Title

Nanometer-Scale Chemistry of a Calcite Biomineralization Template: Implications for Skeletal Composition and Nucleation

Journal

Proceedings of the National Academy of Sciences of the United States of America, 113(46)

ISSN

0027-8424

Authors

Branson, Oscar
Bonnin, Elisa A
Perea, Daniel E
et al.

Publication Date

2016-11-15

DOI

10.1073/pnas.1522864113

Peer reviewed

Nanometer-Scale Chemistry of a Calcite Biomineralization Template: Implications for Skeletal Composition and Nucleation

Oscar Branson[a,1,2,3], Elisa A. Bonnin[b], Daniel E. Perea[c], Howard J. Spero[a,3], Zihua Zhu[c], Maria Winters[c], Bärbel Hönisch[d], Ann D. Russell[a], Jennifer S. Fehrenbacher[a,e], and Alexander C. Gagnon[b,3]

[a]Department of Earth and Planetary Sciences, University of California, Davis, CA 95616; [b]School of Oceanography, University of Washington, Seattle, WA 98195; [c]Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, WA 99352; [d]Department of Earth and Environmental Sciences, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964; and [e]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331

Edited by Patricia M. Dove, Virginia Tech, Blacksburg, VA, and approved September 23, 2016 (received for review November 19, 2015)

Plankton, corals, and other organisms produce calcium carbonate skeletons that are integral to their survival, form a key component of the global carbon cycle, and record an archive of past oceanographic conditions in their geochemistry. A key aspect of the formation of these biominerals is the interaction between organic templating structures and mineral precipitation processes. Laboratory-based studies have shown that these atomic-scale processes can profoundly influence the architecture and composition of minerals, but their importance in calcifying organisms is poorly understood because it is difficult to measure the chemistry of in vivo biomineral interfaces at spatially relevant scales. Understanding the role of templates in biomineral nucleation, and their importance in skeletal geochemistry requires an integrated, multiscale approach, which can place atom-scale observations of organic-mineral interfaces within a broader structural and geochemical context. Here we map the chemistry of an embedded organic template structure within a carbonate skeleton of the foraminifera *Orbulina universa* using both atom probe tomography (APT), a 3D chemical imaging technique with Ångström-level spatial resolution, and time-of-flight secondary ionization mass spectrometry (ToF-SIMS), a 2D chemical imaging technique with submicron resolution. We quantitatively link these observations, revealing that the organic template in *O. universa* is uniquely enriched in both Na and Mg, and contributes to intraskeletal chemical heterogeneity. Our APT analyses reveal the cation composition of the organic surface, offering evidence to suggest that cations other than Ca^{2+}, previously considered passive spectator ions in biomineral templating, may be important in defining the energetics of carbonate nucleation on organic templates.

biomineralization | templating | foraminifera | geochemistry | paleoceanography

The formation of calcium carbonate biomineral skeletons is integral to the survival of diverse groups of marine organisms (1) and is a major component of the global carbon cycle (2). Ancient $CaCO_3$ skeletons preserve a chemical archive of the oceanographic conditions they were formed in, and provide one of our most comprehensive geochemical records of past climate (3). Despite the role that biomineralization plays in evolution, global chemical cycling, and our understanding of past climate, we lack a complete mechanistic understanding of biomineral growth that can predict how it responds to environmental change, and affects shell geochemistry.

Calcification is particularly complex because it is affected by both environmental and biological factors. During biomineralization, organisms control the flow of ions, manipulate aqueous speciation, and use organic–mineral interactions to modulate the morphology, composition, and growth rate of skeletal minerals (4). Of these processes, the role of organic–mineral interactions is the least understood, and has the most potential to revise our understanding of the balance between environmental and biological control during skeletal growth. The formation of many carbonate biominerals begins around organic templating structures (5), where atomic-scale interactions between organic surfaces and mineral growth processes define the gross architecture of the biomineral (4). At a mechanistic level, these organic–mineral interactions facilitate heterogeneous nucleation processes, and allow an organism to overcome chemical and physical barriers to crystal nucleation and growth (4, 5). Organic–mineral interactions may also influence the composition of biominerals (6, 7). Many $CaCO_3$ biominerals exhibit complex, systematic intraskeletal chemical variability (e.g., refs. 7–10), which cannot be explained by external environmental factors. Given the use of biomineral geochemistry as a tool to explore past climate change, it is important to understand the mechanisms controlling these patterns. This is particularly relevant today, as technological advances have improved the spatial resolution and sensitivity of geochemical analyses to the point where intraskeletal chemical variability can directly influence geochemical paleoclimate records (8, 11). It is critical to know how organic components affect skeletal growth and composition, and whether these organic-associated features can be distinguished from environmental signals.

Laboratory-grown mineral templating experiments show that the charge and structure of organic additives can affect crystal morphology (4, 12), polymorph (13), composition (14), orientation and nucleation rate (15). In vitro studies have also revealed that

Significance

Many marine organisms build complex $CaCO_3$ shells, which record an archive of past climate in their trace chemistry. Organic–mineral interactions are a crucial, poorly understood aspect of shell formation, which may alter shell composition and bias climate records. We map the chemistry of an organic template preserved within a type of shell that is widely used in studies of past climate. We find that the organic template embedded within this shell is enriched in Na and Mg, and can locally influence shell composition. Atom-scale maps of template chemistry offer a detailed view of the chemical interactions at an organic mineral template, and suggest that elements other than Ca^{2+} may be important in defining the energetics of $CaCO_3$ nucleation during biomineralization.

Author contributions: O.B., D.E.P., H.J.S., and A.C.G. designed research; O.B., E.A.B., D.E.P., Z.Z., M.W., and A.C.G. performed research; D.E.P. and Z.Z. contributed new analytic tools; O.B., E.A.B., D.E.P., Z.Z., M.W., and A.C.G. analyzed data; B.H., A.D.R., and J.S.F. were essential in culturing foraminiferal specimens for the study; and O.B., H.J.S., and A.C.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]O.B. and A.C.G. contributed equally to this work.

[2]Present Address: Research School of Earth Sciences, Australian National University, Acton ACT 2601, Australia.

[3]To whom correspondence may be addressed. Email: oscarbranson@gmail.com, hjspero@ucdavis.edu, or gagnon@uw.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1522864113/-/DCSupplemental.

Ca-specific recruitment to the template surface is an important primary step in carbonate nucleation (16), and it is commonly proposed that similar interactions are also crucial within mineralizing organisms (5). However, the importance of the interactions observed during in vitro experiments to biomineralization is poorly understood, because the chemistry of in situ organic templating surfaces in calcifying organisms is extremely hard to observe at spatially relevant scales. Observations of the atomic-scale chemistry of organic–mineral interfaces could tell us which chemical species and interactions are most crucial for nucleation within biominerals, and help build a detailed mechanistic understanding of biomineralization. However, to fully evaluate the role of organics in nucleation and shell geochemistry, we need multiscale analyses that can place atom-scale measurements of template chemistry in the context of skeletal-scale compositional patterns. If we can link these scales, we can be confident that our atom-scale observations capture the chemical patterns of templating processes, and test whether these organic components exert a measurable effect on shell geochemistry.

The planktic foraminifera *Orbulina universa* (Fig. 1) provides an ideal system for studying mineral-associated organics. Foraminifera are a cosmopolitan group of marine protozoans with a fossil record that extends back to the earliest Cambrian period (17), offering one of the most comprehensive geochemical archives of past climate and ocean chemistry (3). *O. universa* has been extensively used as a model organism for studying foraminiferal biomineralization and geochemistry (e.g., ref. 11). The $CaCO_3$ "test" (shell) contains a single, embedded organic structure known as the primary organic sheet (POS*; Fig. 1 *B* and *C*; 18, 19), which is produced during the initial stage of skeletal formation, and is composed of polysaccharides and acidic amino acids (20, 21). Little is known about the specifics of POS structure and formation, but studies of analogous structures in other organisms (22, 23) suggest that the polysaccharide component provides structural integrity to the organic template, and the acidic amino acids guide calcite nucleation (24). The fluids surrounding the POS are thought to originate from seawater, and are modified by unknown mechanisms before calcification (6, 7, 19). In *O. universa*, nucleation and mineral growth occurs on both sides of the POS, encasing the organic templating structure within the biomineral (18). This embedded POS can be located within the test (shell, Fig. 1), and offers an ideal target for a multiscale investigation of organic template chemistry in *O. universa*.

We use atom probe tomography (APT, ref. 25) to measure the atom-scale chemistry of the interface between the POS and the carbonate mineral skeleton in *O. universa*, and provide a compositional map of an organic template surface in a carbonate biomineral. This technique precisely determines the 3D position and identity of individual ions evaporated from a solid sample to within ~2 Å [~0.2 nm (25)], using a combination of field ion microscopy and time-of-flight mass spectrometry. Recent developments have allowed the application of APT to a range of natural materials (26–29) but carbonate minerals, which are by far the most abundant and diverse biomineral group produced in the plant and animal kingdoms, remain challenging targets for APT analysis. The structural heterogeneity, high evaporation field, and low thermal and electrical conductivity of carbonate biominerals make them a difficult material for analysis by routine APT techniques. We have optimized sample preparation and APT analysis parameters (*SI Appendix*, sections 3.1 and 3.2), and obtained APT data from a biomineral $CaCO_3$–organic interface. These analyses are given context by 2D chemical maps of the entire shell wall, collected using time-of-flight secondary ionization mass spectrometry (ToF-SIMS). ToF-SIMS is similar to the more established nano-SIMS technique used in previous studies (9, 30), but uses a time-of-flight mass separation device instead of a magnet to simultaneously measure a wide range of masses.

*The POS has previously been called the "primary organic membrane", but because it is not a membrane in the biological sense (i.e., a phospholipid bilayer), we adopt the "POS" nomenclature advocated by de Nooijer et al. (19).

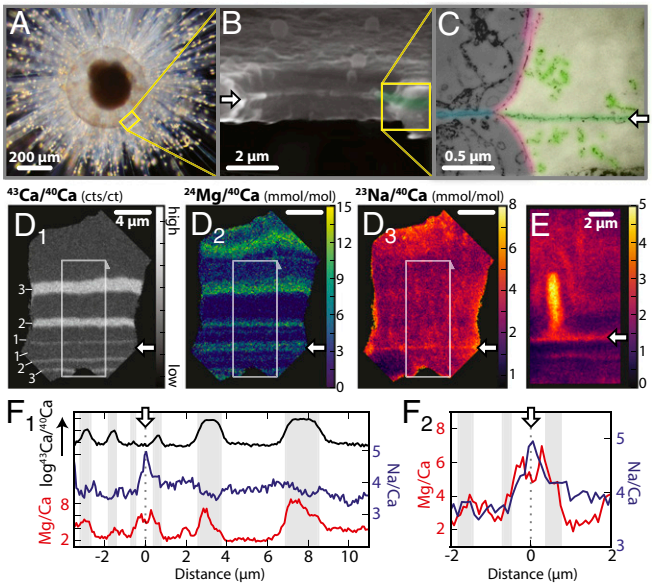


Fig. 1. The spherical shell of *O. universa* is constructed around a POS and is chemically heterogeneous. (*A*) Light micrograph of *O. universa* with a newly calcified spherical chamber surrounded by calcite spines and cellular material. Bright points on the spines are symbiotic dinoflagellates. (*B*) SEM of a 1-d-old calcite sphere, broken to expose a wall cross-section and shell pore. The white arrow identifies the embedded POS, which is visible as a raised ridge in the cross-section. Calcite precipitation occurs on both sides of the POS. (*C*) TEM of the cross-section of an agar-mounted decalcified shell wall (paler, *Right*) and pore (darker, *Left*). The POS (green horizontal layer, white arrow) is visible within the decalcified shell. The pore has an organic lining (red), and "pore plate" (blue) contiguous with the POS. For a detailed discussion of these features, see ref. 18. The POS appears to be a complex, branched structure rather than a continuous sheet, with numerous electron-dense regions of similar structure to the POS extending ~500 nm either side of the POS (light green). SEM and TEM images were generated following methods in ref. 18. (*D* and *E*) ToF-SIMS $^{43}Ca/^{40}Ca$, $^{24}Mg/^{40}Ca$, and $^{23}Na/^{40}Ca$ maps of cross-sections through a mature, 4-d-old shell wall. This foraminifera was moved between ^{43}Ca enriched seawater at night, and "natural" seawater during the day, resulting in three ^{43}Ca enriched bands on each side of the POS. The position of the POS (white arrows) is bracketed by these symmetric ^{43}Ca "labels" (numbered in D_1; *SI Appendix*, section 2.4), which identify the origin of calcification and constrain the location of the POS (white arrow). The region associated with the POS has elevated Na and Mg. *E* captures the base of a Na-rich calcite spine within the shell wall of a different specimen. The spines originate at the POS (18), further identifying the Na-rich band in D_3 as associated with the POS. Profiles in F_1 are extracted from white boxes in *D*, and enlargements in F_2 highlight the coincident Mg/Ca and Na/Ca maxima that are uniquely present at the POS. The double Mg maxima either side of the Na maximum in this specimen is present in ~33% of specimens, and is caused by the POS being between two, close high-Mg bands. A more typical signal can be seen in *SI Appendix*, section 2.4.

Results and Discussion

ToF-SIMS: Submicrometer 2D Maps. ToF-SIMS maps of skeletal chemistry across a wide range of elements (up to 72 amu) were collected using a 312-nm Bi^{+} analysis beam with a step size of 78.1 nm. This allows the detection of intraskeletal features at a minimum resolvable distance of 370 nm, beyond the resolution limits of previous techniques (e.g., ref. 31). Among the wide range of elements analyzed by ToF-SIMS, Mg and Na alone were found to exhibit systematic heterogeneity perpendicular to the growth axis of the shell (Fig. 1*D*). Magnesium variability confirms previous observations of intratest Mg heterogeneity in *O. universa* (8, 11, 31). Sodium variability is of lower magnitude, except for a single fine band that is uniquely enriched in both Na and Mg (Fig. 1 *D* and *F*). This feature is present in all specimens analyzed, occurring only once in each specimen. Within this band Na and Mg are elevated by 49 ± 20% and 147 ± 85% (n = 18),

respectively, relative to the adjacent mineral, with a full-width-at-half-peak-maxima (FWHM) of 673 ± 272 nm (n = 18). In most specimens, this band contains the highest concentration of Na in the whole skeleton, whereas the Mg concentration is within the range of other skeletal variations. Other elements known to exhibit systematic banding in *O. universa* (e.g., S:7, 9; B:10) were not detected by ToF-SIMS.

The position of the POS in ToF-SIMS maps is constrained by culturing foraminifera with a ^{43}Ca pulse-chase label, which produces symmetric layers of ^{43}Ca-enriched calcite on either side of the POS (Fig. 1D_1 and *SI Appendix*, section 2.4). The coincident Mg and Na enriched band is consistently located within the innermost ^{43}Ca bands, associated with the location of the POS (Fig. 1 *D* and *F*). Additional evidence that this feature is associated with the POS is provided by ToF-SIMS maps which contain the base of a "spine" (Fig. 1*E*), a calcite structure which is known to originate at the POS and protrude through the shell wall [Fig. 1*A* (18)]. These maps reveal that the spines are significantly enriched in Na, and terminate at the Na and Mg enriched band within the skeleton.

Sodium variability away from the POS is of lower amplitude, broader, and not clearly associated with sharp Mg maxima. Because these variations are unlikely to be associated with organic components (32), further discussion is outside the scope of this study and will be described in detail elsewhere. The ability to identify the location of the POS by the presence of simultaneous Mg and Na maxima using ToF-SIMS offers a way to map the location of organic components in foraminifera, and allows us to specifically target this important structure in our APT analyses.

APT: Subnanometer 3D Maps. Using APT, we map the 3D chemistry of the interface between an embedded organic structure and the adjacent calcite within a single foraminiferal specimen. Within our APT data volume, calcite is identified by its high Ca content relative to the organic layer, which was low in Ca but rich in C and H (Figs. 2 and 3). The organic layer is further distinguished from calcite by the types of C-containing ions emitted from each material during analysis (*SI Appendix*, section 3.3.4). The organic material yielded significantly fewer ions than calcite. This is expected, given the dependence of field evaporation on the bond strength of atoms (33), and prevalence of strong covalent bonds within organic material. However, it is likely that organic-hosted cations would be ionically bound and ionize relatively easily, making cation measurements from the organic more accurate than nonmetals.

In the APT reconstruction (Fig. 2), the calcite and organic regions are separated by a continuous, broadly planar interface that is defined by an abrupt change in Ca, C, and H over ∼2 nm. This interface is operationally identified by a 50% Ca isoconcentration surface (Fig. 2*A*), which is slightly curved and exhibits some nonsystematic roughness, manifest as surface undulations with a 5–10-nm period, and 1–2-nm amplitude. Compositional profiles across this boundary (proxygrams, Fig. 3) reveal an abrupt increase in Na and Mg concentration in the organic layer, compared with the calcite. Magnesium is variable throughout the organic layer with maxima ∼3 and ∼6 nm into the organic material. Sodium has a distinct maximum on the organic side of the interface, and a secondary maximum at ∼5 nm into the organic (Fig. 3*B*). The variability of Na and Mg within the organic layer may be taken to suggest an ∼5-nm periodicity, which could relate to aspects of the formation or structure of the organic layer, which are poorly constrained. At its surface maximum Na/Ca reaches a ratio of ∼0.6 counts/count, which corresponds to more than one Na for every two Ca detected at the surface of the organic layer. The lower evaporation rate of the organic layer can lead to the evolution of local topography during data collection, and deflect interfacial ions away from the organic toward the calcite (34). The compositional patterns on the organic side of the interface may therefore be conservatively constrained to the surface of the organic material, whereas ions on the calcite side may originate from either calcite or organic parts of the sample. Together with the ∼0.2-nm resolution of APT, this allows us to conservatively constrain the location of this Na maximum to the outer 1–2 nm of the organic layer, rather than at the interfacial boundary, or within the layer of calcite immediately adjacent to the organic surface.

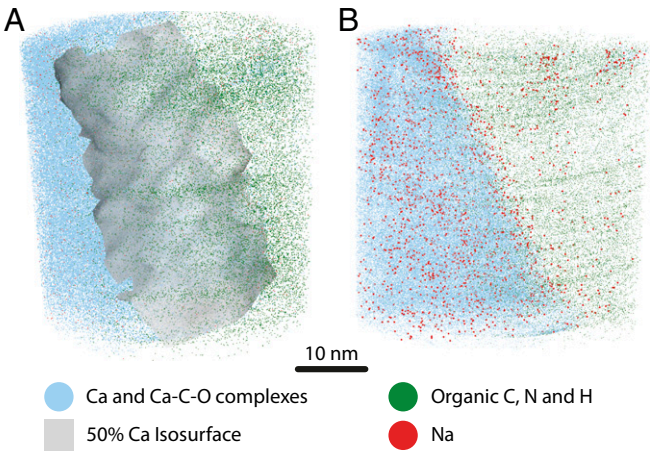


Fig. 2. Three-dimensional APT reconstruction of a planar organic-mineral templating surface within foraminiferal calcite. (*A*) The APT reconstruction captures the interface between calcite (Ca-rich) and an organic (Ca-poor) region. The structure of this predominantly planar interface is highlighted by a 50% Ca-concentration isosurface. (*B*) When viewed in plane with the interface, Na appears elevated at the interface. The symbols for Na have been enlarged 2× for emphasis. Note that whereas Na appears more abundant on the calcite side of the interface, this reflects the higher ionic yield from calcite than the organic layer, and by atomic fraction Na is more abundant in the organic layer, as highlighted in Fig 3.

The calcite layer is compositionally homogeneous at the scale of APT analyses, with 4.6 ± 0.7 counts/kcounts Mg/Ca and 7.6 ± 0.1 counts/kcounts Na/Ca, and no interface-specific compositional patterns. To investigate the influence of instrumental bias on APT-derived compositions, we compared the composition of calcite measured by both ToF-SIMS and APT in the same specimen. In ToF-SIMS, calcite adjacent to the POS contains ∼6.6 mmol/mol Mg/Ca and ∼3.6 mmol/mol Na/Ca (Fig. 1), suggesting that the uncalibrated APT analyses underestimate Mg/Ca by a factor of ∼0.7 and overestimate Na/Ca by a factor of ∼2 in calcite. Whereas this provides a rough estimate of instrumental bias, it is not possible to use this comparison with directly calibrate APT data, because ToF-SIMS analyses are unable to resolve calcite chemistry on the scale of APT analyses, and templating interactions may influence the chemistry of the initial calcite layer. The major element composition of the organic layer is uniform, with a measured stoichiometry of approximately $C_4O_3H_2$. However, the low yield of nonmetal elements in both the calcite and the organic layers likely biases this measurement.

Connecting Spatial Scales. The organic material observed by APT is elevated in both Na and Mg, and the only region in ToF-SIMS data consistently elevated in both Na and Mg is at the POS. This qualitative correspondence suggests that the organic layer observed in APT data could be the POS, but differences in the scale of the analyses preclude a direct quantitative comparison. Given the possible occurrence of non-POS organics within foraminifera (35), we test whether the organic layer captured by APT is compositionally consistent with the POS-associated chemical signal observed by ToF-SIMS using a quantitative model.

Because APT and ToF-SIMS integrate over different spatial scales, the magnitude of both Mg and Na signals will be different between the techniques, but the relative enrichment of Mg and Na within one method must match the other. To test this, we calculate the simultaneous Na and Mg enrichment that would be observed by ToF-SIMS when measuring an organic layer embedded in a calcite substrate, based on APT-measured compositions of the two materials. This model tests the hypothesis that the POS-associated

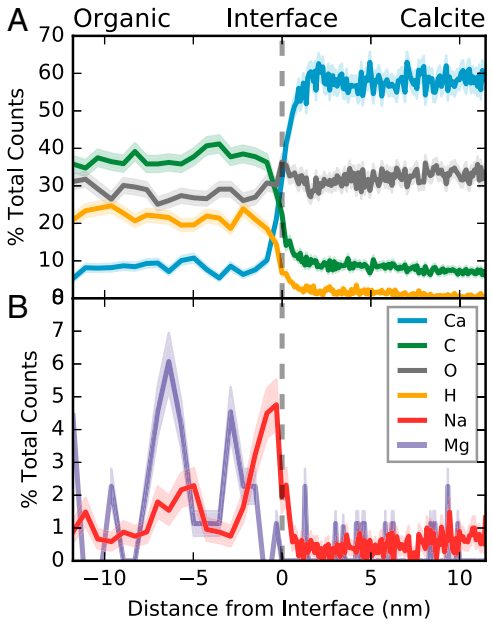


Fig. 3. Compositional profiles through the APT reconstruction reveal Na enrichment in the organic surface. (*A*) Chemical profiles ("proxygrams") across the biomineral interface in Fig. 2 quantify sharp changes in major elements across an ∼2-nm interface, defining the transition between the calcite and organic layers (dashed line). Profiles are presented as percent of collected ions, with profile bins defined using a constant number of ions per bin. Shaded error envelope calculated based on counting statistics. (*B*) Both Na and Mg are elevated in the organic material relative to the calcite. This enrichment is not centered on the interface, but instead occurs within the first few nanometers on the organic side of the interface.

signal observed in ToF-SIMS is derived from a material of similar composition to the ∼10-nm organic layer measured by APT. The inputs to our model are the volume-normalized APT counts of Ca, Na, and Mg from each material, the ToF-SIMS beam size (312 nm), and the width of the simulated POS. The POS width is unconstrained, and is tuned to match the absolute magnitude of Na and Mg enrichment observed in ToF-SIMS. Whereas POS width changes the absolute amplitude of the Na and Mg enrichment, it does not change their relative magnitude. We may therefore accept our hypothesis if our APT-based model can simultaneously predict both the Mg and Na enrichment observed in ToF-SIMS using only a single POS width.

Because the APT ion ratios are uncalibrated, we model the relative magnitude of the organic-driven increase above the adjacent background calcite. This approach allows for variable ionization efficiency between Ca, Mg, and Na in APT. We assume that this bias is similar for the three metals between organic and calcite layers. Within these constraints, our model accurately predicts both the Na and Mg enrichment observed by ToF-SIMS (Fig. 4 and *SI Appendix*, section 4), demonstrating that the POS-associated signal observed in ToF-SIMS is consistent with an organic material of the same average composition as the organic layer measured by APT. This provides an important test of the consistency between APT and ToF-SIMS data, and connecting these atom- and submicrometer-scale observations constitutes a significant step forward in correlative chemical microscopy. The strong correspondence between these scales demonstrates that the Mg and Na maxima in the ToF-SIMS data are likely caused by the same organic material measured by APT. This reveals that the POS can contribute to intraskeletal chemical patterns in foraminifera.

Within our model, the absolute amplitude of the predicted signal depends upon the thickness of the POS. This offers an estimate of POS width that is independent of the relative Na and Mg enrichment used to evaluate our hypothesis. The amplitudes of the ToF-SIMS Na and Mg maxima are consistent with an ∼130-nm-thick POS, in agreement with previous POS width estimates [∼100 nm, Fig. 1 (18)]. In contrast, the width of the POS-associated ToF-SIMS signal is more consistent with a POS thickness of ∼700 nm. Oblique sample sectioning could broaden the linear features in ToF-SIMS analysis, but is insufficient to explain the approximately sixfold difference between both previous transmission electron micrograph (TEM) analyses and ToF-SIMS peak widths (*SI Appendix*, section 1.2). This discrepancy may be accommodated if the POS is not a single laminar organic structure, as previously proposed (18, 19), but a dispersed organic framework that is concentrated around the origin of calcification. Our suggestion of a more complex POS geometry is supported by careful examination of *O. universa* TEM and SEM images [Fig. 1*C* (18)]. Whereas sample preparation procedures may alter or displace the delicate structures in these images, they show that the POS is branched and multifaceted, implying a larger effective width than a single, laminar structure. In line with this observation, we also observed two smaller organic structures in APT samples extracted away from the POS (*SI Appendix*, section 3.4).

Organics and Foraminiferal Geochemistry. Systematic variations in the intrashell concentrations of Mg, S, and B are widely recognized in *O. universa* and other foraminifera (Mg:8, 31, 32; S:7, 9;

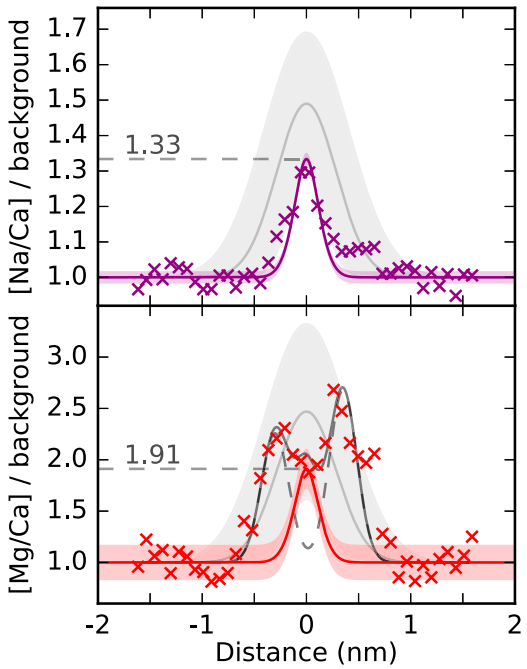


Fig. 4. Observed POS-specific Na and Mg maxima in ToF-SIMS can be explained by APT measurements, suggesting that both techniques are imaging the same structure. ToF-SIMS and APT cannot be directly compared because of differences in spatial scale. Instead, the relative magnitude of the POS Na and Mg enrichment measured by ToF-SIMS (crosses) are compared with APT measurements of POS composition (solid lines) in the same specimen using a quantitative model. The shaded gray envelopes show the mean and SD of the POS-associated signal in 18 other ToF-SIMS specimens, and highlight a degree of variability in Na and Mg enrichment between individuals. Modeled intensities are based on APT measured compositions, a beam width of 312 nm, and a POS thickness of 130 nm. The double Mg peaks bracketing the POS can be explained by two Gaussians with 312-nm FWHM (dashed peaks), consistent with two narrow bands of higher-Mg calcite ∼300 nm either side of the POS. At this distance, the contribution of these peaks to the POS location is negligible. The combination of these Gaussian peaks and the modeled POS signal (solid black line) describe the POS-associated Mg maxima in this specimen (*SI Appendix*, section 4.5).

B:10), but the mechanisms driving this variability are poorly constrained. Hypotheses range from physiological processes altering the chemistry of calcification (11), to the inclusion of organic components within the structure (6, 7), with many models in between. Our atom-scale characterization of POS composition "scales up" to quantitatively explain the POS-associated compositional patterns observed in a 2D map of the same specimen, demonstrating that the POS can drive localized enrichments in both Na and Mg in *O. universa*. However, because there is only one POS in *O. universa* (18), other processes must account for the intrashell chemical heterogeneity seen elsewhere in the shell, like Mg/Ca "banding." This is consistent with X-ray analyses of Mg enrichment away from the POS in *O. universa*, which reveal that non-POS Mg is not hosted in organic structures (32). Thus, the presence of organic components can explain an important aspect of, but not all, intrashell Mg and Na heterogeneity in *O. universa*.

The correspondence between our APT observations and ToF-SIMS data demonstrates that the POS can contribute to intraskeletal chemical heterogeneity and may impact paleo reconstructions based on highly resolved analytical techniques. This raises the important question of whether organic components may influence the bulk geochemistry of *O. universa*, and affect paleoceanographic records. The similarity between ToF-SIMS observations of POS composition in 19 *O. universa* specimens offers an important level of confidence that observations from our single APT specimen may be more broadly applicable to *O. universa* as a whole, and allow us to evaluate the importance of organics to the overall geochemistry of *O. universa*.

The Mg/Ca of foraminifera is a well-established proxy for past ocean temperature (36), and Na/Ca is being considered as a proxy for ocean salinity (37). We calculate the contribution of the POS to bulk Mg/Ca and Na/Ca in *O. universa* based on our APT measurements, by extending the APT-based model applied to predict our ToF-SIMS observations to the bulk scale (*SI Appendix*, section 4.6). The contribution of the POS decreases logarithmically as shell thickness increases, and will add >0.5 °C to a Mg/Ca-derived temperature estimate at a shell thickness below ∼4.6 μm, and increase a salinity estimate by >1 practical salinity unit below ∼2.5 μm. Because thin-shelled foraminifera are rarely preserved in marine sediments, the POS is unlikely to be important in paleoceanographic applications of *O. universa*. However, the POS should be carefully considered when creating paleoproxy calibrations from cultured foraminifera, where thin-shelled specimens may be included in analysis. Furthermore, the majority of paleoceanographic analyses measure the chemistry of multichambered foraminifera, which may contain multiple organic layers (7, 19). If these organic layers are all enriched in Mg and Na, they may contribute significantly to bulk geochemistry. Quantifying the magnitude of these possible effects requires further investigation and characterization of the composition and frequency of organic layers within multichambered foraminifera. To this end, the POS-associated Na and Mg signal identified in this study may offer a promising new way to map the number and location of organic layers within the skeletons of more complex, multichambered foraminifera.

Pattern to Mechanism: Insights into Nucleation Processes. Given the presence of a POS-associated Na and Mg maximum in all 19 *O. universa* specimens analyzed, and the quantitative correspondence between our APT and ToF-SIMS datasets, it is likely that our APT measurement is a representative example of POS chemistry in *O. universa*. The distinctive atom-scale surface chemistry patterns may therefore provide new insights into the factors influencing template-directed calcite nucleation during biomineralization. The compositional patterns at the POS surface may be the result of the active or passive interaction of cations with a templating surface, increased trace element uptake into the initial calcite layer by template-induced lattice strain, or the exclusion of non-$CaCO_3$ ions from a crystal nucleating on the surface. The position of the Na maximum ∼1 nm on the organic side of the 50% Ca isosurface, before any appreciable increase in Ca concentration, conservatively constrains the location of the Na maximum to within the organic surface; it is neither at the interphase boundary between the template and mineral nor within the initial mineral layer. This suggests that the Na maximum is either the result of solute ions interacting with the template surface, or the exclusion of ions from crystal nuclei that are taken up into the organic surface. In either case, non-Ca^{2+} ions accumulate at the POS surface before or during heterogeneous nucleation.

The compositional patterns we observe at the surface of the POS may be altered up to the point at which the organic template is "buried" by crystalline carbonate, and can no longer exchange ions with solution. The POS surface Na maximum is enriched in both Ca^{2+} and Mg^{2+}, relative to the abundance of the cations in seawater (78% Ca, 18% Na, 4% Mg, or 39:9:2 vs. 1:50:5 in seawater; Fig. 3). This supports the preferential accumulation of Ca^{2+} at the template surface, as commonly suggested by analogous synthetic studies (5), but also suggests the simultaneous adsorption of non-Ca^{2+} ions. Because the chemistry of the calcification environment is unknown, we cannot distinguish whether these ions accumulate "passively" in proportion to their abundance in solution, or "actively" through ion-specific interactions with the organic surface. Regardless of the driving mechanism, the composition of the POS surface will influence the interaction between the biological template and crystal nuclei.

For nucleation to occur, ions must overcome an energetic barrier to reorganize into a new solid phase. One of the main factors controlling the magnitude of this barrier, and thus the likelihood of nucleation, are the interfacial energies between template, solution, and mineral (38). This is true for both ordered and amorphous calcium carbonate phases (39). Changes in the organic surface charge (38, 40), degree of structural similarity to the crystal (41), and ordering of water molecules adjacent to the surface (42) can each alter the interfacial energy of the template, and change the dynamics of nucleation. The presence of Na and Mg at the organic template surface can influence each of the surface properties that contribute to interfacial energy.

The ability of specific ions to stabilize or disrupt organic–water interfaces by modifying surface hydration or direct ion-pairing interactions is described by the Hofmeister series (42–44). Broadly, an organic–water interface is stabilized by weakly positive or strongly negative ions, and destabilized by strongly positive or weakly negative ions. The major seawater cations are strongly charged, and will tend to exert a destabilizing (chaotropic) influence, increasing the organic–water interfacial energy. The strength of these effects is ion specific, and sensitive to organic surface structure and solute composition (43, 45, 46).

Sodium and magnesium are major components of POS surface chemistry, making up 18% and 4% of the detectable surface "cation load." Given the differences in size and charge density between these ions, each may have a distinct effect on the surface properties of the mineral template, alter the barriers to nucleation, and affect the energetics of biomineralization as a whole (38). Because of its relatively strong chaotropic influence, it is likely that Mg^{2+} will be a particularly potent modifier of surface properties, although the greater abundance of Na^{+} at the surface may make its overall contribution more significant.

Studies have considered the effect of overall ionic strength on in vitro template-facilitated mineral nucleation (38), but the effects of specific solute ions on nucleation energetics proposed here have yet to be explored. Our study highlights the need for inorganic nucleation experiments to investigate the effects of non-Ca^{2+} ions, which are currently thought to be "spectator ions" in nucleation, on the energetics of template-directed nucleation. In particular, determining the magnitude and direction of the proposed ion-specific effects is essential to evaluate their importance to biomineralization. If the abundance of Mg^{2+} and Na^{+} at the POS surface is related to their concentration in seawater, the chemical patterns at the POS surface offer an intriguing connection between changes in seawater composition through time and the energetics of biomineralization.

Methods

Foraminifera Culturing. Specimens of the foraminifera *O. universa* were hand-collected (47), and individual foraminifera were cultured following standard methods (48) at 21 °C on a 12-h/12-h day–night light cycle in ambient seawater of known composition (*SI Appendix*, section 1.1).

ToF-SIMS Data Collection. The terminal spherical calcite chambers of *O. universa*, which grew completely in the laboratory, were chemically cleaned of surface-associated organic material using a hot NaOH H_2O_2 solution, mounted in Araldite epoxy resin and polished to expose the test cross-section, and gold coated for analysis (*SI Appendix*, section 1.2). Polished specimens were analyzed in a TOF.SIMS5 (IONTOF GmbH, Germany) ToF-SIMS with a 25 keV Bi^+ ion beam with a pulsed beam current of 0.66 pA, and a measured beam width of 312 nm (FWHM). Data were collected between 0 and 78 amu, and specific peaks corresponding to $^{23}Na^+$, $^{40}Ca^+$, $^{43}Ca^+$, and $^{24}Mg^+$ were identified. Resulting maps and profiles are the sum of 1,600 analyses collected over 2 h for each location (*SI Appendix*, section 2.1).

APT Data Collection. APT specimens were extracted from the polished foraminiferal test and sharpened through focused ion beam lift-out techniques (49) using a Helios NanoLab 600 (*SI Appendix*, section 1.3). Specimens were extracted from within ~1 μm of the location of the POS, as determined by the position of ^{43}Ca-enriched bands (Fig. 1). APT data were collected using a Cameca LEAP 4000X-HR instrument at 44 K with a 50-pJ 355-nm laser pulsed at 160–200 Hz. Data were reconstructed with Cameca's Integrated Visualization and Analysis Software, using measurements of pre- and post-analysis SEM images of the sample to constrain the volume.

APT Data Processing. Ion identities were assigned to Da ranges based on sample composition (*SI Appendix*, section 3.3.1). The biomineral interface was defined as the 50% Ca concentration isosurface (*SI Appendix*, section 3.3.4). Proxygram profiles for every spectral peak and associated background regions were taken normal to the interface, and imported into Python for analysis (*SI Appendix*, section 3.3.5). After background subtraction, proxygram profiles were rebinned to maintain a constant total ion count in each bin, making counting statistics comparable. Na counts were calculated by subtracting predicted $^{46}Ca^{2+}$ from the 23-Da peak, calculated based on the natural abundance of Ca isotopes and the counts of $^{40}Ca^{2+}$. Mg counts were calculated based on the $^{26}Mg^+$ and $^{25}Mg^{2+}$ peaks (*SI Appendix*, section 3.3.2). Atomic compositions were calculated from the background-corrected counts, based on the composition assigned to each peak. Further detail on sample preparation, analysis, and processing can be found in the *SI Appendix*.

ACKNOWLEDGMENTS. We thank Adam Wallace for insightful discussions. This research was supported in part by US National Science Foundation Awards 1420689 (to A.C.G.), 1061676 (to H.J.S.), 1261519 (to A.D.R. and J.S.F.), and 1232987 (to B.H.). APT and ToF-SIMS analysis was supported through User Proposal 48564 to A.C.G., part of a Special Science Call at the Environmental Molecular Sciences Laboratory, a Department of Energy Office of Science User Facility sponsored by the Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory.

1. Knoll A (2003) Biomineralization and evolutionary history. *Rev Mineral Geochem* 54(1):329–356.
2. Elderfield H (2002) Climate change. Carbonate mysteries. *Science* 296(5573):1618–1621.
3. Katz ME, Cramer BS, Franzese A, Hönisch B (2010) Traditional and emerging geochemical proxies in foraminifera. *J Foraminiferal Res* 40(2):165–192.
4. de Yoreo JJ, Wierzbicki A, Dove PM (2007) New insights into mechanisms of biomolecular control on growth of inorganic crystals. *CrystEngComm* 9(12):1144–1152.
5. Gilbert PUPA, Abrecht M, Frazer BH (2005) The organic-mineral interface in biominerals. *Rev Mineral Geochem* 59(1):157–185.
6. Bentov S, Erez J (2006) Impact of biomineralization processes on the Mg content of foraminiferal shells: A biological perspective. *Geochem Geophys Geosyst* 7(1):Q01P08.
7. Erez J (2003) The source of ions for biomineralization in foraminifera and their implications for paleoceanographic proxies. *Rev Mineral Geochem* 54:115–149.
8. Eggins S, de Deckker P, Marshall J (2003) Mg/Ca variation in planktonic foraminifera tests: Implications for reconstructing palaeo-seawater temperature and habitat migration. *Earth Planet Sci Lett* 212(3):291–306.
9. Paris G, Fehrenbacher JS, Sessions AL, Spero HJ, Adkins JF (2014) Experimental determination of carbonate-associated sulfate $\delta^{34}S$ in planktonic foraminifera shells. *Geochem Geophys Geosyst* 15(4):1452–1461.
10. Branson O, Kaczmarek K, Redfern SAT, Misra S (2015) The coordination and distribution of B in foraminiferal calcite. *Earth Planet Sci Lett* 416:67–72.
11. Spero HJ, et al. (2015) Timing and mechanism for intratest Mg/Ca variability in a living planktic foraminifer. *Earth Planet Sci Lett* 409:32–42.
12. Orme CA, et al. (2001) Formation of chiral morphologies through selective binding of amino acids to calcite surface steps. *Nature* 411(6839):775–779.
13. Falini G, Albeck S, Weiner S, Addadi L (1996) Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science* 271(5245):67–69.
14. Wang D, Wallace AF, De Yoreo JJ, Dove PM (2009) Carboxylated molecules regulate magnesium content of amorphous calcium carbonates during calcification. *Proc Natl Acad Sci USA* 106(51):21511–21516.
15. Pouget EM, et al. (2009) The initial stages of template-controlled $CaCO_3$ formation revealed by cryo-TEM. *Science* 323(5920):1455–1458.
16. Smeets PJM, Cho KR, Kempen RGE, Sommerdijk NAJM, De Yoreo JJ (2015) Calcium carbonate nucleation driven by ion binding in a biomimetic matrix revealed by in situ electron microscopy. *Nat Mater* 14(4):394–399.
17. Pawlowski J, et al. (2003) The evolution of early Foraminifera. *Proc Natl Acad Sci USA* 100(20):11494–11498.
18. Spero HJ (1988) Ultrastructural examination of chamber morphogenesis and biomineralization in the planktonic foraminifer *Orbulina universa*. *Mar Biol* 99(1):9–20.
19. de Nooijer LJ, Spero HJ, Erez J, Bijma J, Reichart GJ (2014) Biomineralization in perforate foraminifera. *Earth Sci Rev* 135:1–11.
20. Robbins LL, Brew K (1990) Proteins from the organic matrix of core-top and fossil planktonic-foraminifera. *Geochim Cosmochim Acta* 54(8):2285–2292.
21. Stathopolos L, Tuross N (1994) Proteins and DNA from modern planktonic-foraminifera. *J Foraminiferal Res* 24(1):49–59.
22. Gotliv BA, Addadi L, Weiner S (2003) Mollusk shell acidic proteins: In search of individual functions. *ChemBioChem* 4(6):522–529.
23. Drake JL, et al. (2013) Proteomic analysis of skeletal organic matrix from the stony coral Stylophora pistillata. *Proc Natl Acad Sci USA* 110(10):3788–3793.
24. Addadi L, Weiner S (1985) Interactions between acidic proteins and crystals: Stereochemical requirements in biomineralization. *Proc Natl Acad Sci USA* 82(12):4110–4114.
25. Miller MK, Forbes RG (2014) *Atom-Probe Tomography* (Springer, New York).
26. Valley JW, et al. (2014) Hadean age for a post-magma-ocean zircon confirmed by atom-probe tomography. *Nat Geosci* 7(3):219–223.
27. Gordon LM, Joester D (2011) Nanoscale chemical tomography of buried organic-inorganic interfaces in the chiton tooth. *Nature* 469(7329):194–197.
28. Perea DE, et al. (2016) Atom probe tomographic mapping directly reveals the atomic distribution of phosphorus in resin embedded ferritin. *Sci Rep* 6:22321–22329.
29. Pérez-Huerta A, Laiginhas F, Reinhard DA, Prosa TJ, Martens RL (2016) Atom probe tomography (APT) of carbonate minerals. *Micron* 80:83–89.
30. Gagnon AC, Adkins JF, Erez J (2012) Seawater transport during coral biomineralization. *Earth Planet Sci Lett* 329-330:150–161.
31. Eggins S, Sadekov AY, de Deckker P (2004) Modulation and daily banding of Mg/Ca in *Orbulina universa* tests by symbiont photosynthesis and respiration: a complication for seawater thermometry? *Earth Planet Sci Lett* 225(3-4):411–419.
32. Branson O, et al. (2013) The coordination of Mg in foraminiferal calcite. *Earth Planet Sci Lett* 383:134–141.
33. Gault B, et al. (2006) Design of a femtosecond laser assisted tomographic atom probe. *Rev Sci Instrum* 77(4):043705.
34. Devaraj A, Colby R, Vurpillot F, Thevuthasan S (2014) Understanding atom probe tomography of oxide-supported metal nanoparticles by correlation with atomic-resolution electron microscopy and field evaporation simulation. *J Phys Chem Lett* 5(8):1361–1367.
35. Cuif J-P, Dauphin Y, Nehrke G, Nouet J, Perez-Huerta A (2012) Layered growth and crystallization in calcareous biominerals: Impact of structural and chemical evidence on two major concepts in invertebrate biomineralization studies. *Minerals* 2(1):11–39.
36. Lea DW, Mashiotta TA, Spero HJ (1999) Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing. *Geochim Cosmochim Acta* 63(16):2369–2379.
37. Wit JC, de Nooijer LJ, Wolthers M, Reichart GJ (2013) A novel salinity proxy based on Na incorporation into foraminiferal calcite. *Biogeosciences* 10(10):6375–6387.
38. Giuffre AJ, Hamm LM, Han N, De Yoreo JJ, Dove PM (2013) Polysaccharide chemistry regulates kinetics of calcite nucleation through competition of interfacial energies. *Proc Natl Acad Sci USA* 110(23):9261–9266.
39. Hu Q, et al. (2012) The thermodynamics of calcite nucleation at organic interfaces: Classical vs. non-classical pathways. *Faraday Discuss* 159(0):509–523.
40. Walker JBA, Heywood BR, Mann S (1991) Oriented nucleation of $CaCO_3$ from metastable solutions under Langmuir monolayers. *J Mater Chem* 1(5):889–890.
41. Mann S, et al. (1993) Crystallization at inorganic-organic interfaces: Biominerals and biomimetic synthesis. *Science* 261(5126):1286–1292.
42. Lo Nostro P, Ninham BW (2012) Hofmeister phenomena: An update on ion specificity in biology. *Chem Rev* 112(4):2286–2322.
43. Hess B, van der Vegt NFA (2009) Cation specific binding with protein surface charges. *Proc Natl Acad Sci USA* 106(32):13296–13300.
44. Chaudhuri RG, Paria S (2009) Dynamic contact angles on PTFE surface by aqueous surfactant solution in the absence and presence of electrolytes. *J Colloid Interface Sci* 337(2):555–562.
45. Collins KD (2004) Ions from the Hofmeister series and osmolytes: Effects on proteins in solution and in the crystallization process. *Methods* 34(3):300–311.
46. Zavitsas AA (2001) Properties of water solutions of electrolytes and nonelectrolytes. *J Phys Chem B* 105:7805–7817.
47. Huber BT, Bijma J, Spero HJ (1996) Blue water scuba collection of planktic foraminifera. Methods and techniques of underwater research. *Proceedings of the American Academy of Underwater Sciences Scientific Diving Symposium, Smithsonian Institution, Washington DC*, eds Lang MA, Baldwin CC (American Academy of Underwater Sciences, Mobile, AL), pp 127–132.
48. Lea DW, Spero HJ (1992) Experimental determination of barium uptake in shells of the planktonic Methods and techniques of underwater research. *Geochim Cosmochim Acta* 56(7):2673–2680.
49. Miller MK, Russell KF, Thompson K, Alvis R, Larson DJ (2007) Review of atom probe FIB-based specimen preparation methods. *Microsc Microanal* 13(6):428–436.

Supplements to: *Nanometer-Scale Chemistry of a Calcite Biomineralization Template: Implications for Skeletal Composition and Nucleation*

O Branson E A Bonnin D E Perea H J Spero Z Zhu
M Winters B Hönisch J S Fehrenbacher A D Russell A C Gagnon

DOI:10.1073/pnas.1522864113

Contents

1 Sample Preparation

1.1 Foraminifera Culturing

Foraminiferal culture techniques have been described in detail elsewhere (for example, Lea and Spero [1992]). Briefly, juvenile specimens of the foraminifera *Orbulina universa* are hand collected using blue-water diving techniques [Huber et al., 1996]. These specimens are cultured in controlled laboratory conditions at constant temperature under a 12-hr/12-hr day-night cycle in seawater of known composition. Each foraminifera is hand fed every other day with individual *Artemia*. During culture, the juvenile foraminifera develop a spherical terminal chamber, by producing a primary organic sheet (POS) that forms the template for calcite growth on both its inner and outer surfaces. The outer side of the POS is more heavily calcified than the inner side.

The calcite of two specimens was periodically 'labelled' with ^{43}Ca to constrain the origin of calcification, and the position of the POS within the mature skeleton. Specimens of *O. universa* were allowed to precipitate $CaCO_3$ in natural seawater during the day, and in ^{43}Ca enriched seawater at night to label 'night' calcite over the duration of a multi-day laboratory culture experiment. Because biomineralization is continuous and occurs on both sides of the POS in *O. universa* [Spero et al., 2015], this night-time ^{43}Ca label generates a predictable number of ^{43}Ca-enriched calcite bands on both sides of the POS.

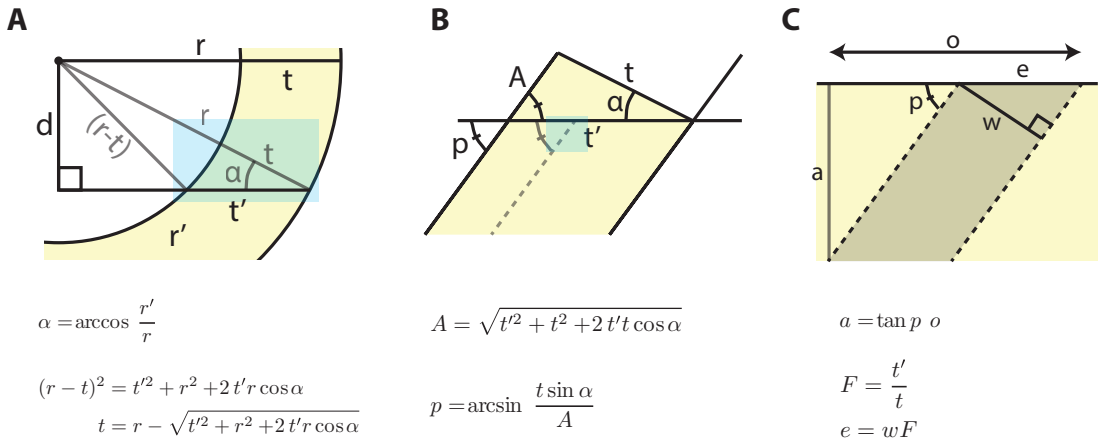
1.2 Sample Mounting and Polishing

The final calcite spheres of cultured *O. universa* specimens were mounted in Araldite epoxy resin and polished to exposed a cross section using successively finer diamond polishing products (15 - 0.125 μm; Buehler). The polished surface was sputter-coated with 10nm of gold.

As far as possible, samples were polished until the exposed cross-section was normal to the sphere wall. Oblique sectioning of the sphere by the polished surface will distort chemical patterns along the radial axis. Furthermore, because ToF-SIMS analyses penetrate beneath the sample surface during analysis, laminar features bisecting the surface at an angle may be measured at depth and appear broader than they actually are. The extent of laminar feature distortion due to oblique sample sectioning is calculated from measurements of pre- and post-polishing sphere radius, and post-polishing wall thickness (Fig 1). Based on these calculations, the POS intersected the polished surface at a mean angle of $68\pm10^{\circ}$, which corresponds to a sample 'stretch' factor (F, see Fig 1) of 1.1 ± 0.1 (i.e. features appear 1.1 times broader than they actually are in the direction of the growth axis).

1.3 FIB Preparation of APT Specimens

APT specimens were extracted from the polished cross sections using a FEI Helios Nanolab Focused Ion Beam (FIB) instrument (Fig 2). First, a protective layer of Pt was deposited along a strip of calcite ~20 μm long and ~2 μm wide spanning and parallel to the region thought to contain the POS. The calcite underneath this protected area was excavated from



$$\alpha = \arccos \frac{r'}{r}$$

$$(r-t)^2 = t'^2 + r^2 + 2\,t' r \cos\alpha$$

$$t = r - \sqrt{t'^2 + r^2 + 2\,t' r \cos\alpha}$$

$$A = \sqrt{t'^2 + t^2 + 2\,t' t \cos\alpha}$$

$$p = \arcsin \frac{t \sin\alpha}{A}$$

$$a = \tan p \; o$$

$$F = \frac{t'}{t}$$

$$e = wF$$

Figure 1: The spherical *O. universa* test may be approximated as two concentric spheres. The angle of sample sectioning may therefore be calculated based on pre- and post- polishing measurements of our specimens. In each diagram, the shaded blue region is enlarged on the right. (A) The radius (r) of the outer sphere is defined by the overall sphere size, and the distance between the inner and outer spheres is the shell thickness (t). Embedding and polishing the shell exposes a plane bisecting these spheres, with a measurable radius (r') and thickness (t'). In the ideal case r' = r, and the section will be perfectly normal to the shell wall, but in some cases r' < r, and the section will be oblique. Resulting maps will be distorted, and laminar features will be artificially broadened. The extent of this distortion may be calculated as a distortion factor (F), which can be calculated based on our measurements of r, r' and t'. (B) Within the constraints of our analyses (ToF-SIMS penetration depth $< 1\mu$m, F < 1.4), the intersection of the shell with the resin can be approximated as a linear system, and the angle between the POS and resin (p) calculated. (C) From this, it is possible to calculate the width of a laminar feature of known thickness that bisects the surface (e), and the depth that an analysis must penetrate (a) to make a thin laminar feature appear a given width (o).

the sample surface, yielding a $20\ \mu\mathrm{m}$ long strip of calcite, with a triangular cross section and a maximum width of $\sim 2\ \mu\mathrm{m}$. The strip was lifted from the sample surface using an Omniprobe micro-manipulator and transferred to a custom sample rotation stage, which allows rotation of the sample along the long axis, to correctly orient features in the APT sample. The rotated sample was then transferred back to the micro-manipulator, and $\sim 2\ \mu\mathrm{m}$ sections of the material were cut from the strip and fixed to sharpened Si APT substrate pillars with a Pt weld. The resulting specimens were then sharpened to a fine point using sequentially finer annular milling FIB patterns. As the sample tip became finer, the energy of the FIB beam was reduced to minimize Ga implantation and sample damage. Ga implantation is directly observed in the APT analysis, allowing damaged regions to be excluded during analysis.

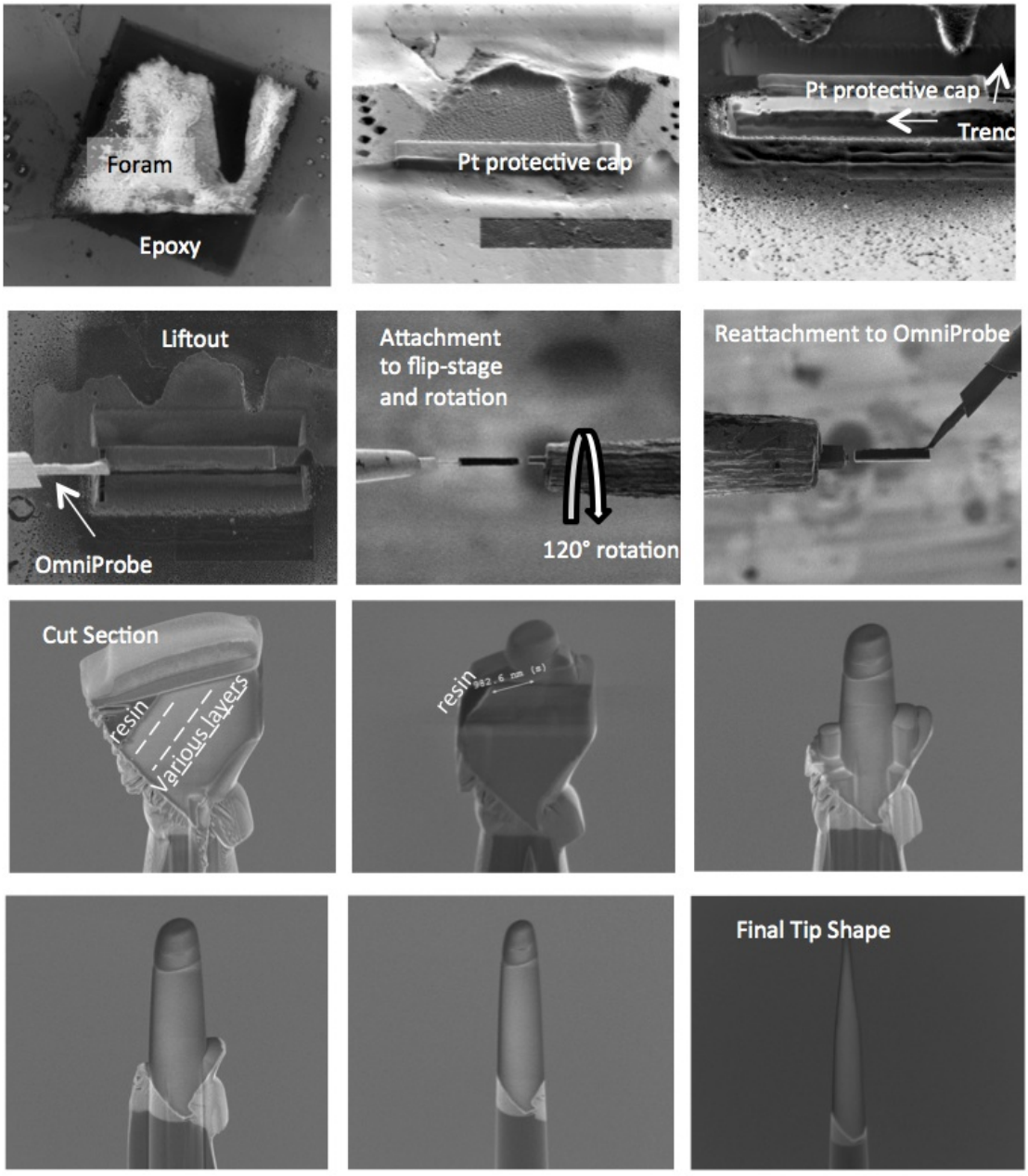


Figure 2: SEM images of FIB sample preparation process starting from a polished section of a foraminiferal shell (roughly 20 microns from inner to outer surfaces) and ending with a roughly 1 micron tall tip ready for APT analysis.

2 Time of Flight Secondary Ionization Mass Spectrometry (ToF-SIMS)

2.1 Data Collection

ToF-SIMS analysis was conducted using a TOF.SIMS5 (IONTOF GmbH, Germany) time-of-flight mass spectrometer. Positive secondary ions were generated through the use of a 25 keV Bi^+ ion beam with a pulsed beam current of 0.66 pA. The Bi^+ beam was focused to have a Gaussian full-width-at-half-maximum (FWHM) of approximately 300 nm. The precise beam FWHM was 312 nm, determined from the distance between the $16-84^{\text{th}}$ percentiles of the signal change in a linear profile collected across a sharp edge on a 1000 mesh Cu grid standard. This corresponds to a minimum resolvable distance of $\sim$370 nm, following the Rayleigh Criterion of a 26.2% signal decrease between adjacent maxima.

Prior to analysis, the region of interest was pre-sputtered at a 25x25 μm field of view for 60 seconds to clear away the gold coating and reach stable beam conditions. For each chemical map, a total of 3200 measurements were collected from 20x20 μm areas, with a step size of 78 nm (256x256 pixels). Data were collected for all masses between 0-78 amu.

After collection, the maps were analyzed for spatial shift and corrected for translational movement during analysis. The intensity of the beam profile was monitored over the course of each analysis, to ensure stable beam conditions. In most measurements, the beam was unstable for the first half of the analysis (up to scan $\sim$1600). These data were excluded from analyses. The counts at each pixel corresponding to Na, Ca, and Mg collected under stable beam conditions ($\sim$1600 maps) were summed to provide chemical maps of the sample. Ca was relatively homogeneous throughout the structure, showing that all observed Mg/Ca and Na/Ca banding are results of changes in the measured Mg and Na intensity.

2.2 Data Calibration

ToF-SIMS Na/Ca and Mg/Ca were calibrated using six carbonate standards (Table 1). Previous measurements of these standards by isotope dilution ICP-MS [Gabitov et al., 2013] revealed a degree of Mg and Sr heterogeneity between grains. We therefore calibrated our ToF-SIMS profiles against Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) analyses of Mg/Ca and Na/Ca taken from the same grain surfaces measured by ToF-SIMS. Analyses were performed using a Teledyne Photon Machines Analyte G2 193 nm excimer laser with a HelEx dual-volume laser ablation cell coupled to an Agilent 7700x quadrupole-ICP-MS. The LA-ICP-MS measured Mg/Ca and Sr/Ca of standards were all within range of published isotope-dilution ICP-MS measurements [Gabitov et al., 2013].

2.3 Data Processing

Matrices containing raw ToF-SIMS counts of Ca, Na and Mg were imported into python for analysis using the numpy and scipy packages. Masks were created to remove counts originating from the mounting resin by smoothing the Ca image with a 3 pixel Gaussian filter, and manually specifying a minimum count threshold based on the Ca count histogram

Table 1: The names, abbreviations and M/Ca in mmol/mol of the six standards used in this study, as determined by LA-ICP-MS.

Standard	Mg/Ca (mmol/mol)	Na/Ca (mmol/mol)
Blue Calcite (BCC)	4.23 ± 0.07	0.015 ± 0.003
UC Berkeley Calcite 135 (CAL-135)	0.108 ± 0.008	0.009 ± 0.001
Hebrew University Aragonite (HUJ)	0.0463 ± 0.0004	1.08 ± 0.02
LAS-20 Calcite (LAS-20)	43.00 ± 0.53	0.48 ± 0.03
Rensselaer Polytechnic Calcite (RPI)	0.402 ± 0.006	0.011 ± 0.002
UC Irvine Calcite (UCI)	3.52 ± 0.08	0.012 ± 0.002

to separated 'calcite' and 'resin' regions within the map. Next Mg/Ca and Na/Ca ratios were calculated from the masked matrices, and calibrated against two or more of the carbonate standards (Table 1), measured during the same ToF-SIMS analysis session. The mean Na/Ca and Mg/Ca ratios of entire analysis regions was within error of LA-ICP-MS measurements of foraminifera cultured under similar conditions (2.6±0.2 mmol/mol Na/Ca, and 6±1 mmol/mol Mg/Ca), demonstrating the efficacy of our calibration.

Profiles of ^{43}Ca/Ca, Mg/Ca and Na/Ca were extracted along the growth axis of the shell, normal to the laminar compositional patterns. This was achieved by rotating the ToF-SIMS matrices using linear interpolation to re-grid the data, and extracting sub-matrices from an region of interest (ROI) within the rotated image. This sub-matrix was then converted to a concentration profile by calculating the per-row mean normal to the growth axis.

2.4 POS Identification

Several lines of evidence identify the location of the POS within our ToF-SIMS maps:

1. The introduced periodic ^{43}Ca 'label' present in the calcite which symmetrically brackets the origin of calcification (Section 1.1).

2. Skeletal Mg/Ca increases naturally each night [Eggins et al., 2004, Spero et al., 2015], offering a similar, though less controlled periodic marker.

3. ToF-SIMS maps of a calcite spine, which terminates at the POS.

In the two specimens containing the ^{43}Ca label, the origin of calcification can be constrained to between the innermost ^{43}Ca bands (Fig 3A). In the specimen analyzed by APT, this corresponds to a $\sim 1.17\ \mu m$ wide band. At approximately the center of this band, we observe a single, distinct Na/Ca maximum, bracketed by double Mg/Ca maxima. This double Mg/Ca maxima was observed in ∼33% of specimens (Fig 3 A & B), as opposed to the remainder of specimens which contained a single distinct Mg maximum coincident with the Na maximum (Fig 3C). This feature can explained if chamber formation is initiated in the late afternoon, causing two distinct layers of high-Mg 'night' calcite [Eggins et al., 2004, Spero et al., 2015] to be deposited on either side of the POS in close proximity to the origin of calcification. Because calcification occurs on both sides of the POS, the origin of

calcification and the position of the POS may be constrained to the $\sim$600 nm between these bracketing Mg/Ca maxima. These maxima are located $\sim$1 FWHM on either side of the POS, and will therefore have minimal contribution to the POS Mg signal. Given the analysis spot (FWHM = 312 nm) and step (78 nm) sizes, the majority of analysis spots between these adjacent Mg/Ca maxima will measure the POS. Thus, in the specimen analyzed by APT, we are able to constrain the position of the POS to within the resolution limits of the instrument. In both ^{43}Ca labeled specimens, the innermost ^{43}Ca bands contained simultaneous Mg/Ca and Na/Ca maxima. These elements did not coincide systematically anywhere else in the cross sections outside the origin of calcification, and appear to be uniquely associated with the position of the POS.

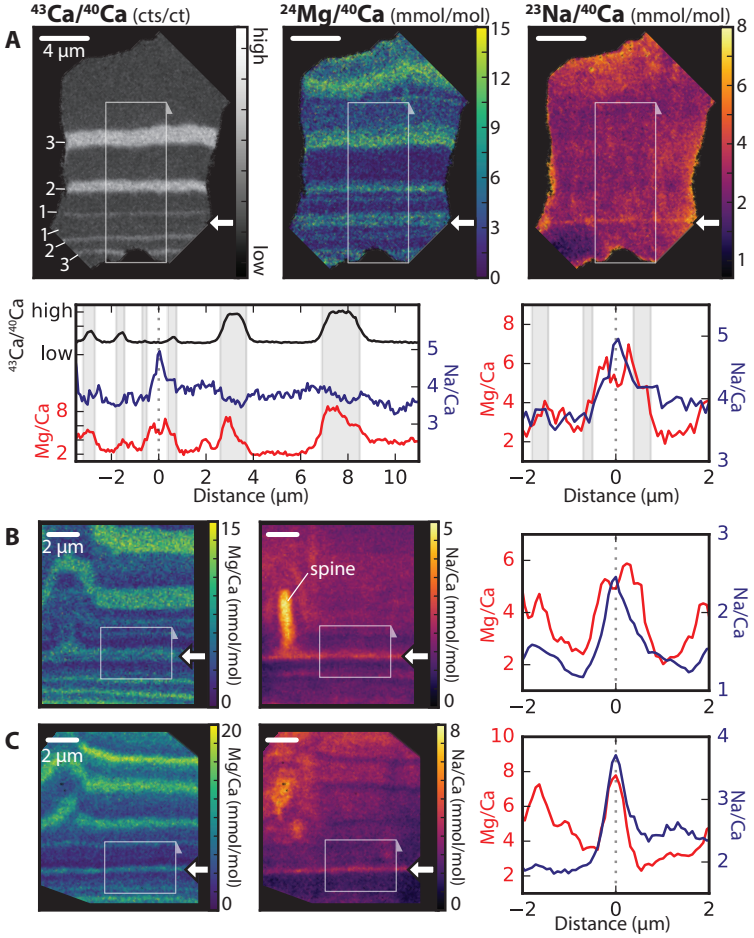


Figure 3: ToF-SIMS maps reveal the distinctive chemistry of the POS. Maps of ^{43}Ca/Ca, Mg/Ca and Na/Ca in three *O. universa* specimens, and the chemistry of the POS. The POS position is marked by a white arrow on the right side of all images, and a dotted vertical line in all profiles. Profiles are extracted from white boxes on the images, with the direction of the distance axis marked by an arrow. Profiles on the right highlight the POS-associated Na/Ca and Mg/Ca enrichment. (A) Maps of *O. universa* specimen analyzed by APT, containing a ^{43}Ca/Ca 'night' calcite label, allowing the identification of the POS. (B) A specimen containing a Na-rich spine originating at the POS. Both (A) and (B) have double Mg maxima either side of the Na maximum, due to the POS being enveloped by two high-Mg bands. This was present in $\sim$1/3 of specimens analyzed, and a more typical POS-signal is shown in (C).

In the 17 specimens *without* a ^{43}Ca label, the location of the POS was identified using the naturally occurring nightly Mg/Ca maxima throughout the shell, which offer a less precise estimate of POS location (Fig 3B&C). In all specimens the region identified using Mg/Ca periodicity contained a simultaneous Mg/Ca and Na/Ca maxima, similar to that observed associated with the POS in the ^{43}Ca labeled specimens. This suggests that the simultaneous Mg/Ca and Na/Ca maxima observed in all specimens is uniquely associated with the origin of calcification, and the position of the POS.

Finally, the location of the POS is independently confirmed by ToF-SIMS maps which contain the base of a 'spine', a calcite structure which is known to originate at the POS and protrude through the shell wall [Spero, 1988]. Our ToF-SIMS maps reveal that these spines are significantly enriched in Na, and that they originate from the POS-associated Na and Mg enriched band (Fig 3B).

In combination, these localization techniques point to a single high Na and Mg enriched band at the origin point of calcification, associated with the POS. Coincident Na and Mg maxima in *O. universa* may provide a convenient intra-skeletal tracer for POS location, which may be applied in future studies targeting the POS.

Our maps also reveal systematic 'banded' Na/Ca variability (Fig 3 B & C), which is distinct from the POS signal. The magnitude of this variability is significantly less than the POS Na/Ca maximum. Away from the POS, the Na/Ca maxima rarely coincides with Mg/Ca maxima, and Na/Ca often appears anti-correlated with the well-characterized distribution of Mg/Ca found throughout the skeleton [Fig 3; Eggins et al., 2003, Spero et al., 2015] Because X-ray analyses of *O. universa* have revealed that Mg away from the POS is not hosted in organic components [Branson et al., 2013], this suggests that non-POS chemical heterogeneity is unlikely to be driven by organic material.

2.5 POS Measurement

The mean Na/Ca and Mg/Ca signal in concentration profiles normal to the POS were fit with a combination of a Gaussian function and a linear background to precisely determine the width, amplitude and position of the POS-associated signal (Fig 4). The Na and Mg maxima were found to have a mean FWHM of 673 ± 272 nm. Estimates of POS width from previous studies, and our 'scaling-up' model suggest it is ~ 100 nm wide. Because of oblique sample sectioning (Section 1.2), the sub-surface analysis of material by the ToF-SIMS beam could cause this broadening. However, to measure a ~ 100 nm feature as a ~ 670 nm feature, the ToF-SIMS beam would have to penetrate 2.2 μm into the sample surface, which is double the maximum estimated analysis depth of $\sim 1 \mu\mathrm{m}$. It is therefore unlikely that this broadening can be explained by the geometry of sample sectioning.

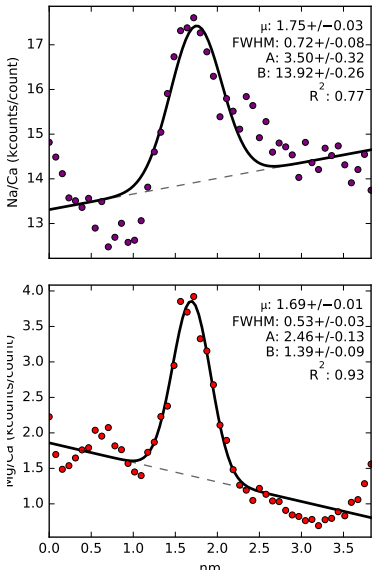


Figure 4: ToF-SIMS measured Na/Ca and Mg/Ca maxima associated with the POS, fitted with a linear background and a Gaussian peak. Statistics noted on the plots: μ = peak position, FWHM = peak Full Width at Half Maximum, A = height of the Gaussian at μ, B = the background at μ, and the R^2 of the fit.

3 Atom Probe Tomography (APT)

3.1 Specimen Optimization

The thermal insulating properties of calcite slow the dissipation of heat from the laser pulses, which can lead to heterogeneous evaporation, changes in complex ion formation, and thermal artifacts in the resulting ToF spectra as ionization harmonics are generated by reflected energy within the sample. These complex effects are determined by subtleties of sample geometry and are challenging to correct for post-analysis. A broad tip with a high shank angle provided the 'cleanest' ToF spectra, because it provided maximal area for heat to dissipate down the sample (Fig 2). This tip geometry also minimized the rate of catastrophic fracture in the tips.

In addition, a variety of tip geometries and sample coatings were investigated in attempts to optimize run success. The electrical insulating properties of calcite make the sample tips prone to fracture under a high applied field, as the hydrostatic stresses induced by the field exceed the strength of the sample tip. Conductive sample coatings (Pt, Cr) were investigated, but differences between the ionization fields between coating and specimen led to the evolution of tip topography that introduced complex artifacts in the samples which rendered the data unusable (Fig 5). Future research to identify a sample coating with a similar evaporation field to calcite has the potential to significantly improve the likelihood of successful analysis. In our study, the highest quality data were collected from uncoated samples. Sample fracture at the start of analysis was minimized by manually ramping up sample potential voltage very slowly, while constantly monitoring ionization rates and patterns.

3.2 Data Collection

Sample tips were analyzed using a Cameca Local Electrode Atom Probe (LEAP 4000X-HR) running in pulsed laser mode (at 44 K, with a 50 pJ 355 nm UV laser pulsed at 160-200 Hz). Laser pulse rate was varied throughout the run to minimize thermal artifacts in the sample spectrum. Specimen voltage was increased manually to 4 keV, or until sufficient ion yield was obtained to allow fine sample alignment relative to the detector and local electrode. After alignment, voltage was automatically controlled to maintain an ion yield of 0.2%. Typically the voltage increased throughout the run, but did not exceed 7 keV. On average, ions were excited every 1033 laser pulses, with 1.13 ions per excitation (19.5 % multi-hits). The mass resolving power is 1270 for $^{40}Ca^{2+}$, and 365 for Na^{+}. Mass resolving power is defined as $M/\Delta M$ using peak width at full width and half maximum height (FWHM) to measure ΔM.

Approximately 10% of ∼30 specimens analyzed yielded usable datasets ($> 1 \times 10^6$ counts), although explosive ionization events driven by structural heterogeneities reduced our success rate to ∼5%. In contrast, ∼90% of samples fractured catastrophically in the initial stages of analysis.

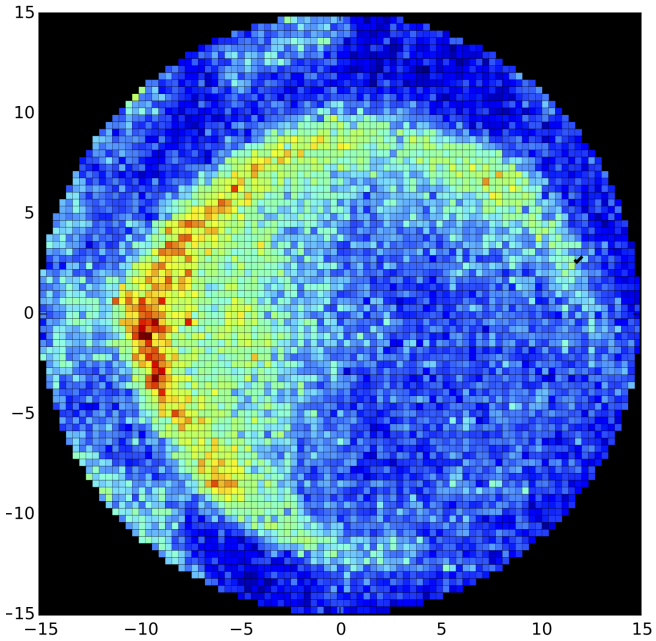


Figure 5: Total detector hits during the analysis of a Cr coated specimen, demonstrating the complications associated with some of the coatings we experimented with. The ring represents anomalously high counts, associated with the deflection of ions by topography at the Cr/calcite boundary.

3.3 Data Processing

Raw LEAP data were processed using a combination of Cameca's IVAS$^{\mathrm{TM}}$ software and python scripts in the IPython environment [Perez and Granger, 2007, Section 5].

3.3.1 Spectral Processing

In IVAS, ToF peaks were corrected for the hemispherical shape of the peak tip by minimizing their full width at half maximum, within the constraints of a physical instrument model. The ToF spectra were then scaled to known peak positions to account for variations in sample-detector distance.

Peaks in the ToF-SIMS spectra were approximately identified by hand, then precisely characterized in python using an asymmetric pseudovoigt function, adapted from Stancik

and Brauns [2008]:

$$\gamma_x(x) = \frac{2\gamma_0}{1 + \exp\left[a(x - x_0)\right]} \tag{1}$$

$$G(x) = \frac{A}{\gamma_x}\sqrt{\frac{4\ln 2}{\pi}}\exp\left[-4\ln 2\left(\frac{x - x_0}{\gamma_x}\right)\right] \tag{2}$$

$$L(x) = \frac{2A/\pi\gamma_x}{1 + 4[(x - x_0/\gamma_x)]^2} \tag{3}$$

$$y(x) = fL(x) + (1 - f)G(x) \tag{4}$$

Where γ denotes peak's full width at half max (FWHM), and varies as a function of x to impart asymmetry to the peak. The degree of asymmetry is determined by the variable *a*. The balance between Lorentzian (*L(x)*) and Gaussian (*G(x)*) components is determined by the coefficient *f*. The instantaneous FWHM (γ_0) determines the overall peak width.

Peak and background windows were chosen based on these fits, and used for all subsequent analysis of bulk and proxygram compositions. The 'on-peak' window was specified as 1 FWHM either side of the peak center, and two 1 FWHM wide 'background' windows were placed either side of the peak 3 FWHM away from the peak center. Background and peak windows were examined manually before proxygram creation, and the position of background regions were adjusted to avoid adjacent peaks and significant peak tails. Where the peak tail was unavoidable (e.g. $^{40}Ca^{2+}$), both background regions were moved to before the peak.

These peak assignments were used to background correct ToF spectra and 'proxygram' composition profiles extracted from IVAS. First, each peak extracted in the proxygram was background corrected, assuming that the background was linear in the peak region:

$$C(d) = C_{pk} - C_{bg} * \frac{n_{pk}}{n_{bg}} \tag{5}$$

where n is the number of bins of peak and background counts captured by their respective windows. The error of the background based on the standard deviation of the background region is taken into account during background subtraction. Background corrected counts are converted to atom counts by decomposing complex ions, based on the compositions assigned to the peaks (Section 5.3.2).

ToF peaks were assigned compositions based on the most likely combination of constituent ions known to be present in the sample (Table 2 and Fig 6). The bio-carbonate samples yielded numerous complex and multiply charged ions, and it was not possible to identify all peaks. Of the total 45 ToF peaks, six could have been produced by multiple likely compositions, so were excluded from analysis. These six unidentified peaks accounted for 1.45% of all on-peak ToF counts, and 98.55% of on-peak counts were assigned ionic identities.

Table 2: All peaks observed in the ToF spectra, and their ionic identities

Da	Composition	Da	Composition
1.0	H^+	27.9	$^{12}C^{16}O^+$
2.0	H_2^+	29.0	$^{12}C^{16}OH^+$
3.0	H_3^+	30.0	$^{12}C^{16}O_3^{2+}$
6.0	$^{12}C^{2+}$	30.5	*Unknown*
6.5	$^{12}CH^{2+}/(^{13}C^{2+})$	31.0	$^{31}P^+$
7.0	$^{14}N^{2+}$	32.0	$^{16}O_2^+$
12.0	$^{12}C^+/^{24}Mg^{2+}$	36.0	$^{40}Ca^{16}O_2^{2+}$
12.5	$^{25}Mg^{2+}$	37.0	*Unknown*
13.0	$^{12}CH^+/^{26}Mg^{2+}$	40.0	$^{40}Ca^+$
14.0	$^{12}CH_2^+/^{14}N^+$	42.0	$^{42}Ca^+$
15.0	$^{12}CH_3^+/^{14}NH^+$	43.9	$^{12}C^{16}O_2^+$
16.0	$^{16}O^+/^{14}NH_2^+$	44.9	*Unknown*
17.0	$^{16}OH^+/^{14}NH_3^+$	48.0	$^{40}Ca^{16}_2O^{2+}$
18.0	$^{16}OH_2^+/^{14}NH_4^+$	55.9	$^{40}Ca^{16}O^+$
18.5	*Unknown*	57.0	$^{40}Ca^{16}OH^+$
19.0	$^{16}OH_3^+$	59.0	*Unknown*
20.0	$^{40}Ca^{2+}$	59.9	$^{12}C^{16}O_3^+$
21.0	$^{42}Ca^{2+}$	61.9	*Unknown*
22.0	$^{44}Ca^{2+}$	68.8	Ga^+
23.0	$^{46}Ca^{2+}/^{23}Na^+$	69.9	$^{40}Ca^{12}_2C^{16}O_3^{2+}$
24.0	$^{48}Ca^{2+}/^{24}Mg^+$	71.9	$^{40}Ca^{16}O_2^+$
25.0	$^{25}Mg^+$	99.9	$^{40}Ca^{12}C^{16}O_3^+$
26.0	$^{26}Mg^+$		

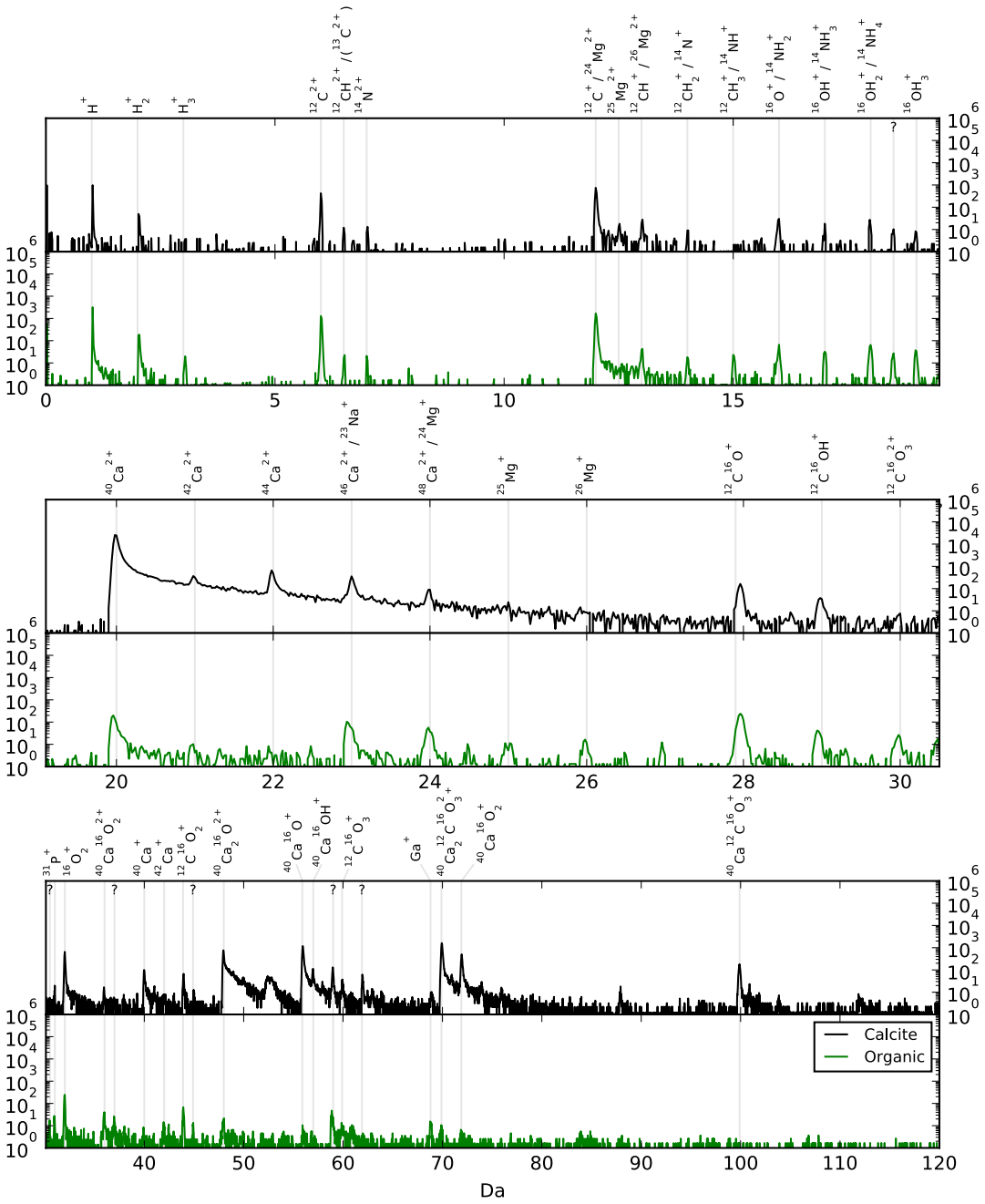


Figure 6: ToF-SIMS pectra from calcite and organic layers, with peak identities (Table 2) marked.

3.3.2 Na and Mg calculation

The main Na and Mg peaks suffer from interferences with major element peaks in the Time-of-Flight spectrum (Table 3). These interferences may be accounted for by either subtracting calculated intereference counts in the case of Na, or by extrapolation from counts of the minor isotope in the case of Mg (Fig 7). For a detailed evaluation of interference correction, see Section 5.3.3.

Table 3: Possible ions of Na and Mg and their Interferences.

Da	Isotope of Interest	Relative Abundance	Interference
23.0	$^{23}Na^{+}$	100%	$^{46}Ca^{2+}$
12.0	$^{24}Mg^{2+}$	79%	$^{12}C^{+}$
12.5	$^{25}Mg^{2+}$	10%	Unknown
13.0	$^{26}Mg^{2+}$	11%	$^{12}CH^{+}$
24.0	$^{24}Mg^{+}$	79%	$^{48}Ca^{2+}$
25.0	$^{25}Mg^{+}$	10%	Unknown
26.0	$^{26}Mg^{+}$	11%	Unknown

Sodium does not have any interference-free peaks, so must be calculated by subtracting calculated interference counts. Interference counts for the 23 Da peak are calculated by estimating $^{46}Ca^{2+}$ counts from $^{40}Ca^{2+}$ counts and the natural abundance of Ca isotopes:

$$^{23}\mathrm{Na}_{\mathrm{c}}^{+} = {}^{23}\mathrm{Da}_{\mathrm{c}} - {}^{46}\mathrm{Ca}_{\mathrm{c}}^{2+} \tag{6}$$

$$^{46}\mathrm{Ca}_{\mathrm{c}}^{2+} = {}^{40}\mathrm{Ca}_{\mathrm{c}}^{2+} \times A^{46}\mathrm{Ca}/A^{40}\mathrm{Ca} \tag{7}$$

Where the relative abundance (A) of $^{46}Ca^{2+}$ and $^{40}Ca^{2+}$ are 0.004% and 96.941%, respectively.

Total Mg counts may be extrapolated from both ^{25}Mg and ^{26}Mg counts, which have no known interferences with major sample constituents. In the case of singly-charged Mg, both $^{25}Mg^{+}$ and $^{26}Mg^{+}$ are evaluated. It is possible that unknown complex ion interferences may contribute to these peaks, so the lesser of the two peaks is used to calculate total Mg counts, following:

$$^{\mathrm{T}}\mathrm{Mg}_{\mathrm{c}}^{+} = {}^{26}\mathrm{Mg}_{\mathrm{c}}^{+} \times 100/A^{26}\mathrm{Mg} \tag{8}$$

Where the $A^{26}\mathrm{Mg}$ is the percentage abundance of ^{26}Mg (11%). For doubly charged Mg, significant interference peaks are present at 12 and 13 Da (Table 3), so $^{25}Mg^{2+}$ counts at 12.5 Da are used to estimate total Mg^{2+} counts:

$$^{\mathrm{T}}\mathrm{Mg}_{\mathrm{c}}^{2+} = {}^{25}\mathrm{Mg}_{\mathrm{c}}^{2+} \times 100/A^{25}\mathrm{Mg} \tag{9}$$

Where the $A^{26}\mathrm{Mg}$ is the percentage abundance of ^{25}Mg (10%). Single and double charged Mg counts may then be considered independently, or summed to calculate total Mg counts.

For Na, the same correction procedure is applied to both bulk APT spectra for calcite and organic components, and during proxygram analysis. For Mg, the sum of $^{25}Mg^{2+}$ and

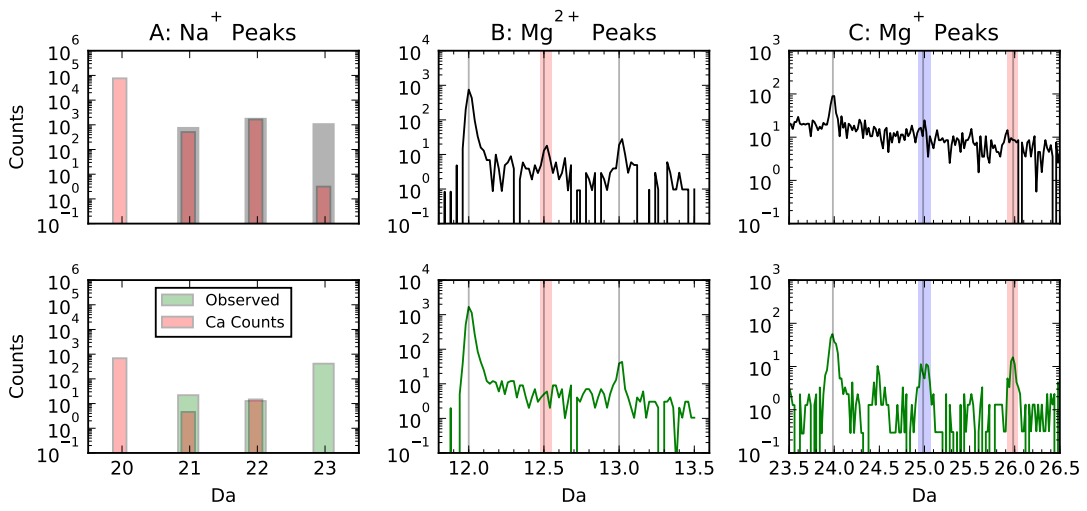


Figure 7: The correction methods used to deconvolve Na and Mg counts from interference peaks. (A) Predicted Ca counts for the 21, 22 and 23 Da peaks are based on measured $^{40}Ca^{2+}$ counts and the natural abundance of Ca isotopes. Na is calculated as the difference between the measured 23 Da peak, and the predicted $^{46}Ca^{2+}$ counts. (B & C) Total Mg counts are calculated from minor $^{25}Mg^{2+}$ and $^{Mg}Ca^{+}$ (highlighted in red). In (C), where both $^{25}Mg^{+}$ and $^{26}Mg^{+}$ peaks are evident, the smaller of the two is used, as the larger is more likely to be affected by unconstrained interferences. For a full evaluation of Mg estimation methods, see Section 5.3.3.

$^{26}Mg^{+}$ are used when dealing with whole-region mass spectra. However, only the $^{26}Mg^{+}$ peak is used in proxygram calculations. This is because the $^{25}Mg^{2+}$ peak at 12.5 Da is negligible in the spectrum extracted from the entire organic region (Fig 7), so considering this peak in proxygrams, where each bin considers 800 counts, would contribute more noise to the proxygram than Mg counts (Section 5.5).

3.3.3 3D Reconstruction

The entire data volume was reconstructed to a 3D point cloud, constraining the tip shape and total volume analyzed to physical measurements of sample tip. The tip shape and volume was determined from measurements of pre- and post-analysis SEM images of the tip, corrected for SEM view angle (Fig 8). This provided a tip radius estimate which evolves with z, and defines the reconstruction shape (Fig 9). The volume of the reconstruction was further altered by tuning the mean ion size and the image scaling factor so the reconstructed volume matched the volume of material removed during LEAP analysis.

The image compression factor is an empirical parameter that is used to account for the compression of electric field lines during analysis of a needle shaped specimen. The mean ion size is used to determine the spatial arrangement of ions in the reconstruction. By default, the ionic radius of the most common ion is used (in this case, Ca). However, the high proportion of complex ions from our samples rendered this approach inaccurate, so this parameter was increased to expand the reconstructed volume to match the volume of material removed by the LEAP. Final reconstruction parameters can be seen in Table 4. The

specimen radius parameter is used to determine the *z* evolution of atom position assignment: *xy* are measured by the detector, and *z* position is increased incrementally during position assignment, based on ion number.

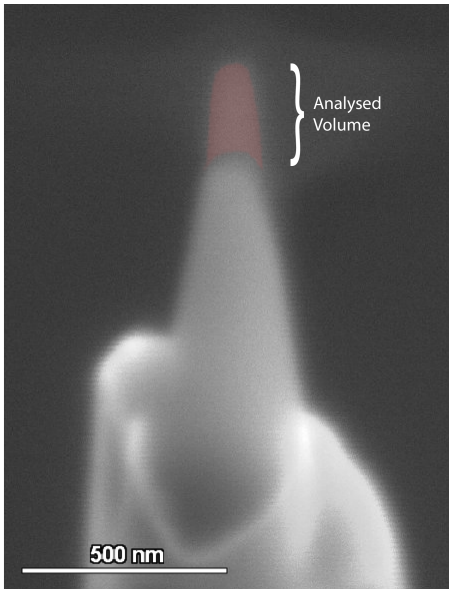


Figure 8: Overlaid pre- and post-analysis SEM images of a sample tip, with the analyzed region highlighted in red.

3.3.4 Interface Identification

Two materials are evident in our reconstructed data, as identified by differences in: (1) the abundance of Ca, C and H (2) changes in complex ions, particularly in C-bearing (3) the evaporation field. In calcite, C tends to ionize as CO_x^{n+} and $Ca_xCO_y^{n+}$ complexes, reflecting the $CaCO_3$ origin of these ions. In contrast, C^{n+} and CHxn+ ions are more prevalent in organic material, reflecting the typical bonding environment of C in organic compounds. This C ionization pattern agrees with the findings of previous APT analyses of organic materials [Gordon and Joester, 2011]. Calcite is further characterized by abundant Ca and a relatively low evaporation field, while the organic is characterized by low Ca, high elemental C and a relatively high evaporation field. The interface between the materials is defined as the 50% Ca isosurface through a 1x1x1.5 nm *xyz* sample mesh, after complex ion decomposition.

Table 4: Instrumental and reconstruction parameters for the specimen presented in the main manuscript.

	Parameter	Value
Instrument Parameters		
	k-factor	3.30
	Detector Efficiency	0.36
	Image Compression Factor	1.0
Materials Parameters		
	Primary Element	Ca
	Evaporation Field	18.0 V/nm
	Atomic volume	0.0435 nm^3
Specimen Parameters		
	Initial Tip Radius	14.0 nm
	Radius Evolution	see Fig 9
	Estimate Shank Half Angle	28.5 degrees

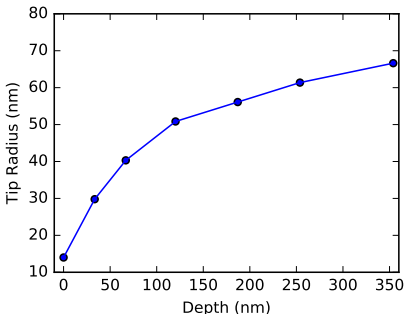


Figure 9: The evolution of tip radius used in the reconstruction, as measured from Fig 8. The depth axis is corrected for an SEM view angle of 56°.

3.3.5 Proxygram extraction and processing

Proxygrams are concentration profiles created normal to a defined interface. In its raw form, it provides the ion counts of all regions of the ToF spectrum specified in a 'range' file, which contained 'on-peak' and adjacent 'background' regions for each ToF peak, as determined by a peak fitting algorithm (Section 3.3.1).

Proxygrams were exported with a fine bin size (0.01 nm), and re-binned such that:

$$\sum_{d}^{d+w} A = n \tag{10}$$

Where bin width (w) at a given position in the proxygram (d) varies to maintain a constant total atom (A) count (n). This reveals relative atomic composition across the proxygram, normalizing for differences in ion yield between the materials.

For a more practical, in-depth description of this procedure, see Section 5.

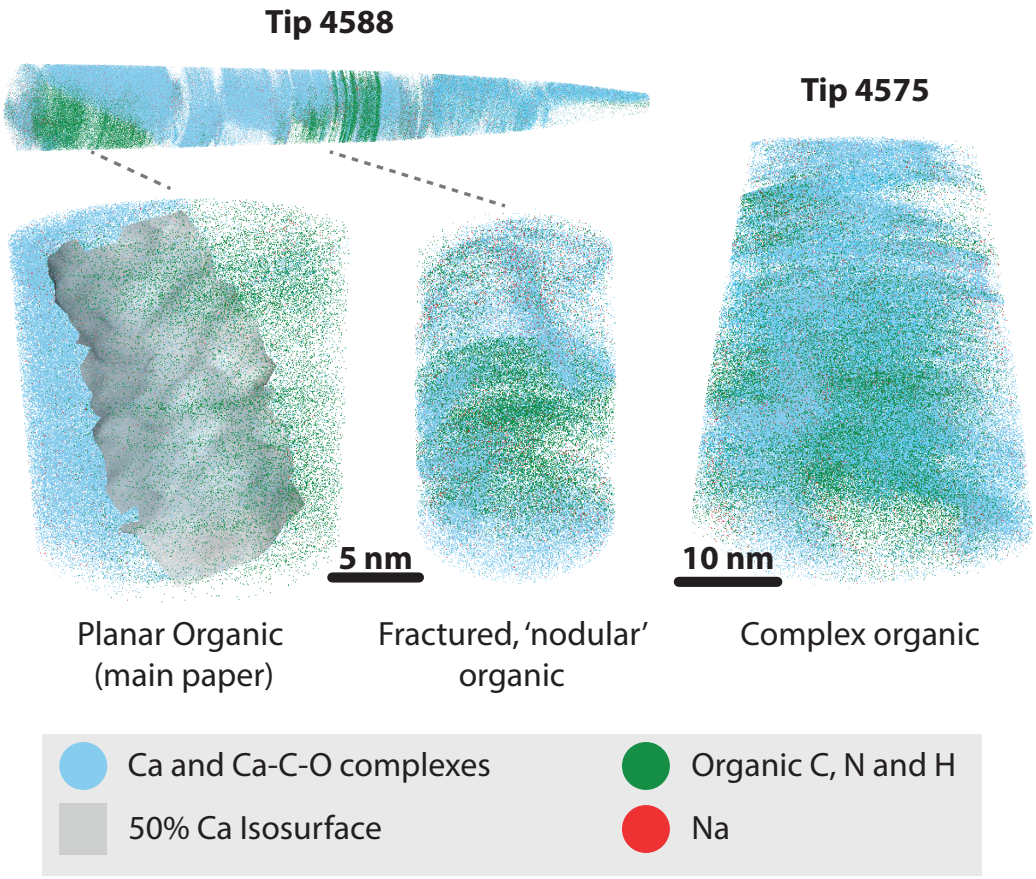


Figure 10: Three organic regions were captured across two APT specimens.

3.4 Secondary Organic Regions

In addition to the large, planar organic-calcite interface region, we observe two other 'nodular' calcite-embedded organic regions in separate APT specimens (Fig 10). Although these nodular regions were too small for robust compositional analysis, they may be non-laminar extensions of the POS, or previously undocumented non-POS mineral-associated organics (Main Paper Fig 1C). These non-POS organics suggest that in *O. universa*, organic-mineral interactions may play a role in biomineralization beyond that of the POS.

4 Scale-Up Model

4.1 Context

We have measured the atom-scale composition of calcite and an embedded organic using APT, and sub-micron-scale skeletal chemistry using ToF-SIMS. We have observed that both the organic observed by APT, and the region associated with the position of a known organic layer in the skeleton measured by ToF-SIMS are both enriched in Mg and Na, compared to the adjacent calcite. This qualitative similarity suggests that the signal observed by ToF-SIMS may be caused by the same organic that we have captured by APT. We seek to determine whether the organic measured by APT can quantitatively account for the Na and Mg enrichment observed in ToF-SIMS However, these data cannot be directly compared because:

1. The spatial scale of the measurement are not comparable.
2. It is not possible to calibrate the APT measurements.
3. Differences in ionization mechanics do not allow direct comparison of raw counts.

To overcome these barriers, we use a quantitative model to simulate the magnitude of the compositional signal we would expect if we observed a calcite-embedded organic layer by ToF-SIMS, using the compositions of the calcite and organic components measured by APT.

4.2 Hypothesis

Our model tests the hypothesis that:

> *The Ca, Mg and Na counts measured in calcite and organic materials by APT match the POS-associated Mg/Ca and Na/Ca enrichment seen in ToF-SIMS.*

If this hypothesis is true, it follows that the POS-associated chemical signal observed by ToF-SIMS originates from the same organic material observed by APT.

Testing this hypothesis with our model is not immediately straightforward. Because the two techniques integrate over different spatial scales, the magnitude of both Mg and Na signals will be different between APT and ToF-SIMS. However, the relative ratio of Mg to Na within one method must match the ratio of Mg to Na in the other method. To accept our hypothesis we must show that the (Na/Ca)/(Mg/Ca) in the APT-measured organic is the same as in the ToF-SIMS signal. For the reasons outlined above it is not possible to directly compare these datasets, so we must use a quantitative model.

We may accept our hypothesis if we are able to simulate *both* the Mg/Ca and Na/Ca enrichments observed in ToF-SIMS with a single model parameterization. If matching Na/Ca and Mg/Ca enrichments requires two separate model parameterizations, the hypothesis must be rejected.

4.3 Parameterization

The model involves two steps:

1. Simulate the proportion of a ToF-SIMS analysis beam that would illuminate organic material as it scanned across a calcite-embedded organic layer.

2. Use an end-member mixing model to translate the 'proportion organic material' to a compositional signal, based on APT-measured Ca, Mg and Na counts in the two materials.

4.3.1 Simulated Analysis Beam

We simulate a ToF-SIMS analysis beam scanning across a calcite-embedded organic layer of specified thickness. In mathematical terms, this is reduced to 2 dimensions, and accomplished by convolving a Gaussian function with a 'tophat' function to calculate the proportion of the ToF-SIMS analysis beam that measures the organic and calcite materials (Figure 11).

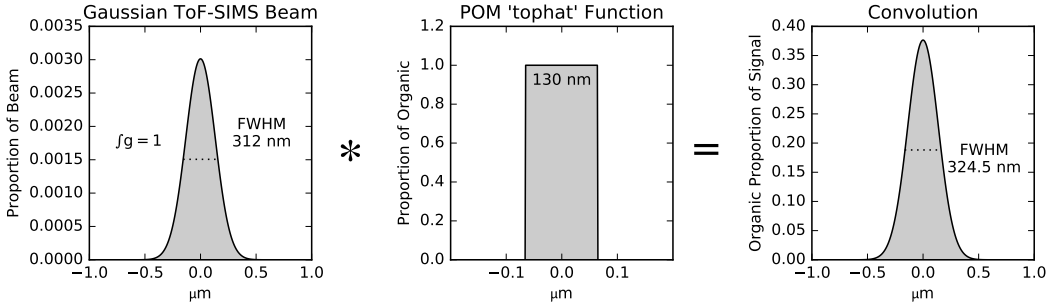


Figure 11: The calculation of the organic signal intensity and shape by convolution. (A) The Gaussian ToF-SIMS beam shape with a FWHM of 312 nm and an integral of 1. (B) The tophat function describing the organic layer, shown here with a total width of 130 nm. (C) The convolution of the gaussian and tophat functions, whose intensity is equal to the proportion of the ToF-SIMS signal that comes from the organic layer, and whose width is a function of both organic layer width and beam FWHM.

This convolution is performed numerically using numpy.convolve, which produces a signal whose intensity corresponds to the proportion of organic material measured by the analysis beam (see Appendix 4.7: Convolution Calculation).

4.3.2 End-Member Mixing Model

Once the proportion of organic illuminated by the simulated analysis beam is determined, this is applied to a simple end-member mixing model. The M/Ca ratio 'seen' by the simulated analysis beam is defined by the total counts of M and Ca from each material:

$$\frac{M}{Ca}_{mix} = R_{mix} = \frac{{}^{M}n_O + {}^{M}n_C}{{}^{Ca}n_O + {}^{Ca}n_C} \tag{11}$$

Where *n* is the number of counts, the superscript denotes the ion (M = non-Ca), the subscript denotes *C*alcite or *O*rganic, and $\frac{M}{Ca}$ is expressed as R. From this:

$$R_{mix}\ {}^{Ca}n_O - {}^{M}n_O = {}^{M}n_C - R_{mix}\ {}^{Ca}n_C \tag{12}$$

And, dividing through by ${}^{Ca}n_x$:

$${}^{Ca}n_O(R_{mix} - R_O) = {}^{Ca}n_C(R_C - R_{mix}) \tag{13}$$

$$\frac{{}^{Ca}n_O}{{}^{Ca}n_C} = \frac{(R_C - R_{mix})}{(R_{mix} - R_O)} \tag{14}$$

$$\left(\frac{{}^{Ca}n_O}{{}^{Ca}n_C}\right) R_{mix} - \left(\frac{{}^{Ca}n_O}{{}^{Ca}n_C}\right) R_O = R_C - R_{mix} \tag{15}$$

$$R_{mix}\left(\frac{{}^{Ca}n_O}{{}^{Ca}n_C} + 1\right) = R_C + \left(\frac{{}^{Ca}n_O}{{}^{Ca}n_C}\right) R_O \tag{16}$$

$$R_{mix} = \frac{R_C + \left(\frac{{}^{Ca}n_O}{{}^{Ca}n_C}\right) R_O}{\left(\frac{{}^{Ca}n_O}{{}^{Ca}n_C} + 1\right)} \tag{17}$$

With knowledge of the relative Ca content of the calcite and organic, and the proportion (p) of each material 'seen' by the simulated analysis beam, $\frac{{}^{Ca}n_O}{{}^{Ca}n_C}$ may be calculated as:

$$\frac{{}^{Ca}n_O}{{}^{Ca}n_C} = \frac{[Ca]_O}{[Ca]_C}\frac{p_O}{p_C} \tag{18}$$

Where the proportions of calcite (p_C) and organic material (p_O) are calculated by convolution (Section 4.3.1) or via a geometric model describing bulk calcification (Section 4.6). Thus, the convolved organic layer signal may be converted to a compositional Mg/Ca or Na/Ca signal following:

$$R_{mix} = \frac{R_C + \left(\frac{[Ca]_O}{[Ca]_C}\frac{p_O}{p_C}\right) R_O}{\left(\frac{[Ca]_O}{[Ca]_C}\frac{p_O}{p_C} + 1\right)} \tag{19}$$

Compositional data are derived from APT spectra extracted from each material within the reconstruction, >1 nm from the 50% Ca isoconcentration surface (Table 5). These spectra integrate over the entire material and do not take into account the spatial variability evident in the proxygram. However, the increase in counting statistics inherent in these whole-region spectra offer higher confidence in the calculated Mg and Na content of the two materials, compared to the proxygram data.

The $\frac{[Ca]_O}{[Ca]_C}$ is calculated using volume-normalized counts, to account for the different volumes of materials captured in the APT volume. The R values are calculated from the raw counts of Na, Mg and Ca from each material, because the count ratio is independent of material volume. The parameters used in equation 19 can be seen in Table 6.

Table 5: Calcite and organic compositions measured by APT.

	Calcite (22797 nm^3)		Organic (17097 nm^3)	
Element	Counts	Counts nm^{-3}	Counts	Counts nm^{-3}
Ca	125971.1 ± 43.8	5.52578 ± 0.00192	1544.0 ± 17.7	0.09031 ± 0.00104
Mg	557.0 ± 95.5	0.02443 ± 0.00419	517.7 ± 54.1	0.03028 ± 0.00316
Na	940.3 ± 16.5	0.04125 ± 0.00072	405.4 ± 4.9	0.02371 ± 0.00029

Table 6: Parameters used in model (Eqn 19).

Parameter	Value
$\frac{[Ca]_O}{[Ca]_C}$	0.01634 ± 0.00019
${}^{Mg}R_C$	0.00442 ± 0.00076
${}^{Mg}R_O$	0.33530 ± 0.03525
${}^{Na}R_C$	0.00746 ± 0.00013
${}^{Na}R_O$	0.26256 ± 0.00437

4.4 Assumptions

This model relies on two key assumptions:

1. **The ionization efficiencies of Ca, Mg and Na are similar between calcite and organic materials in APT**. This is likely given the dependence of field evaporation on bond strength, and the tendency for these species to be ionically bound in both calcite and organic, which are both predominantly made of C, O or (in the case of organic) other low-Z elements.

2. **The entire organic layer has the same average composition as the ~10 nm layer measured by APT**. This untestable, but given the absence of any other data we have used this simple assumption.

4.5 Interpreting Model Output

The immediate output of the model is the counts M / count Ca signal associated with scanning the simulated beam across the simulated organic layer. This cannot be directly compared in terms of 'composition' because the APT data cannot be calibrated, or in terms of 'raw counts' because the ionization mechanisms of the techniques are fundamentally different. Furthermore, the efficiencies of Ca, Na and Mg are likely to be different in APT and ToF-SIMS. To overcome this, we consider the results of the simulation in terms of 'increase above background', where the organic-derived Mg/Ca or Na/Ca signal is normalized to the adjacent calcite region. This accounts for differences in the ionization efficiency between Ca, Na and Mg in APT, as long as the ionization efficiencies of each element remains relatively constant between the organic and calcite regions of the APT specimen.

A further consideration in interpreting the model output is the double Mg maxima either side of the POS position observed in our specimen (Fig 12A). These maxima are most likely

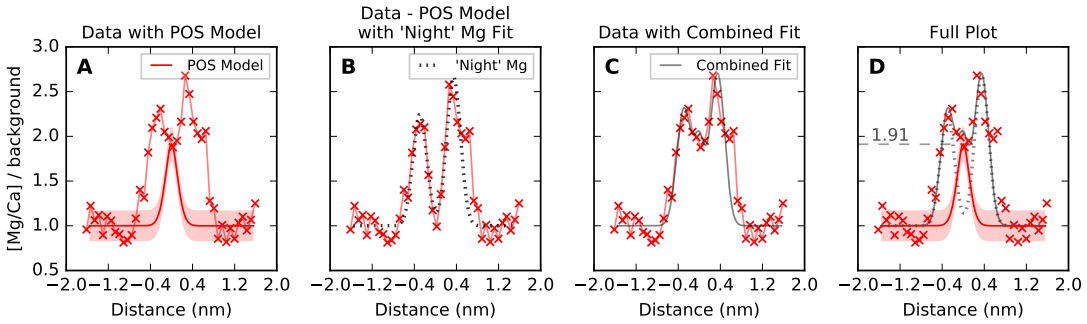


Figure 12: The components of the fit to the POS-associated Mg maximum. (A) Shows the ToF-SIMS profile data compared to the modelled POS signal, which under-estimates the data. (B) Shows the difference between the ToF-SIMS profile and the POS model, which is described well by two 312 nm Gaussian peaks. This is consistent with these peaks being caused by two thin layers of higher-Mg calcite approximately 300 nm away from the POS. The contribution of this Mg signal to the POS is negligible. (C) Shows the combination of the POS model and the adjacent Gaussians compared to the data. (D) Shows the complete fit, including both components.

associated with two layers of high-Mg 'night' calcite deposited on either side of the POS shortly after the start of calcification (~ 300 nm away from the POS). The distance between these peaks is $\sim$2 times greater than the minimum resolvable distance of ToF-SIMS, so they can be treated as independent Mg enrichment signals, which are not associated with the POS. Because the distance of these peaks is ~ 1 beam FWHM away from the POS location, their contribution to POS counts should be negligible, and the POS signal should have little effect on these peaks. Therefore, these peaks may be included alongside the modelled POS signal to quantitatively explain the POS-associated Mg maxima in this specimen (Fig 12B). To achieve this, two Gaussians may be added to the modelled POS Mg signal, described by:

$$y = G(x, amplitude, center, fwhm) \tag{20}$$
$$= G(x, h, 0 - d, fwhm) + G(x, h \times dh, 0 + d \times dd, fwhm) \tag{21}$$

Where d and h are the distance from the POS and the peak height, and dd and dh are factors that allow the outer Mg maxima to be further from the POS and more intense, as expected based on the assymmetry of foraminiferal calcification. These parameters are fit to the data after subtraction of the modelled POS signal (Fig 12B). The double peak background was best fit by $d = 301.8 \pm 19.2$, $dd = 1.34 \pm 0.15$, $h = 1.26 \pm 0.12$, $dh = 1.16 \pm 0.09$. The multiplicative factors dd and dh are consistent with the magnitude of calification assymmetry in *O. universa*.

Within this framework, we are able to reproduce both the POS-associated Mg/Ca and Na/Ca maxima observed by ToF-SIMS using a simulated POS width of 130 nm (Fig 12 C & D). Because our model is able to reproduce *both* signals with a single parameterization, we may accept our hypothesis. This tells us that the POS-associated signal observed in ToF-SIMS can be caused by a 130 nm organic layer of similar composition to the 10 nm organic layer observed by APT. Alongside the qualitative similarity between APT and ToF-SIMS results, the sampling location of the APT specimen, and in context of what is known about organic layers in foraminiferal calcite, this offers compelling evidence that the organic

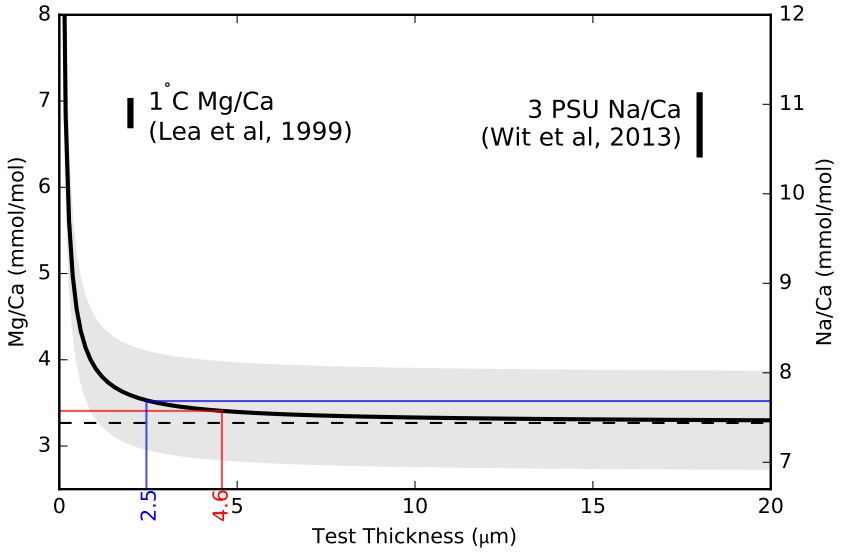


Figure 13: The effect of the POS on bulk foraminifera composition as a function of shell thickness. Red and blue lines indicate the shell thickness below which there is a greater than 0.5 C increase in Mg/Ca derived temperature, or a greater than 1 PSU increase in Na/Ca derived salinity, respectively.

measured by the APT is the POS, and that the POS is driving the simultaneous Na/Ca and Mg/Ca enrichment observed in ToF-SIMS.

Our hypothesis is tested by the ability of our model to simultaneously reproduce both the Na/Ca and Mg/Ca enrichments from a single parameterization. The POS width used in the parameterization does not influence the test of the hypothesis, and offers independent information on the amount of organic material required to cause the observed ToF-SIMS signal. If the entire organic layer has the same average composition as the organic layer measured by APT, this is consistent with a POS width of ∼130 nm. This is in line with previous estimates of POS width ($\sim$100 nm).

4.6 Extension to Bulk Geochemistry

The end member mixing model (Section 4.3.2) can be extended to explore the contribution of organic layers to the overall skeletal chemistry of *O. universa*. This is achieved by using a simple geometric model of *O. universa* calcification to determine the p_O and p_C terms in Equation 18.

The *O. universa* chamber is described as two concentric spheres, a set distance (t) apart, containing a POS of known thickness (t_{POS}). At the start of calcification $t = t_{POS} = 130nm$, and as calcite is added to the sphere $t > t_{POS}$. Thus, the volume of the POS is

constant, while the volume of calcite increases cubically with shell thickness:

$$S(r) = \frac{4}{3}\pi r^3 \tag{22}$$

$$V_{POS} = S(r_0 + t_{POS}/2) - S(r_0 - t_{POS}/2) \tag{23}$$

$$r_i = r_0 - t_{POS}/2 - \frac{t}{1+f} \tag{24}$$

$$r_o = r_0 + t_{POS}/2 + \frac{ft}{1+f} \tag{25}$$

$$V_{CAL} = S(r_i) - S(r_i) - V_{POS} \tag{26}$$

Where the volume (V) of each material is determined by the subtraction of two spheres whose radii (r) differ by a set distance. For the POS, the difference between the radii is determined by POS thickness (t_{POS}), which is evenly distributed either side of the initial radius (r_0). For calcite, the radii change with shell thickness (t), with a greater change on the outside of the shell (r_i) than the inside (r_i) to simulate the asymmetry of calcification. The asymmetry is described by a factor (f) that is the rate of external thickening divided by the rate of internal thickening. Thus, p_O and p_C are then calculated as:

$$p_O = \frac{V_{POS}}{V_{POS} + V_{CAL}} \tag{27}$$

$$p_C = \frac{V_{CAL}}{V_{POS} + V_{CAL}} \tag{28}$$

Using typical size and geometry parameters for an *O. universa* sphere ($r_0 = 200$ μm, $f = 3.0$) and the POS width estimated from the ToF-SIMS scale-up model ($t_{POS} = 130$ nm), this can be used to predict the contribution of the POS to bulk foraminiferal composition, as a function of shell thickness (Fig 13).

4.7 Convolution Calculation

The p_O and p_C terms (Eqn 18) for the simulated ToF-SIMS analysis beam are calculated numerically in python.

```python
# python 3.5
import numpy as np

def gauss(x, *p):
    """
    Gaussian function.
    Parameters:
        x: array-like
        *p: parameters unpacked to A, mu, sigma
            A: area
            mu: centre
            sigma: width

    Note:
        FWHM = 2 * sigma * np.sqrt(2 * np.log(2))
    """
    A, mu, sigma = p
    return A * np.exp(-(x - mu)**2 / (2 * sigma**2))

# define parameters
o_width = 130  # width of organic layer
beam_fwhm = 312  # width of gaussian
```

```
window = 3  # width (n * fwhm) of the convolution

# generate distance (x) scale for convolution
ds = np.linspace(-n * beam_fwhm, n * beam_fwhm, 2 * n * beam_fwhm)

# calculate sigma for input into gaussian function
beam_sigma = beam_fwhm / (2 * np.sqrt(2 * np.log(2)))

# generate gaussian beam function
beam = gauss(ds, 1, 0, beam_sigma)
# normalise gaussian so integral = 1
nbeam = beam / beam.sum()

# generate organic layer function (tophat)
org = np.zeros(o_width * 3)
org[o_width : 2 * o_width] = 1.

# convolve beam with organic layer
c = np.convolve(nbeam, org, mode='same')

# The convolution of the gaussian whose integral is equal to 1
# and the tophat function ('c') is the proportion of the total
# illuminated area that contains organic material as the beam
# scans across the organic layer.
```

5 Data Processing Code

5.1 Data

Data Availablility All data required to reproduce our analysis and raw APT data files are available as an additional supplement to to main manuscript.

Data Format Data files are in .csv format. Proxygram data contain distance and ion count columns, mass spectral data contain Da (mass/charge ratio) and count data.

5.2 Python Environment

All analyses are performed using the Jupyter Notebook environment. Non-base python packages are:

```
ipython==4.2.0
lmfit==0.9.3
matplotlib==1.5.1
numpy==1.10.2
pandas==0.18.1
scipy==0.16.1
uncertainties==2.4.6.1
```

To initialise the analysis environment, import the following:

```
import re
import time
import numpy as np
import pandas as pd
pd.options.mode.chained_assignment = None
import uncertainties.unumpy as un
import matplotlib.pyplot as plt

from scipy.optimize import curve_fit

%matplotlib inline

# For readability, helper functions for these analyses
# are stored in 'APT_helper_fns.py', which should be
# placed in the working directory
import APT_helper_fns as apt

# uncertainties helper functions
def nom(a):
    try:
        return un.nominal_values(a)
    except:
        return a

def err(a):
    try:
        return un.std_devs(a)
    except:
        return a
```

5.3 Compositional Analysis

5.3.1 Load Spectra and Peak Ranges

```
# load mass spectra from calcite and organic
# calcite volume
mscal = pd.read_csv('data/R31_04588-v02-calcite_v22797.csv')
mscal.columns = ['da', 'uc', 'c', 'bkg']
# organic volume
msorg = pd.read_csv('data/R31_04588-v02-organic_v17097.csv')
msorg.columns = ['da', 'uc', 'c', 'bkg']

# volume of two regions (nm3), noted in end of file name
calvol = 22797.
orgvol = 17097.

# load range file
ranges = apt.read_rrng('data/interface_bkg_corr_checked.RRNG')

# isolate pre-, pos- and peak- windows from range file
pre = ranges.iloc[["pre" in i for i in ranges.name.tolist()],:]
post = ranges.iloc[["post" in i for i in ranges.name.tolist()],:]
pk = ranges.iloc[["_" in i for i in ranges.name.tolist()],:]

pk.index = np.arange(len(pre))
pre.index = np.arange(len(pre))
post.index = np.arange(len(pre))

post.loc[:,'n'] = post.index.values
pre.loc[:,'n'] = pre.index.values
pk.loc[:,'n'] = pk.index.values

# combine 'pre' and 'post' regions into single backgrounds file
bkgs = pd.merge(pre, post, on='n', suffixes=('_pre', '_post'))
```

5.3.2 Identify peak and background counts

Counts within the peak and background windows are summed, and the background error is determined.

Background error is based on the standard deviation of the background region, and summed alongside the background region:

$$err_{bkg} = \sqrt{\sum err_i^2 + err_{i+1}^2 ... err_n^2} \tag{29}$$

where n is the number of background points. The background, and its associated error, is then corrected for differences in the number of points captured by each region, and subtracted from the summed peak counts, along with the associated error:

$$c_{corr} = c_{pk} - \frac{n_{pk}}{n_{bg}} c_{bg} \tag{30}$$

where c denotes total counts, and n denotes the number of histogram bins captured by the peak window.

```
# Calcite calculations
pk.loc[:,'cal_pk_n'] = pk.apply(lambda x: apt.npoints(x, mscal), axis=1)
pk.loc[:,'cal_bkg_n'] = (pre.apply(lambda x: apt.npoints(x, mscal), axis=1) +
                         post.apply(lambda x: apt.npoints(x, mscal), axis=1))

pk.loc[:,'cal_uc'] = pk.apply(lambda x: apt.sig_sum(x, mscal), axis=1)
pk.loc[:,'cal_bkg'] = bkgs.apply(lambda x: apt.bkg_sum(x, mscal), axis=1)
pk.loc[:,'cal_bkg_std'] = bkgs.apply(lambda x: apt.bkg_std(x, mscal), axis=1)

# package the standard deviations
pk.loc[:,'cal_ubkg'] = un.uarray(pk.loc[:,'cal_bkg'], pk.loc[:,'cal_bkg_std'])
# calculate background multiplication factor to account for uneven numbers of measured points
pk.loc[:,'cal_bkg_f'] = pk.loc[:,'cal_pk_n'] / pk.loc[:,'cal_bkg_n']
# background subtraction
pk.loc[:,'cal_c'] = pk.loc[:,'cal_uc'] - pk.loc[:,'cal_ubkg'] * pk.loc[:,'cal_bkg_f']

# Organic calculations
pk.loc[:,'org_pk_n'] = pk.apply(lambda x: apt.npoints(x, msorg), axis=1)
pk.loc[:,'org_bkg_n'] = (pre.apply(lambda x: apt.npoints(x, msorg), axis=1) +
                         post.apply(lambda x: apt.npoints(x, msorg), axis=1))
```

```
pk.loc[:,'org_uc'] = pk.apply(lambda x: apt.sig_sum(x, msorg), axis=1)
pk.loc[:,'org_bkg'] = bkgs.apply(lambda x: apt.bkg_sum(x, msorg), axis=1)
pk.loc[:,'org_bkg_std'] = bkgs.apply(lambda x: apt.bkg_std(x, msorg), axis=1)

# package the standard deviations
pk.loc[:,'org_ubkg'] = un.uarray(pk.loc[:,'org_bkg'], pk.loc[:,'org_bkg_std'])
# calculate background multiplication factor to account for uneven numbers of measured points
pk.loc[:,'org_bkg_f'] = pk.loc[:,'org_pk_n'] / pk.loc[:,'org_bkg_n']
# background subtraction
pk.loc[:,'org_c'] = pk.loc[:,'org_uc'] - pk.loc[:,'org_ubkg'] * pk.loc[:,'org_bkg_f']
```

5.3.3 Na and Mg Interference Corrections

Na is calculated by subtracting predicted $^{46}Ca^{2+}$ counts from the 23 Da peak, based on $^{40}Ca^{2+}$ counts.

Mg is calculated by two approaches:

1. Subtraction of predicted $^{46}Ca^{2+}$ counts from the 24 Da peak
2. Prediction of total Mg from the minor ^{25}Mg and ^{26}Mg peaks

The following outlines the evaluation of these different Mg selection methods, and the selection of the Mg correction method outlined in Figure 7.

Note: The natural abundances of Mg isotopes are .79, .10 and .11 for 24, 25 and 26 Mg, respecitvely. Total Mg counts may therefore be calculated by either:

$$^{T}Mg_c = {}^{24}Mg_c\frac{1}{.79} = {}^{25}Mg_c\frac{1}{.10} = {}^{26}Mg_c\frac{1}{.11} \tag{31}$$

Ca subtraction from Mg^+ and Na^+ peaks. The counts of less abundant Ca isotopes are predicted from $^{40}Ca^{2+}$.

```
# Specify Ca isotope abundances:
caiso = {'40Ca': 96.941000000000003,
         '42Ca': 0.64700000000000002,
         '43Ca': 0.13500000000000001,
         '44Ca': 2.0859999999999999,
         '46Ca': 0.0040000000000000001,
         '48Ca': 0.187}
# isolate Ca 2+ peaks
opks = pk.loc[(pk.centre >= 19.5) & (pk.centre <= 24.5),
              ['centre', 'cal_c', 'org_c']].copy()

opks
```

	centre	cal_c	org_c
15	19.9657	76146.1+/-4.7	680.7+/-2.9
16	20.9783	531.0+/-21.2	20.0+/-2.7
17	21.9745	1600.0+/-17.7	6.0+/-4.0
18	22.9785	943.4+/-16.5	405.5+/-4.9
19	23.9763	214.0+/-10.2	197.0+/-3.9

```
# predict Ca peak intensity from 40Ca counts
opks.loc[:,'ca_abund'] = [caiso[p] for p in ['40Ca','42Ca','44Ca','46Ca','48Ca']]
opks.loc[:,'cal_pred'] = (opks.loc[:,'ca_abund'] *
                          opks.loc[15, 'cal_c'] /
                          opks.loc[:,'ca_abund'].max())
opks.loc[:,'org_pred'] = (opks.loc[:,'ca_abund'] *
                          opks.loc[15, 'org_c'] /
                          opks.loc[:,'ca_abund'].max())

# plot code ommitted
```

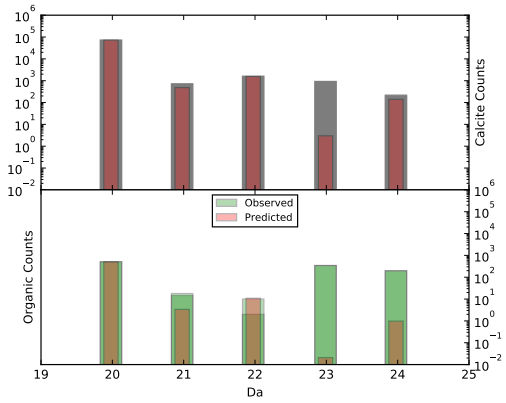


Figure 14: The measured (green/grey) counts for ToF-peaks between 19.5-24.5 Da from organic and calcite spectra, compared to the predicted counts (red) based on the $^{40}Ca^{2+}$ peak. For measured counts, the difference between the two bars (e.g. at 22 Da in organic) represents the standard deviation of the counts.

```
# subtract predicted from observed counts
opks.loc[:,'org_sub'] = opks.loc[:,'org_c'] - opks.loc[:,'org_pred']
opks.loc[:,'cal_sub'] = opks.loc[:,'cal_c'] - opks.loc[:,'cal_pred']

# isolate Ca subtracted Na and Mg counts
cacorr = opks.loc[[18, 19],['centre','cal_c', 'cal_pred', 'cal_sub', 'org_c', 'org_pred', 'org_sub']]

# correct for total Mg counts
cacorr.loc[19,'cal_tMg'] = cacorr.loc[19, 'cal_sub'] * 1 / .79
cacorr.loc[19,'org_tMg'] = cacorr.loc[19,'org_sub'] * 1 / .79

print('Ca Subtracted\n' +
      '  Na counts\n    Calcite: {:.1f}\n    Organic: {:.1f}'.format(*cacorr.loc[18,['cal_sub', 'org_sub']])
     +
      '\n  Mg counts\n    Calcite: {:.1f}\n    Organic: {:.1f}'.format(*cacorr.loc[19,['cal_tMg', 'org_tMg'
     ]]))
```

```
Ca Subtracted
  Na counts
    Calcite: 940.3+/-16.5
    Organic: 405.4+/-4.9
  Mg counts
    Calcite: 85.0+/-13.0
    Organic: 247.7+/-5.0
```

Mg^+ counts from $^{25}Mg^+$ and $^{25}Mg^+$ peaks. Peaks at 24, 25 and 26 Da are evident in both materials, corresponding to monovalent Mg ions. The peak at 24 suffers from a significant interference from $^{48}Ca^{2+}$, and cannot be directly used to measure $^{24}Mg^+$. There may be polyatomic interferences at 25 and 26 Da, but these are impossible to constrain. A conservative estimate of Mg counts is therefore obtained from the lesser of the 25 and 26 peaks.

```
# Specify Mg isotope abundances:
mgiso = {'24Mg': 0.7899,
         '25Mg': 0.1000,
         '26Mg': 0.1101}

# isolate peaks of interest (24 & 25 Da)
```

```
mg1 = pk.loc[(pk.centre >= 24.5) & (pk.centre <= 26.5),
             ['centre', 'cal_c', 'org_c']].copy()

# calculate total Mg counts based on their intensity
mg1.loc[:,'org_tMg'] = mg1.loc[:,'org_c'] * [1/mgiso['25Mg'], 1 / mgiso['26Mg']]
mg1.loc[:,'cal_tMg'] = mg1.loc[:,'cal_c'] * [1/mgiso['25Mg'], 1 / mgiso['26Mg']]

mg1.loc[:,['centre','cal_c', 'cal_tMg','org_c', 'org_tMg']]

print('25Mg+ count predicted\n   Calcite: {:.1f}\n   Organic: {:.1f}'.format(*mg1.loc[45,['cal_tMg', 'org_tMg
    ']]) +
      '\n26Mg+ count predicted\n   Calcite: {:.1f}\n   Organic: {:.1f}'.format(*mg1.loc[46,['cal_tMg', '
    org_tMg']]))
```

```
25Mg+ count predicted
   Calcite: 300.0+/-82.7
   Organic: 460.0+/-33.1
26Mg+ count predicted
   Calcite: 154.5+/-52.7
   Organic: 418.2+/-27.5
```

Mg^{2+} counts from $^{25}Mg^{2+}$ peak Peaks at 12 and 13 Da are clear in both materials, and calcite has a clear 12.5 Da peak. The 12 and 13 peaks suffer from interference by $^{12}C^{+}$ and $^{12}CH^{+}$, and cannot be reliably assigned to $^{24}Mg^{2+}$ or $^{26}Mg^{2+}$. Total Mg counts are therefore estimated using the 12.5 Da peak, which is attributable to $^{25}Mg^{2+}$.

```
# isolate 12.5 peak
mg2 = pk.loc[(pk.centre >= 12.3) & (pk.centre <= 12.7),
             ['centre', 'cal_c', 'org_c']].copy()

# calculate total Mg counts
mg2.loc[:,'cal_tMg'] = mg2.loc[:,'cal_c'] * 1 / mgiso['25Mg']
mg2.loc[:,'org_tMg'] = mg2.loc[:,'org_c'] * 1 / mgiso['25Mg']

print('25Mg++ count predicted\n   Calcite: {:.1f}\n   Organic: {:.1f}'.format(*mg2.loc[44,['cal_tMg', '
    org_tMg']]))
```

```
25Mg++ count predicted
   Calcite: 257.0+/-47.7
   Organic: 57.7+/-42.7
```

Calculation of Mg Counts A calculated Mg count estimate must include both single and double charged ions. Double charged ions may only be estimated by prediction from the 12.5 peak. Singly charged ions may be estimated by Ca subtraction, or prediction from the 25 or 26 Da peak.

```
# compile all Mg estimates
mgpred = mg1.append(cacorr.loc[19,['centre', 'cal_tMg', 'org_tMg']])

# bring in Mg2+
mgpred.loc[:,'cal_tMg2'] = mg2.loc[:,'cal_tMg'].values
mgpred.loc[:,'org_tMg2'] = mg2.loc[:,'org_tMg'].values

# calculate sum
mgpred.loc[:,'cal_sMg'] = mgpred.loc[:,'cal_tMg'] + mg2.loc[:,'cal_tMg'].values
mgpred.loc[:,'org_sMg'] = mgpred.loc[:,'org_tMg'] + mg2.loc[:,'org_tMg'].values

# plot code omitted.
```

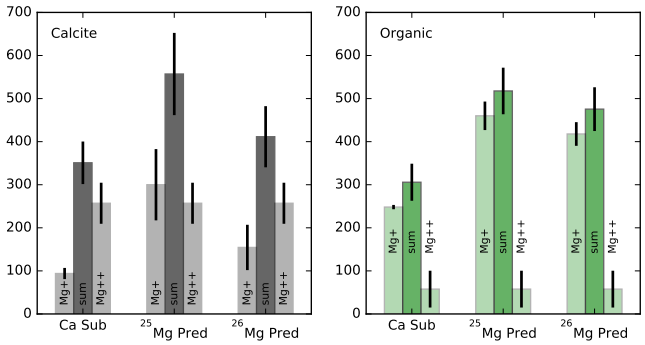


Figure 15: The total Mg counts estimated via three different methods. The central, darker bar is the sum Mg counts, bracketed by the Mg^{+} and Mg^{2+} component peaks. All peaks contain the Mg^{2+} counts predicted from the 12.5 Da peak, and differ in their Mg^{+} component: in 'Ca Sub' the single charge contribution is calculated by subtraction of the interfering $^{48}Ca^{2+}$ counts, and in '^{25}Mg Pred' and '^{26}Mg Pred' they are extrapolated from the $^{25}Mg^{+}$ or $^{26}Mg^{+}$ peaks, respectively.

```
print('Ca Sub\n    Calcite: {:.1f}\n    Organic: {:.1f}\n'.format(*mgpred.loc[19,['cal_sMg', 'org_sMg']]) +
      '25Mg Pred\n    Calcite: {:.1f}\n    Organic: {:.1f}\n'.format(*mgpred.loc[45,['cal_sMg', 'org_sMg']]) +
      '26Mg Pred\n    Calcite: {:.1f}\n    Organic: {:.1f}\n'.format(*mgpred.loc[46,['cal_sMg', 'org_sMg']]))
```

```
Ca Sub
   Calcite: 342.0+/-49.4
   Organic: 305.4+/-43.0
25Mg Pred
   Calcite: 557.0+/-95.5
   Organic: 517.7+/-54.1
26Mg Pred
   Calcite: 411.6+/-71.0
   Organic: 475.8+/-50.8
```

Observations:

- *Ca Sub*: In both materials, the Ca subtracted $^{24}Mg^{+}$ peak gives a lower total Mg count than the other estimates of Mg^{+}. The propagated error on the Ca subtraction is relatively small, but this does not reflect uncertainties in the shape of peaks - e.g. the $^{40}Ca^{2+}$ peak has a large tail, compared to the other Ca^{2+} peaks. This could introduce large uncertainties in this background correction, particularly where the intereference peak is a significant portion of the observed peak (i.e. this matters more for $^{48}Ca^{2+}$, than $^{46}Ca^{2+}$; Fig 14).

- *Mg Pred* peaks: In calcite, *^{25}Mg Pred* and *Ca Pred* estimates are within error of each other. *^{25}Mg Pred* is larger, possibly suggesting a minor polyatomic interference at 26 Da? In organic, *^{25}Mg Pred* and *^{26}Mg Pred* are similar, and both larger than the *Ca sub* estimate. The same Mg^{2+} count estimate is used, so differences are caused by variation in the Mg^{+} estimate.

Choice of Mg Correction. The ^{26}*Mg Pred* is used to estimate Mg counts throughout because:

1. The *Ca sub* method deals with the subtraction of peaks with complex, variable shapes, and non-equivalent selection regions of the peaks may bias the subtraction in unconstrained ways. Because ^{48}Ca makes a significant contribution to the 24 Da peak, this may be important here (as opposed to for Na, where the peak is almost entirely Na counts).

2. In both calcite and organic the 26 Da peak yields the lowest counts of the two minor Mg peaks, and is the most conservative.

3. In both calcite and organic, the ^{26}*Mg Pred* value is within error of another estimate (^{25}*Mg Pred* in organic, *Ca Pred* in calcite). This gives confidence that it is representative of each material, while allowing the use of the same Mg correction method in both materials.

5.3.4 Apply Mg and Na Corrections

```
# Na counts
na = pd.DataFrame({'name': 'R58_Na1',
                   'cal_c': cacorr.loc[18, 'cal_c'],
                   'org_c': cacorr.loc[18, 'org_c'],
                   'centre': 23.},
                   index=[18.1])

# Mg counts
mg = pd.DataFrame({'name': 'R61_Mg1',
                   'cal_c': mgpred.loc[46, 'cal_sMg'],
                   'org_c': mgpred.loc[46, 'org_sMg'],
                   'centre': 24.},
                   index=[19.1])

# Add these to the count dataframe
pk = pd.concat([pk, na, mg])

# remove the Na and Mg counts from the minor Ca peaks
pk.loc[[18, 19],'cal_c'] = cacorr.loc[[18,19], 'cal_c']
pk.loc[[18, 19],'org_c'] = cacorr.loc[[18,19], 'org_c']
```

5.3.5 Calculate Material Compositions

First, complex ions must be decomposed into their constituent elements so that they may be counted correctly.

```
# Find which elements are present
nmsc = [re.sub('R[0-9]+_?|Name|p[0-9]+','',n) for n in pk.name]
ccomps = [apt.composition(n) for n in nmsc]
elcont = apt.composition(''.join(nmsc)).keys()

# sum the counts of each element in each material.
# 1. create an empty dataframe for counts of all these elements
comp = pd.DataFrame(0, index=elcont, columns=['cal','org'])

# 2. populate the dataframe with counts for each element
for e in elcont:
    M = []
    for c in ccomps:
        if e in c.keys():
            M.append(c[e])
        else:
            M.append(0)
    for n in np.arange(len(M)):
        comp.loc[e,'cal'] += pk.cal_c.iloc[n] * M[n]
        comp.loc[e,'org'] += pk.org_c.iloc[n] * M[n]

```

```
# 3. calculate total counts
comp['total'] = comp.loc[:,'cal'] + comp.loc[:,'org']

# 4. calculate percentage of total counts
cpc = comp.copy()
cpc.loc[:,'org'] = cpc.loc[:,'org'].values / np.sum(cpc.loc[:,'org'].values)
cpc.loc[:,'cal'] = cpc.loc[:,'cal'].values / np.sum(cpc.loc[:,'cal'].values)
cpc.loc[:,'total'] = cpc.loc[:,'total'].values / np.sum(cpc.loc[:,'total'].values)

cpc = cpc * 100

# 5. calculate M/Ca ratios
rats = comp / comp.loc['Ca',:]
rats.drop('Ca', inplace=True)

# Na/Ca and Mg/Ca ratios in cts/kcts Ca
rats.loc[['Mg','Na'],:] * 1E3
```

	cal	org	total
Mg	4.4+/-0.7	313.0+/-28.9	9.5+/-0.8
Na	7.43+/-0.13	189.0+/-2.4	10.45+/-0.13

5.4 Data Quality Checks

5.4.1 On-Peak Counts

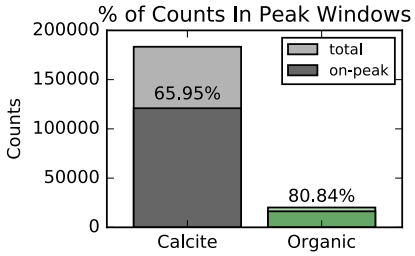


Figure 16: This offers an estimate of the rough 'signal/background' ratio for the materials. A lower proportion of counts in calcite are 'on-peak', because of the large $^{40}Ca^{2}+$ tail, which gives a high background.

5.4.2 Expected vs. Observed Efficiency

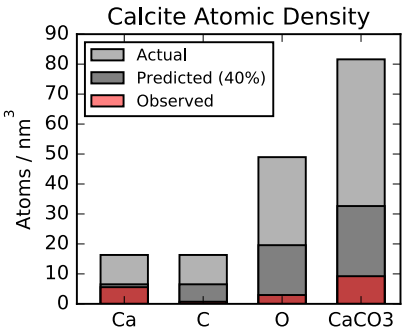


Figure 17: This plot shows the actual atoms/nm^3 for Ca, C and O in calcite, the predicted density based on the nominal 40% efficiency of the LEAP 4000, and the measured atomic density of calcite by APT. Ca recovery is close to the nominal instrument efficiency, while C and O recovery is much lower. Note that the ratio of O:C is ~ 3: C and O loss is in line with CO_3 stoichiometry, perhaps implying the mid-flight decomposition of CO_3 groups into neutral products?

5.5 Proxygram Processing

```
# import raw proxygram data: raw count data from each of the ranges defined in the *.rrng* file above, for
      0.01nm bins perpednicular to the organic/mineral interface (50% Ca isosurface).
pg = pd.read_csv('data/R31_04588-v02-artefactGone_interface40Ca32_full-bkg-proxygram-Mg25_12nm_0.01
      bin_ioncounts.csv')

# separate pre, post and pk proxygrams
nms = pg.columns.tolist()[1:-1]
cpre = ['Distance (nm)'] + [n for n in nms if 'pre' in n]
cpost = ['Distance (nm)'] + [n for n in nms if 'post' in n]
cpk = ['Distance (nm)'] + [n for n in nms if 'post' not in n and 'pre' not in n]

pgpre = pg[cpre]
pgpost = pg[cpost]
pgpk = pg[cpk]

# rename columns so they match
pgpre.columns = [pgpre.columns[0]] + [n.replace('pre','') for n in pgpre.columns[1:].tolist()]
pgpost.columns = [pgpost.columns[0]] + [n.replace('post','') for n in pgpost.columns[1:].tolist()]

# combine background counts
pgbkg = pgpre + pgpost

# make column names the same as peaks
o = pgbkg.columns.tolist()
n = pgpk.columns.tolist()
nc = ['Distance (nm)']
for oi in o:
    for ni in n:
        if oi + "_" in ni:
            nc.append(ni)
pgbkg.columns = nc

# apply background correction
pgc = pgpk.subtract(pgbkg, axis=1)
# replace distance scale
pgc['Distance (nm)'] = pg['Distance (nm)']
# get total count and dstep data back
pgc['Sample Count'] = pg['Sample Count']
# pgc['dstep'] = pg['dstep']
```

5.5.1 Na Peak Overlap Correction

Corrected by subtracting predicted $^{46}Ca^{2+}$ counts from the 23 Da peak, as above.

```
# isolate 20 and 23 Da proxygrams
nac = pgc.loc[:,['Distance (nm)', 'R46_Ca1', 'R58_Ca1']].copy()
# predict Ca46 counts
nac.loc[:,'Ca_pred'] = nac.R46_Ca1 * caiso['46Ca'] / caiso['40Ca']
# calculate Na counts
nac.loc[:,'p1Na1'] = nac.loc[:, 'R58_Ca1'] - nac.loc[:, 'Ca_pred']
# calculate remaining Ca counts
nac.loc[:,'R58_Ca1'] = nac.loc[:, 'Ca_pred']

# transfer predicted Na counts to pgc
pgc.loc[:,'p1Na1'] = np.nan  # create empty column
pgc.loc[:,['R58_Ca1', 'p1Na1']] = nac.loc[:,['R58_Ca1', 'p1Na1']]
```

5.5.2 Calculation of Mg Proxygram

Peak Choice. Because the 12.5 Da peak is not obvious in the mass spectrum of the total organic region (Fig 7), it is likely that this peak will be undetectable in a proxygram, where the total number of ions in each bin is orders of magnitude less than the entire region. Therefore, we investigate the proxygrams for the 'peak' and 'background' regions independently, to evaluate whether or not to include the 12.5 Da peak in proxygram Mg calculation.

```
bkgcheck = apt.rebin(pg, 800, 'Sample Count')

# plot code omitted
```

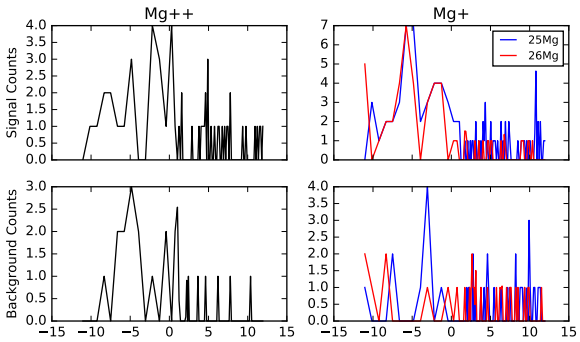


Figure 18: Large regions of the Mg++ proxygram have a greater background than signal. Using the 12.5 Da peak to calculate Mg counts in the proxygram simply adds noise. The 25Mg+ and 26Mg+ signal counts both show the same pattern within the organic layer, giving confidence that they do represent Mg counts. The $^{26}Mg^{+}$ has the lower background, and will be used in the proxygram

Only the $^{26}Mg^{+}$ peak will be used for generating the Mg proxygram.

5.5.3 Mg Extrapolation from $^{26}Mg^{+}$

```
# find names of peaks
nms = pk.loc[(pk.centre < 26.5) & (pk.centre > 25.5), 'name'].values
nms = np.concatenate([['Distance (nm)'], nms])

# isolate 26+ peak
mgc = pgc.loc[:,nms]

# calculate Mg+ counts
mgc.loc[:,'mg1'] = mgc.loc[:,'R142_Mg26'] * 1 / mgiso['26Mg']

# remove old Mg peaks labelled as 'Mg' in pgc
pgc = pgc.loc[:,[c for c in pgc.columns if 'Mg' not in c]]

# propagate changes to main proxygram dataframe
pgc.loc[:,'p2Mg1'] = mgc.loc[:,'mg1']
```

5.5.4 Decompose Complex Ions

```
# Find which elements are present
nms = [nm for nm in pgc.columns.tolist() if nm not in ['Distance (nm)', 'Sample Count', 'dstep']]
nmsc = [re.sub('R[0-9]+_?|Name|p[0-9]+','',n) for n in nms]
ccomps = [apt.composition(n) for n in nmsc]
elcont = apt.composition(''.join(nmsc)).keys()

# Find total counts of all elements
pgc['total_atoms'] = 0

for e in elcont:
    M = []
    for c in ccomps:
        if e in c.keys():
            M.append(c[e])
        else:
            M.append(0)
    # Multiply ion count by stoichiometry to get total
    pgc.loc[:,'t'+e] = 0
    for n in np.arange(len(M)):
        pgc.loc[:,'t'+e] += pgc.loc[:,nms[n]].values * M[n]
    pgc['total_atoms'] += pgc['t'+e]
```

5.5.5 Atomic Yield Differences

Calcite yields 10x more atoms than the organic material (Fig 19).

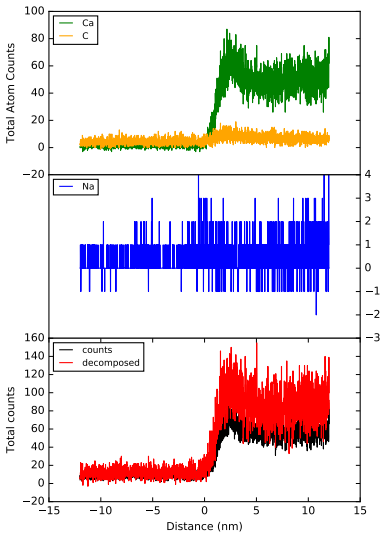


Figure 19: The current proxygram, with a 0.01 nm bin width.

This is most likely a combination of a real atomic density differences, and a disparity in the evaporation field of the two materials, but it is impossible to determine which factor is most important. Organic materials are expected to have a higher evaporation field, as there are a higher proportion of covalent bonds in the material, c.v. calcite. **To assess the best way to deal with this, we must consider the relative atomic density of organic and calcite.**

```
# 1. Calculate the atomic density of calcite.
calden = 2.71 # g/cm3
# NB 1cm3 = 1E21 nm3
calMW = 100 # CaCO3
calmol = calden / calMW # moles of CaCO3 in 1 cm3
avo = 6.0221413E23 # avogadro's number
calatoms = 5 * calmol * avo # atoms in 1cm3 of CaCO3 (*5 for stoichiometry)
calnm3 = calatoms / 1E21 # atoms in 1 nm3 of CaCO3
# calnm3

# 2. Calculate the average atomic density of organic compounds
# import a database of organic compounds
com = pd.read_csv('data/organic_compounds.csv', encoding="ISO-8859-1")
# a list of 307z organic compounds from http://www.kayelaby.npl.co.uk/chemistry/3_3/3_3.html
robbins = pd.read_csv('data/Robbins_compounds.csv') # amino acids idenfied from foraminiferal tests (Robbins
    & Brew, 1990)

# import a periodic table
els = pd.read_pickle('data/elements.pkl') # scraped from elements.com

# define a function for calculating the molecular weight of the organics
def MWcalc(formula):
    comp = re.findall('[A-Z]{1}[a-z]?[0-9]?[0-9]?[0-9]?',formula)
    mass = 0
    tn = 0
    for c in comp:
        n = re.findall('[0-9]+',c)
        if n == []:
            n = 1
        else:
            n = int(n[0])
        esub = els.loc[els.element == re.findall('[A-Za-z]+',c)[0],:]
        mass += np.sum(esub.mass * esub.abundance / 100) * n
        tn += n
    return mass, tn

# clculate the molecular weight of all the organics
tmp = list(zip(*com.formula.map(MWcalc)))
com['mw'] = tmp[0]
```

```
com['n'] = tmp[1]
# calculate the atomic density
com['mol'] = com.density/com.mw
com['at'] = com.mol * avo * com.n
com['nm3'] = com['at'] / 1E21

# repeat for the Robbins & Brew compounds
tmp = list(zip(*robbins.formula.map(MWcalc)))
robbins['mw'] = tmp[0]
robbins['n'] = tmp[1]
# calculate atomic density
robbins['mol'] = robbins.density/robbins.mw
robbins['at'] = robbins.mol * avo * robbins.n
robbins['nm3'] = robbins['at'] / 1E21

# plot code omitted
```

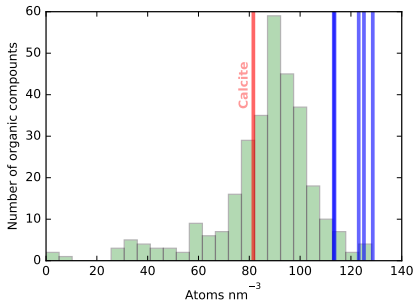


Figure 20: The atomic density of a range of organic compounds (green histogram), amino acids known to be associated with foraminiferal calcite (blue lines) and calcite (red line).

The majority of organic molecules are **more** dense than calcite, particularly the amino acids identified as abundant in foraminiferal organics (Fig 20). Note, however, that these organic density estimates will be derived from crystallized or dried organics, and therefore may not reflect the *in-vivo* density of the material. Based on these data, we would expect an approximately similar atomic density from the organic and calcite materials - not 10x fewer from the organics, as we observe. **Therefore, we will normalise the data to total atom count, to normalise for differences in evaporation field between the materials**.

5.5.6 Rebin Proxygrams

The proxygrams are rebinned to maintain constant total atom count (i.e. normalise for atomic yield differences).

To perform this normalisation, and to produce statistically useful data, it is further useful to re-bin the proxygram data, such that each bin contains the same number of total atoms. This effectively normalises for differences in atomic yield, and provides similar count statistics for data on either side of the interface, allowing their direct comparison, at the cost of spatial resolution.

```
# isolate data of interest
pgs = pgc[['Distance (nm)'] + ['t' + e for e in elcont] + ['total_atoms']]

pga = apt.rebin(pgs, 800, 'total_atoms') # Note: rebinning function is in APT_helper_fns.py

# plot code omitted
```

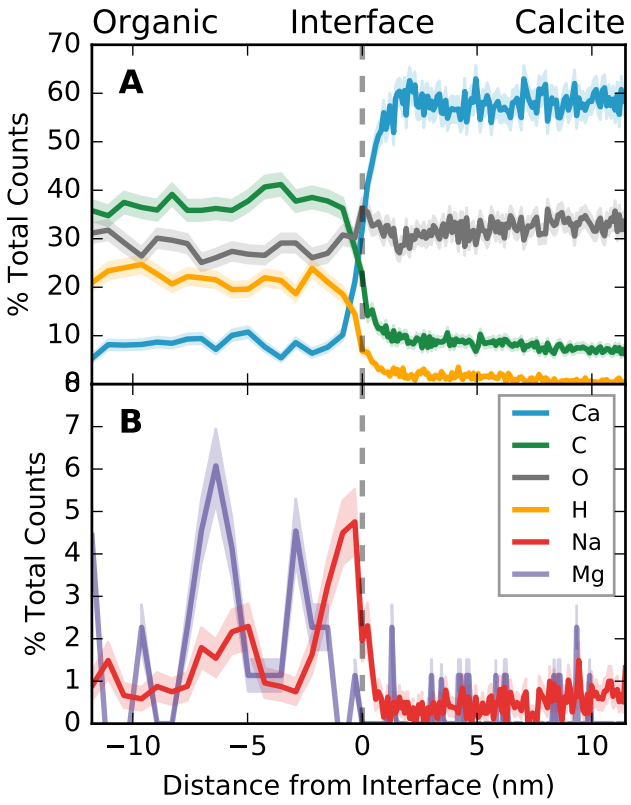


Figure 21: The rebinned proxygrams.

5.6 APT Helper Functions

The following contains a number of functions used numerous times in the above code. If you wish to exectute any of the above, copy the functions below into a text file, name it 'APT_helper_fns.py', and place it in your python working directory.

```
# helper functions for APT analysis

import numpy as np
import pandas as pd
import re


def read_rrng(f):
    """
    Loads an IVAS .rrng file.

    Returns
    -------
    ranges : pandas.DataFrame
    """
    import re

    rf = open(f, 'r').read()

    # extract relevant field using a regex find
    ranges = re.findall('(Range[0-9]+)=+(\d+.\d+) +(\d+.\d+) +Vol:(\d+.\d+) +Name:(R[0-9]+.*) +Color:([A-Z0
      -9]{6})',rf)
    ranges = pd.DataFrame(ranges, columns=['range', 'rmin', 'rmax', 'vol', 'name', 'color'])
    ranges.iloc[:, 1:4] = ranges.iloc[:, 1:4].astype(float)
    # define range centres and width
    ranges['centre'] = (ranges.rmin + ranges.rmax) / 2
    ranges['width'] = ranges.rmax - ranges.rmin

    return ranges


# Functions for getting per-range counts from spectra.
def npoints(pk, df):
    # return number of points measured
    return len(df.loc[(df.da >= pk.rmin) & (df.da <= pk.rmax)])


def sig_sum(pk, df):
    # return sum of points in single interval
    return np.nansum(df.loc[(df.da >= pk.rmin) & (df.da <= pk.rmax), 'c'])


def bkg_sum(pk, df):
    # return sum of points in 'pre' and 'post' background regions
    pre = df.loc[(df.da >= pk.rmin_pre) & (df.da <= pk.rmax_pre)]
    post = df.loc[(df.da >= pk.rmin_post) & (df.da <= pk.rmax_post)]
    return np.nansum(pre.c) + np.nansum(post.c)


def bkg_std(pk, df):
    # return standard deviation of background sum.
    sub = df.loc[(df.da >= pk.rmin_pre) & (df.da <= pk.rmax_pre) |
                 (df.da >= pk.rmin_post) & (df.da <= pk.rmax_post)]
    std = np.std(sub.c)
    n = len(sub)
    return np.sqrt(np.sum([std**2] * n))


# function to identify elemental contents from range name
def composition(comp):
    ions = {}
    els = ['Th', 'Pb', 'Tl', 'Hg', 'Au', 'Pt',
           'Ir', 'Os', 'Re', 'Ta', 'Hf', 'Lu',
           'Yb', 'Tm', 'Er', 'Ho', 'Dy', 'Tb',
           'Gd', 'Eu', 'Sm', 'Nd', 'Pr', 'Ce',
           'La', 'Ba', 'Cs', 'Xe', 'Te', 'Sb',
           'Sn', 'In', 'Cd', 'Ag', 'Pd', 'Rh',
           'Ru', 'Mo', 'Nb', 'Zr', 'Sr', 'Rb',
           'Kr', 'Br', 'Se', 'As', 'Ge', 'Ga',
           'Zn', 'Cu', 'Ni', 'Co', 'Fe', 'Mn',
           'Cr', 'Ti', 'Sc', 'Ca', 'Ar', 'Cl',
           'Si', 'Al', 'Mg', 'Na', 'Ne', 'Be',
           'Li', 'He', 'W', 'I', 'Y', 'V', 'K',
           'S', 'P', 'F', 'O', 'N', 'C', 'B', 'H']
    for e in els:
```

```
        if e in comp:
            ion = re.findall(e+"[0-9]?", comp)[0]
            comp = comp.replace(ion, "")
            if ion[-1].isalpha():
                ion += "1"
            ions[re.findall("[A-Za-z]+", ion)[0]] = \
                int(re.findall("[0-9]+", ion)[0])
    return ions


# Function to re-bin proxygram data
def rebin(pg, sc=800, axis='total_atoms', bins=None):
    tmp = pg.cumsum() # cumulative sum of all counts along distance axis
    tmp['Distance (nm)'] = pg['Distance (nm)'] # reapply original distance scale, to preserve space dimension
    tmp.index = tmp[axis] # set index to cumulative sample count, for interpolation.
    if bins is None:
        bins = np.arange(0, tmp[axis].max(), sc) # if no specific bins are provided, create evenly spaced
     bins 'sc' apart.
#     new = tmp.reindex(bins) # make new index from 0 to max counts, at intervals of [sc]
    new = pd.DataFrame(np.nan, index=bins, columns=tmp.columns) # create a blank dataframe with the new count
      intervals as the index.
    new.loc[0] = 0
    new.loc[0]['Distance (nm)'] = tmp['Distance (nm)'].iloc[0] + (tmp['Distance (nm)'].iloc[0] - tmp['
     Distance (nm)'].iloc[1])

    tmp = pd.concat([tmp,new]) # join the new index to the cumulative sum dataframe
    tmp = tmp.sort_index() # sort the merged dataframe so values are ready to interpolate
    tmp = tmp.interpolate(method='values') # interpolate the values onto the new count axis
    tmp = tmp.loc[new.index] # only select the new ones
    tmp = tmp.groupby(tmp.index).first().dropna() # get rid of any pesky double values (arise if any count %
     cm == 0)
    ds = tmp['Distance (nm)'] # record distance scale of interpolated values, before back-transformation

    tmp = tmp.diff() # diff the data, to get back from cumulative sum to histogram space
    tmp['dstep'] = tmp['Distance (nm)'] # record the diffed distance scale as a step size variable
    tmp['Distance (nm)'] = ds # reinsert interpolated distance scale
    tmp = tmp.iloc[1:] # chuck NaN values
    return tmp
```

References

O. Branson, S. A. T. Redfern, T. Tyliszczak, A. Y. Sadekov, G. Langer, K. Kimoto, and H. Elderfield. The coordination of Mg in foraminiferal calcite. *Earth Plan Sci Lett*, 383: 134–141, 2013.

S. Eggins, P. de Deckker, and J. Marshall. Mg/Ca variation in planktonic foraminifera tests: implications for reconstructing palaeo-seawater temperature and habitat migration. *Earth Plan Sci Lett*, 212(3):291–306, 2003.

S. Eggins, A. Y. Sadekov, and P. de Deckker. Modulation and daily banding of Mg/Ca in Orbulina universa tests by symbiont photosynthesis and respiration: a complication for seawater thermometry? *Earth Plan Sci Lett*, 225(3-4):411–419, 2004.

R. I. Gabitov, A. C. Gagnon, Y. Guan, J. M. Eiler, and J. F. Adkins. Accurate Mg/Ca, Sr/Ca, and Ba/Ca ratio measurements in carbonates by SIMS and NanoSIMS and an assessment of heterogeneity in common calcium carbonate standards. *Chem Geol*, 356:94–108, 2013.

L. M. Gordon and D. Joester. Nanoscale chemical tomography of buried organic–inorganic interfaces in the chiton tooth. *Nature*, 469(7329):194–197, 2011.

B. T. Huber, J. Bijma, and H. J. Spero. Blue water SCUBA collection of planktic foraminifera. In M. A. Lang and C. C. Baldwin, editors, *Methods and Techniques of Underwater Research*. 1996.

D. W. Lea and H. J. Spero. Experimental determination of barium uptake in shells of the planktonic foraminifera Orbulina universa at 22°C. *Geochim Cosmochim Ac*, 56(7): 2673–2680, 1992.

F. Perez and B. E. Granger. IPython: A system for interactive scientific computing. *Computing in Science & Engineering*, 9(3):21–29, 2007.

H. J. Spero. Ultrastructural examination of chamber morphogenesis and biomineralization in the planktonic foraminifer *Orbulina universa*. *Mar Biol*, 99(1):9–20, 1988.

H. J. Spero, S. M. Eggins, A. D. Russell, L. Vetter, M. R. Kilburn, and B. Hönisch. Timing and mechanism for intratest Mg/Ca variability in a living planktic foraminifer. *Earth Plan Sci Lett*, 409(C):32–42, 2015.

A. L. Stancik and E. B. Brauns. A simple asymmetric lineshape for fitting infrared absorption spectra. *Vib Spectrosc*, 47(1):66–69, 2008.