Supplementary information to Fossil and present-day stromatolite ooids contain a meteoritic polymer of glycine and iron.

[1*]Julie E M McGeoch, [2]Anton J Frommelt, [3]Robin L Owen, [4]Gianfelice Cinque, [5]Arthur McClelland, [6]David Lageson and [7]Malcolm W McGeoch

[1]Department of Molecular and Cellular Biology, Harvard University, 52 Oxford St., Cambridge MA 02138, USA & High Energy Physics Div, Smithsonian Astrophysical Observatory Center for Astrophysics Harvard & Smithsonian, 60 Garden St, Cambridge MA 02138, USA.
[2]LRL-CAT, Eli Lilly and Company, Advanced Photon Source, Argonne National Laboratory, 9700 S. Cass Avenue, Lemont, IL, 60439
[3,4]Diamond Light Source, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK.
[5]Center for Nanoscale Systems, Harvard University, 11 Oxford St, LISE G40, Cambridge, MA 02138, USA.
[6]Department of Earth Sciences, 226 Traphagen Hall, P.O. Box 173480 Montana State University, Bozeman, MT 59717.
[7]PLEX Corporation, 275 Martine St., Suite 100, Fall River, MA 02723, USA.
*Corresponding author. E-mail: Julie.mcgeoch@cfa.harvard.edu

Section S1: Current Status of Structural Knowledge on Hemoglycin

The structure of hemoglycin has revealed itself in a diverse set of measurements spanning from 2015 up to the present day. Some aspects, listed below, need further confirmation. To date, hemoglycin has only been available in micro-gram quantities, and even then has not been of guaranteed purity. It crystallizes in many forms that appear to be determined by the presence and nature of additional substances that have accreted within the several types of open lattice that it can form. We summarize here the principal findings on structure. Figure S1.1 depicts the four structural ingredients which combine in different sets to make all crystal and lattice types. There is a "core unit" of mass 1494Da that comprises two antiparallel 11-residue polyglycine strands, with hydroxylation of two residues on each strand, closed out by iron atoms. Prominent in the mass spectrum of meteoritic extracts (McGeoch at al. 2021a, 2021b) is a 1638Da variant in which the core unit has two additional Fe atoms, plus two additional oxygen atoms. Beyond these are silicon connector atoms that exhibit two-fold or fourfold valence bonding to oxygen at the ends of the 1494 and 1638Da "rods". One structural example is the tetrahedral junction shown in Figure S1.1 that is repeated throughout a three-dimensional lattice that hemoglycin forms, with vertices related to each other in diamond 2H symmetry (McGeoch et al., (2021a), and present MS).
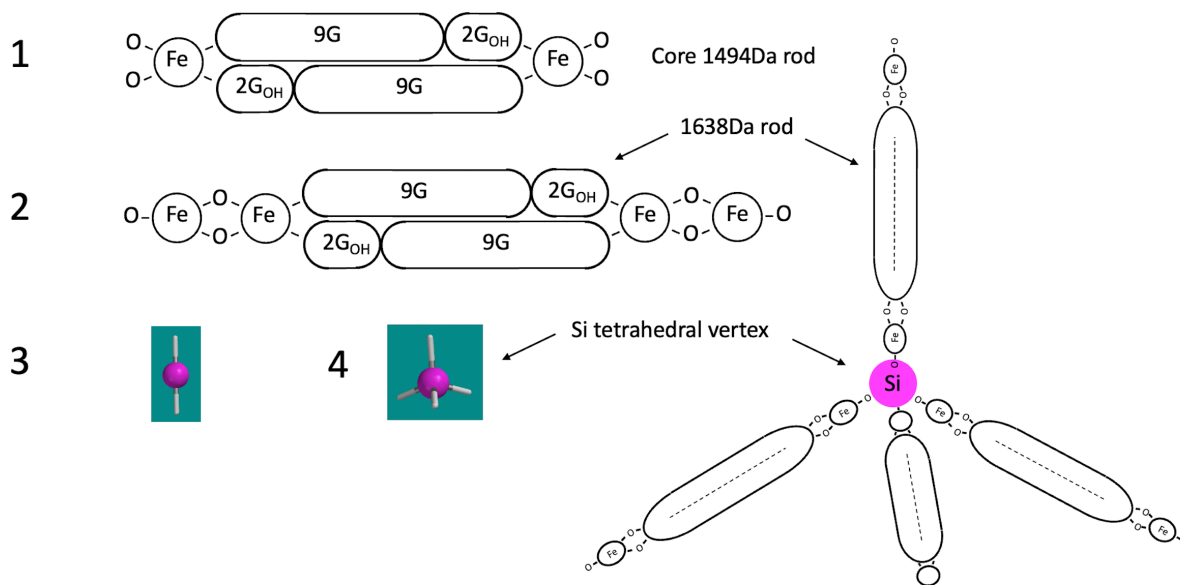
FIGURE S1.1. The four building blocks of hemoglycin extended structures. 1. a 1494Da core unit; 2. a 1638Da rod containing 4 Fe atoms; 3. Si atoms with two bonds; 4. Si atoms with 4 bonds. As an example, the tetrahedral junction of the 3-D hemoglycin lattice is shown.

In Figure S1.2, three 1494Da "core units" of hemoglycin are shown hydrogen-bonded edge to edge, another example of an assemblage based on components in Figure S1.1. The entity shown has hydrogen on each terminal oxygen to serve as a place-holder for oxygen-oxygen, or oxygen-silicon bonds in extended two- and three- dimensional lattices. The structure in Figure S1.2 is based upon equilibrium geometry B3LYP energy minimization within each unit, followed by MMFF energy minimization as a trimeric sheet.

Structural findings are now listed:

1. Polymers containing glycine and hydroxyglycine, ranging from 2 to >20 residues, were first observed in MALDI mass spectrometry within an extract of the Allende meteorite (McGeoch et al., 2015).

2. A 4641Da molecule was observed in MALDI in both Allende and Acfer 086 extracts (McGeoch et al., 2017). The details of this peak complex were different in each case, and there was zero signal at the same mass in a volcano control processed in the same way. Higher multiples up to 4 units of 4641Da were present, indicating that the species was connecting into larger assemblies. Lower mass fragments contained glycine and hydroxyglycine, with high isotope enhancements that confirmed extraterrestrial origin. This was the first observation of an organized entity as opposed to random-length polymers. Later, once the core unit was established, the 4641 entity was identified as a triskelion of 1494Da core units connected by silicon bridges (McGeoch et al., 2021b).
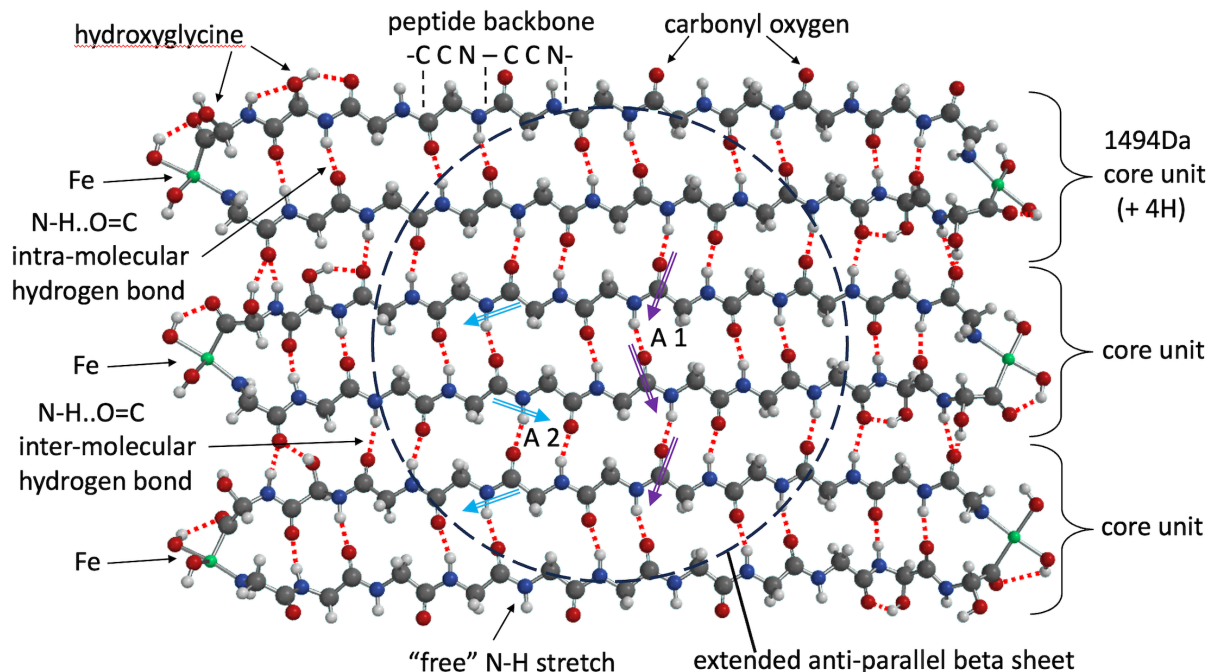
Figure S1.2. Array of 3 hemoglycin "core units" bonded in a sheet by inter-molecular hydrogen bonds. Iron atoms green, hydrogen white, carbon grey, nitrogen blue and oxygen red. Hydrogen bonds red dotted lines. Energy minimized in MMFF. Amide I dipoles (A1) and Amide II dipoles (A2) shown, on average perpendicular to each other.

3. In MALDI (McGeoch et al. 2021b) an extract of Acfer 086 yielded dominant 1566Da and 1638Da peaks (seen at 1567 m/z and 1639 m/z in mass spectra). For the first time iron was detected in the molecule via its ^{54}Fe content that added subordinate peaks at (-1) and (-2) mass units below the principal "mono-isotopic" peak. The amplitude of these subordinate peaks indicated, for example, that the 1566Da entity contained three Fe atoms, while the 1638Da entity contained four Fe atoms. When the principal peaks were subjected to MS/MS analysis, in which they were broken into fragments, the larger m/z peaks yielded lower mass peaks from the original spectrum, indicating the simultaneous presence of a progressive series that comprised a basic or "core" unit with various additions. Crystals of extracts from the Orgueil meteorite that potentially contained a purer and more uniform molecular species (this MS), yielded the 1494Da core unit peak predicted in McGeoch et al (2021b).

4. A FIB/TOF/SIMS experiment (McGeoch et al., 2023b) on Allende and Acfer 086 micron-scale particles revealed that ^{15}N was enriched at a cometary level in components with the same fragmentation pattern as a polymer amide backbone of –CCN-CCN-. This ^{15}N/^{14}N ratio was used, together with total isotope enrichment from mass spectrometry, to determine the D/H ratio of hemoglycin to be 25 times terrestrial (McGeoch et al. 2023b) and, when applied to data in the present MS, as high as 40 times terrestrial.

5. X-ray diffraction of a fiber crystal of Acfer 086 (McGeoch et al., 2021a) revealed a set of diffraction orders with a 4.8 nm fundamental spacing. This matched the spacing of Fe atoms in

computed hemoglycin lattices with silicon junctions (McGeoch et al., 2021b) confirming the length of hemoglycin, which had previously only been calculated using quantum chemistry. Later, a crystal from the Sutter's Mill meteorite yielded an even more extended "ladder" of high diffraction orders (up to N=12) and gave an almost identical Fe-Fe spacing. Such high diffraction orders are seen because Fe atoms act as "markers" that have much higher X-ray scattering than the H,C,N,O constituents of the polymer strands. Furthermore, several Fe atoms can be in close proximity at the vertices of extended polymer lattices, enhancing the effect.

6. Following the prediction of a chiral 480nm absorption in hemoglycin, a Sutter's Mill crystal exhibited the predicted absorption (McGeoch et al., 2022). The absorption was associated with the Fe-glycine bond region, being strong at 480nm only when a) there was a glycine C-terminus next to Fe, and b) that glycine residue carried an 'R' chirality hydroxylation on its alpha carbon atom. The existence of the absorption simultaneously confirmed the presence of Fe in the molecule and the location of at least one hydroxylation adjacent to the Fe atom.

7. In the present MS the structure of the three-dimensional space-filling lattice (McGeoch et al., 2021a) was confirmed in 2-Angstrom X-ray scattering of a present-day ooid. We report 18 higher order diffraction rings that confirm the diamond 2H structure with its tetrahedral lattice junction (Figure S1.1) and again show an inter-vertex spacing of 4.9 nm, consistent with the calculated length of the 1638Da rods, including Si atoms that join the rods.

8. Infrared (IR) absorption, first presented in the present MS, confirms that hemoglycin crystals contain anti-parallel beta sheets, as first suggested in (McGeoch et al. 2017). In both ooid and meteorite the Amide I band was split indicating an <u>extended</u> anti-parallel beta sheet, as shown in Figure S1.2 (absorbance data in the main text Fig.8). This splitting would not be expected in solitary core units. For the first time, via Fourier transform IR spectrometry, the secondary structure of this space polymer has been seen and is found to be in accord with the above identification of an amino acid polymer of glycine.

9. Observation of a visible fluorescence induced by X-rays (McGeoch et al., 2023a) confirmed that the 480nm absorption had a counterpart in molecular fluorescence. The fluorescence of ooids in the X-ray beam comprised the same set of components as for an Orgueil meteorite crystal, although their amplitudes were different, reflecting a different chemical history (present MS).

Further work is needed to:
a) determine the location of hydroxyl groups beyond the proven set on residues with C terminus adjacent to Fe atoms. They could be on the second residue distant from Fe, in view of possible photon-induced reactions associated with the visible absorption, or on the N terminus residue adjacent to Fe.
b) study silicon, which is believed to be present at most types of hemoglycin lattice junction, bonded to oxygen atoms on the terminal Fe and imparting tetrahedral symmetry to the junctions in the diamond 2H three-dimensional lattice. (Figure S1.1). However, the mass of Si has only appeared as a necessary part of the mass fit to the 4641 Da triskelion (McGeoch et al., 2021b). It is desirable to confirm the presence of silicon more generally at the junctions of 2D and 3D lattices because it is apparently necessary, in addition to rod lengths, to provide the 4.9nm inter-vertex

spacing that is repeatedly observed in diffraction from 2D and 3D lattices. The core unit itself as shown in Figure S1.1 is 4.0nm in length between iron atoms.

References for all S sections

McGeoch JEM and McGeoch MW (2015). Polymer amide in the Allende and Murchison meteorites. *Meteorit and Planet Sci.* **50**, 1971-1983.

McGeoch JEM and McGeoch MW (2017). A 4641Da polymer of amino acids in Acfer-086 and Allende meteorites. https://arxiv.org/pdf/1707.09080.pdf

McGeoch JEM and McGeoch MW (2021a). Structural organization of space polymers. *Phys. Fluids* **33**, 067118.

McGeoch MW, Dikler S and McGeoch JEM (2021b). Meteoritic proteins with glycine, iron and lithium, *arXiv*:2102.10700.

McGeoch JEM and McGeoch MW (2022). Chiral 480nm absorption in the hemoglycin space polymer: a possible link to replication. *Scientific Reports* **12**, 16198.
https://doi.org/10.1038/s41598-022-21043-4

McGeoch MW, Owen RL, Jaho S and McGeoch JEM (2023a). Hemoglycin visible fluorescence induced by X rays. *J. Chem. Phys.* **158**, 114901.

McGeoch MW, Samoril T, Zapotok D and McGeoch JEM (2023b). Polymer amide as a carrier of ^{15}N in Allende and Acfer 086 meteorites. Under review at *International Journal of Astrobiology*.

McGeoch JEM and McGeoch MW (2024), Polymer amide as a source of the cosmic 6.2 micron emission and absorption. arXiv:2309.14914 [astro-ph.GA]. Mon. Not. Roy. Astr. Soc. DOI: 10.1093/mnras/stae756.

McGeoch MW and McGeoch JEM. Hexagonal Cladding of a Hemoglycin Vesicle in X-ray Diffraction. Unpublished.

XQuartz 2.8.5, X window system, 2003-2023 X.org Foundation, Inc.

Section S2: Lattice analysis via higher order diffraction

In X-ray analysis of meteorite polymers of amino acids we observe high order diffraction "ladders" in thin samples that are sheet-like (McGeoch et al., 2022), or fibers apparently composed of rolled up sheets (McGeoch et al., 2021a). In addition, there is evidence for a truly three-dimensional lattice of the diamond 2H structure that represents the maximum volume that a 3D lattice of identical "rods" can enclose (McGeoch et al., 2021a). The rods in this case are polymers of glycine comprising anti-parallel glycine chains of 11-residue length closed at each end by an iron atom with hydroxylation of glycine residues adjacent to Fe, termed hemoglycin. X-ray diffraction from such lattices is from two main features: a) Fe atoms in groups at vertices, the groups spaced from

each other by the 49A length of a rod and b) nano-crystals of any substance filling the lattice spaces.

Stromatolites are dominantly formed of ooids, which are small ovoid, predominantly calcium carbonate, grains of typical length a few hundred microns. Following the mass spectrometry finding (this MS) that the 1494Da core unit of hemoglycin was present in both the Orgueil meteorite and fossil stromatolite, an X-ray study was made of modern day ooids found within a recent stromatolite sample from Shark Bay, Australia. Two X-ray wavelengths, 0.979 Angstroms and 2.066 Angstroms, were used on APS beam line 31-1D-D. These lay respectively at higher energy and lower energy than the Fe K edge absorption at 7.1keV (1.74 Angstroms). Differences in X-ray scattering between these two wavelengths would potentially give information on the disposition of Fe atoms within the ooid. The diffraction patterns were indeed markedly different. At the shorter wavelength, a set of rings between 3.86 and 1.61 Angstroms, listed in Table S2.1, indicated that both the calcite and aragonite forms of calcium carbonate were present. No significant rings were observed at "d" spacing larger than 3.85 Angstroms.

Table S2.1. Ooid (Shark Bay) diffraction rings on APS at 0.979 Angstroms indicating the presence of calcite and aragonite, labeled w, m, s, vs for weak, medium, strong, very strong. The run was at nominal room temperature.

Strength	"d" spacing	Aragonite (a) Calcite (c)
w	3.85	c
s	3.41(5)	a
s	3.29	a
vs	3.02	c
s	2.71(5)	a
s	2.49	c, a
w	2.42	-
m	2.38	a
s	2.34(5)	a
m	2.28(5)	c
w	2.20	-
w	2.11(5)	-
w	2.09(5)	c
s	1.98(5)	a
s	1.91(5)	c
vs	1.88(5)	a
s	1.82(5)	-
s	1.74(5)	a
s	1.72(5)	-
w	1.61	c

In contrast, at 2.066 Angstroms, (Figure S2.1) an intense new set of large "d" spacing rings (central) was superimposed upon a weak version of the calcite and aragonite rings (outside), the latter having "d" spacing less than 3.86 Angstroms. The new set of 18 large "*d*" spacing rings

ranged from 4.808 to 11.54 Angstroms, <u>when interpreted as first order</u> (Figure S2.1 and data Table S2.2). We show here that these come from higher order diffraction off Fe groupings at the 5nm - spaced junctions of the diamond 2H open lattice structure already identified in meteoritic work. The present analysis shows that there is a slight distortion to the lattice involving a relative increase of the trigonal axis.
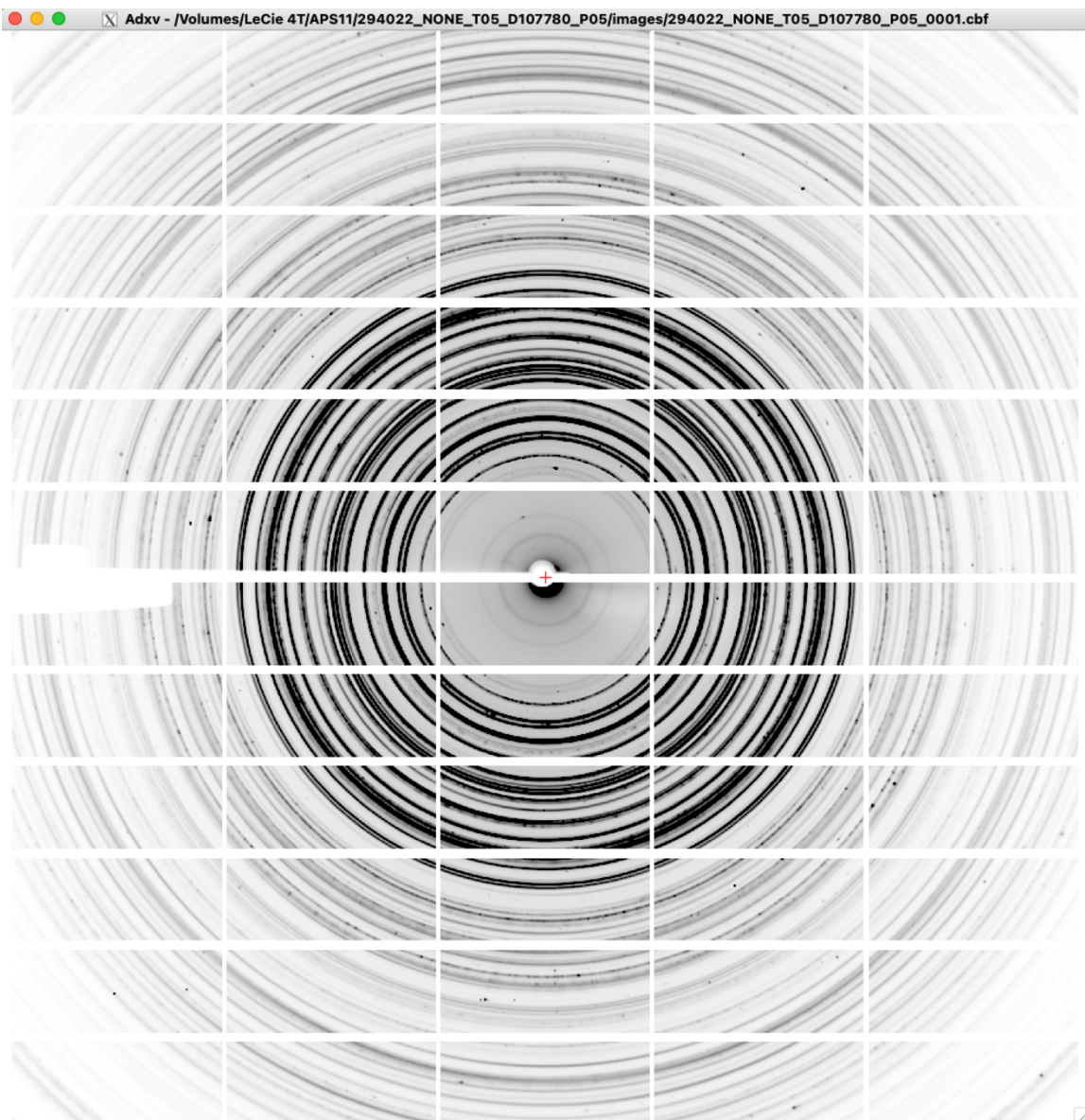
Figure S2.1 X-ray diffraction at 2.066 Angstrom from ooid in present era stromatolite (Shark Bay) showing dark lattice rings between 11.54 and 4.81 Angstroms from the center outward, plus calcium carbonate rings in a faint outer pattern.

Data reduction to lattice "d" spacings from the set of high order rings

The inner rings in Figure S2.1 had the "d" parameters listed in Table S2.2. These were calculated automatically in XQuartz (X.org foundation) on the assumption that they were in first order diffraction according to $2d\sin\Theta = n\lambda$ where n is the diffraction order, Θ is one half of the deflection angle and λ is the wavelength. In this data the wavelength was 2.066 Angstroms.

The left hand column of Table S2.2 contains the set of 18 rings that represented larger "d" spacing than 4 Angstroms. Their relative intensities are shown in the scan of Figure S2.2, part A. These rings did not match either the calcite or aragonite values, but were reminiscent of the ladders of high order diffraction previously seen in hemoglycin lattices. In McGeoch et al (2021a) the ladder contained orders 2 through 5, with a fitted first order lattice parameter of 48.38 ± 0.2 Angstroms. In McGeoch et al., 2022 the ladder contained orders 2 through 12, with a fitted first order parameter of 49.03 ± 0.18 Angstroms. Quick inspection of the present 18 ring set yielded a fit to 49.0 Angstroms in 5th 6th 7th and 9th orders as follows:

5 x 9.823 Angstroms = 49.11;
6 x 8.127 Angstroms = 48.76 ;
7 x 7.022 Angstroms = 49.15
9 x 5.445 Angstroms = 49.00

When several additional sequences also fitted calculated inter-vertex distances in the (h = 49 Angstrom, diamond 2H) 3D lattice, a thorough programmed search was performed in which each of the 18 data values was assessed as a divisor of trial "d" spacings in the range 30 Angstroms to 140 Angstroms, rising in 0.05 Angstrom increments. The number of such divisors found in each 0.05 Angstrom region was plotted in the range 30 to 140 Angstroms in Figure S2.3, after binomial smoothing and raising to the second power to accentuate the longer sequences.

This discovery routine yielded about 10 principal candidates for "d" spacings, the main ones being listed in Table S2.1 together with the relevant diffraction orders and the accuracy of a match (always better than 0.5%). Initially not all of these strong candidates matched calculated vertex-to-vertex spacings for the ideal diamond 2H lattice. A mathematical model was then constructed to handle a non-ideal lattice with a constant rod length "h" and variable deviations from the tetrahedral angle, that is, changes to the angle α where ($90 + \alpha$ = 109.471221 deg. is the tetrahedral angle). Changes to α represent axial stretch or compression of the lattice. To anticipate the results, we are able to fit the principal "d" spacing candidates with an increase of α by 4.5 deg., computed results shown by red bars in Figure S2.3.
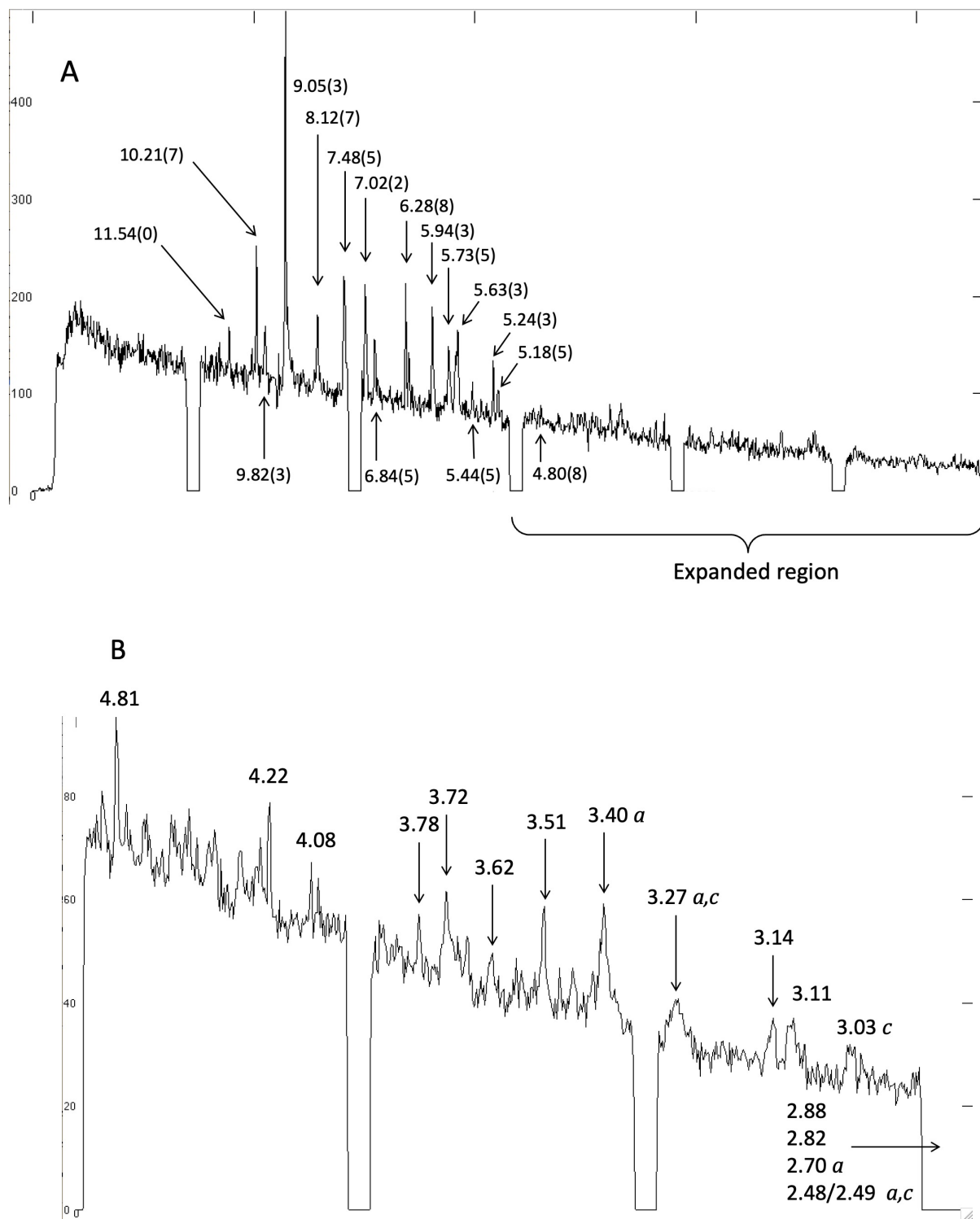
FIGURE S2.2 A: Vertical intensity scan of Figure S2.1. Note that the 7.12(8) peak is obscured by a detector grid. B: Scan of expanded region. Additional peaks 2.88-2.48 Angstroms were measured on the diagonal.
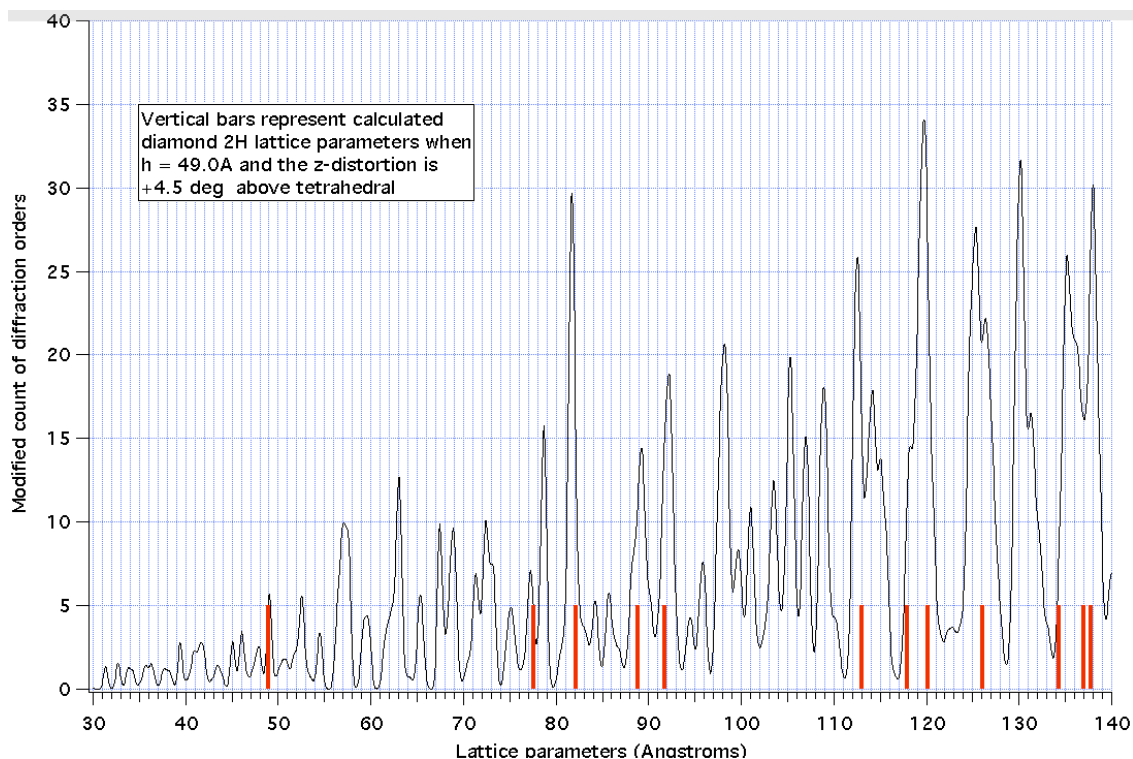
Figure S2.3. (thin black line) The results of a computer search for higher order clusters or loci, represented as the square of the number of "hits" within any 0.05 Angstrom band, with binomial smoothing applied. The red bars are calculated "d" spacings for a slightly distorted diamond 2H lattice, discussed below.

Table S2.2 Ooid diffraction rings in first order (left hand column). Higher order fits (top row) listed as diffraction order in bold with percentage mis-match.

Angstroms	49.0	81.65	92.05	112.75	119.9	126.25
4.80(8)		**17**, 0.1%			**25**, 0.2%	
5.18(5)						
5.24(3)			**17**, 0.1%			**24**, 0.3%
5.44(5)	**9**, 0%	**15**, 0.2%			**22**, 0.1%	
5.63(3)				**20**, 0.1%		
5.73(5)					**21**, 0.4%	**22**, 0.1%
5.94(3)			**15**, 0.1%	**19**, 0.1%		
6.28(8)		**13**, 0.4%		**18**, 0.4%	**19**, 0.3%	**20**, 0.4%
6.57(8)						
6.84(5)			**13**, 0.1%			
7.02(2)	**7**, 0.3%			**16**, 0.3%	**17**, 0.4%	**18**, 0.1%
7.12(8)						
7.48(5)				**15**, 0.4%	**16**, 0.1%	
8.12(7)	**6**, 0.5%	**10**, 0.1%	**11**, 0.3%			
9.05(3)		**9**, 0.2%				**14**, 0.3%
9.82(3)	**5**, 0.2%					
10.21(7)		**8**, 0.1%		**11**, 0.3%		
11.54(0)						

10

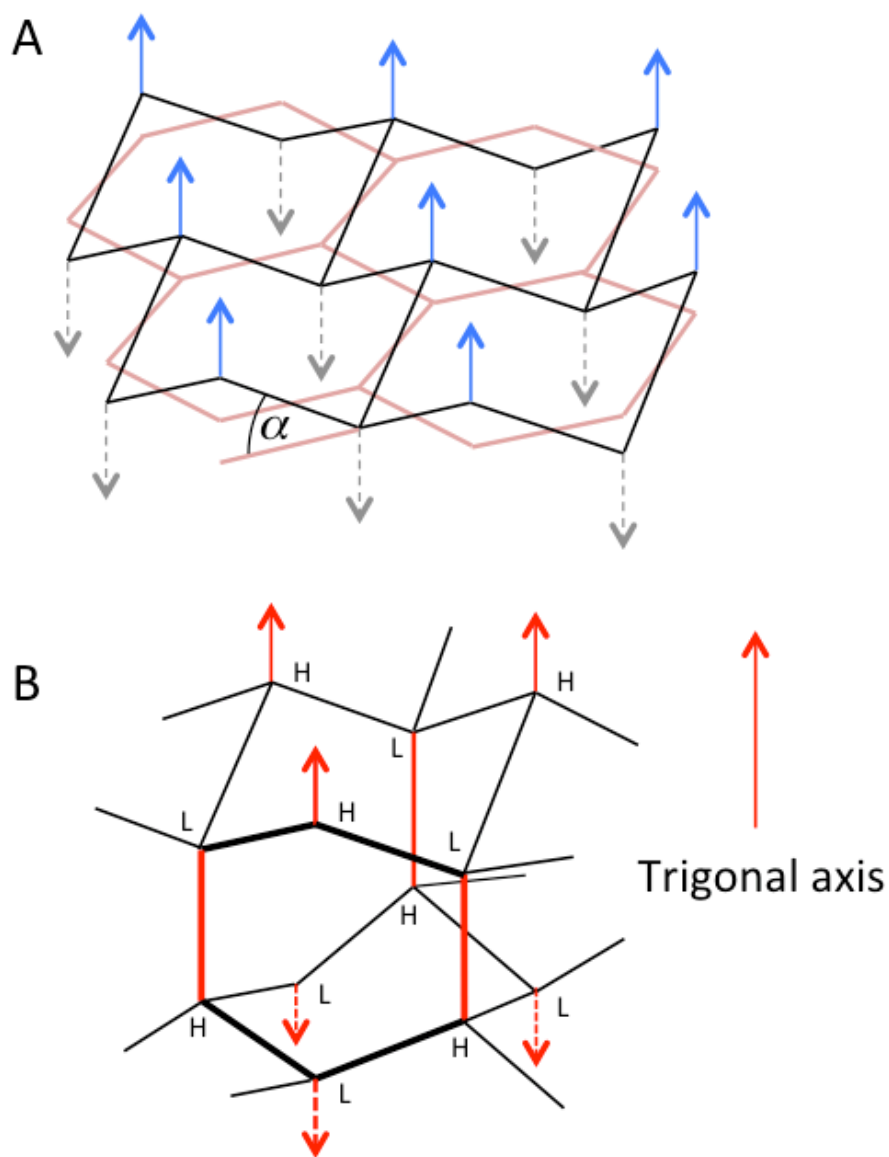**Mathematical construction of the diamond 2H lattice**

**Figure S2.4. Reproduced from (McGeoch et al. 2021a, Figure 10). Part A: a plane lattice of hexagons (beige) is distorted to have alternating vertices above or below the plane, with the quasi-hexagon sides making angle α with the plane. Part B: two layers spaced along the vertical trigonal symmetry axis. High vertices in the lower layer connect to low vertices in the next layer up. Each connecting rod is identical.**
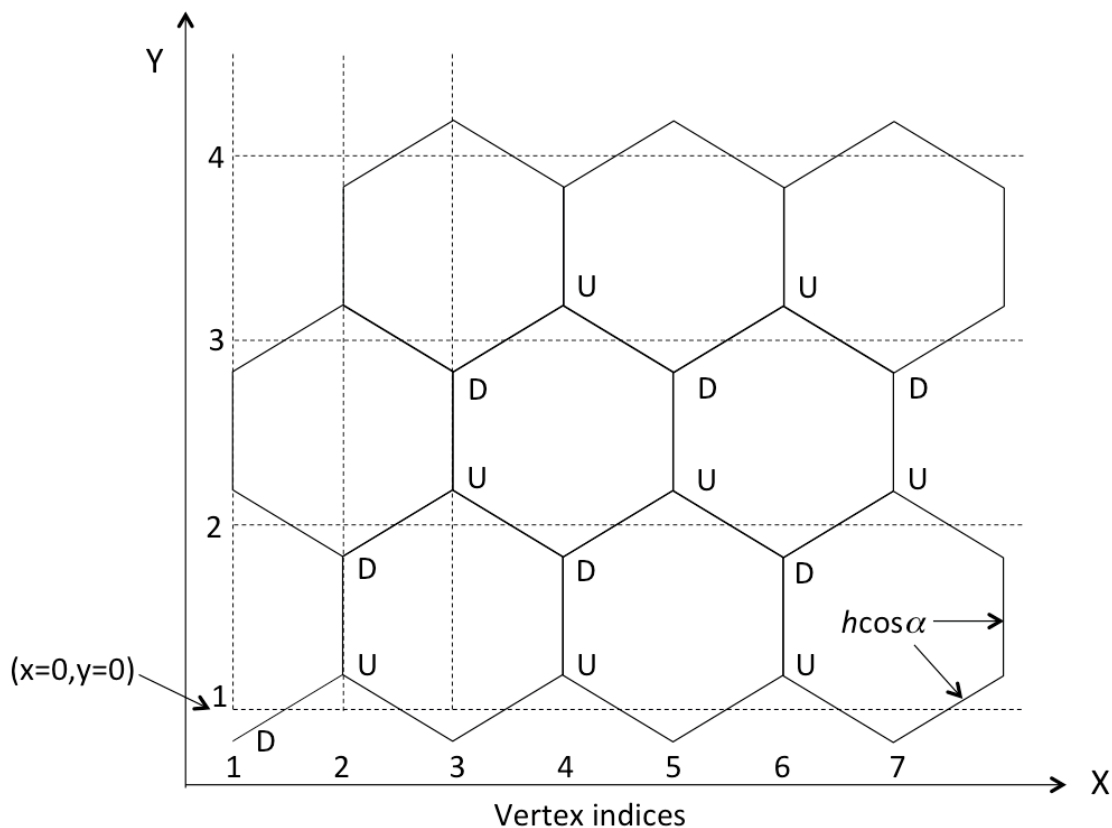
Figure S2.5. View down the trigonal axis of the diamond 2H structure showing the structure's hexagonal projection with apparent length $h\cos\alpha$ of the connecting rods. Vertices are coded by X and Y numerals. Layers of the hexagonal projection are superimposed in a stack coming out of the page. From one layer to the next higher layer there are interchanged U and D characters that represent axial displacements of vertices upward and downward (along the z-axis). The layers are connected from a high point U (in a lower layer) to a low point D (in the layer above) by axial rods of length h. Each layer is coded by a Z numeral.

Rather than seek a unit cell, which is conceptually difficult for this structure, we created (x,y,z) coordinates for each lattice vertex as a function of the numbers in an intuitive labeling system illustrated in Figure S2.5, which is a view down the trigonal symmetry axis. This takes as a starting point the hexagonal projection layers that lie in alignment vertically above and below each other to fill space, then adds or subtracts height from alternating vertices around the hexagons in the manner discussed in McGeoch et al. (2021a). We need to have twice the number of "x" values as "y" and "z" values to uniquely specify the vertices of an approximately cubic lattice volume (Figure S2.5). In Figure S2.5 we are looking down upon fore-shortened sides to the hexagons, of apparent length $h\cos\alpha$, where h is the true length of a connecting rod and α is the angle that each rod takes relative to the horizontal plane (illustrated in Fig. S2.4). When $\sin\alpha = 1/3$ the exact tetrahedral symmetry exists at every junction in the structure (McGeoch et al., 2021a). Here we allow angle α to be variable, giving access to the parameters of axially distorted structures.

Where *J,K,L* are the indices (coding numerals) along the *x,y,z* axes, the position of any vertex is given by

$$x = (J-1)\frac{\sqrt{3}}{2}h\cos\alpha$$

$$y = (K-1)\frac{3h}{2}\cos\alpha + (-1)^{J}.(-1)^{K-1}.\frac{h\cos\alpha}{4}$$

$$z = (L-1)(h+h\sin\alpha) + (-1)^{J}.(-1)^{(K-1)}.(-1)^{(L-1)}.\frac{h\sin\alpha}{2}$$

Here for clarity in the superscripts we replaced the indices J_X, J_Y, J_Z used in the program listed below by *J,K,L*.

With this apparatus we can consider index ranges suitable for a cube of side N quasi-cells (defined in [2]):

$$1 < J < 2N$$

$$1 < K < N$$

$$1 < L < N$$

A central point in the lattice is chosen, say at coordinates (N, N/2, N/2) with N even, then nearby lattice distances can be evaluated, for example in ranges for J , K , L

$$N - 4 < J < N + 4$$

$$N/2 - 3 < K < N/2 + 3$$

$$N/2 - 3 < L < N/2 + 3$$

When run with a maximum spacing cut-off of 140 Angstroms, this produces 69 spacings, many of which are duplicates. Further ordering and pruning to yield unique "*d*" values reduces this to a short list. In the comparison of data to theory that follows we chose a cutoff of 140Angstroms. When the exact tetrahedral angle is used, there are 9 distinct lattice "d" values less than 140 Angstroms. If an axial distortion is applied the degeneracy is broken and 12 values now appear under 140 Angstroms. Figure S2.6 plots the changes to these values as the deviation of α from tetrahedral (α = 19.471221 deg.) ranges through -5 to +5 degrees (the 49.0 Angstrom line, not shown, does not vary with α).
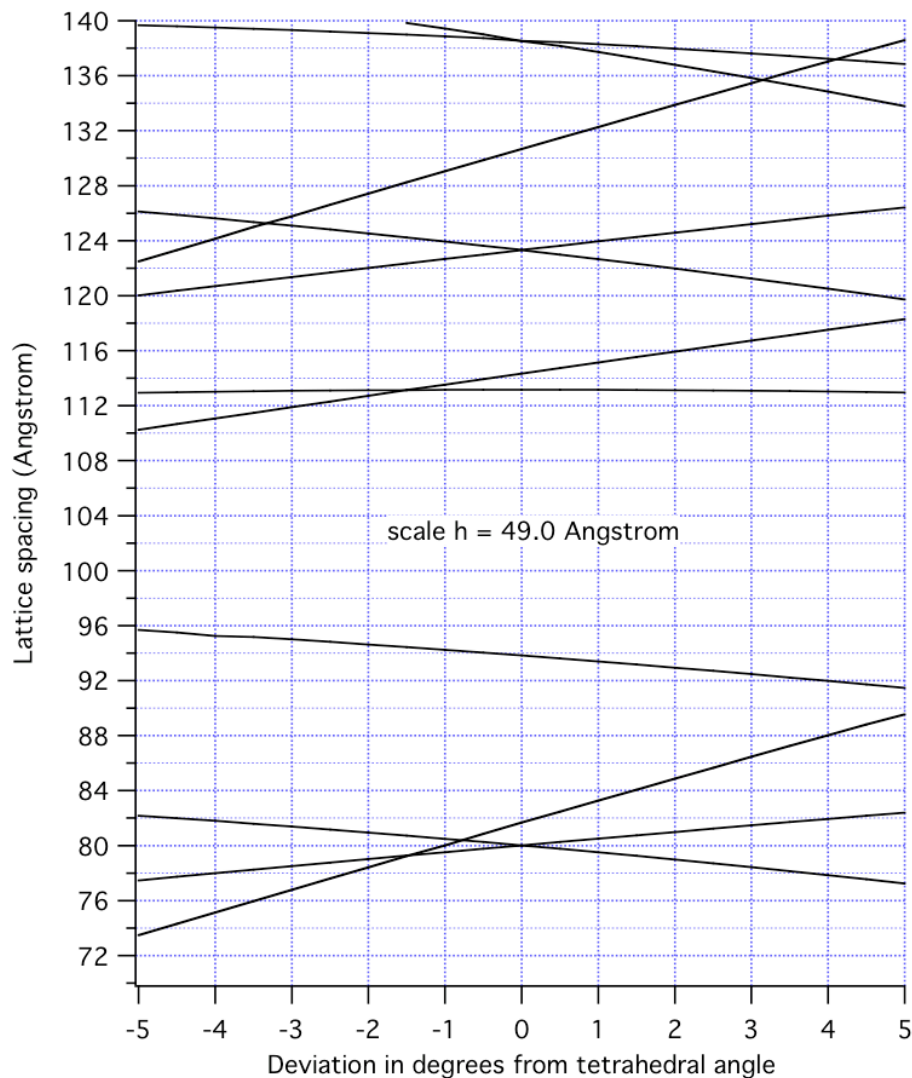
Figure S2.6. Variation of lattice spacings under 140Angstrom when angle α is varied above and below the tetrahedral angle of 19.471221 degrees (relative to horizontal). The 49.0 Angstrom fundamental spacing does not vary with angle and is not shown on this plot. These results may be scaled to rod length X Angstroms by multiplying by X/49.0 .

Fit including lattice distortion to observed set

It is found that a lattice spacing of 49.0 ±0.2 Angstroms combined with an angular deviation of +4.5 deg. gives the best fit to the reduced data of Figure S2.3 (fit shown as red bars). Only two major peaks in Figure S2.3 were not covered, at 98 and 130 Angstroms. Minor peaks were potentially due to random coincidences. The 98 Angstrom peak happens to be exactly two times the 49.0 Angstrom rod length. It can be produced automatically by doubling the diffraction orders {5,6,7 and 9} of column 2 in Table S2.2, therefore its presence is expected. The 130 Angstrom peak is not covered at +4.5 deg. in Figure S2.6, however it is covered precisely by the 0 deg. deviation non-distorted lattice. It could be that some parts of the lattice are un-distorted. The peak at 63 Angstrom is expected as a corollary of the 126.25 Angstrom peak which contains orders 14,

18, 20, 22, 24 that may be divided by 2 to give the valid orders 7, 9, 10, 11, 12, however no new lattice parameter is implied.

S2 Discussion
The use of wavelengths spanning the Fe K edge absorption allowed the iron-dependent aspects of ooid diffraction to stand out due to decreased absorption at the longer wavelength. It also appears to have been important to have reduced diffraction order numbers by use of a factor of two longer wavelength, contributing to a simpler analysis. It is considered, in conjunction with the ooid fluorescence data (main text), that the identified lattice is that of hemoglycin, which in prior x-ray work (McGeoch et al. 2021a, 2022) has shown a "rod" length of 4.9nm. A clear analysis without competing structures suggests that this is the dominant and only iron-containing lattice within the ooid. A relatively small iron content of 0.15 wt % can produce this lattice, in conjunction with a glycine content of approximately 0.87 wt %. This considers a calcite-filled lattice with density 2.71. The density of aragonite varies from 2.93 to 2.95.

The lattice is filled with many calcium carbonate crystals, of a size that could be as small as the "cells" within the diamond 2H structure, i.e. between 5nm and 8nm across. Nano-crystals of nickel at this scale were observed in X-ray diffraction of a lattice sample from the Acfer 086 meteorite (analysis unpublished). Such fillings will vary depending upon the mix of atoms available as the lattice forms.

It is probable that a small degree of lattice distortion can be induced by any given crystal type within the cells. The cell volume is maximized for exact tetrahedral symmetry (McGeoch et al. 2021a) but very little energy is required to induce an axial lattice distortion to accommodate a particular filling type of crystal. As the lattice is lengthened slightly along the trigonal axis the $h\cos\alpha$ projection decreases slightly, possibly providing an energy minimum for the combined system of lattice plus enclosed calcium carbonate crystals.

Section S2 Appendix. Program to find diamond 2H lattice spacings

```
'Programmed in qb64 for mac - qb64 is a generic Basic language
PRINT "Diamond Lattice"
PRINT "calculates x,y,z coordinates in a cube of side N"
PRINT "output is in a new file called dspace in the same folder as the program"
DIM Vx(22), Vy(11), Vz(11), set(15), fit(15), testall(100), testran(100), testwin(15)

INPUT "Variation (deg) from alpha at tetrahedral vertex ", dalpha
INPUT "number of quasi-cells (even and less than or equal to 10) ", N
INPUT "inter-vertex distance h (Angstroms) ", h

DIM dspace AS STRING
OPEN "dspace" FOR OUTPUT AS #1

pi = 3.14159265359
alp = (19.471221 + dalpha) * pi / 180
Salp = SIN(alp)
```

```
Calp = COS(alp)

Nx = 2 * N
Ny = N
Nz = N

Fx = SQR(3) * h * Calp / 2
Fy1 = 3 * h * Calp / 2
Fy2 = h * Calp / 4
Fz1 = h + h * Salp
Fz2 = h * Salp / 2

' set up central point N, N/2, N/2
N2 = N / 2

x0 = (N - 1) * Fx
y0 = (N2 - 1) * Fy1 + (-1) ^ N * (-1) ^ (N2 - 1) * Fy2
z0 = (N2 - 1) * Fz1 + (-1) ^ N * (-1) ^ (N2 - 1) * (-1) ^ (N2 - 1) * Fz2

'generate vertices around central point and distances from center

index = 1
FOR Jx = N - 4 TO N + 4
Vx(Jx) = (Jx - 1) * Fx

FOR Jy = N2 - 3 TO N2 + 3
Vy(Jy) = (Jy - 1) * Fy1 + (-1) ^ Jx * (-1) ^ (Jy - 1) * Fy2

FOR Jz = N2 - 3 TO N2 + 3
Vz(Jz) = (Jz - 1) * Fz1 + (-1) ^ Jx * (-1) ^ (Jy - 1) * (-1) ^ (Jz - 1) * Fz2

test = SQR((x0 - Vx(Jx)) ^ 2 + (y0 - Vy(Jy)) ^ 2 + (z0 - Vz(Jz)) ^ 2)
IF test > 140 THEN GOTO 16
testall(index) = test
index = index + 1

PRINT #1, USING "     .####^^^^"; Jx, Jy, Jz, test
16 'continue

NEXT Jz
NEXT Jy
NEXT Jx

PRINT "index = ", index
CLOSE #1
```

```
'rank list of "d" spacings in vector testall
rank = 1
FOR k = 1 TO index
trial = testall(k)

FOR m = 1 TO index
IF trial > testall(m) + 0.0001 THEN rank = rank + 1
NEXT m

testran(rank) = trial
rank = 1
NEXT k

'winnow testran to obtain unique set
q = 1
FOR p = 1 TO index
IF testran(p) > 0 THEN testwin(q) = testran(p) ELSE GOTO 20
q = q + 1
20 'continue
NEXT p

FOR g = 1 TO q - 1
PRINT testwin(g)
NEXT g
END
```

Section S3: X-ray diffraction on crystals and ooids

S3.1 Basic x-ray diffraction observations on fossil stromatolite and meteoritic crystals

In 2021 and 2022 at APS diffraction was studied in a range of crystals derived as described above from fossil stromatolite No. 1 Wyoming (2.1Gya), provenance in Section S 3.2.

In all except one crystal the dominant features were rings, and not spots, indicating generally multiple crystalline nature (Table S3.1). The stromatolite and Acfer 086 data is from beam line 31-1D-D, APS, at a wavelength of 0.9793 Angstroms. The vesicle hexagon data is from Diamond Light Source, at 1.000 Angstroms.

Table S3.1: Rings measured on various crystals (Angstroms). All detectable rings are listed: gaps indicate no data. Sample to detector distances (mm) noted.

Stromatolite 272807 fossil No. 1 (400mm)	Stromatolite 275049 fossil No. 1 (329mm)	Stromatolite 289349 fossil No. 1 (190mm)	Rolled-up Hemoglycin sheet Acfer 086 [2]	Vesicle hexagon [9] Sutter'sMill
-	6.28(5) ± 0.021	6.277	6.31	-
5.25	5.26 ± 0.016	5.251	5.29	-
4.70	4.78 ± 0.019	4.784	4.80	-
-	4.20 ± 0.011	4.196	-	-
4.03	4.05 ± 0.018	4.064	-	4.085 -only
3.68	3.69 ± 0.013	-	-	-
3.45	3.49(5) ± 0.011	3.454 ± 0.043	3.52	-
-	3.34(5) ± 0.011	3.351 ± 0.034	-	-
3.06	3.09(3) ± 0.004	3.097 ± 0.033	-	-
2.89	2.89 ±0.007	2.892 ± 0.023	-	-
2.80	-	-	-	-
2.77	-	-	-	-
2.71	2.67(5) ± 0.005	-	-	-
2.62	-	-	-	-
2.51	2.54 ±0.007	-	-	-
2.40	2.40 ± 0.005	-	-	-
2.29	-	-	-	-
2.25	-	-	-	
2.19	2.195 ± 0.005	-	-	
2.14	-	-	-	
2.06	-	-	-	
2.01	2.015 ± 0.005	-	-	
1.94	-	-	-	
1.88	-	-	-	
1.84	-	-	-	
1.80	1.805 ± 0.005	-	-	

The first five entries in column 3 under 289349 were processed differently via scanning and subtraction of a relatively large background, so as to find accurate peak positions. The errors for these five points are systematic and of the order of 0.03 Angstroms. All other errors are one standard deviation of multiple ring measurements. Although the crystals in columns 2 and 3 do not appear to be externally similar at all, there is a very good lattice parameter match across all the major rings. Crystal 275049 was thin, purplish and well-shaped whereas crystal 289349 was a white "blob" with much greater depth. A high featureless background in the central part of the 289349 pattern could be due to amorphous material within its lattice. The first five entries under 289349 were, as a set, very much more intense than the subsequent entries.

We searched our meteorite and stromatolite data for matches to the above pair. In the 4th column we list the observed rings from a fiber crystal of Acfer 086 (McGeoch et al., 2021a) the first two of which were identified with the Fe-Fe spacing at the four-way connection of hemoglycin rods. In the fifth column we list the single dominant spacing of Fe atoms at the three-way junction of hexagonal sheets of Sutter's Mill hemoglycin that formed a vesicle (McGeoch et al., unpublished). Although the latter may be a fortuitous match, these prior data resemblances to the stromatolite dominant rings could point to the presence of Fe-Fe spacings within fossil stromatolite similar to those of meteorite extract crystals.

S 3.2 X-ray diffraction observations on present day and fossil ooids.
At Diamond Light Source a series of x ray diffraction runs compared the crystal powder diffraction patterns of **ooids** from the following stromatolite sources:

Sample Sources: Present-day stromatolites are supplied by Andrew Knoll of the Museum of Comparative Zoology and Organismic and Evolutionary Biology (OEB) Harvard.
Present-day stromatolite details are:
1. Shark-Bay Western Australia – collected by Elso Barghoorn 1971- estimated to be 2000-3000 years old.
2. Sample K-05 SS-1 from San Salvador Island Bahamas – collected by Andrew Knoll in 2005 –- A modern mineralized microbialite.

Fossil stromatolite details are:
1. 2.1Ga stromatolite No. 1 from Medicine Bow region, Wyoming – collected by David Lageson of Montana State University.
2. 2.1Ga stromatolite No. 2 from Medicine Bow region, Wyoming – collected by David Lageson of Montana State University.

X-ray diffraction results for ooids in the two present day stromatolites and in a second Wyoming sample, No. 2, are compared in Table S3.2. No ooids could be found in Fossil stromatolite No. 1.

Table S3.2 Comparison of present day and fossil ooid diffraction patterns on Diamond Light Source. All mounted samples are multiple crystals, giving rings. The number of x ray runs that were averaged is given. Almost uniformly the rings identify as aragonite, with one possible calcite ring.

Ooids, x ray 1.000A, all data in cryo-stream at 100K, "d' in Angstroms				
Shark Bay No acid n = 2	Shark Bay Acid treated n = 2	2nd Fossil Stromatolite n = 3	San Salvador n = 1	Assignment a = aragonite c = calcite
	16.605			
	11.65			
		10.458		
10.325			10.32	
	8.605			
	8.246			
	7.424			
		7.11	7.093	
	6.905			
	6.376			
	5.505			
		4.38		
4.194		4.194		
3.791	3.819	3.812		
	3.493			
3.388	3.393	3.387		a
3.263	3.263	3.262	3.265	a
	3.213			
2.988	2.988	2.993	3.00	
2.860	2.865			
2.725		2.70		a
2.690		2.689	2.698	a
2.480	2.47	2.472	2.475	a
2.40		2.407		
2.363	2.358	2.367	2.37	a
2.328	2.328	2.324	2.325	a
2.248	2.245			
2.185		2.182		
2.065		2.08		
1.978	1.968	1.974	1.975	a
1.877	1.875	1.876	1.878	c , a
1.808	1.808	1.807	1.81	
	1.738	1.737	1.74	a
		1.719	1.72	
1.60				