methods, spiking experiments are essential to assess the robustness and effectiveness of detection. They may not suffice however, as it is uncertain whether DNA added extraneously is able to reach the same binding sites as does any DNA present in the sample. Martian living organisms might for instance have the ability to have their DNA bound so strongly, or contain that DNA within such strong membrane structures as to make extraction as fragments larger than nucleotides impossible by any of our methods.

DNA or DNA-like molecules still have a great advantage over other biomarkers since they can be amplified, fostering their sensitive detection, even of single molecules. Conventional PCR-based methods used in microbial ecology and astrobiology also have their limitation, because they require previous sequence information in relation to the use of defined primers. With respect to this limitation, other amplification methods, such as WGA, provide alternatives (Vora et al., 2004, Chapter 5). The results described in Chapter 5 indicate that in particular pWGA is suitable to detect organisms present at low numbers, while MDA is more suitable for fragmented DNA. Evidence of extinct life is likely only to remain in highly fragmented form, perpetuating the relevance of MDA notwithstanding its limitation of bias.

The identification of pWGA as a suitable method for primer-free amplification of nucleic acids, also opens the way to start evolving the polymerase in such a way that it is also capable of recognizing and replicating deviating nucleic acids. This begins to address a strong

limitation of the amplification methods used in this thesis for life detection on Mars: these methods rely on the information molecules to be almost identical to the DNA we know.

Considering that diverse nucleobases have been found in meteorites and proven to be of extraterrestrial origin (Callahan et al., 2011; Martins et al., 2008), life may have evolved outside our planet. If that extraterrestrial life would have colonized both Mars and Earth, then Martian DNA might still be similar to terrestrial DNA. The panspermia theory could support the idea that life from Mars and Earth could have been interchanged through meteorites (Fajardo-Cavazos et al., 2007; Stöffler et al., 2007) again allowing for similar DNA on Mars and Earth. Then again Martian biochemistry could be similar to terrestrial biochemistry, possibly comprising hereditary molecules with alternative nucleobases and/or backbones: even with a common origin of life it is not unlikely that the nucleic acids on the two planets would have diverged from each other. Ideally, one would therefore like to be able to amplify and detect hereditary molecules different from standard DNA. Even on Earth alternative nucleobases are found in living organisms. Examples of known alternative terrestrial nucleobases that can pair with at least one of the common nucleobases include: 5-methylcytosine (a common modified base in DNA; Ehrlich and Wang, 1981), hypoxanthine (occurs naturally in tRNA and can also be derived from a mutation in adenine; Sun and Lee, 2010), and 2,6-diaminopurine (found in S-2L cyanophage DNA to substitute for adenine completely; Kirnos et al., 1977, Cheong et al., 1988). Inosine (hypoxanthine nucleotide) occurs naturally in the brain (Zai et al., 2009) and it is known as an 'universal base' because it is able to pair with all common bases (Sun and Lee, 2010), possibly being suitable for amplification of extraterrestrial nucleic acids, without the reliability required for subsequent detection however.

The relative ease of amplification has led to intensive research on the detection of life by using nucleic acid amplification. Here it should be realized that this amplification relies on

enzymes and that enzymes are usually quite specific. The same enzymes would not reliably replicate polymers with even minor differences from standard nucleic acids. Moreover, when we set such amplification up in the laboratory it is extremely unlikely that we supply the low molecular weight substrates required for this amplification. The latter are the monomers that are part of the information containing polymers likely to be present in life. We might be better served therefore by a focus on the isolation of polymers from environments where we expect new life forms and the identification of monomers which we should then supply in attempts to amplify the polymers.

These thoughts drive home the severe limitation to much of the present drive towards detecting new life forms, be it on Mars or Earth: they may be too much based on the assumption that the unknown life forms also contain essentially the DNA and RNA as presently understood life forms do. This in fact is almost a *contradictio in terminis*.

Alternative DNA molecules or even non-DNA based life may in fact be as close as our own skin or in our own intestines but could have escaped detection until now since current amplification techniques only target known molecules. In addition, more than 99% of microorganisms are uncultivable (Kaeberlein et al., 2002) and would go undetected by colony forming assays. A study suggesting that arsenic substituted for phosphorous in DNA of a bacterium which tolerates elevated arsenic concentrations (40 mM; Wolfe-Simon et al., 2011) has been refuted in meanwhile (Reaves et al., 2012). Known backbone analogues of DNA, although artificial, include peptide (PNA), locked (LNA), glycol (GNA), and threose nucleic acids (TNA) (Ehrenfreund et al., 2006; Karkare and Bhatnagar, 2006). Recently, it was hypothesised that a TNA world may have preceded the RNA world (Yu et al., 2012). TNA has not yet been detected in living organisms. Most of the aforementioned nucleic acids are recognised by terrestrial DNA polymerases, although incorporation of common nucleotides may be required (Karkare and Bhatnagar, 2006; Veedu et al., 2008). Thus, it is expected that

terrestrial polymerases can amplify nucleic acids containing alternative backbones or nucleobases. It is an issue however whether they would produce multiple identical copies comparable to the original template. The standard detection methods used here, such as PCR and either WGA method could fail in such cases.

The answer for the detection of alternative life forms may reside in the advancing of directed evolution experiments to improve DNA polymerases in terms of relaxing their substrate specificity. DNA polymerases can be optimized to detect and repair damaged DNA (e.g. d'Abbadie et al., 2007) which is important in the case of ancient DNA but also for Mars because if life existed on Mars and was based on DNA, this life may have gone extinct (Isenbarger et al., 2008) and if DNA still exists at some locations, it might be degraded. In addition, DNA polymerases can also be engineered to replicate molecules alternative to RNA and DNA, like XNA (xeno nucleic acids) (Pinheiro et al., 2012) or TNA (Chaput and Szostak, 2003). Directed evolution experiments and design of DNA polymerases is based on the introduction of mutations (e.g. Pinheiro et al., 2012). Methods to introduce these mutations include gene shuffling, cassette mutagenesis, error-prone PCR or staggered extension process PCR. Techniques to identify polymerases with the required functions are phage or ribosomal display and complementation (Xia et al., 2002).

It is not impossible that alternative life forms exist on Earth, in the sense that they store their hereditary information in molecules other than DNA and RNA. In fact this may even be more likely than that such life exists on Mars. Therefore, terrestrial isolated niches should be considered as possible targets in the search for life forms with alternative genetic molecules. Interesting locations include the subglacial Lake Vostok (in Antarctica). The covering ice of this lake is a paleoclimatic record of 420 000 years, although its water may have been isolated for >15 million years (Christner et al., 2006). Ice from above the lake (which was once water from the lake) was analysed and showed that the lake is a source of

bacterial carbon even if this is an extremely cold, high pressure, potentially high oxygen, low nutrient and dark environment (Christner et al., 2006). 16S rRNA-gene analysis identified bacteria most closely related to *Proteobacteria* (Christner et al., 2001; Christner et al., 2006). However, one cannot ascertain if there were more microorganisms that were missed, i.e. not detected. A sequence most closely related to a thermophilic *Proteobacterium* which until now had only been found in hot springs, suggests that a geothermal system exists underneath Lake Vostok (Bulat et al., 2004).

Other isolated environments such as the subsurface should also be considered possible targets (such as gas and oil reservoirs; Head et al., 2003; Head et al., 2006) for detection of new life forms on Earth, since these have been isolated for long periods of time but have carbon sources available. Thomas Gold has also claimed that subsurface life may have evolved and be widespread in other planetary bodies even if surfaces are inhospitable (Gold, 1992).

Even though the study suggesting that arsenic substituted for phosphorus in the DNA of a bacterium (Wolfe-Simon et al., 2011) has been refuted (Reaves et al., 2012), the fact that this bacterium is able to grow in such elevated arsenic concentrations (40 mM) is by itself extraordinary. Even on Earth, life has been found in the most unexpected places. Not only extremophile bacteria but also eukaryotes have been found in exotic locations such as the Mariana Trench, the deepest seafloor on Earth (~11 km deep) (e.g. Yayanos et al., 1981; Takai et al., 1999; Takami et al., 1997), or in a South African gold mine (3.2 km deep) (Takai et al., 2001; Borgonie et al., 2011). Recently, scientists from NASA have found a massive bloom of phytoplankton under the Arctic Ocean (north Alaskan ice), a finding claimed to be equivalent to discover a rainforest in a desert (Arrigo et al., 2012). Even on Earth, life is still being found in places once considered inhabitable.

Life should always be complex. One characteristic of life is structural and chemical complexity (Nealson et al., 2002). Therefore, one should also design alternative mechanisms to detect complexity. The search for structural complexity can be performed for example with computer created algorithms that are able to identify complex structures (which later can be characterized chemically) or by searching for chemical disequilibrium in terms of redox reactions in environments (and then look for structural and chemical complexity) or even by the detection of non-random movement which is very characteristic of life forms (Nealson et al., 2002). However, even if these approaches may help in the detection of any life forms, they might still be very limited when applied in low biomass environments.

***Final conclusions***

All the knowledge obtained in this study was directed to the aim of finding optimal ways to assert whether there is life on Mars. The results may however also be applied in the search for life in other planetary bodies (e.g. Europa, Titan or Enceladus) than Mars, and possibly also aid in finding clues on the origin of life on Earth. These results can be used in forthcoming sample return missions.

The possibility that other existing life forms might even not be based on carbon cannot be excluded. We, as humans, will always be limited by what our current knowledge allows us to consider. Thus, there is the risk we may fail to detect life even if we encounter it. However, we can always work with the available tools by improving them or by developing new tools. In addition, knowledge is not stagnant. We are in an age where knowledge is rapidly advancing and with synthetic life being created in the laboratory (Gibson et al., 2010). As this text is being written, the discovery of a particle consistent with Higgs Boson was announced (CERN, 2012). With the certainty that there are still many exciting scientific findings to be made and that the sky is not the limit for humankind, the study presented in

this thesis provides **a step towards the molecular assessment of whether there is life on Mars.**