A Picture of Disorder in Hydrous Wadsleyite – Under the Combined Microscope of Solid-State NMR Spectroscopy and Ab Initio Random Structure Searching

David McKay,1 Robert F. Moran,1 Daniel M. Dawson,1 John. M. Griffin,2 Simone Sturniolo,3 Chris J. Pickard,4 Andrew J. Berry5 and Sharon E. Ashbrook1,*

1 School of Chemistry, EaStCHEM and Centre of Magnetic Resonance, University of St Andrews, St Andrews, KY16 9ST, UK
2 Department of Chemistry and Materials Science Institute, Lancaster University, Lancaster, LA1 4YB, UK
3 Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0QX, UK
4 Department of Materials Science and Metallurgy, University of Cambridge, Cambridge, CB3 0FS, UK
5 Research School of Earth Sciences, Australian National University, Canberra, ACT 2601, Australia

Supporting Information

S1. Referencing of DFT-calculated NMR parameters
S2. AIRSS-generated semi-hydruszous wadsleyite structures
S3. k-means clustering of AIRSS-generated hydrous wadsleyite structures
S4. Additional plots of structural and NMR parameters for semi-hydruszous wadsleyite
S5. The enthalpic stability of Si-vacant fully-hydruszous wadsleyite
S6. Choice of Mg site combinations for AIRSS study of fully-hydruszous wadsleyite
S7. The Smyth fully-hydruszous wadsleyite model
S8. Additional NMR experiments
S9. Experimental FTIR and simulated IR spectra of fully-hydruszous wadsleyite
S10. References
S1. Referencing of DFT-calculated NMR parameters

To compare experimental chemical shift values to those calculated using DFT, the calculated isotropic chemical shielding values, $\sigma_{\text{iso}}^\text{calc}$, must be converted to a corresponding calculated shift, $\delta_{\text{iso}}^\text{calc}$, by

$$
\delta_{\text{iso}}^\text{calc} = \sigma_{\text{ref}} - \sigma_{\text{iso}}^\text{calc},
$$

(S1.1)

where $\sigma_{\text{ref}}$ is a reference shielding. The most straightforward way to determine $\sigma_{\text{ref}}$ is to compare the experimental shift and calculated shielding for a simple model compound where the experimental shift is accurately determined and the structure understood. Though this approach can work well, in some cases uncertainties in the experimental measurements of the NMR parameters and in the structural models (and therefore calculated shielding values), and the limitations of the DFT approximation employed can lead to errors in the computed chemical shifts. If a material contains more than one crystallographically distinct species, an alternative approach is to consider only the relative shift/shielding differences between the peaks (i.e., determining a reference shielding from either one, or the average of several shift/shielding values within the selected reference material. A more robust approach, however, is to compare the experimental shift and calculated shielding for multiple materials. Although this approach alleviates the effect of any experimental errors or structural uncertainties, it cannot correct for any errors in the chosen exchange-correlation functional. Therefore, to mitigate any issues of chemical transferability it is preferable that the set of materials used to generate the computed chemical shifts should be as similar to the system of interest as possible.

Tables S1.1 and S1.2 show the $^{17}$O and $^{29}$Si $\sigma_{\text{iso}}^\text{calc}$ and $\delta_{\text{iso}}^\text{exp}$, the latter obtained from the literature, for three polymorphs of Mg$_2$SiO$_4$ forsterite ($\alpha$-Mg$_2$SiO$_4$), wadsleyite ($\beta$-Mg$_2$SiO$_4$) and ringwoodite ($\gamma$-Mg$_2$SiO$_4$). In each case, the initial geometries of the structural models were optimized prior to the calculation of NMR parameters using CASTEP version 8. Figures S1.1a and S1.1b plot $\delta_{\text{iso}}^\text{exp}$ against $\sigma_{\text{iso}}^\text{calc}$ for the three Mg$_2$SiO$_4$
polymorphs, and show good linear correlations ($R^2 = 0.9928$ and $0.9979$ for the $^{17}$O and $^{29}$Si NMR parameters, respectively), although it should be noted that in each case the gradient of the line varies from the ideal value of 1. Therefore, the equations shown in Figures S1.1a and S1.1b were used to reference the $^{17}$O and $^{29}$Si calculated isotropic chemical shielding values, respectively, for the calculations in the main text.

**Table S1.1** $^{17}$O calculated isotropic chemical shielding ($\sigma_{iso}^{calc}$) and experimental isotropic chemical shift ($\delta_{iso}^{exp}$) for three polymorphs of Mg$_2$SiO$_4$, forsterite ($\alpha$-Mg$_2$SiO$_4$), wadsleyite ($\beta$-Mg$_2$SiO$_4$) and ringwoodite ($\gamma$-Mg$_2$SiO$_4$).

| Compound       | $\sigma_{iso}^{calc}$ (ppm) | $\delta_{iso}^{exp}$ (ppm) |
|----------------|----------------------------|-----------------------------|
| Forsterite$^1$ | 203.69                     | 48$^{2-4}$                  |
|                | 184.66                     | 64$^{2-4}$                  |
|                | 188.00                     | 61$^{2-4}$                  |
| Wadsleyite$^5$| 214.01                     | 38$^{4,6}$                  |
|                | 171.94                     | 76$^{4,6}$                  |
|                | 181.37                     | 66$^{4,6}$                  |
|                | 183.48                     | 65$^{4,6}$                  |
| Ringwoodite$^7$| 183.12                     | 63$^1$                      |

**Table S1.2** $^{29}$Si calculated isotropic chemical shielding ($\sigma_{iso}^{calc}$) and experimental isotropic chemical shift ($\delta_{iso}^{exp}$) for three polymorphs of Mg$_2$SiO$_4$, forsterite ($\alpha$-Mg$_2$SiO$_4$), wadsleyite ($\beta$-Mg$_2$SiO$_4$) and ringwoodite ($\gamma$-Mg$_2$SiO$_4$).

| Compound      | $\sigma_{iso}^{calc}$ (ppm) | $\delta_{iso}^{exp}$ (ppm) |
|---------------|-----------------------------|-----------------------------|
| Forsterite$^1$| 387.40                      | $-61.9^8$                   |
| Wadsleyite$^5$| 406.11                      | $-79.0^{4,9,10}$            |
| Ringwoodite$^7$| 408.23                      | $-82.0^4$                   |
Figure S1.1. Plot of (a) $^{17}$O and (b) $^{29}$Si $\delta_{\text{iso}}^{\text{exp}}$ against $\sigma_{\text{iso}}^{\text{calc}}$ for three polymorphs of Mg$_2$SiO$_4$, forsterite ($\alpha$-Mg$_2$SiO$_4$), wadsleyite ($\beta$-Mg$_2$SiO$_4$) and ringwoodite ($\gamma$-Mg$_2$SiO$_4$).

Finding a procedure to reference calculated $^1$H $\sigma_{\text{iso}}^{\text{calc}}$ is more challenging as there are few hydrated high-pressure silicate minerals with accurate experimental chemical shift information and well-defined structures. As a result, calculated $^1$H isotropic chemical shieldings were referenced by using experimental and calculated NMR parameters corresponding to hydrous wadsleyite itself. Conclusions drawn from this paper, as well as results from a previous AIRSS-based study into the hydration of wadlseyite,\textsuperscript{11} have been used to assign several of the resonances in the $^1$H MAS spectrum of fully-hydrated wadsleyite containing 3.3 wt% H$_2$O, presented by Griffin et al.\textsuperscript{10} From this, and the previous AIRSS investigation, it was concluded that the resonances that appear at 3.4, 6.7 and 8.6 ppm, correspond to protonation of O1, O3 and O4 oxygen sites, respectively. The calculated $^1$H isotropic chemical shielding values for the AIRSS-generated semi- and fully-hydrated wadsleyite structures (which contain 1.65 and 3.3 wt% H$_2$O, respectively)
believed to be responsible for the $^1$H resonances at 3.4 and 8.6 ppm (O1 and O4 protonation) were identified and used to reference the $^1$H isotropic chemical shifts that are quoted in the main text. The structures selected represent the lowest enthalpy examples of a particular protonation motif for a specific combination of Mg vacancies, i.e., the lowest enthalpy structure that has two Mg3 vacancies or an Mg1 and an Mg3 vacancy. Tables S1.3 and S1.4 show $\sigma_{iso}^{calc}$ and $\delta_{iso}^{exp}$ for the selected semi- and fully-hydrated wadsleyite structures that were used to reference calculated $^1$H $\sigma_{iso}^{calc}$. The average $^1$H $\sigma_{iso}^{calc}$ is quoted where more than one of a particular protonation arrangement is observed, i.e., when two or four protons are all located on O1 sites, the average $^1$H $\sigma_{iso}^{calc}$ is used.

**Table S1.3** $^1$H calculated isotropic chemical shielding ($\sigma_{iso}^{calc}$) and experimental isotropic chemical shift ($\delta_{iso}^{exp}$) for the semi-hydrated wadsleyite structures chosen to reference $^1$H $\sigma_{iso}^{calc}$ for structures with this hydration level.

| Mg site vacancy | Protonation motif | Protonation site | $\sigma_{iso}^{calc}$ (ppm) | $\delta_{iso}^{exp}$ (ppm) | Relative enthalpy / eV |
|----------------|------------------|-----------------|-----------------------------|-----------------------------|------------------------|
| Mg3            | 2 × H–O1         | O1              | 25.38                       | 3.4                         | 0.00                   |
| Mg1            | 2 × H–O4         | O4              | 18.67                       | 8.6                         | 0.33                   |

**Table S1.4** $^1$H calculated isotropic chemical shielding ($\sigma_{iso}^{calc}$) and experimental isotropic chemical shift ($\delta_{iso}^{exp}$) for the fully-hydrated wadsleyite structures chosen to reference $^1$H $\sigma_{iso}^{calc}$ for structures with this hydration level.

| Mg site vacancy | Intervacancy distance / Å | Protonation motif | Protonation site | $\sigma_{iso}^{calc}$ (ppm) | $\delta_{iso}^{exp}$ (ppm) | Relative enthalpy / eV |
|----------------|---------------------------|------------------|-----------------|-----------------------------|-----------------------------|------------------------|
| Mg3/Mg3        | 7.13                      | 4 × H–O1         | O1              | 26.05                       | 3.4                         | 0.00                   |
| Mg1/Mg3        | 2.89                      | 2 × H–O1, 1 × H–O3, 1 × H–O4 | O4              | 20.63                       | 8.6                         | 0.37                   |
Figures S1.2a and S1.2b plot $\delta_{\text{iso}}^{\text{exp}}$ against $\sigma_{\text{iso}}^{\text{calc}}$ for the selected AIRSS-generated semi- and fully-hydrated wadsleyite structures. The equations shown in Figures S1.2a and S1.2b were used to reference the $^1\text{H} \sigma_{\text{iso}}^{\text{calc}}$ for the semi- and fully-hydrated wadsleyite structures, respectively.

Figure S1.2. Plot of $^1\text{H} \delta_{\text{iso}}^{\text{exp}}$ against $\sigma_{\text{iso}}^{\text{calc}}$ for the selected AIRSS-generated (a) semi- and (b) fully-hydrated wadsleyite structures used to reference the $\sigma_{\text{iso}}^{\text{calc}}$ isotropic chemical shielding values quoted in the main text.
S2. AIRSS-generated semi-hydrous wadsleyite structures

Figure S2.1 shows the relative enthalpy plotted against structure rank for the 223, 245 and 819 AIRSS-generated structures with an Mg1, Mg2 and Mg3 vacancy, respectively. A plot with all three Mg vacant structure series plotted on the same set of axes is shown in Figure S2.2, displaying the relative enthalpy difference between the three sets of structures. From Figure S2.2 it can be seen that the lowest enthalpy structures come from the series of structures with an Mg3 vacancy, with the highest enthalpy ground state resulting from hydrogen incorporation charge balanced by the removal of an Mg2 cation.

**Figure S2.1.** Plot showing the relative enthalpy of all AIRSS-generated structures with an (a) Mg1, (b) Mg2 and (c) Mg3 vacancy.
Figure S2.2. Plot showing the relative enthalpy of all AIRSS-generated structures, with structures with Mg1, Mg2 and Mg3 vacancies shown in green, red and blue, respectively.
S3. k-means clustering of AIRSS-generated hydrous wadsleyite structures

A k-means clustering approach was adopted for the selection of smaller subsets of AIRSS-generated structures for further study. Scripts were developed in Python 2.7, using the structural and statistical functionality of the Soprano library,\textsuperscript{12} which extends the Atomic Simulation Environment (ASE) library\textsuperscript{13} for the analysis for computed crystal structures. k-means clustering, a method of “unsupervised machine learning”, involves the subdivision of a dataset into $k$ groups, called “clusters”, which are determined to be similar according to a normalized set of pre-defined parameters, known as “genes”. In an iterative process, $k$ data points are first chosen at random and neighboring data points become “members” of these clusters, thereafter the central data point of each cluster is set to that closest to the cluster mean and the cluster members refined according to their distances to the mean data point. In terms of a set of chemical structures, the end goal is to segregate this into $k$ sets of like structures from which one representative structure can be used for further studies and, which due caution, the remaining cluster members can be discarded.

**Semi-hydrous wadsleyite clustering**

For clustering, genes were constructed based on: the relative enthalpy, $\Delta H$, with respect to the lowest enthalpy structure in the series of semi-hydrous structures with a Mg$^3$ vacancy; hydroxyl O type; vacancy type; H⋯H distance; H⋯vacancy distance; O–H bond length; OH⋯O hydrogen-bond length and the magnitude of the combined hydroxyl orientation vector, $\|\overrightarrow{OH}\|$ (where $\overrightarrow{OH}$ is defined as the sum of unit vectors placed along the two O–H bonds and hence $0 \leq \|\overrightarrow{OH}\| \leq 2$, going from anti-parallel to parallel hydroxyls). The vacancy position was formally defined as the fractional coordinates of the Mg cation removed from anhydrous wadsleyite, a reasonable approximation as the unit cell was fixed and atoms heavier than H moved little during geometry optimization. Genes based on $\Delta H$, O type and $\|\overrightarrow{OH}\|$ were found to cluster structures well, where, for a given small enthalpy range, a range of possible values of $\|\overrightarrow{OH}\|$ was found and changing O type has a large effect on $\Delta H$. The alternative genes of H⋯H distance, H⋯vacancy
distance, O–H bond length and OH⋯O hydrogen bond length exhibited consistent values over larger enthalpy ranges and so were not useful in clustering. The required number of clusters, $k$, was then determined by plotting the sum of squared errors within each cluster against $k$, a plot which will reach unity at $k = N$, where $N$ is the number of structures in the series. A maximum value of $k$ was chosen such that the sum of squared errors within each cluster was well converged (see Figure S3.1, in which the series of AIRSS-generated structures with an Mg3 vacancy is used as an illustrative example).

Figure S3.1. Convergence plots of $k$-means clustering quantities, where the sum of squared errors is plotted against $k$ for the series of Mg3-vacant semi-hydrous wadsleyite structures. Plots show (a) the convergence of $\Delta H$, (b) the convergence of $|\text{OH}|$ and (c) the convergence of O type. Dashed lines represent the chosen maximum value of $k$ for this dataset.
The quality of clustering was assessed by plotting each cluster as a Gaussian function against values contained in that cluster according to a specific gene:

\[ f(x|\mu, \sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (S3.1) \]

\[ f(x|\mu, \sigma) = e^{-\frac{(x-\mu)^2}{(0.1+\sigma)^2}}, \quad (S3.2) \]

where \( x \) is a value of \( \Delta H \), \( \| \bar{O} \| \) or O type of a given structure, \( \mu \) is the mean of that quantity for all structures in the cluster to which that structure belongs and \( \sigma \) is the standard deviation of that value in the cluster. Equation S3.1 was used for \( \Delta H \) and \( \| \bar{O} \| \) genes and S3.2 was used for the O-type gene, since a value of \( \sigma = 0 \) was possible and precluded visualization. In the case of the \( \Delta H \) gene, the aim was to have narrow Gaussians at low \( \Delta H \) and broad Gaussians at high \( \Delta H \) (see Figure S3.2a), since the low \( \Delta H \) structures would be better candidates for study. A weighting was applied to the O-type gene, due to its discreet nature, such that each cluster contained structures with a single common O type. This is illustrated in Figure S3.2c, where the Gaussians are generally narrow.
Figure S3.2. Gaussian representations of (a) $\Delta H$, (b) $|\text{OH}|$ and (c) O-type quantities found within each of 19 $k$-means clusters found for the set of Mg3-vacant semi-hydrous wadsleyite structures.

Clusters were then used to inform the selection of structures. First, the structure with lowest $\Delta H$ was selected. Clusters were then pre-filtered based on a $\Delta H$ cut off, which was set at the $\Delta H$ value where the energy profile changed from a plateau to a rapidly increasing enthalpy (e.g., chosen as 1.4 eV for the Mg1- and Mg2-vacant series and 2.0 eV for the Mg3-vacant series; see Figure S3.3, in which the Mg3-vacant series is used as an illustrative example).
Figure S3.3. Plot of $\Delta H$ vs structure rank for Mg3-vacant structures (blue circles), showing selected structures (white crosses) and cluster mean enthalpy cut off (2.0 eV) and percentile enthalpy cut off (2.2 eV) (dashed lines).

A cluster was discarded if the mean value of $\Delta H$ lay above this cut-off value. Then the median $\Delta H$ structure was taken from each remaining cluster. The width of the clusters in the $\Delta H$ and $\|\mathbf{O}\|$ genes were then used to select additional structures. A cluster was judged to be broad if its width covered > 5% of whole series, in which case, structures at the 5th and 95th percentile were also selected, providing these were below an energy cut off, set to 0.2 eV above the mean $\Delta H$ cut off. This resulted in the selection of a maximum of five structures per cluster (see Figure S3.4).
Figure S3.4. Plot of $\Delta H$ against $||\text{O}||$ colored by cluster number for Mg3-vacant structures, showing selected structures (white crosses).

The robustness of the method for structure selection was tested by verifying the selected structures against the series of structures with a Mg3 vacancy studied in previous work,\textsuperscript{11} where the aim was to manually select structures for their uniqueness. Indeed, it was seen that the fewest structures were selected where we had found structures to be identical (or very similar), normally forming plateaus in $\Delta H$, and more structures were selected where more structural diversity was found, such as at higher energy. This method reduced the numbers of structures to 32, 7 and 49 structures with Mg1, Mg2 and Mg3 vacancies, respectively. It should be noted that, owing to the statistical nature of this structure selection approach, the particular structures chosen, and the number of structures chosen, can vary between runs of the script. It is therefore advisable to run the script multiple times to assess the effect this has; \textit{i.e.}, whether different structures are selected, but are merely different examples of the same structure, or whether a different set of structures is found. In the latter case, selection parameters should be modified.
**Fully-hydrous wadsleyite clustering**

Clustering scripts used to select structures from AIRSS-generated fully-hydrous wadsleyite series were based on those developed for semi-hydrous wadsleyite with the addition of a gene for intervacancy distance, which was weighted (as with O type, described above) to ensure each cluster contained a set of structures with a single intervacancy distance. Alterations were made to existing genes to allow the combination of two vacancies and four protons to be handled: the H⋯H distance and H⋯vacancy distance genes were modified to consider the average of the shortest distance for each H and for the $\|\text{OH}\|$ gene, $\|\text{OH}\|$ was redefined as the sum of unit vectors placed along all four O–H bonds, giving $0 \leq \|\text{OH}\| \leq 4$, going from anti-parallel to all parallel hydroxyl arrangements. Other genes, based on relative enthalpy, O–H bond length, OH⋯O hydrogen-bond length, hydroxyl O type and vacancy type remained unchanged between semi- and fully-hydrous systems. Analysis and validation of the clusters was directly analogous.
S4. Additional plots of structural and NMR parameters for semi-hydrous wadsleyite

The plots shown in Figure S4.1 highlight the variation in $^1\text{H} \delta_{\text{iso}}$ and $^2\text{H} C_Q$ with the hydroxyl bond distance for the 58 semi-hydrous wadsleyite structures with a $\Delta H < 1.0$ eV. Figure S4.1a and S4.1b show there are well-defined regions of $^1\text{H} \delta_{\text{iso}}$ and $^2\text{H} C_Q$ for Mg–OH and Si–OH environments, with Figure S4.1a showing that hydrogen atoms in an Mg–OH environment have $^1\text{H} \delta_{\text{iso}} < 4.5$ ppm, with Si–OH environments appearing exclusively above 5 ppm.

**Figure S4.1.** Plots of calculated covalent H–O bond length against (a and c) $^1\text{H} \delta_{\text{iso}}$ and (b and d) $^2\text{H} C_Q$, coloured by (a and b) protonation environment (*i.e.*, Mg–OH or Si–OH) and (c and d) protonation site for the 58 semi-hydrous wadsleyite structures with $\Delta H < 1.0$ eV selected by $k$-means clustering.

The plots shown in Figure S4.2 highlight the variation in $^1\text{H} \delta_{\text{iso}}$ and $^2\text{H} C_Q$ with the hydrogen-bond distance for all 88 semi-hydrous wadsleyite structures. Figure S4.2a and S4.2b show there are reasonably well-defined regions of $^1\text{H} \delta_{\text{iso}}$ and $^2\text{H} C_Q$ for Mg–OH and
Si–OH environments. Figure S4.2c and S4.2d show that the hydrogen-bond distance for the O2–H hydroxyl is noticeably larger than the hydrogen-bond distances for protons on either O3 or O4 sites. Figure S4.2c and S4.2d also show that the hydrogen-bond distance, as well as the $^1H\delta_{iso}$ and $^2H\ C_Q$ for protons on O3 or O4 sites are very similar and cover the same range, meaning no clear distinction between a protonated O3 or a protonated O4 site can be made based on the hydrogen-bond distances.

**Figure S4.2.** Plots of calculated hydrogen-bond length against (a and c) $^1H\delta_{iso}$ and (b and d) $^2H\ C_Q$ colored by (a and b) protonation environment (i.e., Mg–OH or Si–OH) and (c and d) protonation site for all 88 fully-optimized AIRSS-generated semi-hydrous wadsleyite structures.
Fully-hydrous wadsleyite models with Si$^{4+}$ vacancies were produced by changing the AIRSS protocol such that four H atoms were added to a unit cell of β-Mg$_2$SiO$_4$ with a single Si atom removed (producing 251 structures). This resulted in the enthalpy profile is shown in Figure S5.1a, with the structures falling into two well-defined regions, where structures ranked 1 to 142 have $0 \leq \Delta H_{Si} \leq 0.04$ eV and those ranked 143 to 251 have $0.27 \leq \Delta H_{Si} \leq 2.77$ eV ($\Delta H_{Si}$ is defined as relative to the lowest enthalpy Si-vacant structure). The set of structures with lower $\Delta H_{Si}$ contain all four H on the tetrahedron of O atoms that surround the Si vacancy. For the lowest enthalpy structure, at $\Delta H_{Si} = 0.0$ eV (Figure S5.1b), each O ($1 \times$ O2, $1 \times$ O3 and $2 \times$ O4) is protonated and involved in H-bonding interactions. It was found that, rather than each OH species receiving and donating a single H-bond, this ground-state structure exhibits 3 such species and one O4–H⋯O4H species that receives two hydrogen bonds, from O3–H and O2–H hydroxyls. The second O4–H hydroxyl donates a H-bond to O3, giving an O4–H⋯O3H interaction. Structures with higher $\Delta H_{Si}$ (ranked from 143 onwards) feature at least one H atom outside the vacant tetrahedron.

**Figure S5.1.** (a) Plot of $\Delta H_{Si}$ against structure rank for 251 AIRSS-generated Si-vacant fully-hydrous wadsleyite structures (3.3 wt% H$_2$O) and (b) the protonation arrangement found in the fully-optimized ground-state Si-vacant structure.
In order to compare hydration mechanisms featuring Si vacancies against those with Mg vacancies, a balanced isodesmic reaction was set up interchanging these materials. The formula for a unit cell of wadsleyite structures with Si vacancies (I) is Mg$_{16}$Si$_7$O$_{32}$H$_4$, with that of fully-hydrous Mg-vacant structures (II) given by Mg$_{14}$Si$_8$O$_{32}$H$_4$. Therefore, stoichiometric balance is achieved by addition of SiO$_2$ to I and 2 MgO to II, as in Equation S5.1. Polymorphs of SiO$_2$ and MgO, stishovite and akimotoite respectively, were chosen to reflect the temperature and pressure used under the conditions of hydrous wadsleyite synthesis (14-15 GPa, 1100-1200 °C).

\[
\begin{align*}
\text{Mg}_{14}\text{Si}_8\text{O}_{32}\text{H}_4 + 2\text{MgO} & \rightarrow \text{Mg}_{16}\text{Si}_7\text{O}_{32}\text{H}_4 + \text{SiO}_2 \\
\Delta H & = +1.16 \text{ eV}
\end{align*}
\]  

(S5.1)

Equation S5.1 shows that formation of Si-vacant fully-hydrous wadsleyite (and stishovite) from Mg-vacant fully-hydrous wadsleyite (and akimotoite) has an enthalpic penalty of $\Delta H = +1.16$ eV, indicating that Si vacancies are unlikely to occur experimentally in hydrous wadsleyite. Furthermore, synthetic conditions used to prepare samples of hydrous wadsleyite (~3 wt% H$_2$O) contained an excess of Si (mimicking mantle conditions), which shifts the equilibrium to the left. Based on the above, Si vacancies were discounted from the remainder of the present study.
S6. Choice of Mg site combinations for AIRSS study of fully-hydrous wadsleyite

To assess the structure(s) of $\beta$-$\text{Mg}_2\text{SiO}_4$ at 3.3 wt% $\text{H}_2\text{O}$ containing Mg vacancies, a list of eleven Mg site combinations was constructed (see Table 2 in the main text and Figure S6.1, below). The enthalpies of semi-hydrated wadsleyite structures suggested Mg3 vacancies were likely to produce the ground-state structure, so a focus was placed on this site, giving Mg3Mg3 site combinations at 2.8 Å (two series with the Mg3⋯Mg3 vector parallel to either the a or b directions), 4.0 Å, 5.0 Å and 7.1 Å, Mg1Mg3 site combinations at 2.9 Å and 6.4 Å and Mg2Mg3 site combinations at 5.1 Å and 6.3 Å. In addition, Mg1Mg1 and Mg2Mg2 site combinations, both with intervacancy distances of 7.6 Å, were considered. Each of these site combinations formed a starting geometry for an AIRSS search for candidate hydrous structures, wherein four $\text{H}^+$ were added (two $\text{H}^+$ within 3 Å) of each Mg vacancy and all atomic positions were relax within a fixed unit cell.

![Figure S6.1](image)

**Figure S6.1.** A unit cell of anhydrous wadsleyite ($\beta$-$\text{Mg}_2\text{SiO}_4$, *Imma*), showing Mg site vacancy combinations and corresponding intervacancy distances in Å.
The enthalpic stability of the theoretical structural model of hydrous wadsleyite containing 3.3 wt% H₂O put forward by Smyth, in comparison to the AIRSS-generated structures presented here, was investigated. This model, which was generated using simple ionic constraints, consists of protonated O1 sites, with the hydroxyls aligning parallel to the c axis, charge balanced by the removal of Mg2 site cations, as shown in Figure S7.1a. The geometry of this hydrous wadsleyite model was optimized prior to the prediction of the solid-state NMR parameters. From Figure S7.1b, which shows the relative enthalpy of this structure, alongside the 199 k-means clustered AIRSS-generated structures, the Smyth model is enthalpically unstable, with a relative enthalpy of 2.45 eV, far higher than any of the cluster-selected structures with two Mg2 site vacancies. The highly unstable nature of this model, relative to the AIRSS-generated structures, could be due to the hydroxyl orientation, which hinders the formation of favourable hydrogen bonding arrangements.
Figure S7.1. (a) Structural model of fully-hydrous wadsleyite put forward by Smyth.\textsuperscript{12} (b) Plot showing the relative enthalpy of the Smyth model (~2.4 eV) and the 199 AIRSS-generated fully-hydrous wadsleyite structures, colored according to the combination of Mg vacancies present.
S8. Additional NMR experiments

The plot of calculated \(^1\)H \(\delta_{\text{iso}}\) against \(^{29}\)Si \(\delta_{\text{iso}}\) for motifs G - J in Figure 9c (main text) shows four groups of points with reasonably well-defined regions of chemical shift, corresponding to the four different types of H···Si interactions. In contrast, the \(^1\)H-\(^{29}\)Si CP HETCOR spectrum published by Griffin et al.\(^{10}\) shows two regions of significant intensity, corresponding to MgOH (more specifically MgOH···SiO) and Si–OH environments, with no significant signal observed for the two additional sets of \(^1\)H-\(^{29}\)Si correlations shown in Figure 9c. However, lower intensity signals are present in this region of the spectrum, as seen in Figure S8.1, which shows three \(^1\)H-\(^{29}\)Si HETCOR spectra of fully-hydrated wadsleyite (containing ~3% wt water), acquired with three different CP contact times.

**Figure S8.1.** \(^1\)H-\(^{29}\)Si CP HETCOR spectra of \(\beta\)-Mg\(_2\)SiO\(_4\) (~3 wt% H\(_2\)O), acquired using CP contact times of (a) 1 ms, (b) 2 ms and (c) 5 ms. Cross sections, taken parallel to \(F_1\) at \(^1\)H \(\delta\) of ~3.5 ppm (top), ~6.5 ppm (middle) and ~8 ppm (bottom) are also shown.
Figure S8.2 shows $^1$H DQ MAS spectra of fully-hydrous wadsleyite (containing ~3% wt water), acquired using 1-3 loops of BABA dipolar recoupling. No significant differences in the cross peaks are observed, only in the absolute signal intensity.

Figure S8.2. $^1$H (14.1 T, 30 kHz) DQ MAS spectra of $\beta$-Mg$_2$SiO$_4$ (~3 wt% H$_2$O), acquired with (a) 1, (b) 2 and (c) 3 loops of rotor-synchronized BABA dipolar recoupling.
S9. Experimental FTIR and simulated IR spectra of fully-hydrous wadsleyite

The hydrous wadsleyite sample used in NMR experiments herein and by Griffin et al.\textsuperscript{10} was also studied using Fourier-transform infrared (FTIR) experiments, providing an opportunity to compare simulated FTIR frequencies and intensities from the set of AIRSS-generated structures (i.e., G-J) used to rationalize the NMR data with experimental measurements. FTIR spectra were recorded using a Bruker Hyperion 2000 microscope and Tensor27 spectrometer. Spectra were recorded using unpolarized light from clumps of powder loosely supported on a TEM grid. A square aperture equal to the size of the grid was used to define the acquisition area. Each spectrum represents the average of a large number of randomly oriented an inhomogeneously distributed crystals. The sample environment was purged with dry air. Spectra have a resolution of 2 cm\textsuperscript{-1} and are the average of 64 scans. Computed structures were first reoptimized, relaxing all atomic positions and lattice parameters, in CASTEP 8, using norm-conserving pseudopotentials, using GEOM\_ENERGY\_TOL and ELEC\_ENERGY\_TOL values of $1 \times 10^{-9}$ eV per atom to achieve highly-converged forces on all atoms. Reoptimized structures were then subjected to phonon calculations using the finite displacement method, with a displaced phonon k-point grid at 1 1 1. Simulated IR data were visualized using the dos.pl utility, with the application of 10 cm\textsuperscript{-1} Gaussian line-broadening. Figure S9.1a shows an overlay of the resulting spectra, representing models of pure crystals of fully-hydrous $\beta$-Mg\textsubscript{2}SiO\textsubscript{4} at 3.3 wt\% H\textsubscript{2}O with protonation motifs G-J. The blue line, arising from ground-state hydrous structure G, shows an intense feature at 3343 cm\textsuperscript{-1} which reproduces the most intense feature in the present experimental spectrum (see Figure S9.1b) and in other experimental spectra of hydrous wadsleyite.\textsuperscript{14-16}
Figure S9.1. (a) Overlay of IR spectra simulated from computed structures with motifs G-J showing the O–H region from 2500-4000 cm\(^{-1}\) with 10 cm\(^{-1}\) Gaussian line-broadening. (b) Overlay of simulated IR spectra with intensities adjusted using proportions of structures G-J suggested from \(^1\)H MAS NMR spectra and experimental FTIR spectrum of hydrous β-Mg\(_2\)SiO\(_4\) (~3 wt% H\(_2\)O).
S10. References

1. Smyth, J. R.; Hazen, R. M. The crystal structures of forsterite and hortonolite at several temperatures up to 900°C. *Am. Mineral.*, 1973, 58, 588-593.

2. Ashbrook, S. E.; Berry, A. J.; Wimperis, S. Three- and five-quantum $^{17}$O MAS NMR of forsterite Mg$_2$SiO$_4$. *Am. Mineral.*, 1999, 84, 1191-1194.

3. Ashbrook, S. E.; Berry, A. J.; Wimperis, S. $^{17}$O Multiple-Quantum MAS NMR Study of High-Pressure Hydrous Magnesium Silicates. *J. Am. Chem. Soc.*, 2001, 123, 6360-6366.

4. Ashbrook, S. E.; Berry, A. J.; Hibberson, W. O.; Steuernagel, S.; Wimperis, S. High-resolution $^{17}$O MAS NMR spectroscopy of forsterite ($\alpha$-Mg$_2$SiO$_4$), wadsleyite ($\beta$-Mg$_2$SiO$_4$), and ringwoodite ($\gamma$-Mg$_2$SiO$_4$). *Am. Mineral.*, 2005, 90, 1861-1870.

5. Horiuchi, H.; Sawamoto, H. $\beta$-Mg$_2$SiO$_4$: Single crystal X-ray diffraction study. *Am. Mineral.*, 1981, 66, 568-575.

6. Ashbrook, S. E.; Berry, A. J.; Hibberson, W. O.; Steuernagel, S.; Wimperis, S. High-Resolution $^{17}$O NMR Spectroscopy of Wadsleyite ($\beta$-Mg$_2$SiO$_4$). *J. Am. Chem. Soc.*, 2003, 125, 11824-11825.

7. Sasaki, S.; Prewitt, C. T.; Sato Y.; Ito, E. Single-crystal X ray study of $\gamma$ Mg$_2$SiO$_4$. *J. Geophys. Res.*, 1982, 87, 7829-7832.

8. Magi, M.; Lippman, E.; Samoson, S.; Engelhardt G.; Grimmer, A.-R. Solid-state high-resolution silicon-29 chemical shifts in silicates. *J. Phys. Chem.*, 1984, 88, 1518-1522.

9. Stebbins J. F.; Kanzaki, M. Local structure and chemical shifts for six-coordinated silicon in high-pressure mantle phases. *Science*, 1991, 251, 294-298.

10. Griffin, J. M.; Berry, A. J.; Frost, D. J.; Wimperis, S.; Ashbrook, S. E. Water in the Earth’s mantle: a solid-state NMR study of hydrous wadsleyite. *Chem. Sci.*, 2013, 4, 1523-1538.

11. Moran, R. F.; McKay, D. Pickard, C. J.; Berry, A. J.; Griffin J. M.; Ashbrook, S. E. Hunting for hydrogen: random structure searching and prediction of NMR parameters of hydrous wadsleyite. *Phys. Chem. Chem. Phys.*, 2016, 18, 10173-10181.

12. Sturniolo, S. *Soprano — a library developed by the CCP for NMR Crystallography*, https://ccp-nc.github.io/soprano/.
13. Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütz, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The Atomic Simulation Environment — A Python library for working with atoms, J. Phys.: Condens. Matter, 2017, 29, 273002.

14. Deon, F.; Koch-Muller, M.; Rhede, D.; Gottschalk, M.; Wirth, R.; Thomas, S.-M. Location and quantification of hydroxyl in wadsleyite: New insights. Am. Mineral., 2010, 95, 312-322.

15. Jacobsen, S. D.; Demouchy, S.; Frost, D. J.; Balloran, T. B.; Kung, J. A systematic study of OH in hydrous wadsleyite from polarized FTIR spectroscopy and single-crystal X-ray diffraction: Oxygen sites for hydrogen storage in Earth’s interior. Am. Mineral., 2005, 90, 61-70.

16. Kohn, S. C.; Brooker, R. A.; Frost, D. J.; Slesinger A. E.; Wood, B. J. Ordering of hydroxyl defects in hydrous wadsleyite (β-Mg$_2$SiO$_4$). Am. Mineral., 2002, 87, 293-301.