Clathrate structure determination by combining crystal structure prediction with computational and experimental $^{129}$Xe NMR spectroscopy

Marcin Selent,‡ Jonas Nyman,‡ Juho Roukala, Marek Ilczyszyn, Raija Oilunkaniemi, Peter J. Bygrave, Risto Laitinen, Jukka Jokisaari, Graeme M. Day,* and Perttu Lantto*†

We present an approach for the structure determination of clathrates using NMR spectroscopy of enclathrated xenon to select from a set of predicted crystal structures. Crystal structure prediction methods have been used to generate an ensemble of putative structures of $\alpha$- and $m$-fluorophenol, whose previously unknown clathrate structures have been studied by $^{129}$Xe NMR spectroscopy. The high sensitivity of the $^{129}$Xe chemical shift tensor to the chemical environment and shape of the crystalline cavity makes it ideal as a probe for porous materials. The experimental powder NMR spectra can be used to directly confirm or reject hypothetical crystal structures generated by computational prediction, whose chemical shift tensors have been simulated using density functional theory. For each fluorophenol isomer we find one predicted crystal structure whose measured and computed chemical shift tensors agree within experimental and computational error margins and these are thus proposed as the true fluorophenol xenon clathrate structures.

‡ The authors have made equal contributions to this article.
1 Introduction

Over the past few years, the combined use of solid state nuclear magnetic resonance (NMR) and computational methods has developed into a practical method for determining crystal structures.\[1\] This is due to the sensitivity of NMR chemical shifts to the molecular environment in a crystal, and the reliability of methods for predicting the relative NMR shielding of atoms in different environments. The approach involves first predicting the full set of low energy crystal structures available to a molecule, followed by chemical shielding calculations, which are matched against the measured chemical shifts of a material. Applications to organic crystals have found that, given a set of predicted crystal structures, the isotropic $^1$H chemical shifts are usually sufficient to identify which of the predicted structures corresponds to the material under investigation.\[2\]

In 1972–73 Kazankin et al.\[3,4\] reported that xenon clathrates can be formed from several monosubstituted phenol compounds. With the exception of phenol,\[5\] hydroquinone,\[5–8\] $p$-cresol\[9\] and $p$-fluorophenol,\[10\] Xe clathrates have not been described in the English literature and have remained relatively unknown in the West. The crystal structures of $o$- and $m$-fluorophenol xenon clathrates have, to our knowledge, never previously been determined. The NMR properties of Xe included in porous structures give a potentially sensitive probe of clathrate structure, which we explore here as a means of determining the structures of the $o$- and $m$-fluorophenol xenon clathrates.

The first $^{129}$Xe NMR experiments were performed in 1951,\[11\] but only a modest number of further studies appeared until the early 1980s when Ripmeester and Davidson investigated enclathrated xenon\[12\] and Ito and Fraissard proposed the use of $^{129}$Xe NMR for probing the properties of zeolites.\[13\] Today, the application range of $^{129}$Xe NMR extends from materials property studies of micro- and mesoporous solids via polymers, liquid crystals, proteins and biosensors to magnetic resonance imaging (MRI) for medical purposes.\[14–17\]

The reason for using xenon to probe the properties of various materials arises from the high sensitivity of the shielding of $^{129}$Xe to its local environment, and that the variability in the shielding stems exclusively from environmental effects.\[18–20\] Another reason is the good NMR receptivity of $^{129}$Xe, which is about 33 times that of $^{13}$C. A disadvantage in some cases may be the long $^{129}$Xe spin-lattice relaxation time $T_1$ which is, for example, several minutes in xenon clathrate hydrate.\[12\] The application range of $^{129}$Xe NMR and MRI further widened when the so-called optical pumping or hyperpolarization technique was invented,\[21\] enabling the increase of the $^{129}$Xe polarization by up to four or five orders of magnitude and making experiments possible with very small amounts of xenon gas.

The work of Ripmeester and Davidson\[12\] revealed the potential of $^{129}$Xe NMR in studies of clathrates. Namely, the $^{129}$Xe NMR spectrum of xenon clathrate hydrate consists of two broad resonance lines: one at 152 ± 2 ppm and the other at 242 ± 2 ppm downfield (smaller shielding) from the low pressure gas peak. The former resonance arises from xenon in large cages, and the latter from xenon in small cages. Thus, the $^{129}$Xe chemical shift relates to the size of cages.\[22\] Furthermore, the resonance at 152 ppm displays a cylindrically symmetric powder lineshape with a chemical shift anisotropy of 32 ± 3 ppm; the resonance lineshape relates to the shape of the cage accommodating xenon atoms. The line intensities in turn relate to the occupation number. An illustrative example is the distribution of xenon atoms in the alpha cages of the NaA zeolite.\[23\] The $^{129}$Xe NMR spectrum of xenon in NaA consists of several distinct signals, the chemical shift being determined by the occupation number of a cage.
Modeling of $^{129}$Xe NMR in clathrate cages dates back to the studies of chemical shifts\cite{24} and NMR line shapes\cite{25} inside clathrate hydrate structures I and II, which were modeled as molecular clusters extracted from the clathrate structures. Similar quantum chemical cluster models were also applied in studies of Xe inside fullerene cages, \cite{26,27} for which both relativistic\cite{28,29} as well as dynamical and environmental\cite{29,30} effects were studied. Current cluster modeling makes use of recent advances in the development of relativistic quantum chemistry methods, which have enabled very demanding studies of large heavy-element systems such as cryptophanes\cite{31} and self-organizing metallo-supramolecular cages.\cite{32}

Diffraction is the usual method for crystal structure solution. However, structure determination of clathrates by diffraction methods can be hindered by their instability. For the materials studied here, we are unable to obtain a powder X-ray diffraction (PXRD) pattern from the $m$-fluorophenol clathrate and could only obtain a low resolution diffraction pattern from the $o$-fluorophenol clathrate, from which structure determination would not be possible. Therefore, we cannot combine $^{129}$Xe NMR with PXRD experiments as usually done for identifying new clathrate phases.\cite{33,34} However, the exceptional sensitivity of xenon’s chemical shift anisotropy to its environment should enable the distinction between different candidate clathrate structures. We investigate this hypothesis by comparing experimentally observed and quantum chemically modeled $^{129}$Xe NMR isotropic and anisotropic shift parameters in clathrate structures obtained by computational crystal structure prediction (CSP).

CSP has until recently been focused primarily on predicting the single thermodynamically stable structure, or possibly a few low-energy polymorphs. The lowest energy crystal structures available to a given molecule are, in all but very rare exceptions,\cite{35} close-packed, leaving no room for the inclusion of guest molecules. There have been few reports of the prediction of porous molecular crystals\cite{36-38} or solvates.\cite{39-41} The empty host molecule frameworks of observed inclusion structures have been shown to often exist as local minima on the lattice energy surface, albeit sometimes at relatively high energies compared to close-packed alternative structures.\cite{41-43} This suggests an efficient approach to the discovery of inclusion structures: searching for stable, empty frameworks using CSP methods and subsequently inserting the guest. This method should be particularly suited to weakly interacting guests such as Xe, where inclusion is expected to leave the host framework relatively unperturbed. The approach should be more efficient than searching the dramatically larger multi-component phase space defined by the host and guest together.

In this study, we have performed CSP calculations for $o$- and $m$-fluorophenol with the specific aim to predict realistic xenon clathrate complexes for which calculated $^{129}$Xe NMR parameters can be compared to measured spectra. Advanced first principles density functional theory (DFT) electronic structure calculations of the $^{129}$Xe NMR shielding tensors, with a proper inclusion of electron correlation as well as relativistic, periodic, and dynamical effects, have been carried out for a set of predicted structures. A comparison of the simulated isotropic and anisotropic Xe chemical shift parameters with experimentally observed solid state $^{129}$Xe NMR data allows us to identify the crystal structure of these elusive clathrates.
2 Experimental section

2.1 Experimental NMR

2.1.1 Preparation of samples

Clathrate samples were prepared from commercially available o- and m-fluorophenol (Aldrich, 98% assay) and used directly without purification. The substances were transferred into pyrex glass tubes (4 mm outer diameter, 0.8 mm walls) and connected to a volume calibrated vacuum line. Two cycles of freeze-thaw were applied to reduce oxygen content. Isotopically enriched $^{129}$Xe gas (Chemgas, 89% enrichment) was then frozen into the evacuated tubes. The amount of transferred gas was controlled by the pressure drop in the vacuum line with a digital pressure gauge (1 hPa precision). The amount of gas inserted into the tubes was chosen to be sufficient for saturation of a 3:1 host-guest clathrate stoichiometry and with excess to pressurize the sample tube so as to maintain clathrate structural stability at the desired experimental temperature of 250 K. The sample glass tubes were flame sealed, cooled in liquid nitrogen and equilibrated for a period of two weeks at ca 243 K prior to NMR measurements.

2.1.2 NMR experiments

The $^{129}$Xe NMR spectrum of xenon in m-fluorophenol was measured at 251 K on a Bruker DSX300WB spectrometer ($^{129}$Xe Larmor frequency 83.03 MHz) using a 7 mm variable angle spinning (VAS) probe head (DOTY Scientific, Inc., USA), without spinning. The static powder spectrum (the axis of the solenoid coil was set perpendicular to the external magnetic field) was observed applying cross-polarization (CP) and proton decoupling. The following acquisition parameters were used: $^{129}$Xe pulse width 5 µs, $^1$H decoupling pulse width 16 µs, mixing time 5.4 ms, repetition time 35 s, strength of the $^1$H decoupling field 34 kHz, and number of collected free induction decay (FID) signals 220. The aim of applying proton decoupling was to diminish the effect of the $^{129}$Xe–$^1$H dipolar coupling on the line width. The $^{129}$Xe chemical shift was measured relative to the isotopic $^{129}$Xe chemical shift in hydroquinone, in which the shift relative to zero-pressure xenon is known to be 222.1 ppm. Temperature calibration was based on the measurement of $^1$H chemical shift difference in an separate methanol sample. Prior to Fourier transformation the FID signal was multiplied by an exponentially decaying apodization function leading to $\simeq 100$ Hz line broadening.

The $^{129}$Xe NMR spectrum of xenon in o-fluorophenol was in turn measured at 253 K on a Bruker DPX400 spectrometer ($^{129}$Xe Larmor frequency 110.70 MHz) using a 5 mm broad band observe (BBO) probe head. $^{129}$Xe chemical shift is given with respect to the signal of zero-pressure xenon gas, which was determined using two samples with known xenon pressure and extrapolation using the second virial coefficient. Temperature calibration was in this case performed using the chemical shift difference in a separate methanol sample. Prior to Fourier transformation, the FID signal was multiplied by an exponentially decaying apodization function leading to 50 Hz line broadening.

The elements of the chemical shift tensors were determined in both cases using DMFIT. The uncertainties were estimated to be ±0.2 ppm, ±0.4 ppm and < 0.04, respectively, in the three adjusted NMR parameters: isotropic chemical shift (CS) $\delta \equiv \delta_{iso} = \sigma_{iso} - \sigma_{ref}$ referenced to the zero-pressure limit Xe gas, chemical shift anisotropy $\Delta \delta = \delta_z - (\delta_{xx} + \delta_{yy})/2$ (CSA) with positive/negative values for prolate/oblate
spheroids along the z-direction and asymmetry parameter $\eta = (\delta_{yy} - \delta_{xx})/(\delta_{zz} - \delta_{iso})$ with the value 0 for an axially symmetric tensor (with respect to z) and 1 for the fully asymmetric case, when $\delta_{yy} = \delta_{iso}$. The isotropic Xe nuclear shielding constant is the trace of the shielding tensor, $\sigma_{iso} = (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3$. The components of the principal axis system (PAS) of the Xe shift tensor follow Haeberlen’s convention:[46] $|\delta_{zz} - \delta_{iso}| \geq |\delta_{xx} - \delta_{iso}| \geq |\delta_{yy} - \delta_{iso}|$. A relatively large contribution to the uncertainties arises from the wavy background in the experimental spectra, which was therefore eliminated.

2.1.3 Powder X-ray diffraction (PXRD) experiments

Samples of both o- and m-fluorophenol were prepared in glass capillaries. Xenon gas was frozen into the evacuated capillaries, which were flame-sealed after submerging in liquid N$_2$. Capillaries containing the samples were kept immersed in liquid N$_2$ prior to mounting in a pre-cooled single crystal diffractometer. Diffraction data were collected at 100 K using graphite monochromated Mo Kα radiation.

2.2 Crystal structure prediction (CSP)

CSP calculations involved five steps: i) the lowest energy conformations of the isolated molecules were identified; ii) hypothetical crystal packings were generated and then lattice energy minimized using a simple force field and rigid-molecule constraints; iii) a subset of the lowest energy predicted crystal structures were re-optimized using an anisotropic atom–atom force field; iv) intramolecular flexibility was introduced, allowing the hydroxyl group to reorient in response to packing forces in each low energy crystal structure; v) porous structures were identified, xenon atoms inserted into the pores and the structures once again re-optimized.

Molecular geometries, energies and charge densities, calculated by DFT using GAUSSIAN 09[47] with the B3LYP[48,49] functional and 6-311G(d,p) basis set, were used throughout the CSP calculations. Two stable, planar conformers of each molecule were identified (Fig. 1) and DFT predicts that their energies are sufficiently close that either could form low energy crystal structures.[50] Therefore, both conformers of both molecules were included in the CSP study.

Hypothetical crystal structures were generated with the rigid DFT-optimized molecular geometries using Monte Carlo simulated annealing[51] with MATERIALS STUDY. Searches were performed in the most commonly observed space groups of known molecular organic crystal structures; 25 space groups were searched with one molecule in the asymmetric unit ($Z' = 1$) and 5 space groups with $Z' = 2$, including all combinations of the two conformers in the asymmetric unit.

Structures with a lattice energy within a 15 kJ/mol window from the lowest energy structure were further refined with an anisotropic, atomic multipole-based intermolecular atom–atom potential model,[52] combined with a DFT treatment of intramolecular energies and geometries. Duplicate crystal structures were removed using the COMPACK[53] program and the resulting unique structures were re-optimized using the CRYSTAL OPTIMIZER[54] program to treat flexibility of the hydroxyl group within each crystal structure. Full details of the conformational analysis and CSP methods are provided in the ESI†.
2.3 Selection of likely clathrate host structures

Lattice energy differences between polymorphs are usually very small and rarely exceed 8 kJ/mol.\[^{55}\] However, since voids in crystal structures are thermodynamically unfavourable,\[^{42,56}\] structures within a larger energy range than in the usual application of CSP to polymorph prediction were considered as putative inclusion frameworks. We considered crystal structures within an energy cutoff of 13 kJ/mol above the global minimum for each molecule.

To guide our selection of potential clathrate host frameworks, we examined the guest-to-host volume ratio \(R_g\) in 31 representative clathrate structures taken from the CSD (see ESI\[^{†}\] for details). The volume ratio \(R_g\) is calculated as

\[
R_g = 100 \cdot \frac{V_g}{V_H} \quad [\%],
\]

where \(V_g\) is the van der Waals volume of the guest molecule and \(V_H\) is an individual void’s contact volume.\[^{57}\]

One xenon atom was inserted at each cavity’s centroid coordinates and the resulting xenon clathrate structures were geometry-optimized as described above, assuming rigid molecules, imposing no space group symmetry and using an ad hoc exp-6 potential for xenon (see ESI\[^{†}\]).

2.4 \(^{129}\text{Xe}\) NMR shielding calculations

The three \(^{129}\text{Xe}\) NMR parameters of the predicted structures were modeled and compared with the experimental data in several stages. In NMR modeling the chemical shift reference is a free Xe atom. In the first step, all predicted clathrate structures were subjected to nonrelativistic (NR) NMR modeling of cluster models. The results from the NR cluster models were used to identify likely candidates for the experimentally observed clathrate structures. These candidates were then further studied using DFT calculations on their fully periodic models. Finally, a few crystal structures that agreed best with experimental NMR results were chosen for more detailed modeling. Full details are given in the ESI\[^{†}\].

2.4.1 Screening of structures by cluster modeling

The NMR parameters of probable clathrate structures were first modeled using clusters consisting of a single xenon-occupied cavity, including the Xe atom and all nearby fluorophenol molecules. Nonrelativistic Xe shielding tensor calculations were performed on the cluster models using TURBOMOLE.\[^{58}\] The BHandHLYP\[^{59,60}\] hybrid functional including 50% of exact Hartree-Fock (HF) exchange (EEX) was chosen based on benchmark calculations on the xenon-benzene system, where high quality nonrelativistic \textit{ab initio} Xe chemical shifts are reproduced reasonably well.\[^{29}\] Likewise, BHandHLYP has been shown to slightly underestimate Xe chemical shift and anisotropy, leaving room for improvements in the modeling by approaching the experimental values from below for Xe-containing molecules.\[^{61-63}\] Xe atoms moving freely inside buckminster fullerences\[^{29,30}\] and in self-organizing metallo-supramolecular cages.\[^{32}\] Those studies show that the typical overestimation of Xe CS and CSA by pure generalized gradient approximation (GGA) DFT functionals, such as PBE\[^{64}\] and BLYP,\[^{59,65}\] is partly compensated when the amount of exact HF exchange is increased in B3LYP (EEX=20%) and BHandHLYP (EEX=50%) hybrid functionals. Calculations have been performed with a mixed basis set denoted as MHA/SVP (see ESI\[^{†}\] for details).
2.4.2 Periodic modeling of likely candidates

Five $o$-fluorophenol and seven $m$-fluorophenols crystal structures with screening level NMR parameters close to the experimental values were chosen for further scrutiny (vide infra). These structures were optimized with respect to both atomic positions and lattice parameters with planewave, periodic DFT with the PBE functional and Tkatchenko–Scheffler (TS) dispersion correction,\(^{66}\) denoted as PBE-TS structures from now on.

Following optimization, NMR shielding tensors were computed with the PBE functional using the gauge-including projector augmented wave (GIPAW) method\(^ {67,68}\) (see Figs. 6 and 7). These periodic DFT results include the scalar relativistic (SR) effects on Xe shielding at the 1-component zeroth-order regular approximation (ZORA)\(^ {69,70}\) level of theory. Details of the periodic calculations with CASTEP\(^ {71,72}\) are in the ESI†.

2.4.3 Detailed modeling of the most probable structures

After periodic NMR calculations of the DFT optimized structures, the most probable clathrate structures were selected for more detailed DFT NMR modeling using the Amsterdam Density Functional\(^ {73,74}\) (ADF). In addition to the NR and SR-ZORA quantum chemistry, ADF provides 2-component spin–orbit SO-ZORA method including also SR effects.\(^ {69,75}\)

As a much wider range of DFT functionals that is currently available in periodic CASTEP can be used in ADF calculations performed with the jepl/TZP mixed basis set\(^ {76}\) (see ESI† for details) on clusters comprising a single cavity, the influence of correlation treatment and the amount of EEX on NMR parameters was tested using PBE, BLYP, B3LYP, and BH\(\text{and}H\)LYP functionals at SR-ZORA level. By using the results in scaling of the periodic PBE results, we obtain estimates of SR periodic BLYP, B3LYP and BH\(\text{and}H\)LYP results. We expect that, as the electron correlation description is improved by an increasing portion of EEX, the calculated NMR parameters for the true crystal structures should approach the experimental results. Hence, the BH\(\text{and}H\)LYP-scaled periodic PBE results provide the best estimation of $^{129}$Xe NMR parameters and lack only contributions due to molecular dynamics and relativistic SO effects. The latter we treat as an additive correction obtained from the difference of the static SR- and SO-ZORA cluster calculations with the BH\(\text{and}H\)LYP functional.

The effect of Xe dynamics at $T = 300$K for $^{129}$Xe NMR shielding parameters was modeled for the few most probable clathrates by canonical NVT Metropolis Monte Carlo (MC-NVT) of Xe motion on a potential energy surface inside a fixed cluster cavity with fixed PBE-TS optimized geometry. The temperature effects on Xe chemical shift and anisotropy were calculated as the difference between the thermally averaged shielding tensor and the reference tensor with Xe at the center of the cage.

3 Results and discussion

3.1 Experimental NMR results

Both isomers formed clathrates with xenon under the experimental conditions and allowed recording of properly shaped powder patterns without active mixing and crushing of crystals during their formation. A powder-like appearance and lack of macroscopic crystallites has been confirmed by visual inspection and fit to spectra. There is no preferential growth imposed by tube walls as this would not allow for proper
random averaging of orientations of crystallites in relation to the magnetic field of the spectrometer.

Axially symmetric powder patterns have been observed for both clathrates (Figure 2), which confirms the existence of highly symmetric voids occupied by xenon atoms. High values of massfluorophenol and mass-chemical shift for both clathrates, \( \delta = 228.5 \) ppm for m-fluorophenol and \( \delta = 256.0 \) ppm for o-fluorophenol, suggest tight cages with only one occupational site. In such cases, where \(^{129}\)Xe NMR powder patterns are expected to reflect only interactions of xenon atoms with the immediately adjacent host molecules, the influence of neighbouring cages with xenon is expected to be negligible.\(^{[77]}\)

The spectrum of m-fluorophenol (Figure 2) resembles that of \( \beta \)-hydroquinone,\(^{[8]} \) \( \beta \)-phenol and p-fluorophenol with one notable difference; it has the largest observed value of \(^{129}\)Xe NMR CSA among known inclusion compounds: \( \Delta \delta = 183.3 \) ppm vs. 161.9, 171 and 164 ppm for hydroquinone,\(^{[78]} \) phenol\(^{[5]} \) and p-fluorophenol,\(^{[79]} \) respectively. Small additional features in the observed spectrum of m-fluorophenol are attributed to amorphous host substance and xenon atoms adsorbed in this amorphous phase.

The \(^{129}\)Xe NMR of o-fluorophenol shows negative CSA (Figure 2), \( \Delta \delta = -47.5 \) ppm, in contrast to all previously known solid state \(^{129}\)Xe NMR powder patterns of phenol clathrates for which positive CSA is observed.\(^{[15-10]} \) Changes in the sign of the CSA can be observed in a group of porous channel-like dipeptides,\(^{[80]} \) where CSA changes from positive to negative with increasing xenon gas pressure and therefore with pronounced \( \text{Xe} \cdots \text{Xe} \) interactions playing a greater role. No changes of the NMR spectrum of o-fluorophenol clathrate have been observed when repeating the experiment with samples of different pressures. The small negative CSA may result from an axially symmetric but oblate environment around the xenon atom in this clathrate.

### 3.2 Powder X-ray diffraction results

A PXRD pattern could not be obtained from m-fluorophenol clathrate, which transformed too quickly to the known high density form when transferred to the diffractometer. This is confirmed by comparison of the measured PXRD pattern to that simulated from the known crystal structure. For o-fluorophenol, we were able to obtain a broad powder pattern that does not correspond to either known crystal structures (see ESI†). The pattern is too broad to index, so it could not be used to determine the structure.

### 3.3 Crystal structure prediction results

CSP resulted in an exceptionally large number of low energy crystal structures for each molecule (Figures 3 and 4).

For o-fluorophenol, \( Z' = 2 \) crystal structures with both molecules in the cis-conformation are energetically favoured. The known\(^{[81]} \) high-pressure polymorph II (CSD refcode QAMWEH01) is located in the search, 0.8 kJ/mol above the global minimum (Figure 3) and with a slightly higher density, as expected for a high pressure polymorph. The known low-temperature polymorph I\(^{[81]} \) (CSD refcode QAMWEH) is a disordered \( Z' = 1.5 \) structure and could not be predicted with the methods employed here.

One crystal structure of m-fluorophenol is known\(^{[81]} \) (CSD refcode QAMTUU) and corresponds to the global lattice energy minimum from the search. The structure is geometrically well reproduced (with a deviation in atomic positions in a 15-molecule cluster taken from predicted and X-ray crystal structures of RMSD\(_{15} = 0.203 \) Å).
Overlays of both matches are included in the electronic supporting information (ESI†). The successful reproduction of the known crystal structures of both molecules amongst the lowest energy predicted structures provides confidence in the sampling of crystal structures and of their lattice energy rankings.

3.4 Selection rules for clathrate structures

Since the inclusion of guest molecules in clathrates can significantly stabilize the structure, using the lattice energy of the empty host alone is not useful for selection of promising structures. Our analysis of known clathrates found that the guest-to-host volume ratio $R_g$ is normally distributed with mean 59% and a standard deviation of 8 percentage units, in good agreement with Rebek’s "55% solution"[82] and providing further evidence for the empirical rule[82,83] that $R_g$ in observed inclusion structures should fall in this limited range.

Of the predicted crystal structures within 13 kJ/mol of the global minimum for o-fluorophenol and m-fluorophenol, 230 and 223 are porous (having cavities $> 10\ \text{Å}^3$). 33 (o-fluorophenol) and 32 (m-fluorophenol) of the predicted structures have cavities of a suitable volume to accommodate xenon as a guest, which we took as those with $R_g$ within three standard deviations ($\pm 3\sigma$) for observed clathrates. These structures are encircled in red in Figs 3 and 4. The lowest energy of these structures are 4.5 and 6.7 kJ/mol above the global minima for o- and m-fluorophenol, respectively.

3.5 Computational NMR results

3.5.1 Screening of structures by cluster models

Examples of cluster models used for initial screening are shown in Fig. 5. As can be seen in Figures 6 and 7, even for a set of structures with similar void volumes and packing energies, the computed NMR tensor parameters for $^{129}$Xe vary widely in our initial screening using cluster models of the predicted structures. This demonstrates the exceptional sensitivity of the $^{129}$Xe NMR parameters to small differences in cavity size and shape.

Calculated NMR parameters for the 33 o- and 32 m-fluorophenol clathrate structures obtained with nonrelativistic DFT cluster models are tabulated in the ESI†, see Tables S4 and S5. Many of these structures, including all predicted $Z' = 2$ structures, have chemical shift tensors with significant asymmetry; these voids are clearly not compatible with the experimental spectra, in which $\eta < 0.04$. At this stage, only the structures with axially symmetric chemical shift tensors (green squares in Figures 6 and 7) are kept as possible candidates for the experimental structures.

Optimisation of the clathrate structure after insertion of the xenon into the host allowed the structures to relax in response to guest insertion. Despite very small structural changes, this significantly affected calculated NMR parameters, see Figure S7 in the ESI†.

Since the screening level of theory is expected to underestimate both CS and CSA by a few tens of ppm, we focused on structures with both properties of the correct sign and underestimated, also including a few structures with overestimated CS for m-fluorophenol and slightly positive CSA for o-fluorophenol for higher level calculations. This excluded all but five o-fluorophenol and seven m-fluorophenol structures; hereafter, we will refer to these as structures oF_A to oF_E (in order of increasing energy, Fig. 3) and mF_A to mF_G (Fig. 4). All of these candidate clathrate structures
have space group symmetry $R\bar{3}$.

3.5.2 NMR chemical shift parameters from periodic DFT calculations

NMR shielding tensors were calculated by periodic DFT for the five plausible $o$-fluorophenol and seven $m$-fluorophenol clathrate structures. We found that periodic modeling is necessary and that the geometry relaxation with periodic PBE-TS shifts the CS and CSA towards larger magnitudes, corresponding to more elliptic, prolate ($\Delta\delta > 0$) and oblate ($\Delta\delta < 0$) $m$- and $o$-fluorophenol cavities, respectively.

NMR parameters from SR GIPAW/PBE calculations at the PBE-TS optimized geometries are tabulated in Tables S6 and S7 in ESI† as well as displayed in Figures 6 and 7 as red diamonds connected with arrows to the corresponding screening cluster result (green squares). Here, we recall that the periodic results set upper limits on $^{129}$Xe CS and CSA, as the PBE functional (and all pure GGA DFT functionals) overestimates the magnitudes of both parameters.\cite{29,32,61–63} Although the disagreement in CS and CSA of some of the candidate structures now makes them unlikely (such as $oF_C$, $oF_E$ and $mF_G$), all were kept for further analysis with more detailed NMR calculations.

3.5.3 Results of detailed NMR calculation of the most likely candidates

Typically, the crystal lattice effect, i.e. the difference of $^{129}$Xe NMR parameters between the periodic and cluster models of the same PBE-TS optimized clathrate, increases both CS and CSA as seen in Tables 1 and 2. However, the extent of the change is very case-specific and unforeseeable, which makes periodic modeling essential, in one form or another, in the present search of a best candidate of an unknown clathrate structure.

The ADF/SR-ZORA results for PBE-TS optimized cluster models were used to extract correction factors (see details in ESI†) by which the periodic, SR GIPAW/PBE results were scaled in order to obtain periodic estimates of $^{129}$Xe NMR parameters with BLYP, B3LYP and BHandHLYP functionals. The estimated periodic data for the five and seven candidates of $o$-fluorophenol and $m$-fluorophenol clathrates are listed in Tables 1 and 2 as well as displayed in Figures 8 and 9, respectively. The best BHandHLYP-scaled periodic estimates bring one clathrate structure for each of the fluorophenol isomers into excellent agreement with the experimental NMR: $oF_D$ and $mF_A$. For $o$-fluorophenol, the other CSP candidates in the PBE-TS crystal geometry have quite different periodic BHandHLYP $^{129}$Xe CS and/or CSA values as compared to the experimental ones, while for $m$-fluorophenol the closest alternative candidates overestimate either the CS ($mF_F$) or CSA ($mF_B$). We expect the BHandHLYP functional to provide reasonable approximation for both quantities due to benchmarking against ab initio calculations.\cite{61–63}

In addition, the relativistic SO correction, obtained as the difference between SO- and SR-ZORA cluster calculations with ADF (see Tables 1 and 2), is added to the static BHandHLYP results of all structures (Figures 8 and 9). Due to its different physical origin, the SO effect is case-specific and may either increase or decrease the CS and CSA, while the latter is affected slightly more. The magnitude of the SO correction is, however, smaller than the dynamical correction (see below) and, hence, does not alter the identification of the best clathrate candidates.

The temperature effect of Xe dynamics at $T = 300$K was modeled for the two ($oF_B$ and $oF_D$) and four ($mF_A$, $mF_B$, $mF_E$, and $mF_E$) most relevant clathrate candidates. The dynamical (DYN) correction is added to the SO corrected BHandHLYP
results (BHandHLYP + SO). In all cases, the thermal averaging increases the magnitudes of both CS and CSA of $^{129}$Xe. As seen in Figures 8 and 9, the inclusion of DYN correction confirms the most probable clathrate structure of both isomers. The effect on the simulated spectra is also shown in Figure 2.

It is evident from Figures 8 and 9 (see also Tables S6 and S7 in ESI†) that in all probability the crystal structures for $\alpha$- and $m$-fluorophenol clathrates are $\alpha$F_D and mF_A. For them, the $^{129}$Xe NMR isotropic chemical shift and chemical shift anisotropies at different levels of theory encompass the experimental data. For both structures, the experimental NMR parameters are approached as the computational level is improved. The approximated periodic BHandHLYP results at PBE-TS optimized structures only slightly underestimate the chemical shift, leaving room for improvements by dynamics and relativistic SO corrections. Astonishingly, the most accurate level of theoretical modeling almost quantitatively reproduces the two NMR parameters that are clearly specific for a given clathrate. The remaining differences between experimental and computational data may be attributed to deficiencies in the structure as well as treatments of electron correlation and thermal averaging of the whole system.

Our PXRD results provide further evidence for the CSP,$^{129}$Xe NMR determined structure of the $\alpha$-fluorophenol clathrate; the simulated diffraction pattern from $\alpha$F_D is similar to the PXRD obtained from the clathrate sample (see Figure S13 in ESI†). However, we cannot use PXRD of the present quality to unambiguously distinguish between the candidates structures.

3.6 Description of the proposed clathrate structures

The crystal structures that we propose as the fluorophenol xenon clathrates belong to space group $R3$, with lattice parameters shown in Table 3. Both structures have 3-fold screw axes parallel to the c lattice vector, resulting in 6-membered rings of host molecules held together with strong hydrogen bonds formed by practically ideal OH···O interactions. The hydrogen bonding forms $R_6^6(12)$ graphs,$^{[84]}$ which seems to be a characteristic feature in clathrates of phenol derivatives. Packing diagrams are displayed in Figures 10 and 11. CIF files are included in the ESI†. Both structures have three cavities per unit cell, with volumes 61.4 and 88.5 Å$^3$ each for $\alpha$- and m-fluorophenol respectively, resulting in $R_g$ values of 68.7 and 47.7% for xenon, both within 1.5 standard deviations from the ideal ratio of 59%.

Despite the same hydrogen bonding and crystal symmetry, the voids in the two structures are of quite different geometry. The void in $\alpha$F_D has an oblate shape with ratios of the principal moments of the free volume (calculated from a grid sampling in PLATON)$^{[85]}$ of 1.00:1.00:0.85, with the short dimension oriented along the crystallographic c-axis. In contrast, mF_A has a prolate shape with calculated ratio of dimensions of 1.00:1.00:1.20, elongated along the crystallographic c-axis. Thus, we find that the sign of the CSA relates to the shape of the void, as expected, with an oblate cavity leading to a negative CSA and prolate cavity yielding a positive CSA.

The proposed structure of the $\alpha$-fluorophenol Xe clathrate ($\alpha$F_D) is one of the highest energy and one of the densest structures that we considered as possible clathrates (Figure 3). The stabilisation of this structure must derive from the interactions of Xe with the host structure, which should be large, due to the tight fit of Xe to the host cavities.

In contrast, the most likely structure of the m-fluorophenol Xe clathrate (mF_A) corresponds to the most stable structure on the crystal structure landscape (Figure 4).
that contains cavities suitable for Xe enclathration; this structure’s stability relates, in part, to the stability of the host framework.

4 Conclusions

We have used detailed first principles NMR calculations on candidates from CSP to propose structures for the xenon clathrates of ortho- and meta-fluorophenol. The exceptional sensitivity of the $^{129}\text{Xe}$ chemical shift tensor to its local environment allows comparisons between observed and calculated NMR chemical shift parameters that have been used to directly confirm or reject hypothetical clathrate structures. Based on these comparisons, we propose likely crystal structures for the two clathrates. The proposed structures strongly resemble a previously known β-hydroquinone xenon clathrate $^{[6]}$ and have similar $R^0_6(12)$ hydrogen bonding motifs.$^{[84]}

The unusual ortho-fluorophenol $^{129}\text{Xe}$ NMR powder spectrum – with its negative CSA – was initially thought to suggest a structural motif different from the known and common $R^0_6(12)$ hydrogen-bonded double sandwich. Our results, however, confirm that the $R^0_6(12)$ motif is an important feature for clathrates of phenol derivatives that is present in both materials studied here.

Static solid-state NMR has clear benefits over high-frequency magic-angle spinning NMR, that would only result in isotropic chemical shifts rather than the complete chemical shift tensors. In this study, using isotropic shifts only would not have allowed us to identify the experimental clathrate structures among the predicted candidates.

We believe that the method presented here is a powerful approach for structure determination of porous materials, which is particularly useful in cases where powder diffraction patterns either cannot be obtained, or are insufficient for structure determination. The success of the method is related to the high sensitivity of the Xe NMR chemical shift and chemical shift anisotropy to minute details of the cavity geometry, such that even structurally very similar clathrates can have vastly different chemical shift tensors. However, this sensitivity also requires highly detailed NMR calculations; in addition to the proper treatment of electron correlation and relativistic phenomena the inclusion of explicit crystal lattice effects as well as Xe dynamics was necessary in order to precisely reproduce experimental NMR values. Developments in this area, along with progress in CSP algorithms, open up new possibilities for the prediction and characterization of new porous materials by combining structure prediction with computational and experimental $^{129}\text{Xe}$ NMR spectroscopy.

5 Acknowledgements

The work reported here has been financially supported by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. 307358, ERC-StG-2012-ANGLE (JN, GD), by the Magnus Ehrnrooth Foundation and NGS-NANO (MS), and by the Academy of Finland project nos. 125316, 218191, 255641, and 285666 (PL). Computational resources for the NMR part of the article were partially provided by CSC – IT Center for Science Ltd, Espoo and Kajaani, Finland. Parts of the calculations have been carried out using resources provided by Wroclaw Centre for Networking and Supercomputing (http://wcss.pl). The crystal structure prediction was carried out using the IRIDIS high performance computing facility at the University of Southampton. MS acknowledges great help of Anu
M. Kantola in performing solid state NMR measurements.
References


Figure 1: Conformers of the two fluorophenol isomers. a) cis- and b) trans-o-fluorophenol c) cis- and d) trans-m-fluorophenol.
Figure 2: Experimental (black) and calculated (indigo) $^{129}$Xe NMR powder spectra of $o$-fluorophenol (top) and $m$-fluorophenol (bottom) with $^{129}$Xe gas reference at zero. Calculated spectra for all structures were obtained from BHandHLYP-scaled periodic PBE NMR parameters at PBE-TS optimized geometry (method/structure = BHandHLYP/PBE-TS). For the most probable candidates, $oF_D$ and $mF_A$, results from different periodic modeling levels are displayed: (1, green) PBE/CSP, (2, orange) PBE/PBE-TS, (3, indigo) BHandHLYP/PBE-TS and (4, red) BHandHLYP/PBE-TS with effects due to Xe dynamics at $T = 300\, \text{K}$ (see text for details).
Figure 3: The crystal energy landscape of $o$-fluorophenol. Each point represents one crystal structure and is coloured according to the molecular conformation ($cis$-$trans-o$-fluorophenol refers to $Z'=2$ structures containing both conformers). Structures with a lattice energy within 13 kJ/mol from the lowest energy structure and having cavities of suitable size for xenon absorption (see text) are encircled in red. The experimentally known high-pressure polymorph (CSD refcode QAMWEH01) is encircled in green. The labels A–E correspond to structures $oF_A$ to $oF_E$ in the text.
Figure 4: The crystal energy landscape of $m$-fluorophenol. Each point represents one crystal structure and is coloured according to the molecular conformation (cis-trans-$m$-fluorophenol refers to $Z'=2$ structures containing both conformers). Structures with a lattice energy within 13 kJ/mol from the lowest energy structure and having cavities of suitable size for xenon absorption (see text) are encircled in red. The experimentally known stable polymorph (CSD refcode QAMTUU) is encircled in green. The labels A–G correspond to structures $mF_A$ to $mF_G$ in the text.
Figure 5: Examples of clusters used in the initial screening NMR calculations of the predicted clathrate structures. The clusters shown correspond to the structures selected as best models of the observed clathrates: a) $o$-fluorophenol structure $oF_D$, b) $m$-fluorophenol structure $mF_A$. Atoms are colored by element: carbon (grey); oxygen (red); hydrogen (white); fluorine (yellow).
Figure 6: $^{129}$Xe NMR parameters computed at screening (nonrelativistic DFT/BHandHLYP/MHA) level for cluster models of the predicted clathrate structures of o-fluorophenol. Structures with asymmetry $\eta > 0.04$ are shown as blue crosses, while structures with symmetric chemical shift tensors ($\eta \leq 0.04$) are shown as green squares. The experimental data (expt, black filled circle) is included for comparison. Structures that were considered for further study are connected with the black arrows to the periodic GIPAW/PBE calculation of Xe shielding from periodic PBE-TS optimized structures of the corresponding crystal (red diamonds). The labels A–E correspond to structures $\sigma F_{A}$ to $\sigma F_{E}$ in the text.
Figure 7: $^{129}$Xe NMR parameters computed at screening (nonrelativistic DFT/BHandHLYP/MHA) level for cluster models of the predicted clathrate structures of $m$-fluorophenol. Structures with asymmetry $\eta > 0.04$ are shown as blue crosses, while structures with symmetric chemical shift tensors ($\eta \leq 0.04$) are shown as green squares. The experimental data (expt, black filled circle) is included for comparison. Structures that were considered for further study are connected with the black arrows to the periodic GIPAW/PBE calculation of Xe shielding from periodic PBE-TS optimized structures of the corresponding crystal (diamonds). The labels A–G correspond to structures $mF_A$ to $mF_G$ in the text.
Figure 8: The periodic GIPAW results for the five most likely $\sigma$-fluorophenol structures optimized at the periodic PBE-TS level of theory. The labels A–E correspond to structures $oF_A$ to $oF_E$ in the text. We expect the correct structure to approach the experimental (expt) $^{129}$Xe NMR parameters, when the computed GIPAW/PBE result is scaled with factors obtained using different pure and hybrid DFT functionals with increasing amount of exact exchange in the series of PBE $\rightarrow$ BLYP(0%) $\rightarrow$ B3LYP(20%) $\rightarrow$ BHandHLYP(50%). The SO correction is added to BHandHLYP values of all structures (blue diamonds). For structures $oF_B$ and $oF_D$, the final points (indigo crosses) include also the effect of Xe dynamics (DYN) at $T = 300$ K.
Figure 9: The periodic GIPAW results for the seven \( m \)-fluorophenol structures optimized at the periodic PBE-TS level of theory. The labels A–G correspond to structures \( mF_A \) to \( mF_G \) in the text. We expect the correct structure to approach the experimental (expt) \( ^{129}\text{Xe} \) NMR parameters, when the computed GIPAW/PBE result is scaled with factors obtained using different pure and hybrid DFT functionals with increasing amount of exact exchange in the series of PBE\( \rightarrow \)BLYP(0\%)\( \rightarrow \)B3LYP(20\%)\( \rightarrow \)BHandHLYP(50\%). The SO correction is added to BHandHLYP values of all structures (blue diamonds). For structures \( mF_A \), \( mF_B \), \( mF_E \), and \( mF_F \), the final point (indigo crosses) include also the effect of Xe dynamics (DYN) at \( T = 300 \) K.
Figure 10: Packing diagram of the PBE-TS optimized $o$-fluorophenol xenon clathrate structure $oF_{D}$ viewed down the $c$-axis. Space group $R3$ and with $R_{6}^{3}(12)$ hydrogen bonds encasing the xenon atoms. See also Figure 5a.
Figure 11: Packing diagram of the PBE-TS optimized $m$-fluorophenol xenon clathrate $mF_A$ viewed down the $c$-axis. Space group $R\bar{3}$ and with $R_8^1(12)$ hydrogen bonds encasing the xenon atoms. See also Figure 5b.
Table 1: Theoretically modelled $^{129}$Xe NMR chemical shift ($\delta$) and anisotropy (CSA, $\Delta \delta$) in ppm for the five most likely $\alpha$-fluorophenol CSP candidates in their PBE-TS optimised structures, unless stated otherwise.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\delta$ PBE</th>
<th>$\Delta \delta$ PBE</th>
<th>$\delta$ BLYP</th>
<th>$\Delta \delta$ BLYP</th>
<th>$\delta$ B3LYP</th>
<th>$\Delta \delta$ B3LYP</th>
<th>$\delta$ BHandHLYP</th>
<th>$\Delta \delta$ BHandHLYP</th>
<th>$\delta$ BHandHLYP $^f$</th>
<th>$\Delta \delta$ BHandHLYP $^f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBE</td>
<td>218.7</td>
<td>-14.3</td>
<td>245.0</td>
<td>-79.3</td>
<td>215.1</td>
<td>4.6</td>
<td>261.5</td>
<td>-38.6</td>
<td>370.5</td>
<td>7.7</td>
</tr>
<tr>
<td>BLYP</td>
<td>226.5</td>
<td>-14.6</td>
<td>253.4</td>
<td>-81.8</td>
<td>222.5</td>
<td>4.1</td>
<td>270.9</td>
<td>-40.2</td>
<td>381.5</td>
<td>7.9</td>
</tr>
<tr>
<td>B3LYP</td>
<td>195.9</td>
<td>-23.1</td>
<td>223.1</td>
<td>-75.6</td>
<td>195.9</td>
<td>0.9</td>
<td>240.8</td>
<td>-36.6</td>
<td>338.1</td>
<td>13.1</td>
</tr>
<tr>
<td>BHandHLYP</td>
<td>165.8</td>
<td>-21.3</td>
<td>190.7</td>
<td>-68.3</td>
<td>162.6</td>
<td>-1.2</td>
<td>209.9</td>
<td>-35.8</td>
<td>288.5</td>
<td>18.3</td>
</tr>
<tr>
<td>BHandHLYP $^b$</td>
<td>145.8</td>
<td>-20.5</td>
<td>175.0</td>
<td>-61.8</td>
<td>143.2</td>
<td>0.9</td>
<td>189.4</td>
<td>-31.6</td>
<td>262.0</td>
<td>14.4</td>
</tr>
<tr>
<td>BHandHLYP $^c$</td>
<td>174.9</td>
<td>-11.6</td>
<td>192.3</td>
<td>-71.2</td>
<td>171.3</td>
<td>6.7</td>
<td>216.7</td>
<td>-31.6</td>
<td>292.0</td>
<td>31.1</td>
</tr>
<tr>
<td>Periodic model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBE $^e$</td>
<td>195.1</td>
<td>-28.1</td>
<td>272.1</td>
<td>-66.4</td>
<td>238.7</td>
<td>7.0</td>
<td>288.8</td>
<td>-26.2</td>
<td>362.1</td>
<td>23.1</td>
</tr>
<tr>
<td>BLYP $^f$</td>
<td>229.8</td>
<td>-38.9</td>
<td>273.2</td>
<td>-86.7</td>
<td>236.4</td>
<td>-15.4</td>
<td>285.0</td>
<td>-46.5</td>
<td>402.7</td>
<td>2.0</td>
</tr>
<tr>
<td>B3LYP $^f$</td>
<td>237.3</td>
<td>-39.1</td>
<td>281.5</td>
<td>-89.3</td>
<td>243.6</td>
<td>-15.9</td>
<td>294.3</td>
<td>-48.1</td>
<td>413.6</td>
<td>2.1</td>
</tr>
<tr>
<td>BHandHLYP $^f$</td>
<td>206.3</td>
<td>-48.0</td>
<td>250.8</td>
<td>-82.9</td>
<td>216.7</td>
<td>-19.3</td>
<td>263.8</td>
<td>-44.5</td>
<td>369.5</td>
<td>7.5</td>
</tr>
<tr>
<td>BHandHLYP $^g$</td>
<td>175.6</td>
<td>-46.3</td>
<td>217.7</td>
<td>-75.4</td>
<td>182.6</td>
<td>-21.5</td>
<td>232.3</td>
<td>-43.6</td>
<td>318.8</td>
<td>12.7</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Scalar relativistic SR-ZORA calculations of a cluster model of one clathrate cavity with ADF code$^{[73,74]}$ using Xe/other=jcpl/TZP basis sets.$^{[76]}$ $^{129}$Xe chemical shift ($\delta = \sigma_{\text{Xe-atom}} - \sigma$) with respect to free Xe atom shielding values ($\sigma_{\text{Xe-atom}}$): PBE/BLYP/B3LYP/BHandHLYP = 5752.2/5752.6/5752.1/5752.0 ppm. $\Delta \delta$ (CSA) defined in principal axis system (PAS) along the unique axis (perpendicular to the plane of O-H rings). Asymmetry parameter $\eta = 0$ in all cavities due to the cylindrical symmetry.

$^b$ As footnote $a$ but at nonrelativistic (NR) level of theory with $\sigma_{\text{Xe-atom}}$(BHandHLYP) = 5643.4 ppm.

$^c$ As footnote $a$ but at relativistic spin-orbit SO-ZORA level of theory with $\sigma_{\text{Xe-atom}}$(BHandHLYP) = 6609.3 ppm.

$^d$ Scalar relativistic GIPAW results obtained with CASTEP code.$^{[71,72]}$ All periodic $^{129}$Xe chemical shifts referenced to CASTEP/PBE value: $\sigma_{\text{Xe-atom}} = 5926.3$ ppm.

$^e$ In the CSP optimised crystal geometry.

$^f$ As footnote $d$ but estimated by scaling GIPAW/PBE PAS components with factors obtained at SR-ZORA level for cluster model resulting in data in footnote $a$.

$^g$ Correction added to result of footnote $f$ due to thermal averaging (AVE) over Xe motion at $T = 300$K by MC-NVT simulation (see text for details).

$^h$ Relativistic spin-orbit (SO) corrections obtained as difference of SO-ZORA and SR-ZORA calculations with ADF code added to result of footnote $g$. 

---

30
Table 2: Theoretically modelled $^{129}$Xe NMR chemical shift ($\delta$) and anisotropy (CSA, $\Delta\delta$) in ppm for the seven most likely m-fluorophenol CSP candidates in their PBE-TS optimised structures, unless stated otherwise.

<table>
<thead>
<tr>
<th>Method</th>
<th>$mF_A$</th>
<th>$mF_B$</th>
<th>$mF_C$</th>
<th>$mF_D$</th>
<th>$mF_E$</th>
<th>$mF_F$</th>
<th>$mF_G$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster model $^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBE</td>
<td>259.6</td>
<td>219.9</td>
<td>272.1</td>
<td>218.5</td>
<td>216.5</td>
<td>256.3</td>
<td>146.6</td>
</tr>
<tr>
<td>BLYP</td>
<td>264.8</td>
<td>222.0</td>
<td>278.7</td>
<td>189.5</td>
<td>224.1</td>
<td>201.3</td>
<td>147.5</td>
</tr>
<tr>
<td>B3LYP</td>
<td>236.9</td>
<td>197.2</td>
<td>240.0</td>
<td>172.1</td>
<td>193.6</td>
<td>182.4</td>
<td>135.5</td>
</tr>
<tr>
<td>BHandHLYP</td>
<td>204.6</td>
<td>174.7</td>
<td>202.4</td>
<td>166.9</td>
<td>159.2</td>
<td>127.8</td>
<td>119.1</td>
</tr>
<tr>
<td>BHandHLYP$^b$</td>
<td>182.3</td>
<td>158.8</td>
<td>189.8</td>
<td>159.5</td>
<td>152.2</td>
<td>147.7</td>
<td>176.7</td>
</tr>
<tr>
<td>BHandHLYP$^c$</td>
<td>211.2</td>
<td>176.3</td>
<td>197.4</td>
<td>162.1</td>
<td>170.1</td>
<td>156.2</td>
<td>186.4</td>
</tr>
<tr>
<td>Periodic model $^d$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBE$^e$</td>
<td>239.7</td>
<td>213.1</td>
<td>428.8</td>
<td>226.8</td>
<td>229.2</td>
<td>271.4</td>
<td>163.4</td>
</tr>
<tr>
<td>PBE</td>
<td>270.6</td>
<td>226.7</td>
<td>295.2</td>
<td>216.0</td>
<td>229.2</td>
<td>271.4</td>
<td>163.4</td>
</tr>
<tr>
<td>BLYP$^f$</td>
<td>275.5</td>
<td>228.9</td>
<td>301.6</td>
<td>217.0</td>
<td>236.7</td>
<td>277.7</td>
<td>164.3</td>
</tr>
<tr>
<td>B3LYP$^f$</td>
<td>247.3</td>
<td>203.4</td>
<td>262.3</td>
<td>199.3</td>
<td>205.7</td>
<td>208.3</td>
<td>144.2</td>
</tr>
<tr>
<td>BHandHLYP$^f$</td>
<td>214.2</td>
<td>180.2</td>
<td>223.8</td>
<td>194.1</td>
<td>179.1</td>
<td>184.5</td>
<td>203.3</td>
</tr>
<tr>
<td>BHandHLYP$^g$</td>
<td>229.3</td>
<td>187.3</td>
<td>229.7</td>
<td>200.4</td>
<td>248.6</td>
<td>171.8</td>
<td>280.7</td>
</tr>
<tr>
<td>BHandHLYP$^h$</td>
<td>235.9</td>
<td>188.9</td>
<td>224.7</td>
<td>195.6</td>
<td>233.8</td>
<td>194.4</td>
<td>247.0</td>
</tr>
<tr>
<td>Experimental</td>
<td>228.5</td>
<td>183.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Scalar relativistic SR-ZORA calculations of a cluster model of one clathrate cavity with ADF code $^{[73,74]}$ using Xe/other=jeclp/TZP basis sets. $^{[76]}$ $^{129}$Xe chemical shift ($\delta = \sigma_{Xe-atom} - \sigma$) with respect to free Xe atom shielding values ($\sigma_{Xe-atom}$): PBE/BLYP/B3LYP/BHandHLYP = 5752.2/5752.6/5752.1/5752.0 ppm. $\Delta\delta$ (CSA) defined in principal axis system (PAS) along the unique axis (perpendicular to the plane of O-H rings). Asymmetry parameter $\eta = 0$ in all cavities due to the cylindrical symmetry.

$^b$ As footnote $a$ but at nonrelativistic (NR) level of theory with $\sigma_{Xe-atom}(\text{BHandHLYP}) = 5643.4$ ppm.

$^c$ As footnote $a$ but at relativistic spin-orbit SO-ZORA level of theory with $\sigma_{Xe-atom}(\text{BHandHLYP}) = 6609.3$ ppm.

$^d$ Scalar relativistic GIPAW results obtained with CASTEP code. $^{[71,72]}$ All periodic $^{129}$Xe chemical shifts referenced to CASTEP/PBE value: $\sigma_{Xe-atom} = 5926.3$ ppm.

$^e$ In the CSP optimised crystal geometry.

$^f$ As footnote $d$ but estimated by scaling GIPAW/PBE PAS components with factors obtained at SR-ZORA level for cluster model resulting in data in footnote $a$.

$^g$ Correction added to result of footnote $f$ due to thermal averaging (AVE) over Xe motion at $T = 300K$ by MC-NVT simulation (see text for details).

$^h$ Relativistic spin-orbit (SO) corrections obtained as difference of SO-ZORA and SR-ZORA calculations with ADF code added to result of footnote $g$. 
Table 3: Unit cell parameters (after PBE-TS geometry optimisation) for the proposed xenon clathrate structures of \( o- \) and \( m- \)fluorophenol

<table>
<thead>
<tr>
<th>Structure</th>
<th>space group</th>
<th>( a ) (Å)</th>
<th>( b ) (Å)</th>
<th>( c ) (Å)</th>
<th>( \gamma ) (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( oF_D )</td>
<td>( R3 )</td>
<td>22.5225</td>
<td>22.5225</td>
<td>5.6879</td>
<td>120</td>
</tr>
<tr>
<td>( mF_A )</td>
<td>( R3 )</td>
<td>22.5526</td>
<td>22.5526</td>
<td>5.6998</td>
<td>120</td>
</tr>
</tbody>
</table>

Figure 12: $^{129}$Xe NMR/CSP crystallography: A new method to determine the structure of porous solids provides the likely ortho- and meta-fluorophenol xenon clathrate structures. They are selected from an ensemble of hypothetical host frameworks generated by crystal structure prediction (CSP) by comparing the $^{129}$Xe powder NMR spectra with detailed NMR shielding tensor first principles calculations.
Keywords: Xe NMR spectroscopy, crystal structure prediction, first principles modelling
The following structure files, in CIF format, accompany this document:

mF_all.cif and oF_all.cif: all predicted crystal structures of both molecules, in order of increasing energy.

mF_hosts.cif and oF_hosts.cif: all predicted crystal structures with voids suitable for hosting a xenon atom (see below for explanation).

mF_clathrates.cif and oF_clathrates.cif: all selected host crystal structures after computational insertion of a xenon atom, followed by lattice energy minimisation (see below for explanation).

mF_A_PBE-TS.cif, mF_B_PBE-TS.cif, mF_C_PBE-TS.cif, mF_D_PBE-TS.cif, mF_E_PBE-TS.cif, mF_F_PBE-TS.cif, and mF_G_PBE-TS.cif. Periodic DFT optimised crystal structures of the 7 predicted clathrate structures of m-fluorophenol with NMR parameters closest to the experimentally measured parameters.

oF_A_PBE-TS.cif, oF_B_PBE-TS.cif, oF_C_PBE-TS.cif, oF_D_PBE-TS.cif, and oF_E_PBE-TS.cif. Periodic DFT optimised crystal structures of the 5 predicted clathrate structures of o-fluorophenol with NMR parameters closest to the experimentally measured parameters.

1 Crystal structure prediction (CSP)

The CSP calculations involve five steps: i) the lowest energy conformations of the isolated molecules are identified; ii) hypothetical crystal packings are generated and the lattice energy is minimised using a simple force field and rigid-molecule constraints; iii) a subset of the lowest energy predicted crystal structures are re-optimised using an anisotropic atom-atom force field; iv) intramolecular flexibility is introduced, allowing the hydroxyl group to reorient in response to packing forces in each low energy crystal structure; v) porous structures are identified, xenon atoms are inserted into the pores and the structures once again re-optimised.

1.1 Molecular conformational analysis

Both molecules have rigid molecular geometries, apart from the orientation of the hydroxyl group. The two compounds o-fluorophenol and m-fluorophenol both have a single dihedral angle (H-O-C-C) and the dihedral angle is likely to have a large impact on the possible crystal structures and their lattice energy. The dihedral angles of both molecules were examined by scanning them in steps of 5° and each conformation was geometry optimised using Gaussian09 version D01 and the B3LYP hybrid functional and Møller-Plesset second order perturbation theory. The cis/0° and trans/180° conformers, where the -OH hydrogen is pointing towards the fluorine, and away from it, respectively, have the lowest energy. In o-fluorophenol the difference in energy between the cis and trans conformers is 13 kJ/mol and in m-fluorophenol, it is only 0.1 kJ/mol.

1.2 Crystal structure generation

A large set of hypothetical crystal structures was generated with the rigid DFT-optimised molecular geometries using Monte Carlo simulated annealing with the POLYMORPH PREDIC-
TOR module in Accelrys' MATERIALS STUDIO software. Searches were performed using space group symmetry in the most commonly observed space groups of known crystal structures in the Cambridge Structure Database (CSD). There are reasons to believe that clathrate structures occur selectively in certain space groups\textsuperscript{2,3}, but no space groups were omitted because of this. Searches with one molecule in the asymmetric unit ($Z' = 1$) were performed in 25 space groups ($P2_1/c$, $P2_12_12_1$, $P\bar{1}$, $P2_1$, $Pbca$, $C2/c$, $Pna2_1$, $Cc$, $C2$, $Pca2_1$, $P1$, $Pbcn$, $P4_12_12_1$, $P2_12_12_1$, $Pc$, $P4_1$, $Fdd2$, $Pccn$, $P2/c$, $P6_1$, $I4_1/a$, $P3_2$, $R3$, $R\bar{3}$, $C222_1$) and were repeated for both planar conformers of both molecules. $Z' = 2$ searches were performed in $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$ and $P2_12_12_1$ with the three possible combinations of conformers for each molecule. Ten simulated annealing cycles starting from different random seeds were deemed to sample the search space adequately for $Z' = 1$, since most lowest energy structures in each space group had been located multiple times. For $Z' = 2$, 10–25 repeats were deemed adequate, depending on the space group.

Hypothetical crystal structures were generated by simulated annealing in the Polymorph mod-
ule (version 6.1, Build date March 7, 2013) in Materials Studio version 6.1.200. Several Monte Carlo simulated annealing calculations were performed starting with different random number seeds, until almost all of the 10 lowest energy crystal structures had been found more than once. This meant approximately 10 runs were required for each space group with \( Z' = 1 \) and 20 runs per space group for \( Z' = 2 \).

The following settings were used:
Search algorithm : MC Simulated Annealing
Maximum number of steps : 10000
Explore torsions : No
Preoptimize structures : No
Steps to accept before cooling : 14
Minimum move factor : 0.1000E-09
Heating factor : 0.02500
Maximum temperature : 150000.0 K
Minimum temperature : 300.0 K

Intermolecular interactions at this stage were modelled using the COMPASS force field with atomic partial charges fitted to the molecular electrostatic potential using the CHelpG method. Optimisation algorithm : Smart
Convergence tolerances:
Energy : 1.5e-005 kcal/mol
Force : 0.001 kcal/mol/Å
Stress : 0.001 GPa
Displacement : 1e-005 Å
Maximum number of iterations : 500
External pressure : 0 GPa
Motion groups rigid : YES
Optimize cell : YES

Forcefield : COMPASS
Electrostatic terms:
Summation method : Ewald
Accuracy : 1e-05 kcal/mol Buffer width : 0.5 Å

van der Waals terms:
Summation method : Atom based
Truncation method : Cubic spline
Cutoff distance : 18.5 Å
Spline width : 1 Å
Buffer width : 0.5 Å

The generated structures were clustered to remove duplicates. Clustering was performed with COMPACK using these settings:
Cluster grouping : Forcefield type
Cutoff : 7.000 Å

Number of bins : 140
Tolerance : 0.1100
### 1.3 Lattice energy minimisation

Structures with a lattice energy within a 15 kJ/mol window from the lowest energy structure were further refined with an anisotropic intermolecular potential model. Lattice energies were calculated from a hybrid method, combining intramolecular DFT energies with an intermolecular force field model:

$$E_{\text{latt}} = E_{\text{DFT-intra}} + E_{\text{atom-atom}}. \quad (1)$$

Intramolecular energies and molecular geometries were derived from DFT. Intermolecular interactions were modelled with an anisotropic model potential of the form:

$$E_{\text{MN-inter}}^{MN} = \sum_{i,k} \left( A^{\kappa \kappa} \exp \left(-B^{\kappa \kappa} r_{ik}\right) - C^{\kappa \kappa} r_{ik}^{-6} + E_{\text{elec}}^{i,k} \right), \quad (2)$$

where $i, k$ are atoms of types $\kappa$ and $\kappa'$ in molecules $M$ and $N$, respectively, separated by the distance $r_{ik}$. The first two terms describe the repulsive and attractive non-electrostatic intermolecular interactions, with $A^{\kappa \kappa}$, $B^{\kappa \kappa}$ and $C^{\kappa \kappa}$ being empirically determined parameters. The final term, describing electrostatic interactions, was calculated from atom-centered multipoles up to hexadecapole on all atoms, obtained from a distributed multipole analysis of the molecular charge density using Gdma. The multipole model accurately describes anisotropic intermolecular interactions, such as hydrogen bonds and $\pi$-stacking. Charge-charge, charge-dipole and dipole-dipole interactions were calculated using Ewald summation, while repulsion-dispersion interactions and all higher multipole-multipole interactions were summed to a cutoff of 15 Å.

A revised version of the W99 exp-6 potential was used, whose re-parametrisation optimised its performance for hydrogen bonding when used together with atomic multipole electrostatics. Fluorine atoms were treated as having an anisotropic repulsion, as described previously. W99 requires all X-H bonds in the DFT-optimised molecular geometry to be foreshortened by 0.1 Å. Using this model potential, crystal structures were lattice energy minimised with molecular geometries kept rigid, using Dmacrys.

Duplicate crystal structures were removed using the COMPACT program, which compares interatomic distances in a cluster of 15 molecules from each crystal structure.

The resulting unique structures were re-optimised using the CRYSTALOPTIMIZER program to treat flexibility of the hydroxyl group. CRYSTALOPTIMIZER minimises the sum of intra- and intermolecular energies by calculating the strain due to crystal packing forces on the hydroxyl dihedral angle and allows this to relax. The minimisation is performed by iteratively computing the intermolecular forces on the flexible dihedral angle in Dmacrys, and optimising the intramolecular geometry in GAUSSIAN under the influence of the packing forces.

CIF-files of all predicted crystal structures are included (mF_all.cif and oF_all.cif). Structures within the CIF are ordered by increasing lattice energy.


2 Crystal structure overlays

Figure S3: Overlay of the lowest energy predicted $m$-fluorophenol crystal structure mF0_P21_3110 with QAMTUU. The match is excellent with $RMSD_{15} = 0.203 \text{ Å}$.

Figure S4: Overlay picture of the predicted structure oF00_P212121_5 with QAMWEH01. The experimental structure was determined at elevated temperature and pressure, making a structure comparison difficult. Re-optimising the predicted structure at the experimental pressure (0.36 GPa) gives a match with $RMSD_{15} = 0.422 \text{ Å}$.
3 Guest to host volume ratios

Molecular volumes were calculated as the sum of atomic fragment contributions using the Molinspiration property calculator available online\(^\text{14}\) and a volume of 42.2 Å\(^3\) was used for xenon. Crystalline cavity volumes were calculated by placing probe particles at positions in a regular grid with 0.2 Å spacing in the unit cell and testing if the particle was within any atom’s van der Waals radius,\(^\text{15,16}\) as implemented in the program PLATON.\(^\text{17}\) Convergence of the calculated void with respect to grid spacing is discussed below. PLATON allows not only the total void volume per unit cell to be determined, but also the volume of each individual pore, and that is the volume used in this work.

Void volumes were calculated with the PLATON program which places a test particle of a certain radius at positions in a regular grid in the unit cell. The calculation is highly sensitive to the particle size and grid spacing. Using a large probe will underestimate the volume and will often result in an erroneous topology of the void space. Small molecules may diffuse through narrow channels between larger cavities. We have chosen a probe size of 1.2 Å radius, corresponding to the kinetic diameter of a H\(_2\) molecule, this is also the default setting in PLATON. A coarse grid will result in inaccurate volumes. A convergence test was carried out on a subset of the known clathrates in CSD.

The relative error of the calculated void volume for different grid spacing was compared to the very fine grid spacing of 0.1 Å.

\[
\text{Error} = 100 \cdot \left( \frac{V_{\text{grid}} - V_{0.1}}{V_{0.1}} \right) \tag{3}
\]

The computational cost prevents the finest grid spacings to be used routinely. The convergence test indicated that the void volume calculation is slowly converging for grid spacings smaller than 0.4 Å. Grid spacings larger than 0.3 Å produces unacceptably large errors. A grid spacing of 0.2 Å was deemed adequate, as it consistently gives volumes with small errors of approximately 5% relative to the 0.1 grid spacing. By extrapolating the curves in the convergence diagram, the true errors relative to a perfect grid spacing of zero can be estimated. The volumes calculated by a 0.2 Å grid spacing are systematically underestimating the true volume by about 8%.

![Figure S5: Convergence test for the void calculation](image-url)
In order to confirm the validity of the 55% rule\textsuperscript{18} for our void calculation method, the CSD was searched for organic clathrate structures (see Table S1). Disordered structures, metallo-organic structures, and hydrates were excluded. In addition, the hydroquinone SO\textsubscript{2} and hydroquinone C\textsubscript{60} clathrates were used. The set of selected structures was chosen only to be large enough to estimate the spread in guest-to-host ratio and to represent a variation in host guest chemistry.

The guest molecules were removed from the structure and a void calculation was performed using a 1.2 Å probe radius and 0.2 Å grid spacing. Guest molecule volumes other than Xe were calculated with the Molinspiration property engine v2013.09, by Molinspiration Cheminformatics 2014. It is available online at http://www.molinspiration.com/cgi-bin/properties. It provides volumes calculated from summing atomic group volumes fitted to vacuum AM1 semiempirical conformations of a large training set. Volumes in Å\textsuperscript{3}. For the volume of xenon, the van der Waals volume 42.2 Å\textsuperscript{3} was used.

![Normal distribution of guest-host volume ratios](image)

Figure S6: The Guest-to-host volume ratios are normally distributed around 58.9% with a standard deviation of 8.0 percentage-points.

Guest to host volume ratios are normally distributed around 60% with a standard deviation of 8 percentage points (see Figure S6).

CIF-files of selected host structures from the CSP sets of structures are included in oF\_hosts.cif and mF\_hosts.cif. These structures, after insertion of Xe and re-optimisation, are included in oF\_clathrates.cif and mF\_clathrates.cif.
<table>
<thead>
<tr>
<th>CSD code</th>
<th># of guests</th>
<th>Guest</th>
<th>Guest Vol</th>
<th>Void Vol</th>
<th>Ratio %</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABAZOT</td>
<td>2</td>
<td>3-pentanone</td>
<td>98.344</td>
<td>327</td>
<td>60.14</td>
</tr>
<tr>
<td>ABUCUV</td>
<td>2</td>
<td>acetonitrile</td>
<td>46.055</td>
<td>158</td>
<td>58.28</td>
</tr>
<tr>
<td>ACAWUW</td>
<td>2</td>
<td>ethanol</td>
<td>54.016</td>
<td>164</td>
<td>65.86</td>
</tr>
<tr>
<td>AJEJII</td>
<td>3</td>
<td>acetonitrile</td>
<td>46.055</td>
<td>193</td>
<td>71.58</td>
</tr>
<tr>
<td>AJEQEL</td>
<td>4</td>
<td>chloroform</td>
<td>70.07</td>
<td>553</td>
<td>50.68</td>
</tr>
<tr>
<td>AJOJEP</td>
<td>1</td>
<td>chloroform</td>
<td>70.07</td>
<td>119</td>
<td>58.88</td>
</tr>
<tr>
<td>ALUSOP</td>
<td>2</td>
<td>chloroform</td>
<td>70.07</td>
<td>275</td>
<td>50.96</td>
</tr>
<tr>
<td>ANIRUL</td>
<td>1</td>
<td>butan-1-ol</td>
<td>87.619</td>
<td>171</td>
<td>51.23</td>
</tr>
<tr>
<td>ASOKOI</td>
<td>1</td>
<td>TCE</td>
<td>100.434</td>
<td>154</td>
<td>65.21</td>
</tr>
<tr>
<td>AVESOJ</td>
<td>4</td>
<td>diethylelamine</td>
<td>91.763</td>
<td>694</td>
<td>52.88</td>
</tr>
<tr>
<td>BABYIN</td>
<td>2</td>
<td>acetonitrile</td>
<td>46.055</td>
<td>171</td>
<td>53.86</td>
</tr>
<tr>
<td>BAFSOQ</td>
<td>2</td>
<td>urea</td>
<td>54.196</td>
<td>154</td>
<td>70.38</td>
</tr>
<tr>
<td>BAZJES</td>
<td>1</td>
<td>acetone</td>
<td>64.74</td>
<td>122</td>
<td>53.06</td>
</tr>
<tr>
<td>BEGZAP</td>
<td>2</td>
<td>propionic acid</td>
<td>72.999</td>
<td>285</td>
<td>51.22</td>
</tr>
<tr>
<td>BERYED</td>
<td>4</td>
<td>methane</td>
<td>28.644</td>
<td>193</td>
<td>59.36</td>
</tr>
<tr>
<td>BICYCLO</td>
<td>2</td>
<td>diethylether</td>
<td>88.345</td>
<td>271</td>
<td>65.18</td>
</tr>
<tr>
<td>BIFQA101</td>
<td>2</td>
<td>benzonitrile</td>
<td>100.903</td>
<td>408</td>
<td>49.46</td>
</tr>
<tr>
<td>BIJVOf</td>
<td>3</td>
<td>acetid acid</td>
<td>46.055</td>
<td>217</td>
<td>63.66</td>
</tr>
<tr>
<td>BOSVEK01</td>
<td>1</td>
<td>sulphur dioxide</td>
<td>41.314</td>
<td>61</td>
<td>67.72</td>
</tr>
<tr>
<td>BUSPAG</td>
<td>1</td>
<td>methylisocyanide</td>
<td>51.287</td>
<td>79</td>
<td>64.92</td>
</tr>
<tr>
<td>CAKGUQ</td>
<td>1</td>
<td>acetonitrile</td>
<td>46.055</td>
<td>85</td>
<td>54.18</td>
</tr>
<tr>
<td>CAVMES</td>
<td>1</td>
<td>diethylether</td>
<td>88.345</td>
<td>271</td>
<td>65.18</td>
</tr>
<tr>
<td>GAGNEJ</td>
<td>2</td>
<td>DMSO</td>
<td>71.433</td>
<td>216</td>
<td>66.14</td>
</tr>
<tr>
<td>GIRBOY</td>
<td>3</td>
<td>xenon</td>
<td>42.2</td>
<td>226</td>
<td>56.01</td>
</tr>
<tr>
<td>HQ-C60</td>
<td>2</td>
<td>C60</td>
<td>539.5</td>
<td>1945</td>
<td>55.46</td>
</tr>
<tr>
<td>HQUACN</td>
<td>1</td>
<td>acetonitrile</td>
<td>46.055</td>
<td>77</td>
<td>59.81</td>
</tr>
<tr>
<td>HQ $\beta$</td>
<td>1</td>
<td>sulphur dioxide</td>
<td>41.314</td>
<td>68</td>
<td>60.75</td>
</tr>
<tr>
<td>ILEZAA</td>
<td>1</td>
<td>xenon</td>
<td>42.2</td>
<td>84</td>
<td>50.23</td>
</tr>
<tr>
<td>JAMKEN</td>
<td>1</td>
<td>xenon</td>
<td>42.2</td>
<td>60</td>
<td>70.33</td>
</tr>
<tr>
<td>LOZNUJ</td>
<td>1</td>
<td>xenon</td>
<td>42.2</td>
<td>55</td>
<td>76.72</td>
</tr>
<tr>
<td>OJITOR</td>
<td>1</td>
<td>xenon</td>
<td>42.2</td>
<td>92</td>
<td>45.86</td>
</tr>
</tbody>
</table>

Table S1: Structures and void properties used in determining selection rules for inclusion hosts. In summary: 58.9 ± 8.0 (1 $\sigma$). TCE = 1,1,2,2-tetrachloroethane
4 Revised Williams99 parameters

All parameters describing interactions between C, N, O and H atoms are described using Williams99 forcefield parameters, apart from hydrogen bond H...A interactions, which have been re-parameterised to work more effectively with the atomic multipole electrostatic model. For H...A interactions, the pre-exponential parameter of the exp-6 model was modified from the Williams99 value. The parameters are given in the following table.

Table S2: Revised H...A parameters in the exp-6 intermolecular model used. C = 0 for all interactions.

<table>
<thead>
<tr>
<th>hydrogen</th>
<th>acceptor</th>
<th>A (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2</td>
<td>N1</td>
<td>149</td>
</tr>
<tr>
<td>H2</td>
<td>N2</td>
<td>166</td>
</tr>
<tr>
<td>H2</td>
<td>N3</td>
<td>163</td>
</tr>
<tr>
<td>H2</td>
<td>O1</td>
<td>129</td>
</tr>
<tr>
<td>H2</td>
<td>O2</td>
<td>105</td>
</tr>
<tr>
<td>H3</td>
<td>N1</td>
<td>70</td>
</tr>
<tr>
<td>H3</td>
<td>N2</td>
<td>118</td>
</tr>
<tr>
<td>H3</td>
<td>O1</td>
<td>127</td>
</tr>
<tr>
<td>H3</td>
<td>O2</td>
<td>133</td>
</tr>
<tr>
<td>H4</td>
<td>N1</td>
<td>141</td>
</tr>
<tr>
<td>H4</td>
<td>N2</td>
<td>77</td>
</tr>
<tr>
<td>H4</td>
<td>N3</td>
<td>56</td>
</tr>
<tr>
<td>H4</td>
<td>N4</td>
<td>112</td>
</tr>
<tr>
<td>H4</td>
<td>O1</td>
<td>34</td>
</tr>
<tr>
<td>H4</td>
<td>O2</td>
<td>198</td>
</tr>
</tbody>
</table>
5 Potential parameters for fluorine

Halogen atoms tend to have an anisotropic van der Waals radius\(^{19}\). To account for this, intermolecular potentials with an anisotropic repulsion term have been developed\(^{10,11}\). A local unit vector \(\mathbf{e}_z\) is defined at each anisotropic site, parallel to the covalent bond joining the halogen to its bonded atom, pointing away from the bond. A second unit vector, \(\mathbf{e}_{ik}\), is the vector between the interacting atoms. DMACRYS describes repulsion anisotropy using a modified exp-6 potential of the form:

\[
V = G \exp (-B^{\kappa}(r_{ij} - \rho^{\kappa}(\Omega_{ik}))) - C^{\kappa}/r^6, \tag{4}
\]

where \(\rho^{\kappa}(\Omega_{ik})\) describes the anisotropy of repulsion, and is defined as:

\[
\rho^{\kappa}(\Omega_{ik}) = \rho_0^{\kappa} + \rho_1^{\kappa}(\mathbf{e}_z \cdot \mathbf{e}_{ik}) + \rho_2^{\kappa}(3[\mathbf{e}_z \cdot \mathbf{e}_{ik}]^2 - 1)/2 + \rho_2^{\kappa}3e_{ik}^2 - 1)/2 \tag{5}
\]

\(\rho_0\) describes the isotropic repulsion, \(\rho_1\) parameters describe a shift of the centre of repulsion and \(\rho_2\) parameters describe a quadrupolar distortion of the atom. Parameters for fluorine were taken from Day’s specifically developed potential for molecule XIII in the 4th blind test of crystal structure prediction\(^{20}\). The parameters were empirically fitted to reproduce the crystal structures of a set of halogenated aromatic molecules. Details are available in the ESI to the 4th blind test paper. The parameters, in input format for DMACRYS are provided below:

```
BUCK  F_01 F_01
     3761.006673 0.240385 7.144500 0.0 70.0
ANIS  F_01 F_01
     0 0 2 0 2 -0.035000
     0 0 0 2 2 -0.035000
ENDS
```
6 Potential parameters for xenon

An exp-6 potential for xenon was parametrised to be used together with the Williams99 force-field\(^8\). The parameters were fitted to reproduce the potential energy surfaces of van der Waals dimers of Xe⋯Xe\(^{21}\), Xe⋯H\(_2\)O\(^{22}\), Xe⋯CH\(_4\)\(^{23}\), Xe⋯N\(_2\)\(^{24}\), Xe⋯O\(_2\)\(^{25}\) and Xe⋯C\(_6\)H\(_6\)\(^{26}\).

### Table S3: Xenon potential parameters

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Fitted value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>8216.849 eV</td>
</tr>
<tr>
<td>B</td>
<td>0.348606 Å(^{-1})</td>
</tr>
<tr>
<td>C</td>
<td>339.32 eV/Å(^6)</td>
</tr>
</tbody>
</table>

Cross terms for heteroatomic interactions are obtained as:

\[
A_{ij} = \sqrt{A_i A_j} \tag{6}
\]

\[
B_{ij} = 0.5 \cdot (A_i + A_j) \tag{7}
\]

\[
C_{ij} = \sqrt{C_i C_j} \tag{8}
\]
7 NMR calculations

7.1 Screening calculations on cluster models

From each predicted clathrate structure, we constructed a cluster model containing a single xenon-occupied cavity. Adjacent fluorophenol molecules with any atoms within 5.3 Å of the Xe atom were included in each cluster. Symmetries of the clusters were not used during geometry optimizations and NMR calculations. Calculations were performed using the BHandHLYP\textsuperscript{27,28} hybrid functional using TURBOMOLE\textsuperscript{29}. The MHA basis set [22s17p14d2f/15s13p11d2f] was used for Xe, which has been tested and found well-converged for nonrelativistic Xe NMR shielding.\textsuperscript{30–32} For other atoms (F, C, O, H) def2-SVP basis sets\textsuperscript{33} were used; this basis set combination is denoted MHA/SVP. The MHA/SVP basis set combination with the BP86\textsuperscript{27,34} functional was used in steepest descent DFT optimizations of hydrogen atom positions with fixed heavy atom coordinates. The influence of optimizing the hydrogen positions was examined and found not to be significant for the calculated NMR results (see below), since the clusters were derived from DFT-optimised molecules with already accurate hydrogen positions.

7.2 Periodic modelling of likely candidates

Five o-fluorophenol and seven m-fluorophenols crystal structures with screening level NMR parameters close to the experimental values were chosen for further scrutiny. These structures were optimised with respect to both atomic positions and lattice parameters using planewave, periodic DFT with the PBE functional\textsuperscript{35} and Tkatchenko-Scheffler (TS) dispersion correction\textsuperscript{36} using Castep\textsuperscript{37,38} and denoted as PBE-TS structures from now on. Optimizations were carried out using ultrasoft pseudopotentials\textsuperscript{39} with 380 eV cutoff and Γ-point centered Monkhorst-Pack\textsuperscript{40} k-point meshes with a spacing of ≤ 0.04 Å\textsuperscript{−1}. Convergence thresholds were 10\textsuperscript{−2} eV Å\textsuperscript{−1} and 5 · 10\textsuperscript{−4} Å for the ionic residual forces and displacements, respectively.

Following optimization, NMR shielding tensors of these PBE-TS optimised models were computed with the PBE functional using the gauge-including projector augmented wave (GIPAW) method\textsuperscript{41,42} (see Tables S6 and S7 below). Norm-conserving pseudopotentials\textsuperscript{43} were used with a 610 eV cutoff and a more stringent k-point spacing of ≤ 0.03 Å\textsuperscript{−1} in all directions, in order to obtain full shielding tensors accurately. These periodic Castep results include the scalar relativistic effects on Xe shielding at the 1-component scalar-relativistic (SR) zeroth-order regular approximation\textsuperscript{44,45} (SR-ZORA) level of theory.

CIF-files of PBE-TS optimized structures are included (oF_A-E_PBE-TS_structure.cif and mF_A-G_PBE-TS_structure.cif).

The structure of the selected clathrate of o-fluorophenol (oF_D) has space group symmetry $R\bar{3}$ in the CIF file. This is lower symmetry than the symmetry of the host framework ($R\bar{3}$) due to the xenon atom optimising to a position just away from the centre of symmetry. Thermal motion will average over the positions of the xenon atom, giving the $R\bar{3}$ structure.
7.3 Detailed modelling of the most probable structures.

Detailed calculations were performed on clusters comprising of a single cavity, whose geometries were taken from the periodic, PBE-TS planewave optimizations. Calculations were performed using the Amsterdam Density Functional (ADF) code including 1-component scalar-relativistic (SR-ZORA) and 2-component spin–orbit (SO-ZORA) methods. The ADF calculations of NMR shielding tensors employed the finite Gaussian nuclear model with all-electron jcpl and TZP basis sets for Xe and other atoms, respectively.

The influence of the DFT functional and the amount of the exact Hartree-Fock exchange (EEX) was tested by carrying out cluster calculations with pure PBE and BLYP (EEX=0%) as well as hybrid B3LYP (EEX=20%) and BHandHLYP (EEX=50%) functionals. The SR-ZORA cluster results obtained with these functionals in cluster calculations were then used in scaling of the periodic PBE results. Due to cylindrical symmetric cavities of the most probable crystal structures, all Xe shielding tensors are axially symmetric (η = 0) and, hence, the scaling is carried out for the principal axis system (PAS) components, for example:

\[
\sigma_{ii}^{\text{periodic}}(\text{BHandHLYP}) \approx \frac{\sigma_{ii}^{\text{cluster}}(\text{BHandHLYP})}{\sigma_{ii}^{\text{cluster}}(\text{PBE})} \times \sigma_{ii}^{\text{periodic}}(\text{PBE}) \quad (i = 1, 2, 3),
\]

which are then used to compute Xe chemical shift δ and chemical shift anisotropy Δδ.

7.4 Modelling of the effect of Xe dynamics

The effect of Xe dynamics for \(^{129}\text{Xe}\) NMR shielding parameters at \(T = 300\,\text{K}\) was modelled for the few most probable clathrates by canonical NVT Metropolis Monte Carlo (MC-NVT) of Xe motion on a potential energy surface (PES) inside a fixed cluster cavity with fixed PBE-TS optimised geometry. An in-house simulation program was used, in which the Xe atom moves on a PES that is piecewise-linearly interpolated from a nonuniform 3D grid of energies. For each clathrate, the 3D grid consists of at least 343 displacement of the Xe atom around the center of the cavity. At the largest displacement in every direction the concave PES value is more than 5 kJ/mol (600 K) higher than the lowest value. The BHandHLYP functional with Grimme’s D3 dispersion correction was used in TURBOMOLE calculations of the PES. In every 3D grid point also the full \(^{129}\text{Xe}\) NMR shielding tensor was computed at the same level. The above-mentioned def2-SVP basis set was used for light elements of the cluster, while for Xe we used an uncontracted \((27s25p21d4f)\) completeness-optimised (co-r) all-electron basis set, which provides practically the basis set limit nonrelativistic Xe NMR shielding tensors. \(^{129}\text{Xe}\) NMR shielding tensor elements were averaged in MC-NVT simulation of \(10^7\) steps with a maximum 1D step-size of 0.15 Å that gives standard errors smaller that 0.1 ppm for all tensor components. The temperature effects on Xe CS and CSA were then computed as the difference of the PAS values of the averaged tensor with respect to the reference Xe shielding tensor at the center of the cage. Due to cylindrically symmetric cavities, all the Xe shielding tensors are axially symmetric in their PAS and, hence, definition of CSA along the symmetric axis is unambiguous and the asymmetry parameter is practically \(\eta = 0\) for all tensors.
8 Changes in chemical shift of $^{129}$Xe NMR with manipulations on cluster structures

Structures obtained in the CSP process include direct insertion of xenon atom in the lattice voids and further optimization of the periodic structures using empirical potentials. For 11 fluorophenol clathrate CSP-structures with lattice energy close to absolute minimum $^{129}$Xe NMR chemical shift and chemical shift anisotropy (CSA) have been calculated for structures (1) without optimization after Xe insertion, (2) optimised with Xe position at the force-field level used in CSP procedure refinements, and (3) structures as in previous but position of hydrogen atoms optimised with DFT. In these calculations, the MHA/SVP basis set combination with the BP86$^{27,34}$ functional was used in steepest descent optimizations of hydrogen atom positions with other atom coordinates fixed. Results are presented in Figure S7 in arrows connecting results (1) to (3) of the same CSP structure.

Figure S7: Changes in computed $^{129}$Xe NMR parameters due to partial optimization of the CSP clusters.
9 Xe NMR results

In Table S4 we present $^{129}$Xe NMR chemical shifts ($\delta$), anisotropies (CSA, $\Delta\delta$) (in ppm) and tensor asymmetries $\eta$ for 33 CSP-generated $o$-fluorophenol clathrate structures calculated at the screening level of theory.

These calculations were performed on cluster models with nonrelativistic $^{129}$Xe NMR shielding tensors computed with TURBOMOLE using the BHandHLYP hybrid DFT functional with the MHA basis set [22s17p14d2f/15s13p11d2f] for xenon and def2-SVP basis sets for all other atoms.

Table S4: Calculated chemical shifts and CSAs at the screening level of theory for selected $o$-fluorophenol clathrate structures.

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>$\delta$</th>
<th>$\Delta\delta$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>oF_A</td>
<td>oF0_R-3_46_Xe</td>
<td>108.7</td>
<td>-12.9</td>
<td>0.0</td>
</tr>
<tr>
<td>oF_B</td>
<td>oF0_R-3_24_Xe</td>
<td>178.2</td>
<td>-43.6</td>
<td>0.0</td>
</tr>
<tr>
<td>oF_C</td>
<td>oF0_R-3_20_Xe</td>
<td>147.0</td>
<td>15.2</td>
<td>0.0</td>
</tr>
<tr>
<td>oF_D</td>
<td>oF180_R-3_51_Xe</td>
<td>181.4</td>
<td>-10.0</td>
<td>0.0</td>
</tr>
<tr>
<td>oF_E</td>
<td>oF180_R-3_44_Xe</td>
<td>241.4</td>
<td>27.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_17_Xe</td>
<td>116.2</td>
<td>-20.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_38_Xe</td>
<td>180.4</td>
<td>167.2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_6_Xe</td>
<td>240.4</td>
<td>67.4</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_7_Xe</td>
<td>246.1</td>
<td>200.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_11_Xe</td>
<td>291.4</td>
<td>39.7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_19_Xe</td>
<td>304.5</td>
<td>25.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_1_Xe</td>
<td>286.2</td>
<td>196.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_2_Xe</td>
<td>131.7</td>
<td>31.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_3_Xe</td>
<td>286.8</td>
<td>31.2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_9_Xe</td>
<td>226.6</td>
<td>188.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_10_Xe</td>
<td>137.6</td>
<td>177.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_11591_Xe</td>
<td>133.3</td>
<td>47.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_14_Xe</td>
<td>125.9</td>
<td>124.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF180_R-3_15_Xe</td>
<td>161.6</td>
<td>143.9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>oF0_1258_Xe</td>
<td>430.3</td>
<td>-148.8</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>oF0_R-3_39_Xe</td>
<td>72.8</td>
<td>-0.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>oF0_13826_Xe</td>
<td>265.1</td>
<td>69.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>oF0_5586_Xe</td>
<td>264.4</td>
<td>69.4</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_14655_Xe</td>
<td>223.8</td>
<td>-140.9</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_14867_Xe</td>
<td>83.6</td>
<td>-37.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_1752_Xe</td>
<td>77.8</td>
<td>-88.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_1773_Xe</td>
<td>126.8</td>
<td>-36.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_1990_Xe</td>
<td>149.8</td>
<td>-144.8</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_25199_Xe</td>
<td>62.0</td>
<td>35.8</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_2991_Xe</td>
<td>87.7</td>
<td>-28.4</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>oF00_P-1_7665_Xe</td>
<td>98.6</td>
<td>57.7</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>oF00_P21-C_3118_Xe</td>
<td>205.0</td>
<td>-122.0</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>oF00_P21-C_5438_Xe</td>
<td>195.9</td>
<td>134.5</td>
<td>0.1</td>
</tr>
</tbody>
</table>
In Table S5 we report $^{129}$Xe NMR chemical shifts ($\delta$), anisotropies (CSA, $\Delta\delta$) (in ppm) and tensor symmetries $\eta$ for predicted $m$-fluorophenol clathrate structures modelled at the screening level of theory.

Table S5: Calculated chemical shifts and CSAs at the screening level of theory for selected $m$-fluorophenol clathrate structures.

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>$\delta$</th>
<th>$\Delta\delta$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$mF_A$</td>
<td>mF180_R-3_8_Xe</td>
<td>159.7</td>
<td>157.9</td>
<td>0.0</td>
</tr>
<tr>
<td>$mF_B$</td>
<td>mF180_R-3_1_Xe</td>
<td>303.3</td>
<td>170.8</td>
<td>0.0</td>
</tr>
<tr>
<td>$mF_C$</td>
<td>mF0_R-3_11_Xe</td>
<td>134.4</td>
<td>113.1</td>
<td>0.0</td>
</tr>
<tr>
<td>$mF_D$</td>
<td>mF0_R-3_1_Xe</td>
<td>240.0</td>
<td>125.8</td>
<td>0.0</td>
</tr>
<tr>
<td>$mF_E$</td>
<td>mF0_R-3_7_Xe</td>
<td>118.7</td>
<td>117.5</td>
<td>0.0</td>
</tr>
<tr>
<td>$mF_F$</td>
<td>mF180_R-3_5_Xe</td>
<td>313.2</td>
<td>148.1</td>
<td>0.0</td>
</tr>
<tr>
<td>$mF_G$</td>
<td>mF0_R-3_5_Xe</td>
<td>332.6</td>
<td>168.2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF0_R-3_6_Xe</td>
<td>202.1</td>
<td>73.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF0_R-3_9_Xe</td>
<td>146.0</td>
<td>87.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF180_119952_Xe</td>
<td>109.2</td>
<td>49.2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF180_R-3_11_Xe</td>
<td>158.5</td>
<td>42.9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF180_R-3_22_Xe</td>
<td>127.0</td>
<td>52.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF0_R-3_10_Xe</td>
<td>123.7</td>
<td>35.7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>mF180_120304_Xe</td>
<td>233.4</td>
<td>95.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>mF0_18479_Xe</td>
<td>158.6</td>
<td>81.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>mF0_R-3_4_Xe</td>
<td>57.6</td>
<td>-0.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>mF180_119925_Xe</td>
<td>132.7</td>
<td>86.1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>mF180_120263_Xe</td>
<td>130.1</td>
<td>-76.9</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>mF180_R-3_18_Xe</td>
<td>157.1</td>
<td>152.4</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>mF00_P-1_14373_Xe</td>
<td>161.6</td>
<td>-89.4</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>mF00_P-1_256_Xe</td>
<td>104.6</td>
<td>-63.6</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>mF00_P21_4252_Xe</td>
<td>210.4</td>
<td>-100.0</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>mF0180_P-1_4959_Xe</td>
<td>136.5</td>
<td>-98.2</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>mF180180_P-1_11297_Xe</td>
<td>137.3</td>
<td>71.9</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>mF180180_P-1_12542_Xe</td>
<td>287.7</td>
<td>175.3</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>mF180180_P-1_20220_Xe</td>
<td>175.9</td>
<td>126.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>mF180180_P-1_22201_Xe</td>
<td>255.3</td>
<td>122.9</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>mF180180_P21_3636_Xe</td>
<td>200.1</td>
<td>72.7</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>mF180180_P21-C_1194_Xe</td>
<td>181.1</td>
<td>75.3</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>mF180180_P21-C_8331_Xe</td>
<td>164.0</td>
<td>86.7</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>mF180180_P21-C_9677_Xe</td>
<td>141.4</td>
<td>87.6</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>mF180180_P21-C_9808_Xe</td>
<td>214.5</td>
<td>77.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Table S6: Calculated $^{129}$Xe chemical shifts $\delta$ and chemical shift anisotropies $\Delta \delta$ from periodic GIPAW/PBE for five likely $o$-fluorophenol clathrate structures at their PBE-TS optimized geometries. The asymmetry parameter $\eta$ is zero for all structures.

<table>
<thead>
<tr>
<th>Structure</th>
<th>$\delta$ (ppm)</th>
<th>CSA, $\Delta \delta$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$oF_A$</td>
<td>229.8</td>
<td>-38.9</td>
</tr>
<tr>
<td>$oF_B$</td>
<td>273.2</td>
<td>-86.7</td>
</tr>
<tr>
<td>$oF_C$</td>
<td>236.4</td>
<td>-15.4</td>
</tr>
<tr>
<td>$oF_D$</td>
<td>285.0</td>
<td>-46.5</td>
</tr>
<tr>
<td>$oF_E$</td>
<td>402.7</td>
<td>2.0</td>
</tr>
<tr>
<td>Experimental</td>
<td>256.0</td>
<td>-47.5</td>
</tr>
</tbody>
</table>

Table S7: Calculated $^{129}$Xe chemical shifts $\delta$ and chemical shift anisotropies $\Delta \delta$ from periodic GIPAW/PBE calculations for seven likely $m$-fluorophenol clathrate structures at their PBE-TS optimized geometries. The asymmetry parameter $\eta$ is zero for all structures.

<table>
<thead>
<tr>
<th>Structure</th>
<th>$\delta$ (ppm)</th>
<th>CSA, $\Delta \delta$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$mF_A$</td>
<td>270.6</td>
<td>226.7</td>
</tr>
<tr>
<td>$mF_B$</td>
<td>295.2</td>
<td>216.0</td>
</tr>
<tr>
<td>$mF_C$</td>
<td>229.2</td>
<td>224.8</td>
</tr>
<tr>
<td>$mF_D$</td>
<td>271.4</td>
<td>163.4</td>
</tr>
<tr>
<td>$mF_E$</td>
<td>272.9</td>
<td>192.4</td>
</tr>
<tr>
<td>$mF_F$</td>
<td>314.1</td>
<td>203.7</td>
</tr>
<tr>
<td>$mF_G$</td>
<td>337.2</td>
<td>242.8</td>
</tr>
<tr>
<td>Experimental</td>
<td>228.5</td>
<td>183.3</td>
</tr>
</tbody>
</table>
10 Powder X-ray Diffraction

10.1 Sample preparation

Samples for powder X-ray diffraction (PXRD) measurements were prepared in similar conditions as the ones used in NMR experiments. Samples were placed into glass capillary tubes with 0.8 mm outer diameter, which were checked to provide low amorphous diffraction background. The capillary tubes were then sealed from one end and transferred to glass reactor tubes (4 mm outer diameter pyrex with 0.8 mm thick walls) connected to a volume calibrated vacuum line. Samples was evacuated, after which xenon was added with the help of liquid nitrogen. The amount of Xe gas was controlled with the pressure drop in the vacuum line and chosen to fill the potential voids with 3:1 (host:guest) stoichiometry and stabilise the clathrates according to available phase diagrams. Samples were flame-sealed and equilibrated for approximately two weeks at 240 K before PXRD experiments. The capillary was kept immersed in liquid nitrogen in order to preserve the clathrate phase before mounting it in pre-cooled single crystal diffractometer.

10.2 Experiments

Diffraction data was collected on a Bruker Nonius Kappa-CCD diffractometer at 100 K using graphite monochromatic Mo Kα radiation (λ = 0.71073 Å; 55 kV, 25 mA). Single frames with static samples were recorded first in order to confirm the presence of a crystalline phase and to obtain initial information about preferred orientation and size distribution of crystallites. The sample was rotated during PXRD measurement (exposure times between 20 to 300 s) in order to simulate randomised powder sample. Data reduction was performed with SUPERGUI diffractometer software (single crystal to powder routine) and thus obtained PXRD patterns further analysed and fitted with MAUD.58

10.3 Results and conclusions

10.3.1 m-fluorophenol

Several samples of m-fluorophenol were prepared, as described above. Unfortunately, no other phase than the known high density stable form (CSD refcode QAMTUU) was observed (Fig. S8). Some outgassing was observed upon melting of powder inside the capillary under a microscope but the amount of gas was rather small suggesting that the amount of actual clathrate was negligible. No change in phase or crystal texture was observed in the PXRD pattern when the temperature was raised from 100 K to 240 K (Fig. S9).
Figure S8: Experimental (orange) and simulated (blue) PXRD patterns for \textit{m}-fluorophenol. The simulated pattern is for known high density stable form (CSD refcode QAMTUU).

Figure S9: PXRD patterns for a sample of \textit{m}-fluorophenol. Exposure at 100 K is obtained with sample freshly recovered from xenon pressurised pyrex tube and the pattern at 240 K is recorded after warming up the sample.
Figure S10: Comparison of simulated PXRD for the known non-clathrate phase of o-fluorophenol (CSD refcode QAMWEH, grey) with experimental PXRD from two separate samples of clathrates (blue, orange).

10.3.2 o-fluorophenol

The PXRD pattern of o-fluorophenol demonstrates the presence of a phase other than the known high density form (CSD refcode QAMWEH). Interestingly, we observe (Fig. S10) that the PXRD patterns are different when the two samples have different history (slightly different pressures, handling during the recovery of capillaries, etc.). One of these patterns (sample 1, Fig. S10) agrees with the known high density form (CSD refcode QAMWEH). The pattern of the second sample (sample 2, Fig. S10) resembles the pattern of the predicted oF,D structure (S11). However, the quality is not sufficient to obtain conclusive confirmation of the structure.

The presence of clathrate was confirmed by strong outgassing upon raising the temperature after the PXRD measurement. In some cases, vigorous decomposition of o-fluorophenol samples could be observed immediately when the temperature was raised during handling the capillaries (mounting in goniometer).
Figure S11: Comparison of the PXRD measured from sample 2 (orange) to the simulated pattern from the predicted structure (blue), oF_D.
Figure S12: Experimental (black) and calculated (indigo) $^{129}$Xe NMR powder spectra of $o$-fluorophenol (top) and $m$-fluorophenol (bottom) with $^{129}$Xe gas reference at zero. Calculated spectra for all structures were obtained from BHandHLYP-scaled periodic PBE NMR parameters at PBE-TS optimized geometry (method/structure = BHandHLYP/PBE-TS). For the most probable candidates, $oF_D$ and $mF_A$, results from different periodic modeling levels are displayed: (1, green) PBE/CSP, (2, orange) PBE/PBE-TS, (3, indigo) BHandHLYP/PBE-TS and (4, red) BHandHLYP/PBE-TS with effects due to Xe dynamics at $T = 300 \text{K}$ (see text for details).
Figure S13: The crystal energy landscape of o-fluorophenol. Each point represents one crystal structure and is coloured according to the molecular conformation (cis-trans-o-fluorophenol refers to $Z' = 2$ structures containing both conformers). Structures with a lattice energy within 13 kJ/mol from the lowest energy structure and having cavities of suitable size for xenon absorption (see text) are encircled in red. The experimentally known high-pressure polymorph (CSD refcode QAMWEH01) is encircled in green. The labels A–E correspond to structures oF_A to oF_E in the text.
Figure S14: The crystal energy landscape of $m$-fluorophenol. Each point represents one crystal structure and is coloured according to the molecular conformation ($cis$-$trans$-$m$-fluorophenol refers to $Z' = 2$ structures containing both conformers). Structures with a lattice energy within 13 kJ/mol from the lowest energy structure and having cavities of suitable size for xenon absorption (see text) are encircled in red. The experimentally known stable polymorph (CSD refcode QAMTUU) is encircled in green. The labels A–G correspond to structures $mF_A$ to $mF_G$ in the text.
Figure S15: $^{129}$Xe NMR parameters computed at screening (nonrelativistic DFT/BHandHLYP/MHA) level for cluster models of the predicted clathrate structures of $o$-fluorophenol. Structures with asymmetry $\eta > 0.04$ are shown as blue crosses, while structures with symmetric chemical shift tensors ($\eta \leq 0.04$) are shown as green squares. The experimental data (expt, black filled circle) is included for comparison. Structures that were considered for further study are connected with the black arrows to the periodic GIPAW/PBE calculation of Xe shielding from periodic PBE-TS optimized structures of the corresponding crystal (red diamonds). The labels A–E correspond to structures $oF_A$ to $oF_E$ in the text.
Figure S16: $^{129}$Xe NMR parameters computed at screening (nonrelativistic DFT/BHandHLYP/MHA) level for cluster models of the predicted clathrate structures of $m$-fluorophenol. Structures with asymmetry $\eta > 0.04$ are shown as blue crosses, while structures with symmetric chemical shift tensors ($\eta \leq 0.04$) are shown as green squares. The experimental data (expt, black filled circle) is included for comparison. Structures that were considered for further study are connected with the black arrows to the periodic GIPAW/PBE calculation of Xe shielding from periodic PBE-TS optimized structures of the corresponding crystal (diamonds). The labels A–G correspond to structures $mF_A$ to $mF_G$ in the text.
Figure S17: The periodic GIPAW results for the five most likely o-fluorophenol structures optimized at the periodic PBE-TS level of theory. The labels A–E correspond to structures $oF_A$ to $oF_E$ in the text. We expect the correct structure to approach the experimental (expt) $^{129}$Xe NMR parameters, when the computed GIPAW/PBE result is scaled with factors obtained using different pure and hybrid DFT functionals with increasing amount of exact exchange in the series of PBE→BLYP(0%)→B3LYP(20%)→BHandHLYP(50%). The SO correction is added to BHandHLYP values of all structures (blue diamonds). For structures $oF_B$ and $oF_D$, the final points (indigo crosses) include also the effect of Xe dynamics (DYN) at $T = 300$ K.
Figure S18: The periodic GIPAW results for the seven $m$-fluorophenol structures optimized at the periodic PBE-TS level of theory. The labels A–G correspond to structures $mF_A$ to $mF_G$ in the text. We expect the correct structure to approach the experimental (expt) $^{129}$Xe NMR parameters, when the computed GIPAW/PBE result is scaled with factors obtained using different pure and hybrid DFT functionals with increasing amount of exact exchange in the series of PBE→BLYP(0%)→B3LYP(20%)→BHandHLYP(50%). The SO correction is added to BHandHLYP values of all structures (blue diamonds). For structures $mF_A$, $mF_B$, $mF_E$, and $mF_F$, the final point (indigo crosses) include also the effect of Xe dynamics (DYN) at $T = 300$ K.
References


