On the edge of habitability and the extremes of liquidity

Hendrik Hansen-Goos,*, Erik S. Thomson, J.S. Wettlaufer

*Institute of Planetary Research, German Aerospace Center (DLR), Rutherfordstraße 2, 12489 Berlin, Germany
bDepartment of Chemistry and Molecular Biology, Atmospheric Science, University of Gothenburg, 41296 Gothenburg, Sweden
cDepartment of Chemistry, New Haven, CT 06520-8109, USA
dNordic Institute for Theoretical Physics (NORDITA), 10691 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 2 January 2013
Received in revised form 17 April 2013
Accepted 19 April 2013
Available online 27 April 2013

Keywords:
Premelting of ice
Deliquescence
Antifreeze proteins
Extrremophiles
Planetary habitability

ABSTRACT

The physical and biological mechanisms that extend the equilibrium domain of liquid water into the ice region of the bulk phase diagram are examined in view of their importance for the enhancement of planetary habitability. The physical phenomena studied are the premelting of ice, which allows for films of liquid water at temperatures well below freezing, and the wetting of hygroscopic salts with the persistence of briny films even for thermodynamic conditions remote from those of bulk liquid water. Organisms are known to produce a variety of frost-suppressing substances, one of which, the anti-freeze protein, is described here. In this article, we provide a synthesis of theoretical and experimental studies whilst extending ideas into new territory as we address the question of habitability.

1. Introduction

There are intrinsic, liquid-enhancing physical mechanisms typically studied in condensed matter and surface physics that, when brought to bear on the study of ice, are seen as propitious for life. These can either enhance liquidity, and thus potential habitat, or can be harnessed directly by biota in order to thwart off the icy deleterious effects of low temperatures such as confinement in the ice lattice or excess osmotic pressure. In this article, we examine the principal physical mechanisms that heighten the presence of liquid water or liquid brines at low temperatures and low atmospheric water content. These are the (a) premelting of ice; a wide class of phenomena of which the most common is curvature enhanced melting known as the Gibbs–Thomson effect, and (b) pre-deliquescence of hygroscopic salts. All of these mechanisms have in common that they are interfacial phenomena with interactions that begin on the sub-nanometer scale and yet have implications up to and beyond planetary scales (Dash et al., 2006). The effects, and their theoretical treatment and experimental evidence, are introduced in Section 2. In Section 3 we provide examples for the bio-relevance of these effects in terrestrial organisms. The foci of Section 4 are the potential implications of these effects on the habitability of extraterrestrial bodies, and we draw together and summarize the issues in Section 5.

2. Physical mechanisms

2.1. Interfacial premelting of ice

The study of ice and the phenomena underlying melting have a long and rich history. Indeed, the threads can be traced to the origins of the scientific method in the early 17th century; Descartes gave a detailed account of the shapes of snow flakes, linking their adhesive properties to musings on the nature of melting (Frank, 1974). The experiments of Faraday (1850), more than two centuries later, focused on the question of how two blocks of ice put into contact freeze together. Such a process had been interpreted using the familiar concept of pressure melting (Thomson, 1861), which is commonly put forward in order to explain why ice is slippery. Pressure melting arises because water has a higher density than ice as reflected in the negative slope of the solid–liquid boundary in the phase diagram. However, we now know that for the rather low pressures generated in common experience pressure melting can only occur at temperatures very close to 0 °C (for a detailed discussion see Dash et al., 2006). Rather than pressure melting being the primary explanation for low friction in ice skating or related phenomena, the confluence of the presence of a premelted liquid layer on the surface of ice at sub-freezing temperatures and frictional heating generated by a slider are recognized as the essential underlying factors.

How does one describe the mechanisms through which microscopic forces can create and maintain a thin layer of liquid water at the surface of a “sub-freezing” block of ice? In order to provide a
reasonably self-contained paper, we outline the mean-field thermodynamic descriptions along the lines of those in Wettlaufer (1999) and Hansen-Goos and Wettlaufer (2010). For concreteness we consider the interface that consists of two half spaces of ice separated by a planar layer of liquid water with thickness $d$, but the same basic approach applies to the cases where one of the half spaces is made of another material or the vapor phase of water. The number of moles of liquid water (solid water; ice) are denoted as $N_i$ (N$_i$), and we assume that the liquid layer contains $N_i$ moles of electrolytic impurities. Hence, we can express the Gibbs free energy $G$ as

$$G(T, p, N_i, N_j, N_k) = \mu_i(T, p) N_i + \mu_j(T, p) N_j + \mu_k(T, p) N_k + R_T N_i \left( \ln \frac{N_i}{N_j} - 1 \right) + I(d),$$

where $\mu_i$, $\mu_j$, and $\mu_k$ are the chemical potentials conjugate to the quantities just described. $R_T = \frac{\ln N_i/N_j}{N_j}$ is the entrophy of mixing and it applies to the cases where one of the half spaces is made of another material or the vapor phase of water. The second term involves an effective interfacial force depression of (impure) bulk liquid water. The term $I(d)$ embodies the contributions to the free energy arising from microscopic forces acting across and at the interface, which we describe below.

Because the solid and liquid phases of water can freely exchange water molecules, the condition for thermodynamic equilibrium is $\partial G/\partial N_i = 0$. Applying this to Eq. (1) gives

$$\mu_i - \mu_j = R_T N_i \left( \ln \frac{N_i}{N_j} - 1 \right) - \frac{\partial I(d)}{\partial N_i},$$

(2)

where the prime denotes differentiation with respect to $d$ and if, and we neglect the influence of the impurities on the molar density of liquid water $\rho_l$, Eq. (2) exploits the relation $d = N_i/\rho_l$. Finally, we make use of the Maxwell relation $\partial \mu_i/\partial T = -\kappa_i$, where $\kappa_i$ denotes the specific entropy of the respective phases. Linearizing the difference in chemical potentials as a function of $T$ near the melting point of pure ice $T_m$ gives $\mu_i - \mu_j = \kappa_i m_0 T_m/T - \rho_l m_0$, where $m_0$ is the latent heat of fusion per mole of ice. Using this in Eq. (2) yields

$$\frac{T_m - T}{T_m} = \frac{R_T N_i}{\rho_l m_0 d} \frac{\partial I(d)}{\partial N_i}.$$

(3)

The first term on the right-hand side corresponds to the colligative effect of the impurities in the liquid layer which leads to a reduction of the freezing point. This effect is based on the entropy of mixing and it is of the most familiar of mechanisms leading to the freezing point depression of (impure) bulk liquid water. The second term involves an effective interfacial force $I(d)$ originating in long-ranged volume-volume interactions such as polarization forces. These manifest themselves as forces per unit area reckoned upon the surfaces defining the interfacial region, and as such depend on the layer geometry and the materials bounding the interface. In the bulk phases such interactions are absent, and $I'' = 0$, whereas if $I''(d) < 0$ ($I'(d) > 0$) these interfacial forces enhance (inhibit) the formation of a premelted liquid layer relative to the layer thickness that follows from the presence of impurities alone. The effect is evident by inspection of the sign of Eq. (3) in the absence of impurities.

Two different kinds of forces are commonly acting in such systems. First, electrostatic interactions that manifest themselves in terms of an entropic repulsive force (i.e. $I'(d) > 0$) are present due to the screening by the ions in the film of the surface charge density $\varphi_i$, at the interface between the ice and the premelted liquid. The electrolytic premelted layer contains an electrical double layer in the vicinity of the interface. The characteristic length scale over which this screening of the surface charge occurs is the Debye length $\kappa^{-1} = \sqrt{e^2 \kappa_i k_B T / (\varepsilon e^2 \varphi_i)}$, where $\varphi_i = N_i/d$, from linearized Poisson–Boltzmann theory. Within this limit, when the distance $d$ between the two surfaces is comparable to $\kappa^{-1}$ the two double layers interact and the contribution to the interfacial free energy $I(d)$ is

$$F_{\text{elec}}(d) = \frac{2 q_i^2}{k_B T} e^{-\kappa d}.$$

As the force $-F_{\text{elec}}(d)$ is always repulsive it promotes the formation of a premelted liquid layer between two grains of ice. Given that $\kappa$ depends on $d$ through $\varphi_i$, the resulting behavior of the layer thickness as a function of impurity concentration $N_i$ and undercooling $\Delta T$ can be quite complex as discussed in several studies (Wettlaufer, 1999; Benatov and Wettlaufer, 2004; Thomson et al., 2010; Hansen-Goos and Wettlaufer, 2010).

The second contribution to $I(d)$ originates in London dispersion forces or van der Waals forces (Dash et al., 2006; French, 2010). As was first discussed by Eisenschitz and London (1930) in the late 1920s, these forces are a consequence of zero-point quantum mechanical fluctuations, and they can be depicted as an interaction between an instantaneous dipole and an induced dipole. In 1948, Casimir and Polder performed one of the first calculations using quantum field theory and determined the interaction between a perfectly conducting plate and an atom in the retarded limit (Casimir and Polder, 1948). In the same year Casimir showed that the van der Waals forces between two perfectly conducting plates separated by vacuum could be calculated using classical electrodynamics (Casimir, 1948). He discovered that, due to the electromagnetic zero-point energy of the modes of the electromagnetic field, an attractive force acts between the plates. Later, using the diagrammatic methods of quantum field theory, Dzyaloshinskii, Lifshitz, and Pitaevskii (DLP) generalized the theory of van der Waals forces (Dzyaloshinskii et al., 1961). Their result can be used to calculate the dispersion forces in a planar geometry between arbitrary materials 1 and 2 interacting via material 3. Material properties enter the theory via the respective dielectric functions $\varepsilon_1$, $\varepsilon_2$, and $\varepsilon_3$, and here we neglect retardation effects in which case the free energy contribution is

$$F_{\text{disp}}(d) = -\frac{A_h}{12 d^2},$$

(5)

where the Hamaker constant is determined from

$$A_h = \frac{3kT}{2} \sum_{n=0}^{\infty} \frac{\sum (\Delta_{123})^{n} n!}{s^n}.$$

(6)

In this expression $\Delta_{mn} = [\varepsilon_m(l_{nm}) - \varepsilon_m(l_{nm})]/[\varepsilon_m(l_{nm}) + \varepsilon_m(l_{nm})]$ and the dielectric functions must be evaluated at imaginary frequencies, where $\varepsilon_m = 2\pi n kT/h$. The prime indicates that the $n=0$ term carries a weight of $\Delta_{123}^3$ (see Bergström, 1997; French, 2010 and references therein).

The fully retarded DLP theory has been applied to the premelting of ice for the ice/ice/water/vapor (surface melting Elbaum and Schick, 1991), the ice/water/substrate (interfacial melting Wilen et al., 1995), and the ice/water/ice (grain boundary Benatov and Wettlaufer, 2004; Thomson et al., 2010) systems, and it is this latter situation that we focus on here. It is found that the surface melting of ice is incomplete due to the crossover in the relative polarizabilities of ice and water in the ultraviolet. Therefore, as $T - T_m$ from below the liquid layer stops growing at a certain maximum thickness of about 3 nm (Elbaum and Schick, 1991). For the grain boundary (GB) geometry we understand that the long-ranged effects are attractive (Benatov and Wettlaufer, 2004) and hence neglect retardation effects. Thus, in the approach developed to test these ideas described next, we use a value of $A_h = 3.3 \times 10^{-22}$ J which has been calculated in Hansen-Goos and Wettlaufer (2010) using the dielectric functions for ice and water from Elbaum and Schick (1991).
2.1.1. Experimental search for grain boundary premelting

To test theoretical descriptions of GB premelting in ice, laser reflectometry measurements in bicrystals have been performed using the method shown schematically in Fig. 1. A low power 2.3 mW, 632.8 nm He–Ne laser was used with an insignificant heating effect on the sample below the resolution of the temperature measurements. The thickness of the premelted layer at the grain boundary can be extracted from the intensity of the reflected laser beam (Thomson et al., 2009b), and the main features of the data are: the range of experimental salt concentrations $C_i$ of the bulk liquid in contact with our ice grain boundaries is 0–5 psu, where psu denotes practical salinity unit, which is approximately 10 times the concentration in percent by weight. The analytical precision of the $C_i$ measurements is $\pm 0.1$ psu. The range of observed grain boundary thicknesses is 1–8 nm, and we have no direct measurement of the surface charge density $\rho$. At zero concentration of electrolyte, a nonzero GB thickness is observed, with different values for GBs of various misorientations. As the solute (NaCl) concentration is increased, there is a trend of increasing GB thickness. In general there is some more pronounced noise in this trend at lower solute concentrations.

In the experimental system shown in Fig. 1 the grain boundary is connected to a reservoir in which a given impurity concentration (sodium chloride) is maintained at a constant value. Hence unlike in treatments where $N_i$ has been assumed constant (Wettlaufer, 1999; Benatov and Wettlaufer, 2004; Thomson et al., 2010; Hansen-Goos and Wettlaufer, 2010), one can take $C_i = N_i / d$ to be constant. This can simplify the result obtained from Eq. (3) when the explicit expression for $f(d) = F_{\text{diss}}(d) + F_{\text{disp}}(d)$ is used. We find that

$$\frac{\rho \Delta T}{T_m} = \frac{R_g T_m N_i}{d} - \frac{\Delta H}{6 \varepsilon d^2} + \frac{\rho \Omega^2}{\kappa d} (1 - \frac{1}{\kappa d}) e^{-\Delta T},$$

(7)

noting that depending on the limits we consider, we can write the second term as $R_g T_m C_i$ and consider $q_i = q_i(C_i)$.

The two limits of the overall theory that capture the experimental trends are described in Thomson et al. (2013). In one limit $q_i$ is very small and there is very slow equilibration of composition along the GB. Here, the effect is principally colligative and hence $d$ increases with $C_i$. The small $q_i$ limit of Eq. (7) for vanishing dispersion forces is represented by Eq. (29) of Hansen-Goos and Wettlaufer (2010). The theory in this limit applies to the case when the solute along the grain boundary (or grain boundary network in bulk polycrystalline ice) is isolated and the absolute undercooling is specified. The total number of moles, and hence areal concentration, $N_i$ of solute is conserved. This is relevant, and has been applied to glaciers and polar ice where the diffusion of impurities is slow enough that the grain boundaries in the interior may be considered to be isolated. In the second limit $q_i$ plays a larger role and we fit its dependence on $C_i$ in a manner consistent with Eq. (7). Here, $d$ also increases with $C_i$ in the GB, which is equilibrated to the bulk fluid reservoir. The dominating effect is the screened Coulomb repulsion between the two sides of the GB.

Small $q_i$, slow equilibration, colligative limit. Because the data are qualitatively consistent with an assumption that the actual concentration $N_i$ along the grain boundary is proportional to the bulk concentration of the reservoir into which the ice grows, one may write $N_i = C_i \ell_{\text{eff}}$ where $\ell_{\text{eff}}$ is a length scale that relates the two quantities. Physically, we have argued that one can think of this length scale as being related to a “funneling” effect from the grain boundary groove as the ice lens is grown into the bulk solution, trapping the total solute across this effective length into the grain boundary. Hence we are then led to consider the possibility that, on the time scale over which the measurement is performed, the bulk concentration in the GB remains greater than that of the reservoir either (a) due to slow diffusion, possibly influenced by finite size effects, or (b) due to effects that lower the chemical potential along the grain boundary in a way not captured in the current theory. This scenario allows us to express the colligative limit of Eq. (7) as follows:

$$\frac{\rho \Delta T}{T_m} = \frac{\rho \Omega^2}{\kappa d} \left[ \frac{\Delta T}{R_g T_m C_i} + \frac{\Delta T^2}{R_g T_m C_i} \right] = \frac{\rho \Omega^2}{R_g T_m C_i} \frac{\ell_{\text{eff}}}{d}.$$

(8)

where $\Delta T$ is the undercooling with respect to the solutally depressed melting temperature. Some algebra along with the further assumption that there is a minimum thickness $d_o$ to the grain boundary due to short range effects leads to

$$d = d_o + \ell_{\text{eff}} \left[ 1 + \frac{\rho \Omega^2}{R_g T_m C_i} \frac{\Delta T}{d_o} \right],$$

(9)

where $d_o$ will depend on the particular grain boundary mismatch. The straight lines in Fig. 2 are the fits using Eq. (9) for which we find an average $\ell_{\text{eff}}$ of approximately 30 nm and used the average $\Delta T^*$ for each experiment.

Large $q_i$, fast equilibration, electrostatic limit. As described above, an experimental motivation for this limit is that the measured GB always extends to the bulk fluid reservoir outside the ice, and the temperature is measured relative to that at the ice-solution edge. Thus, we consider that the solute in the GB is diffusionally connected to an external reservoir of specified bulk solute concentration. In equilibrium, then, one might expect the volume concentration of solute in the grain boundary to be equal to that of the reservoir and replace $N_i / d$ by $C_i$ in Eq. (7) to yield

$$\frac{\rho \Delta T}{T_m} = \frac{\rho \Omega^2}{\kappa d} \left[ 1 - \frac{1}{\kappa d} \right] e^{-\Delta T} - \frac{\Delta H}{6 \varepsilon d^2}.$$

(10)

As already noted, we have no experimental measurements of the interfacial surface charge and we have no direct measure of the composition within the GB. Thus, as described in Thomson et al. (2013) we take as an ansatz that $q_i$ is a function of the concentration $C_i$ and consider a linear dependence of the grain boundary thickness $d$ on the impurity concentration $C_i$. Hence, for an undercooling of 1.5 K we take $d = 3$ nm for $C_i = 0.5$ psu and $d = 4$ nm for $C_i = 1.6$ psu and linearly extrapolate. Now, neglecting the temperature dependence of $q_i$ over the relatively small experimental temperature ranges, we use...
Eq. (10) to calculate \( q_s(C_i) \), the result of which is shown in Fig. 3(a), and we use this in Eq. (10) to determine \( d \) as a function of \( C_i \) for undercoolings from 0.5 to 2.0 K. We compare these theoretical results with the experimental data in Fig. 3(b), in which we see that beyond \( C_i=1 \) psu the general linear increase in the GB thickness with concentration is captured. The inter-experimental variability at low values of \( C_i \) is again seen here; depending on the undercooling \( d \) can either increase or decrease with \( C_i \).

Clearly these measurements provide essential information for the effective medium properties of ice polycrystals. Indeed, the interfaces between grains provide a ready pathway for the transport of soluble impurities and isotopes, and potential refugia for biota. We see that although premelted films at grain boundaries are 1–10 nm thick, they may dominate the volume of liquid through which transport is controlled due to the fact that they dominate the area of polycrystalline boundaries. Next we turn to the premelting effect responsible for the presence of liquid veins in polycrystalline ice where three grains abut and the nodes where four or more grains terminate.

### 2.2. Gibbs–Thomson effect

In the previous section we explicitly dealt with premelting at a planar boundary between surfaces. This implies that the interfacial area \( A \) between the solid and the liquid does not change as molecules are transferred from one phase to the other. Clearly, such is no longer the case if we consider, for example the interfaces between finite domains in a polycrystal or a spherical particle of ice in a supercooled liquid. Assuming that the interfacial free energy \( \gamma_{sl} \) of the ice–water interface is independent of crystallographic orientation (which has to be revisited in more realistic treatments) we can express the Gibbs free energy \( G \) of this ideal spherical system as

\[
G(T, p, N_i, N_l) = \mu_s(T, p) N_i + \mu_l(T, p) N_l + \gamma_{sl} A(N_i),
\]

where \( N_i (N_l) \) is the number of moles of ice (water), and below we describe how the interfacial area \( A \) depends on \( N_i \). Importantly, we can amend the free energy in Eq. (1) with the term \( \gamma_{sl} A(N_i) \) to capture all of the premelting effects in one description but to isolate the effects of geometry we treat it separately here. The condition
for thermodynamic equilibrium is the same as in the previous section: \( \partial G / \partial N_i = \partial G / \partial N_i \). Applying this to Eq. (11) gives

\[
\mu_i - \mu_i = \frac{\partial A}{\partial N_i}
\]

(12)

where \( \mu_i - \mu_i = q_m^i \Delta T / T_m \) with the same arguments as in Section 2.1. In order to evaluate \( \partial A / \partial N_i \) we assume that molecules are attached or detached from the surface \( \partial B \) of the body \( B \) consisting of the solid phase with equal probability, irrespective of the location on \( \partial B \). This motivates the introduction of the parallel surface of \( B \) defined by all the points that are at distance \( \epsilon \) from \( \partial B \). For any smooth surface \( \partial B \) and as long as \( \epsilon \) is sufficiently small, we can apply Steiner’s formula from integral geometry (Santaló, 1976), which implies that the body defined by the surface parallel to \( \partial B \) has

\[
V(\epsilon) = V + A \epsilon + C \epsilon^2
\]

(13)

\[
A(\epsilon) = A + 2C \epsilon + C \epsilon^2,
\]

(14)

where we have introduced the volume \( V \) and the surface integral of mean curvature \( C \) of body \( B \). The derivative required in Eq. (12) is now obtained through the chain rule as

\[
\frac{\partial A}{\partial N_i} = \frac{\partial A}{\partial \epsilon} \left( \frac{\partial V}{\partial \epsilon} \right)^{-1} \frac{\partial V}{\partial N_i} = \frac{2C}{\rho \epsilon} A
\]

(15)

where we have used that \( \rho_i = N_i / V \). Using this in Eq. (12) combined with the approximation for \( \mu_i - \mu_i \) we obtain

\[
\Delta T = \frac{q_m^i T_m}{\rho_i q_m} \left( \frac{1}{r_1} + \frac{1}{r_2} \right)
\]

(16)

for a surface with constant mean curvature \( \kappa = \frac{1}{2} \left( \frac{1}{r_1} + \frac{1}{r_2} \right) \). This result implies that a convex ice–liquid interface (which has positive mean curvature) experiences a local freezing point depression \( \Delta T \). The magnitude of the effect is \( \Delta T = 1 \, ^\circ C \) for a radius of curvature \( r_1 = 25 \, nm \) (\( 1 / r_2 = 0 \)), with \( \rho_i = 50.1 \times 10^3 \, mol/m^3 \), \( q_m = 6.0 \times 10^3 \, J/mol \), and \( \gamma_d = 2.2 \times 10^{-2} \, J/m^2 \), with \( \gamma_d \) from Raymond and DeVries (1977). This curvature induced freezing point depression is called the Gibbs–Thomson effect. In polycrystalline ice, the Gibbs–Thomson effect is responsible for the formation of a network of liquid veins in locations where grain boundaries intersect (see Fig. 4).

2.3. Pre-deliquescence of hygroscopic salts

The theoretical treatment of the interfacial premelting of ice reviewed in Section 2.1 is valid for temperatures that are relatively close to the bulk melting point. This is reasonable considering that experimental work indicates that the liquid layer associated with premelting of pure ice does not persist at any significant thickness for temperatures much colder than \(-20 \, ^\circ C \) (Bluhm et al., 2002; Dash et al., 2006). However, it is a well established fact that liquid water can occur at temperatures well below this threshold albeit with the aide of impurities. Naturally occurring solutions with high salinity close to their eutectic temperatures \( T_{eut} \) can remain liquid down to \(-80 \, ^\circ C \) (Rempel et al., 2001). For example, Möhlmann and Thomsen have calculated \( T_{eut} \) for some materials that form what they term “cryobrines”, i.e. liquid brines at low temperatures (Möhlmann and Thomsen, 2011). Given the fact that halophilic organisms can cope with NaCl concentrations up to 300 g/l (Oren, 2002), cryobrines provide a potential host reservoir for life at low temperatures. Many extreme environments are both very cold and very dry such as the Antarctic dry valleys on Earth or vast areas on Mars (Lobitz et al., 2001). Fortunately, many salts not only form cryobrines but are also hygroscopic, which refers to the ability of a material to take up water vapor from the atmosphere. This process leads to the phenomenon of deliquescence, which is the liquefaction of a hygroscopic salt due to the extraction of water vapor from the atmosphere. Deliquescence adds a second important characteristic to the eutectic temperature \( T_{eut} \) of a material; the deliquescence relative humidity \( DRH \). The \( DRH \) is the threshold value of the relative humidity \( RH \) that must be exceeded in order for the material to liquefy. Because the \( DRH \) varies with temperature \( T \), in order for a material to deliquesc in under given atmospheric conditions both \( T > T_{eut} \) and \( RH > DRH \) must be satisfied.

Whilst acting alone, deliquescence is a powerful mechanism to extend the domain of liquid water to low temperatures and low relative humidities, in many scenarios, such as on Mars, atmospheric water vapor content is very low and hence the \( DRH \) of a given material may only seldom be reached. In analogy to the interfacial premelting of ice, which produces layers of liquid water at temperatures \( T_{eut} \), we investigate the pre-deliquescence effect, which we expect to provide liquid brine layers on hygroscopic materials at relative humidity \( RH < DRH \). As we show next, the two effects are rather closely related.

In analogy with the derivation of Eq. (3) for the thickness of a layer of premelted ice, we begin with the Gibbs free energy \( G \) of the system which consists of a hygroscopic crystal covered with a brine layer of thickness \( d \) in contact with the vapor phase. We assume that the brine layer is at saturation with a given salt concentration \( C_i = N_l / N_r \), where \( N_r (N_l) \) is the number of moles of salt (number of moles of water) in the brine layer. With this assumption it can easily be shown that the energy of ideal mixing used in Eq. (1) can be written as \( g_{mix}(T)N_r \), a form that holds also

![Fig. 4. (a) Sketch of a liquid vein cross section. (b) A photograph of a polycrystal, reprinted from Mader (1992) with permission of the International Glaciological Society.](image-url)
for non-ideal mixtures as long as the pressure dependence of the free energy of mixing is negligible. Hence, we can eliminate \( N_i \) from the expression for \( G \) to obtain

\[
G(T,p,N_i,N_{gi}) = \mu_i(T,p)N_1 + \mu_n(T,p)N_v + g_{\text{max}}(T)N_j + \mathcal{I}(d),
\]

where \( N_i \) is the number of moles of water in the vapor phase in contact with the brine layer. The layer thickness can be approximated by \( d=\Delta f/\rho_i \), where \( \rho_i = 55.5 \times 10^3 \text{ mol/m}^3 \) is the molar density of liquid water. As in Eq. (1), \( \mathcal{I}(d) \) is the interfacial free energy resulting from the long-range intermolecular forces acting between the hygroscopic salt and the brine layer.

For premelting the equilibrium condition \( \partial G/\partial N_i = \partial G/\partial N_{gi} \) exploits the fact that water molecules can be freely exchanged between the ice and the premelted liquid. In the case of pre-deliquescence, this condition must be modified. Clearly, the exchange of water molecules between the liquid brine and the vapor phase imposes the condition for thermodynamic equilibrium \( \partial G/\partial N_i = \partial G/\partial N_{gi} \). Applying this condition to Eq. (17) at \( RH=\text{DRH} \), where \( d=\Delta f \) and the interfacial forces \( \mathcal{I}(d) \) vanish, gives

\[
\mu_i^\text{f} = \mu_i^\text{f} + g_{\text{max}}. \quad \text{However, for} \quad RH < \text{DRH}, \quad \text{we obtain} \quad \mu_i = \mu_i^\text{f} + g_{\text{max}} + \mathcal{I}(d)/\rho_i. \quad \text{At constant temperature} \quad T, \quad \text{we can combine both equations to give}
\]

\[
\mu_i^\text{f} = \mu_i^\text{f} + \mathcal{I}(d)/\rho_i, \tag{18}
\]

in analogy with Eq. (2). In the treatment of premelting the difference in chemical potential was linearized for temperatures \( T-T_i \) and pressures \( p-p_i \), but the term proportional to pressure is about a million times smaller (e.g., Dash et al., 2006) and hence this term was treated as a constant (although this constraint can be relaxed; see Wettlauer, 2010a). In pre-deliquescence we assume \( T \) to be constant and perform the linearization of the interphase chemical potential for water vapor partial pressure \( p \) close to the pressure \( p_D \) at deliquescence. Using the Maxwell relation \( (\partial \mu_i)/\partial p = \rho_i \rho_\text{w} \), we can write

\[
\mu_i^\text{f} = \mu_i^\text{f} + \mathcal{I}(d)/\rho_i, \tag{19}
\]

Considering the fact that \( p_D > p_i \), we can neglect \( \mu_i^\text{f} \) with respect to \( \mu_i - \mu_i^\text{f} \). Using the ideal gas equation of state for the vapor phase, \( p_D = p_D \rho_D \), and that by definition \( RH = p/p_\text{sat} \) and \( DRH = p_D/p_\text{sat} \), where \( p_\text{sat} \) is the water vapor partial pressure at saturation, we can further simplify the equation above as

\[
\mu_i - \mu_i^\text{f} = R_s T \left( \frac{p_D}{p_\text{sat}} - 1 \right) = R_s \left( \frac{RH}{DRH} - 1 \right). \tag{20}
\]

Combining this with Eq. (18) gives our final result

\[
\frac{DRH - RH}{DRH} = - \frac{\mathcal{I}(d)}{R_s \rho_i}, \tag{21}
\]

which implies that a brine layer on the surface of a hygroscopic crystal can be stable for \( RH < DRH \) provided that the interfacial forces are such that \( \mathcal{I}(d) < 0 \). This is in direct analogy with the stabilisation of a film of pure liquid on an inert substrate against its sub-saturated vapor phase by attractive adsorbate/substrate interactions (Elbaum et al., 1995).

The forces acting across the interface are the same as those we discussed for premelting: dispersion and electrostatic interactions. However, we note again that the magnitude and range of these forces depend on the different materials involved; here the hygroscopic salt crystal and the surface brine layer rather than ice and pure or moderately saline liquid water. When retardation effects are neglected, the contribution \( F_{\text{disp}}(d) \) of dispersion forces to the interfacial free energy \( \mathcal{I}(d) \) is again given by Eq. (5). Thus, the strength of dispersion forces in the pre-deliquescence scenario is determined by the Hamaker constant \( A_h \), the calculation of which we have discussed in Section 2.1. In order to calculate \( A_h \) within the framework of DLP theory we need the dielectric functions \( \kappa \) of the solid hygroscopic salt and that of the liquid brine. Here we discuss only the case of sodium chloride (NaCl), but we note that measured dielectric functions can be found in the literature for a variety of materials. For a compilation of experimental results and calculated Hamaker constants see the work of Bergström (1997). For solid NaCl we use the dielectric function measured by Eldridge and Staal (1977) and for NaCl solutions in water we use the measurements of Buchner et al. (1999). These latter measurements show that increasing the concentration of NaCl in water is akin to rescaling \( \kappa \) of pure water by a factor \( f < 1 \). This fact allows us to assess the range of relevant Hamaker constants used in Eq. (6). For a film of pure liquid water on solid NaCl we find \( A_h = -5 \times 10^{-20} \text{ J} \), which is in agreement with Bergström (1997). As the concentration of NaCl increases, dispersion forces strengthen with a maximum corresponding to a Hamaker constant of about \( A_h = -1.5 \times 10^{-20} \text{ J} \).

Electrostatic forces contribute to the interfacial free energy \( \mathcal{I}(d) \) wherein \( F_{\text{elec}}(d) \) is given in Eq. (4). It is instructive to calculate the Debye length \( \kappa^{-1} \) for a saturated NaCl solution in water. At room temperature, \( \kappa^{-1} \text{ (nm)} = 0.304/\sqrt{\lambda} \), where \( \lambda \) is the ionic strength of the solution in mol/l. For a saturated NaCl brine \( \lambda = 6 \text{ mol/l} \), which gives a Debye length of \( \kappa^{-1} = 0.1 \text{ nm} \). In contrast \( \kappa^{-1} = 1 \mu\text{m} \) in pure water (Israelachvili, 2011). Therefore, the Debye length of a saturated NaCl solution is very short, reflecting the availability of ions in the solution to efficiently screen electrostatic forces. However, measurements show surface brine layers several nanometers thick and hence at least an order of magnitude larger than \( \kappa^{-1} \). Therefore, the assumption that a brine film is saturated may be at odds with electrostatic forces controlling the film thickness. Moreover, because dispersion forces alone are insufficient to generate the experimental film thicknesses, we consider the following two explanations. Either the surface brine layer is undersaturated and hence the Debye length \( \kappa^{-1} \) is large enough for electrostatic forces to play a role, or we can treat Eq. (4) as an empirical contribution to the interfacial free energy, absent an interpretation through electrostatic forces. In the latter case, \( \kappa^{-1} \) and the surface charge density \( q_s \) are merely fitting parameters. Because the physical mechanism leading to undersaturation of the brine layer is unclear, the second approach is common (Kuni et al., 1996; Bruzewicz et al., 2011).

Theoretical results from Eq. (21) for NaCl surface brine layer thicknesses \( d \) as a function of relative humidity \( RH \) are shown in Fig. 5. Deliquescence is assumed to occur at \( DRH = 74.5 \% \) where \( d \) diverges to a macroscopic value. The approach, which combines dispersion forces and electrostatic forces (or the equivalent empirical contribution), can be fit to the measured layer thicknesses (Bruzewicz et al., 2011; Hansen-Goos et al., in preparation). The best fit to measurements yields \( \kappa^{-1} = 2 \text{ nm} \), which would correspond to a liquid layer with salinity less than 0.1 mol/l. As mentioned above, dispersion forces alone, even when ions dissolved in the brine layer are taken into account, are insufficient to describe the measured film thicknesses. The fitted value of the surface charge density is 0.05 cm\(^{-2}\). Note that the surface charge is likely caused by a charge dependent release of ions from the surface of the NaCl crystal; numerical and experimental work has shown that Cl\textsuperscript{-} ions tends to be released preferentially into the brine layer (Cabrera-Sanfelix et al., 2007), leaving the surface with a positive net charge. The surface charge density that we determine is consistent with this interpretation, as it corresponds to about 20% of the Cl\textsuperscript{-} ions of the 100 faces of a NaCl crystal being released into the solution.

3. Biological relevance

Our Earthbound scientific experience biases our belief that liquid water is required for the persistence of all known forms of
life. Our search for extraterrestrial life is tied to the search for water in all of its phases. Within our solar system, stable supplies of liquid water are invariably found in close association with ice, and hence our focus here is on the extent to which water persists in this association. The habitability of the many stable subfreezing microscopic water reservoirs can be addressed in a many ways. First, for biological organisms a distinction must be made between the ability of biota to persist in cryo-habitats versus their ability to thrive, develop, or evolve. In addition to simply considering the availability and geometry of liquid reservoirs, the liquid composition and hydraulic conductivity for nutrient cycling must also be considered. Therefore, cryo-habitats that may appear homologous with regard to physical structure might be significantly different in terms of ecological potential.

3.1. Life within and on ice

In the terrestrial context sea ice represents the most significant laboratory for cryospheric life. Some 13% of the ocean surface is seasonally ice covered and thus comprises an enormous ecosystem exceeding those of the deserts, grasslands or tundra taken alone (see Wettlaufer, 2010b, and references therein). That dense populations of bacteria, algae and metazoans colonize sea ice provides compelling evidence that elemental and original forms of life have evolved to thrive at low temperatures and high salinities. The intrinsic mushy layer structure of sea ice, with nearly pure solid ice bathed in brine (Feltham et al., 2006), provides an ideal unfrozen landscape for a range of organisms (Meiners et al., 2002). Moreover, the dynamic-thermodynamic environment of sea ice is propitious for life because it controls transport properties such as the flow permeability (e.g., Wells et al., 2011; Neufeld and Wettlaufer, 2011; Wettlaufer, 2010b). The density of life in oceanic ice is greater than that in the water below, and hence the ecology of sea ice has such a fundamental influence on the food web we dare not attempt to summarize it all here, but refer the interested reader to more thorough reviews (Mock and Thomas, 2005; Wettlaufer, 2010b).

Polar ice sheets differ from lower latitude glacial ice in both spatial scale and temperature; very cold at the surface with a base typically at or near the pressure melting point. Because the amount of liquid is controlled by geometry, composition, and temperature (Section 2.1) the cold polar temperatures place an important constraint on the scale of microscopic liquid inclusions in ice-sheet ice. However, adaptation is evident for bacteria found in ice core samples are smaller than average bacteria (≤0.5 μm compared to ~1 μm) (Srikanta Dani et al., 2012). Premelted veins in ice sheets may reach ~10 μm in the warmest areas but are at least an order of magnitude smaller near the surface, and different microscopic techniques have been used to observed cellular fauna within them (e.g., Fig. 1, Ref. Srikanta Dani et al., 2012). Extrapolating from bulk impurity measurements of ice cores, typical ionic molarities in veins are calculated to be in the range of enriched laboratory media (0.3–3 M, Mader et al., 2006), and are thus rich reservoirs for nutrients. This is primarily due to the extremely small segregation coefficient of ice (~10⁻⁶), ensuring that soluble impurities remain in the liquid. Thus concentrated solutions in veins can remain liquid down to their eutectics, some of which can reach ~−80 °C (Rempel et al., 2001). Concentrated solutions can sustain metabolic activity even in the absence of oxygenation. For example, Wadhams et al. (2004) found stable microbial communities in anaerobic conditions in Arctic glacial beds. Importantly, the overall nutrient cycling necessary for microbes to prosper in ice may be reduced by the metabolic slowing that occurs at low temperatures. Thus, although the absolute density of micro-organisms in ice may be low, because global ice volume is large (~10¹³ km³) it represents a significant repository of extremophile life.

Temperate glaciers are much like polar ice but because they are at their melting points year round they might afford a higher potential for hosting microbial life. Hence, premelted liquid networks are much larger and more developed in temperate ice, allowing for more abundant colonies of psychrophiles. Because of the shorter residence time of temperate ice, and the added possibility of wash-out due to the presence of more mobile liquid, life in temperate ice might be rapidly altered (Lliboutry, 1996). However, because 99% of global glacier volume is contained in the polar ice sheets, the detailed effects of temperate glaciers are unlikely to be globally significant. Nevertheless, when ice retreats (advances) locally preserved microbial communities from ice may be important for recolonizing glaciated terrain (glaciers).

Frozen water saturated soils are riddled with life. Microorganisms in permafrost soil have been shown to maintain a slow metabolism at temperatures as low as ~−40 °C (Price and Sowers, 2004), which has been attributed to thin films of premelted liquid water (Rivkina et al., 2000). The liquid water budget in frozen soils has since been investigated using detailed thermodynamic arguments (Hansen-Goos and Wettlaufer, 2010), but the rigorous coupling between basic physical effects and their important biological implications is a relatively untouched discipline.

3.2. Antifreeze proteins and the role of the Gibbs–Thomson effect

The remarkable ability of organisms to survive under extreme conditions such as the low temperatures in Siberian permafrost (Rivkina et al., 2000), the low humidity in the Atacama desert (Parro et al., 2011), or a combination of these, as found in the Antarctic dry valleys (Cary et al., 2010), is suggestive of extremophile organisms developing strategies to act back on their harsh environments in order to improve living conditions. An important example of this is exopolymERIC substance (EPS) (Wettlaufer, 2010b), which has been
detected in sea ice in the polar regions. In recent campaigns in Antarctica, EPS concentration and composition in brines, ice cores, and interstices in sea ice have been determined and the measured fraction of insoluble EPS has been speculated to cause the surroundings of cells to be hydrophobic (Underwood et al., 2010). Recent laboratory experiments with ice grown from saline solutions have revealed that, in the presence of EPS, bulk ice salinity increases and pore morphologies have a higher fractal dimension as compared to EPS-free samples (Krembs et al., 2011). These findings have been interpreted in the sense that the expression of EPS improves habitability for micro-organisms in sea ice (Krembs et al., 2011). In this work we focus on a different case of freezing protection, which is provided by the antifreeze protein (AFP), demonstrating a strategy that a variety of organisms have devised in response to sub-zero temperatures. These organisms range from Arctic fishes, where moderate protection of a few °C is sufficient, to plants, micro-organisms, and insects, which can require freezing protection of 5 °C or more. AFs have been isolated from various organisms and they have been shown to greatly vary in structure, leading to the definition of multiple types of AFSs, including the so-called hyperactive active AFP that is found in insects (for a recent review see Venketesh and Dayananda, 2008). Several authors have argued that the fundamental differences encountered in the structures of different AFSs support convergent evolution (Davies et al., 2002).

A predominantly colligative mechanism underlying AFS function can be ruled out based on the low concentrations that are required for AFS induced freezing point depression. For example, the hyperactive spruce budworm AFP experiments of Leinala et al. (2002) show that a concentration as low as 100 μmol/l is sufficient to reduce the freezing point by 5.0 °C. This should be compared to a concentration of about 2.5 mol/l that would be required in order to obtain the same freezing point depression based on the colligative effect, i.e., based on mixing entropy of water with the impurities. Compared to NaCl, this AFS is more efficient by a factor of about 10^5 in terms of molar concentration and by a factor 50 in terms of mass concentration.

How can the AFS manifest this impressively strong effect? The underlying mechanism was identified in early work to be a surface effect as opposed to the colligative effect which acts in the bulk (Raymond and DeVries, 1977). A freezing point depression is realized by AFS binding to the ice-water interface as soon as a crystal starts to form in the undercooled liquid. This prevents further growth of the seed crystal. In organisms that protect themselves from frost damage this effect can occur inside the cells and, for higher level organisms, directly in the blood vessels. In the laboratory, isolated AFS is commonly studied in a solution of pure H$_2$O, allowing the control of well defined conditions in terms of undercooling and impurity concentration. Direct visualization of AFS concentration can be achieved by conjugating AFS with a fluorescent protein, a technique that has been employed by Zepeda et al. (2008) and Bar-Dolev et al. (2012). This approach provides insight on how AFSs interact with the ice-water interface, including an observed crystallographic specificity of AFS binding (see Fig. 6). It is important to note that experiments find a thermal hysteresis. Namely, once the AFS solution freezes entirely, thereby incorporating the AFSs, no AFS action other than an insignificant colligative effect is observed upon remelting the sample.

While it has been agreed for decades that the underlying AFS mechanism is chiefly a surface effect, the exact process of growth inhibition is difficult to access experimentally due to the molecular length scales involved. Although recent computer simulations have provided some new insights (Nada and Furukawa, 2012), the general picture of AFS function has changed little since the first explanation was given by Raymond and DeVries (1977). They argued that the ice crystal growth front is pinned by AFSs binding to it, which results in a freezing point depression due to the Gibbs-Thomson effect. Next we will review this so-called “adsorption inhibition mechanism” of AFS action and subsequent refinements. Following this we present some new ideas that result from assuming the AFSs to be randomly distributed at the ice-water interface.

Raymond and DeVries (1977) considered a common growth model of crystals which assumes accretion through the propagation of steps across a crystal face. Let $r$ be the distance between two AFSs that bind to the ice-water interface. As a step propagates on the surface, it is arrested at the loci of the AFSs and is thereby forced to develop a circular shape with a radius $r$. The minimum radius imposed by the AFSs is dictated by the distance $r$ as $r_c = r/2$. According to the Gibbs-Thomson effect (Eq. (16)) the associated undercooling is

$$\Delta T = \frac{\gamma_{AFP} T_m}{\rho c \Delta m r_c},$$

where we have used that one of the principal curvatures equals the curvature of the circle $1/r_c$, and the second principal curvature vanishes in the step-geometry.

While Eq. (22) illustrates a freezing point depression based on curvature induced pinning of the growth front associated with one pair of AFSs, it is not obvious how the undercooling of a large crystal face decorated with many AFSs can be inferred from it. Two difficulties must be overcome. First, the number $N$ of AFSs binding to a crystal face with surface area $A$ has to be determined based on the concentration $c$ of AFS in the solution. It is convenient to introduce the number density $n = N/A$ of AFSs on the given surface. Raymond and DeVries obtained $n$ by assuming that only those molecules that are eventually buried in the ice participate in freezing inhibition. Hence, they can calculated $n = n_{max}$, where $n$ is the partition coefficient of the given AFS in ice and $\sigma$ is a length-scale representative of the size of the AFS. An obvious weakness of this approach lies in the assumption that the concentration of AFS incorporated into the ice during crystal growth (which occurs at high velocities; $-1$ to $30 \mu m/min$ Bar-Dolev et al., 2012) is related to the AFS surface density of during growth inhibition (a stationary state). Alternatively, $n$ can be calculated from different adsorption models, such as the Langmuir adsorption isotherm. The latter would imply that $n = n_{max}(Kc/1 + Kc)$, where $K$ is the Langmuir adsorption coefficient and $n_{max}$ is the maximal surface coverage that can be obtained in the limit of large AFS concentration in the solution (Wang et al., 2012). However, in this approach $K$ and $n_{max}$ must be determined from fitting experimental data for $\Delta T$ and hence this theory lacks a predictive capability. Therefore, we do not pursue this approach here and instead use the route taken by Raymond and DeVries.

The second difficulty is calculating the distance $r$ between two AFSs from $n$, and hence $r_c = r/2$ for use in Eq. (22). The common approach is to assume the AFSs to be distributed on a square lattice so that $l = 1/\sqrt{n}$. However, there is no obvious reason to assume the square lattice distribution, and it is more natural to assume that the AFSs are randomly distributed on a given crystal face. The probability $P(l) \phi$ for the nearest neighbor of a given AFS to be located at a distance between $\rho$ and $\rho + \phi$ was shown by Hertz (1909) to be

$$P(l) \phi = (N-1) \cdot \left( 1 - \frac{2\pi l^2}{A} \right)^{N-2} \cdot \frac{2\pi l^2 \phi}{A}.$$

The first factor takes into account that there are $N-1$ different realizations of a nearest neighbor, the second factor is the probability that an AFS sits at a distance greater than $\rho$, and the last factor is the probability that an AFS rests at a distance between...
leads to corresponding undercooling $\Delta$ with probability $P$, for two nearest neighbor $\text{AFP}$s to sit at a distance larger than $\ell$ is given by

$$P = \int_0^\ell P(\rho) \, d\rho = 1 - e^{-\ell/\alpha^2}.$$  

(25)

It is reasonable to assume that the freezing inhibition at a given undercooling $\Delta T$ is most reliably guaranteed if none of the $\text{AFP}$s on a given crystal face have a nearest neighbor that is further away than the distance $\ell$ specified by Eq. (22), with $r_c = \ell/2$. Because there are $N/2$ pairs of nearest neighbors, the corresponding probability $P$ is

$$P = P^{N/2} = [1 - e^{-\ell/\alpha^2}]^N/2.$$  

(26)

Obviously, we will always have $P < 1$. Hence, no matter how large $\ell$, statistical fluctuations will always insure a non-zero probability for two nearest neighbor $\text{AFP}$s to sit at a distance larger than $\ell$. The corresponding undercooling $\Delta T$ can therefore only be achieved with probability $P$. Now, since a certain probability $P$ is required, we can gain insight into how $\ell$ depends on $n$ and $A$ by using the ansatz $\ell = a\sqrt{n/2\pi}$ in Eq. (26). Assuming that $N$ is large this leads to $P = e^{-\alpha^2}$. One can debate the value for $\alpha$ that guarantees an undercooling $\Delta T$ with an associated distance $\ell$. We argue that $P$ should be significantly larger than 0 but not very close to 1 based on the following logic. Not every instance of a nearest neighbor distance fluctuation to a value larger than $\ell$ necessarily leads to runaway crystal growth. This is because a fluctuation at any given time might occur on a growth front that is already pinned in a previous growth stage by other $\text{AFP}$s as such second nearest neighbors or those just joining the surface from the melt. For simplicity, we choose $\epsilon = 1/2$, which gives $P=0.6$. Using our ansatz for $\ell$ with this value of $\epsilon$ and substituting the definition for $a$ we obtain the final result of our stochastic model as

$$\ell_{\text{rand}} = \sqrt{\frac{\ln nA}{\pi P}}$$  

(27)

which should be compared to the result for a square lattice $\ell_{\text{square}} = 1/\sqrt{n}$. As expected, unless $N=nA$ is very small, we find $\ell_{\text{rand}} > \ell_{\text{square}}$. This illustrates the fact that stochastic fluctuations insure large nearest neighbor distances are likely to occur on sufficiently large crystal faces, unlike in the case of a square lattice were nearest neighbor distances are fixed. It follows from Eq. (22) that the resulting thermal hysteresis $\Delta T$ is smaller for the random distribution than for the distribution on a square lattice.

In Fig. 7 we compare the thermal hysteresis $\Delta T$ from the two models with the measurements of Raymond and DeVries (1977). Note that in the cases where the theory using $\ell_{\text{square}}$ overestimates $\Delta T$ (Fig. 7C and D) the reduction of $\Delta T$ arising from the random distribution of $\text{AFP}$s, which is accounted for in $\ell_{\text{rand}}$, provides an excellent match to experiment. The best agreement is obtained for crystal surface areas $A = 35 \mu m^2$ and $A = 100 \mu m^2$, respectively. In Fig. 7A and B, which are already well described by Raymond–DeVries theory, the corresponding surface areas are much lower ($A = 0.07 \mu m^2$ and $A = 0.05 \mu m^2$). One speculative interpretation is that these $\text{AFP}$s inhibit crystal growth at an earlier stage of crystal formation. In Fig. 7A and B we do not show the result for $\Delta T$ from the present theory at low $\text{AFP}$ concentration $c$. This is because low crystal surface area $A$ low concentration $c$ insures that $N$ is too small for the arguments leading to $\ell_{\text{rand}}$ to apply. This low $c$ range requires an alternative derivation of $\ell_{\text{rand}}$.

Obviously, the approach of randomly distributed $\text{AFP}$s is one possible theory to explain why Raymond–DeVries theory fails in certain cases. As we have noted above, the computation of $\text{AFP}$ number density $n$ on a given crystal phase from the $\text{AFP}$ concentration $c$ in the solution has uncertainty. Moreover, it is known that a number of $\text{AFP}$s tend to bind preferentially to faces with certain crystallographic orientation (Zepeda et al., 2008; Bar-Dolev et al., 2012). In light of the complexity of $\text{AFP}$ binding, which molecular simulations have just begun to shed light on, it is clear that we cannot expect theories as general as the ones presented here to give a full account of $\text{AFP}$ function for any given type of $\text{AFP}$. Such methods of statistical mechanics will have to be informed by further experimental and microscopic insights.

### 3.3. Life in dry environments (pre-deliquescence)

As noted in Section 3.2, an abundance of life has been found in hypersaline, dry climate ecosystems on Earth, such as the Atacama and Namibia deserts, and Antarctica’s dry valleys. In arid regions microbiotic crusts, which are consolidated matrices of flora and fauna, are central to ecosystem dynamics (Evans and Johansen, 1999). The soils can be highly saline due to continuous desiccation, but again clear indications of life exist (Parro et al., 2011). In the past, studies of such habitats have generally described the biogeochemical cycling and botanical distributions specific to the particular stresses of such environments. For example, it is observed that nitrogen fixing by organisms in microbiotic crusts dominate ecosystem nitrogen input (Evans and Johansen, 1999). However, the focus has primarily remained on biological processes, and less attention has been paid to controlling physical parameters, such as comprehensively assessing liquid availability.
Typically the diversity of life in these environments is assessed by exposing field samples to bulk liquid water. Microbes respond rapidly to wetting and quickly exploit the surrounding resources, whereas in the absence of culturing much of the biology is encysted or otherwise hidden from observation. This technique is considered a reasonable simulation of the scenario that such biomes are entirely dependent on sporadic rain events in arid regions (Bamforth, 2004), but clearly may not accurately reflect their full climatology. It is also observed that microbial response is tied to liquid availability, and that different micro-ecosystem states exist along a wetting or wetness continuum (Bamforth, 2004). Different biota thrive at different wetnesses and thus the success of any given microbe does not automatically increase with increasing liquid. This suggests that in the absence of rainwater, deliquescence events at low relative humidities might supply suitable liquid for microbiota in such environments. This idea is further supported by direct observations of microbial cells bound to salt and mineral crystals in oven dried samples of dessert soils (Parro et al., 2011). These soils have an abundance of hygroscopic salts, such as halite, and in environmentally controlled experiments at low temperature (4°C) and relative humidity (75%) they were observed to continuously accumulate water via deliquescence.

Observations clearly show that on Earth biology can thrive at the margins. In fact, the very conditions that force extremophile adaptations provide niche-ecosystems for biological viability. While similar physical habitats may exist extra-terrestrially, it remains to be seen whether the biological complexity necessary to take advantage of such conditions has arisen. At the very least the physical mechanisms we have described can serve as a guide in the search for life away from Earth.

4. Implications for planetary habitability

In Section 3 we discussed the biological implications of the interfacial phenomena introduced in Section 2. Naturally, these examples pertain to carbon-based organisms that are present on Earth; the only planet we know to harbor such life. How do the physical mechanisms responsible for interfacial liquid water and liquid veins influence our assessment of the habitability of planetary bodies other than Earth? Let us first consider the interfacial premelting of ice. Is the effect significant for a place such as Mars with a cold and dry surface? What is the effect of the low atmospheric pressure on Mars, which is below the pressure at the triple point of water in most locations? Theoretical and experimental studies have been carried out for standard pressure conditions and hence we cannot yet provide an unequivocal answer. Clearly the surface conditions in Mars are deeply in the sublimation regime and hence we expect no translation of laboratory conditions to date into planetary conditions as we understand them. However, we cannot exclude that for certain interfacial forces, say between ice and a rocky material, a liquid layer several nanometers thick could be stable, even for pressures...
significantly below that of the triple point of water, and at low temperatures. High pressure conditions allow for momentary existence of liquid water (Wettlaufer, 2010a), and these may have been important in delivering substantial volatiles to the inner solar system and may underlie the recent discovery of polar ice on Mercury (Lawrence, 2012).

Moving to the subsurface of Mars or Mercury is more encouraging. Not only does the regolith provide a shield against harmful UV radiation, but pressure and temperature conditions that are more conducive for the premelting of ice. Due to the lithostatic pressure and the geothermal heat flow, conditions for liquid water can be reached in the Martian subsurface (Mellon and Phillips, 2001). For a regolith with high thermal conductivity, the melting of ice occurs a few kilometers below the surface while melting conditions are reached between 100 and 200 m in areas where thermal conductivity of the subsurface is low (Mellon and Phillips, 2001). We expect premelting to occur at the upper boundary of the melting zone, potentially extending the availability of liquid water to shallower depth and possibly mitigating the effect of obliquity variations on the melting zone. In contrast to permanent liquid water, temporary liquid water on the surface of Mars occurs only at low altitudes where atmospheric pressures exceed 6.1 mbar (Lobitz et al., 2001). Low partial pressures of water vapor in the atmosphere, and seasonal and diurnal sub-freezing temperatures cycles, prevent liquid water from being permanently stable on the surface. However, Lobitz et al. (2001) found areas where both temperatures and pressures are at or near the liquid domain of bulk water and hence premelting effects are expected to significantly enhance liquidity, both geographically and temporally. Jakosky et al. (2003) have used premelting effects at ice grain–dust grain and ice grain–ice grain contacts in order to argue for the habitability of water ice at the surface of the Martian polar cap, which they corroborate with experimental work on organisms isolated from Siberian permafrost. However, while premelting is a suitable explanation for activity of organisms in permafrost on Earth (see also Section 3.1), the mechanism has not been studied quantitatively in the sublimation regime. Because the polar regions of Mars’ surface are not conducive to temporarily liquid water in bulk (Lobitz et al., 2001), premelting is likely a more viable mechanism for subsurface water.

Between the permanent liquid water that might exist in the Martian subsurface and the temporary liquid water on the surface of the planet, other mechanisms may involve the presence of liquid water close to the surface, but for geologically short amounts of time. Numerical modeling shows that impacts can generate hydrothermal systems, and even crater lakes, which covered by an ice sheet, can last for more than 10,000 years (Rathbun and Squyres, 2002). This provides a significant volume of porous rock with liquid water at temperatures suitable for microbial life (Abramov and Kring, 2005). The thermal conductivity of the subsurface material affects the depth of the melting zone, but it also determines the duration of impact induced hydrothermal systems and the corresponding discharge rates (Barnhart et al., 2010). Magmatic intrusions can melt ground ice and trigger hydrothermal systems that can have lifetimes of up to 1 million years over which they generate considerable discharge of liquid water (Gulick, 1998). While we do not expect premelting to have a major impact on discharge rates, we assume that it will increase the volume of rock that becomes habitable subsequent to an impact or magmatic intrusion event, mainly caused by gains of liquid water at the boundary of the induced melting zone.

What are the implications for habitability that we can expect as a consequence of the Gibbs–Thomson effect? This effect gives rise to pockets of undercooled liquid water at contact points of rocky grains in a frozen water–saturated soil (Hansen-Goos and Wettlaufer, 2010). As far as Mars or Mercury are concerned, these pockets are likely to form at the boundary between a melting zone at depth and a frozen soil closer to the surface. A more significant consequence of curvature-induced melting results from the formation of liquid vein networks in polycrystalline ice (see Fig. 4). In particular for bodies with a large fraction of polycrystalline ice, such as the icy moons of Saturn and Jupiter, extended vein networks could be crucial for the functioning of life, ensuring transport of nutrients and waste products and communication between microbial colonies inhabiting pockets of liquid water or liquid brines. Indeed, if such ices have grown from an electrolytic ocean, the mechanisms of mushy layer formation will insure ready conduits and transport paths (Wettlaufer, 2010b). The potential of icy moons to host life underneath their surface, possibly at the interface between the ice cover and an interior ocean has been recognized in the last couple of years (Raulin et al., 2010), but space missions probing in detail the life potential of moons like Europa, Enceladus, or Titan are still in the future.

In Section 3.2 we examined how the Gibbs–Thomson effect can be used to explain the action of antifreeze protein (AFP). In the context of habitability, this example provides an excellent class of strategies that life on another planet could devise to cope with extreme environmental conditions. This is important because it demonstrates that habitability cannot be strictly defined without knowing the life form that is expected to inhabit the planetary body under scrutiny. To go even further, we must acknowledge that the definition of habitability can never be static, because life itself influences environmental conditions, be it locally by reducing the freezing of water via the expression of AFPs, or globally through the creation of an atmosphere caused by microbial metabolism. This complexity indicates that a certain take on exobiology can be problematic. Once we start speculating about life other than those carbon-based forms known from Earth, the line between science and science fiction is crossed all to easily (Simpson, 1964). As long as we are unable to rebuild existing primitive organisms in the laboratory, putative exotic forms of life belong to a realm of entertainment. Pre-deliquescence, the third general physical mechanism that we discussed, is caused by interfacial forces that stabilize liquid brines under temperature and (liquid vapor) pressure conditions where they are unstable in bulk. Since the late 1970s it has been known that on Mars brines are likely to be occasionally stable in bulk (Brass, 1980). In this respect pre-deliquescence has a favorable extraterrestrial reference point for exploration and testing in the subsurface. More recently, these early findings have been corroborated by experimental work under simulated Martian conditions pertaining to evaporation (Altheide et al., 2009; Hanley et al., 2012) and deliquescence (Zorzano et al., 2009) of sulfate, chloride, and chlorate brines. In recent theoretical work, stability of liquid brines on Mars has been determined as a function of geographic location, season, and time of day based on a global climate model (Möhlmann, 2011).

The importance of brines as probably the only providers of a liquid medium for life on or close to the surface of Mars has been addressed in numerous studies. These include discussions of halophilic organisms (Litchfield, 1998), water activity in saline solutions (Davila et al., 2010), and bacterial growth at high concentrations of magnesium sulfate, typical for Martian soils (Crisler et al., 2012). Other studies that expose micro-organisms to simulated Martian conditions do not focus on the role of brines. However, it is to be expected that both samples of Arctic permafrost (Hansen et al., 2009) and rocks with alpine lichen (de Vera et al., 2010) are subject to some brine formation in the interstices of the soil or in cracks of the rock. It is challenging to directly measure the amount of liquid brine in a biological sample mounted in a Mars simulation chamber. It is even more difficult to measure the amount of liquid originating from pre-deliquescence corresponding to salts present in the sample that do not allow for a stable bulk brine.
Obviously, pre-deliquescence of hygroscopic salts increases the amount of liquid medium in a sample and it broadens the stability range of brines on Mars. In order to quantify the magnitude of the effect in theory, assumptions have to be made about the relevant interfacial forces for specific materials as discussed in Section 2.3. Measurements of brine availability due to pre-deliquescence in biological samples could proceed indirectly by correlating biological activity with the amount of liquid brine available, in analogy to the work of Rivkina et al. (2000) in the context of premelting in ice.

On Earth it is difficult to separate relative roles of premelting and pre-deliquescence under extraterrestrial conditions. Places that are warm and dry, such as the Atacama desert in Northern Chile (Parro et al., 2011), allow us to rule out any contribution to and pre-deliquescence under extraterrestrial conditions. Places work of Rivkina et al. (2000) in the context of premelting in ice. activity with the amount of liquid brine available, in analogy to the work of Rivkina et al. (2000) could provide a liquid medium feeding putative Venus-adapted microorganisms, which have been speculated to exist (see Wickramasinghe and Wickramasinghe, 2008 and references therein).

5. Conclusion

We have examined three main physical mechanisms that act to extend the equilibrium domain of liquid water into the solid or vapor phase of the bulk phase diagram. These effects are the interfacial premelting of ice due to long-ranged intermolecular forces, the curvature induced freezing point depression known as the Gibbs–Thomson effect, and the pre-deliquescence of hygroscopic crystals, which is the wetting of a solid with a liquid brine. These phenomena all describe how surface and size effects influence phase transitions and while they have quantitative manifestations that depend on the material under consideration, they are common to all materials. In Section 2.1 we reviewed the theory of interfacial premelting and interpreted experimental results on the grain boundary melting of ice. The relevance of premelting for life on Earth was discussed in Section 3.1 and examples throughout the cryosphere were drawn upon. Because most naturally occurring ice is polycrystalline, the depression of the freezing point due to the convexity of the solid/liquid interface, the Gibbs–Thomson effect, has a special interest (Section 2.2). This effect has implications for life due to the formation of a liquid vein network in polycrystalline ice, but we dwelled on its role in the functioning of the antifreeze protein (AFP), which is part of an active strategy employed by some organisms to fight the cold by attacking ice crystals. We reviewed AFP function and described an improved theory to explain the corresponding freezing point depression (Section 3.2). Finally, in Section 2.3 we introduced a theoretical description of what we call pre-deliquescence. This effect is important for life under extreme conditions in two respects. Owing to the low eutectic temperature of a number of salts, certain brines stay liquid down to temperatures as low as −80 °C. In addition, pre-deliquescence phenomena give rise to thin layers of liquid brines at relative humidities well below the deliquescence point. The possible impact of pre-deliquescence on microbial communities under dry atmospheric conditions was discussed in Section 3.3.

In Section 4 we presented some ideas on the potential of these effects to influence our notions of habitability. The basic theme is the possibility of shifting the boundaries of habitable zones into domains where liquid water is unstable in bulk, but thin interfacial liquid layers or liquid veins might persist. Because the basic effects we have detailed can operate across a wide range of conceivable habitats on other planets or exoplanets, they appear to be useful tools in scouting out territory for searches. They are operative in the subsurface, be it rocky or icy, on suspended particles in a moist or dry atmosphere, or in bodies exhibiting temporary surface liquid water. The key point is that, when defining habitability, we ought to consider the value of the small. A persistent liquid film of 10 nm is not “too thin” for a microbe measuring 1 μm. Indeed, a thirsty 2 m human might view a 2 cm deep pond as an oasis. Thus, the study of habitability might well be the study of potential oases for microbial life.

Acknowledgments

The authors acknowledge the U.S. National Science Foundation Grant No. OPP0440841, the Department of Energy Grant No. DE-FG02-05ER15741, the Helmholtz Gemeinschaft Alliance “Planetary Evolution and Life,” and Yale University for generous support of this research. EST also thanks the University of Gothenburg and the Nordic Top-Level Research Initiative CRAICC for support. JSW thanks the Wenner-Gren Foundation, the John Simon Guggenheim Foundation, and the Swedish Research Council for generous support of his research.

References

Crisler, J.D., Newville, T.M., Chen, F., Clark, B.C., Schneegurt, M.A., 2012. Astrobiology 12, 98.