

# The influence of the crystal sponge framework on guest molecule conformation<sup>†</sup>

Eleanor M. Soper, Simon J. Coles, and Graeme M. Day\*

The crystalline sponge (CS) method has become an important technique for structural elucidation of compounds that are challenging to crystallise. However, the impact of the CS environment on guest molecule conformations has not been systematically studied. We present a computational investigation of the conformations of organic molecules of varying flexibility in a set of experimentally determined CS structures, comparing them to gas phase conformers and, where available, pure and co-crystal structures. Via solid state and molecular density functional theory calculations, we quantify the total relative energy, conformational energy, and intramolecular strain of guest molecules, as well as framework strain. Our results show that while CS structures induce some distortion in guest geometries (total relative energies up to 41 kJ/mol), they generally adopt low-energy conformations, often within 2 kJ/mol of the global energy minimum. Intramolecular strain in CS structures is often lower than in conventional crystal structures, suggesting a more neutral packing environment where molecules are closer to their favoured isolated-molecule geometries. We also observe that multiple guests can influence each other's geometries, even in the absence of direct guest-guest interactions. These findings provide a quantification of conformational distortion that can form the basis for interpreting molecular geometries obtained from CS structures.

## 1 Introduction

The accurate determination of molecular structure is of critical importance in chemical research. Single crystal X-ray diffraction (SCXRD) is considered to be one of the most powerful methods of structure elucidation, giving the ability to determine the absolute configuration and stereochemistry of molecules with relative ease. A major drawback of the technique is that, as the name suggests, a good-quality single crystal needs to be produced. Many synthetic organic or natural products can be difficult to crystallise, produced in minute quantities, or are liquid at ambient temperatures, and as a result are often unsuitable for SCXRD analysis.

The crystal sponge (CS) method, first reported by Fujita and coworkers<sup>1</sup> in 2013, has become an important technique in the structural elucidation of compounds considered "uncrystallisable". The technique removes the necessity of sample crystallisation by utilising the porosity of crystalline metal-organic frameworks (MOFs) to form host-guest complexes with the desired compound, which can then be analysed by SCXRD.<sup>2</sup> The CS method is particularly useful in pharmaceutical research, where knowing the exact structure and stereochemical information of a molecule is of critical importance.

The most successful and generally applicable MOF used for the CS method is the  $[(ZnX_2)_3(\text{tpt})_2 \cdot x(\text{solvent})]_n$  system, (tpt = 2,4,6-tris(4-pyridyl)-1,3,5-triazine ligand; X = I, Br, Cl). The system has several advantageous properties including large, hydrophobic pores (5 x 8 Å) with electron-deficient binding sites offered by the tpt ligand. The ZnX<sub>2</sub> components and pyridyl protons of the tpt ligands provide hydrogen-bond donor and acceptor sites, respectively, which enables reproducible and long-range

ordering of the guests within the MOF pores and absorption of organic and aromatic molecules of a range of sizes. The interpenetrated framework allows for some flexibility and expansion of the pore size to accommodate larger guests.<sup>3,4</sup>

The study of MOFs and host-guest complexes by computational methods is a growing area of research, dominated by studies aiming to model the uptake of gases, drug delivery, and calculation of binding energies and interaction strengths of small molecules in the pores of frameworks by molecular dynamics, Monte Carlo simulations, and DFT methods.<sup>5-9</sup> There have been fewer studies directed at MOFs containing flexible molecules, and specifically CS structures. There are certain challenges that arise due to the size and complexity of these systems, especially when considering the presence of large guest molecules and solvent(s); however, the development of molecular simulation methods and access to more powerful high performance computing has meant that modelling such systems is more accessible, and has led to an increasing number of MOFs being studied computationally.

While there are many examples of computational approaches being applied to MOFs, the  $(ZnX_2)_3(\text{tpt})_2$  structures most commonly used for the CS method are not well-studied from a computational perspective, despite the large subset of >500 CS structures available in the Cambridge Structural Database<sup>10</sup> (CSD). Investigation into the CS method has primarily focused on the determination of novel, otherwise intractable, molecular structures and the improvement of experimental procedures.<sup>3,11-16</sup> At the time of writing, the only study of CS structures using computational methods that the authors are aware of is the work of Cardenal and Ramadhar<sup>17</sup>, who demonstrated the use of dispersion-corrected density functional theory (DFT) to analyse a set of previously-published experimental crystalline sponge structures and obtained binding energies of guests within the pores of the framework.

This study seeks to provide a fundamental understanding of the extent to which the CS structure impacts the geometry of

School of Chemistry, University of Southampton, Southampton, SO17 1BJ, UK; E-mail: e.soper@soton.ac.uk, s.j.coles@soton.ac.uk, g.m.day@soton.ac.uk

<sup>†</sup> Electronic Supplementary Information (ESI) available: Details of crystal and molecular geometry optimisations, conformational searches, calculated strain, conformational, and total relative energies for all molecules, correlation between strain energy and molecular geometry changes.

guest molecules and the type of environment that is present in the pores of the CS. The geometries of guest molecule(s) from the optimised CS structures are compared to the DFT conformer landscape to quantify the intramolecular strain and conformational energetics of the distortion induced by the pore environment. Comparison to pure and cocrystal structures, where available, aids further understanding of the exact nature of the pore, defining if the environment of the CS is similar or different from typical, purely organic crystal structures.

## 2 Methods

### 2.1 Geometry Optimisations

The first step is to obtain geometric and energetic information for the conformations of a molecule in its observed crystal structures. The observed molecular geometries and energies are evaluated after optimisation of the experimentally determined crystal structures, which removes any experimental artefacts and errors in atomic positions remaining after structure refinement.

For consistency, the same level of theory and calculation settings are used for all crystal structure and single molecule energy calculations. All density functional theory (DFT) calculations were performed in the CRYSTAL17<sup>18</sup> software using the POB-TZVP basis set<sup>19,20</sup> and B3LYP functional<sup>21</sup> in combination with the D3 dispersion correction<sup>22</sup>. To account for the basis set superposition error (BSSE) that arises in finite basis sets, the geometric counterpoise (gCP) correction<sup>23</sup> was implemented in all calculations.

The symmetry-independent (unique) guest molecules were extracted from the DFT optimised experimental crystal structures. Single-point energy calculations were performed to obtain the energies of the isolated molecules from their crystalline state,  $E_{crystal}^{mol}$  (see red point in Fig. 1). The crystalline molecular geometry was also used to obtain the associated local energy minimum on the conformational energy surface,  $E_{local}^{opt,mol}$ , by running unconstrained geometry optimisations to relax the molecule to the nearest local energy minimum (blue point marked LM in Fig. 1). The energy difference between the single-point and relaxed molecular geometries of the unique molecules within the crystal structures is the intramolecular strain energy,  $\Delta E_{strain}$  (see Fig. 1 and Eq. 1):

$$\Delta E_{strain} = E_{crystal}^{mol} - E_{local}^{opt,mol} \quad (1)$$

which is an energetic measure of the distortion of the molecule in the CS away from its nearest stable gas phase geometry.

### 2.2 Conformational Searches

The CREST conformer search<sup>24</sup> method was applied to generate complete sets of conformers for each guest molecule studied. To ensure thorough sampling, we generated multiple CREST searches from diverse starting structures to fully search the conformational landscape. Duplicates from the combined searches are removed by clustering based on torsion angles within a maximum energy window of 2 kJ/mol and torsion angle tolerances were set to 10.0° and 5.0° for the single angle maximum and the maximum RMS angle, respectively.

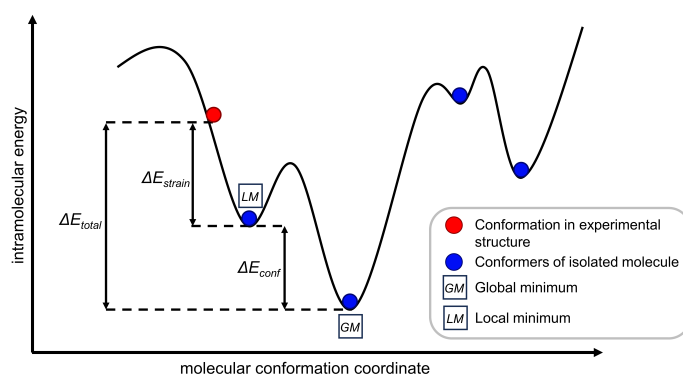


Fig. 1 Schematic of a conformational energy surface: the molecular energy vs a general conformational degree of freedom. The molecular geometry found in the experimental crystal structure is represented as a red point and the conformers of the isolated molecule are represented as blue points. The global minimum conformer, at the lowest point on the potential energy landscape, is indicated by GM. The local minimum conformer of the guest in the experimental structure is indicated by LM.

All unique conformers resulting from the initial search were re-optimised with DFT-D3, applying the same computational parameters as described in Section 2.1. The result should be a complete set of low energy conformers for each molecule. The DFT-optimised conformers were re-clustered to remove duplicates, resulting in the final DFT conformer landscape.

To identify the local conformer of the isolated molecule from the observed crystal structures in the conformational landscape, geometric comparisons with the sets of conformers were performed. The difference in energy between the lowest energy generated conformer (the global energy minimum conformer),  $E_{GM}$ , and the energy of the local conformer corresponding to the crystalline molecular geometry,  $E_{local}^{opt,mol}$ , is referred to as the conformational energy,  $\Delta E_{conf}$  (see Fig. 1 and Eq. 2):

$$\Delta E_{conf} = E_{local}^{opt,mol} - E_{GM} \quad (2)$$

As such, once the energy of the global energy minimum conformer is known, the total relative energy of the isolated molecule from the experimental crystal structure,  $\Delta E_{total}$ , can be found as described in Eq. 3:

$$\Delta E_{total} = \Delta E_{strain} + \Delta E_{conf} = E_{crystal}^{mol} - E_{GM} \quad (3)$$

which is the total energy difference between the molecular conformation in the optimised CS structure and the most stable gas phase molecular geometry.  $\Delta E_{total}$ , and its partitioning into  $\Delta E_{strain}$  and  $\Delta E_{conf}$ , quantify the influence, in energetic terms, of the CS environment on the guest molecule geometry.

### 2.3 Framework Strain

The strain energy of the MOF was also quantified to provide insight into the distortion of the framework that occurs on uptake of the guest molecule. All guest and solvent molecules were removed from the optimised CS structure, leaving the empty  $(ZnX_2)_3(tpt)_2$  framework. The energy of the MOF in the op-

timised CS structure  $E_{MOF}^{crystal}$  was calculated by running a single point energy calculation of the empty framework. An unconstrained geometry optimisation was also performed on the empty MOF structure, applying the same calculation settings as described above. The energy difference between the single-point and relaxed empty MOF structures ( $E_{local}^{opt,MOF}$ ) is the intramolecular strain energy of the framework,  $\Delta E_{MOF}$  (Eq. 4):

$$\Delta E_{MOF} = E_{crystal}^{MOF} - E_{local}^{opt,MOF} \quad (4)$$

Since the space group varies between structures, framework strain energies are quoted per minimum framework unit ( $ZnX_2$ )<sub>3</sub>(tpt)<sub>2</sub>, or 81 atoms, to maintain consistency between structures.

## 2.4 Selection Criteria for Crystalline Sponge Structures

In order to gain insights into the effect of the CS framework on the conformation of guest molecules within the pore, it is important to choose a set of experimental structures that reflect the range of flexibility of guests that have been soaked into CS. While this set must be representative of the full range of different types of analytes, it must also comprise of experimental structures that are of a suitable quality to be a starting point for computational analysis. Making decisions solely based on individual quantitative crystallographic statistics, such as the R-factor, cannot be relied upon as there are many factors, including some qualitative ones, which contribute to the completeness of a CS structure. Therefore, the selection process requires manual filtering to sort through the CS candidates and produce a subset for study. The considerations that have been used in this process, and their justification, are provided in the Supporting Information.

To identify CS structures for this study, the CSD was searched based on the  $ZnX_2$  ( $X = Cl, Br, I$ ) and tpt ligand features of the ( $ZnX_2$ )<sub>3</sub>(tpt)<sub>2</sub> framework. The subset of CS structures were chosen from the resulting CSD hits and selected based on the following criteria: (1) the guest molecule satisfies Lipinski's Rule of Five, (2) no solvent masking methods have been used in the structure refinement, and (3) there is minimal residual solvent accessible void space. The number of rotatable bonds, excluding methyl groups, was also considered to ensure the set of molecules included the full range of possible guest flexibility.

## 2.5 Preparation of the Dataset

CS structures, i.e. those with isolated molecules residing in framework pores, typically have more disorder than conventional organic crystal structures. This can be present in both the framework and also the guest molecules in the pore, which can comprise both the analyte of interest and solvent molecules. The guest molecules can be situated over a symmetry site, or due to low occupancy i.e. not being present in exactly the same location in every pore, found in multiple positions, orientations, or disordered with solvent. Consequently, it is important to consider how disorder affects experimental structures on a case-by-case basis to ensure that the structure taken forward for computational study is as realistic as possible. Full details for dealing with framework,

guest and solvent disorder is described in the Supporting Information.

## 2.6 Structures Chosen for Study

Beyond filtering on quality, the 17 CS structures used in this study were selected primarily to reflect the range of conformation flexibility of guest molecules that have successfully been soaked into this CS framework. The 16 unique guests are detailed in Fig. 2. The flexibility distribution, based on the number of rotatable bonds (excluding methyl groups), can be found in the Supporting Information. A guest-free (solvent-only) structure (CSD Refcode: LABMOT) was also included in the structure set for the purpose of framework strain analysis. If available, the single-component (pure) crystal structures and any co-crystal and/or solvate structures of the guest molecules were also studied to provide comparison.

## 3 Results and Discussion

As many of the chosen structures contain some guest and/or solvent disorder, multiple configurations have been optimised to consider any effects of positional changes in the pores of the CS. The resulting values of  $\Delta E_{strain}$  and  $\Delta E_{conf}$  vary by <1 kJ/mol between configurations; this small energy difference is consistent with the presence of disorder. For discussion, in all further analysis we only include the CS structure configurations with the lowest energy of the guest molecule. These results are summarised in Table 1. The full results are reported in **Table S2**.

The experimental structure CENBUV contains one dapsonone molecule, two disordered n-hexane solvent molecules and two water molecules in the asymmetric unit. Both n-hexane solvent molecules are 2-part disordered with equal occupancies, so four configurations of the structure have been considered to account for all combinations of solvent location. The  $\Delta E_{total}$  of the dapsonone guest in each configuration are within a 2 kJ/mol window of each other (see **Table S2**). Breaking down the  $\Delta E_{total}$  term into its components, we can see this is solely resulting from variation in intramolecular strain energy  $\Delta E_{strain}$ , as all configurations share the same conformation, with a calculated  $\Delta E_{conf} = 1.8$  kJ/mol. Although it is not large, the difference in energies between configurations shows the impact of the proximity of solvent, and therefore sterics, on the total energy of guest molecules.

In the following sections, the impact of the CS environment on the energy and conformation of the guest molecules will be discussed. Any analysis of interactions between the guest and the framework or other guests will be commented on from a qualitative perspective; detailed quantitative discussion, similar to recent work by Carroll<sup>25</sup>, would be required to fully understand the strength of the interactions.

The structure of doxorubicinone (EZEWOY) has unphysically high  $\Delta E_{strain}$  and  $\Delta E_{conf}$  energies. The host framework is well-resolved, however the guest model required a considerable number of restraints and constraints to impose reliable geometry and crystallographic parameters. There is a high likelihood that these have significantly distorted the molecular geometry and resulted in these artificially high calculated energies. For this reason this

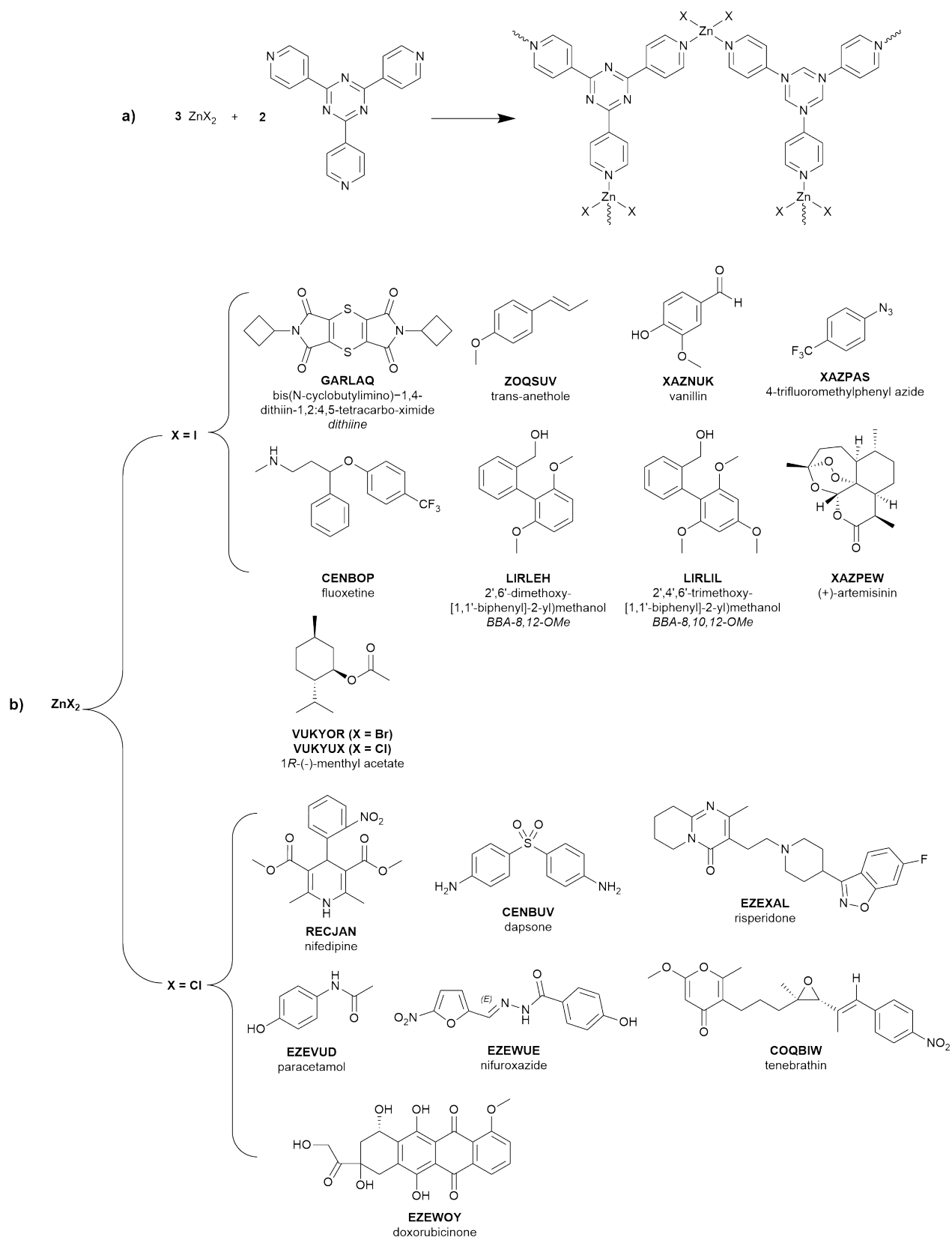


Fig. 2 (a) Schematic for the synthesis of the  $(\text{ZnX}_2)_3(\text{tpt})_2$  crystalline sponges, where  $\text{X} = \text{Cl}, \text{Br}, \text{or I}$ . The zinc (II) ions coordinate to the tpt ligands to form the displayed repeat unit. (b) Chemical diagrams of the guest molecules and associated CSD reference codes of the studied crystalline sponge structures. Each structure includes the guest molecule soaked into the  $(\text{ZnX}_2)_3(\text{tpt})_2$  framework.

structure is an outlier in the group and has not been included further in this analysis. All remaining discussion includes only the remaining 16 CS structures.

### 3.1 Total Relative Energy

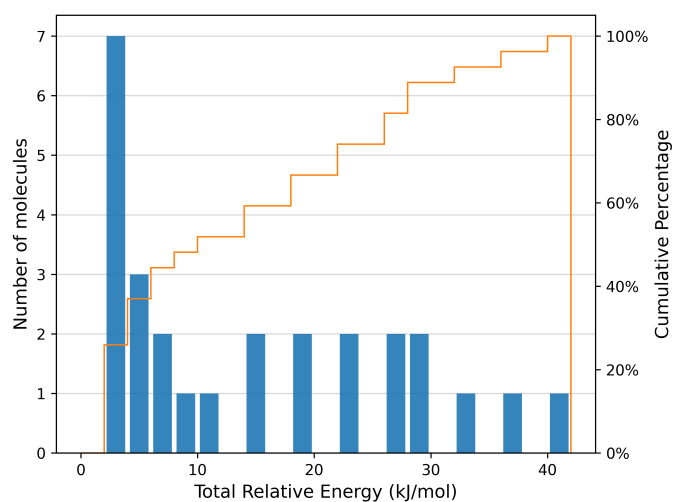


Fig. 3 Histogram detailing the distribution of the total relative energy  $\Delta E_{total}$  of the 27 distinct molecular geometries of the 15 unique guests in the 16 CS structures chosen, with a cumulative percentage of the distribution of molecules as a function of strain energy (orange line). Histogram bin width = 2 kJ/mol.

The distortion of any molecule away from its gas phase geometry involves an increase in the intramolecular energy  $\Delta E_{strain}$  and possibly of the conformational energy  $\Delta E_{conf}$ , if the molecule adopts a conformer other than its most favourable gas phase conformation within the CS. Firstly the total relative energy  $\Delta E_{total}$  of a molecule in a CS crystal structure from its gas phase geometry is analysed. This is calculated as the difference between the energy of the isolated molecule in the geometry from its DFT optimised crystal structure and the gas phase (global minimum) energy, according to Equation 3.

The distribution of  $\Delta E_{total}$  is detailed in Fig. 3 for the 27 distinct guest geometries exhibited in the 16 optimised experimental CS structures outlined in Fig. 2. The distribution includes the total relative energies of each independent molecule found within each crystalline sponge structure.

Molecules which would typically be considered rigid are slightly distorted and therefore have a non-zero total relative energy: we calculate values of  $\Delta E_{total}=2.50$  kJ/mol for 4-trifluoromethylphenyl azide (CSD Refcode: XAZPAS), and 3.99 kJ/mol for (+)-artemisinin (XAZPEW), which has zero rotatable bonds, excluding terminal methyl groups (Fig. 4a). Although the geometric difference is small, this demonstrates the crystal packing effects present in CS structures induce distortions of even the most rigid of molecules. The lowest calculated  $\Delta E_{total}$  is 2.42 kJ/mol for one of the four guest molecules of vanillin (XAZNUK) and one third (9 of 27) of molecules studied have a calculated  $\Delta E_{total}$  less than 5 kJ/mol (approximately 2RT at room temperature). These mostly correspond to the smaller molecules with

few rotatable bonds (Table 1): vanillin, (+)-artemisinin, trans-anethole, 4-trifluoromethylphenyl azide and 1R(-)-menthyl acetate.

The conformations of larger molecules with more flexible degrees of freedom are more substantially influenced by the CS environment. Ignoring the outlier, doxorubicinone, calculated  $\Delta E_{total}$  of guests in CS structures range up to 41 kJ/mol. Examples of the highest energy guests include nifedipine (RECJAN,  $\Delta E_{total}=32.7$  kJ/mol), fluoxetine (CENBOP,  $\Delta E_{total}=33.7$  kJ/mol), and tenebrathin (COQBIW,  $\Delta E_{total}=41.0$  and 36.2 kJ/mol in the two distinct guest sites), which all show significant geometric distortion from their gas phase conformation in the CS (see Figs. 4b, 4c and 5).

From the perspective of understanding the impact of the CS on the conformation of a guest molecule, these results demonstrate that the interactions with guest molecules in the pore environment often play a role in dictating their molecular conformation: there is always some distortion of molecular structures determined using the CS method away from the intrinsically most stable geometry of the isolated molecule. While this can result in small energetic changes in the structures of small molecules, the effect is often large and should be considered when interpreting CS structures.

We analyse  $\Delta E_{total}$  in more depth in the following sections, where the distortion is broken down into conformational change  $\Delta E_{conf}$  and intramolecular strain  $\Delta E_{strain}$ .

### 3.2 Conformational Energy: which conformation is adopted in CS structures?

During crystallisation, molecules are not restricted to their lowest energy conformer, and the balance of inter- and intramolecular interactions can sometimes lead to higher energy conformers.<sup>30</sup> We quantify this with  $\Delta E_{conf}$ , the difference in energy between the conformer that relates to the crystalline molecular geometry and the global minimum energy conformer found by the conformational search process (see Fig. 1). The distribution of conformational energies of guest molecules in CS structures, and comparison to those in pure and cocrystal structures, is detailed in Fig. 6.

For all molecules, a corresponding conformer to that obtained from the experimental crystal structure is found in the generated DFT conformer landscape. Most guest molecules do not adopt the global minimum energy conformer in their CS structures: only 6 out of 16 CS structures contain a guest molecule which adopts the global energy minimum conformer. (+)-artemisinin (CSD Refcode: XAZPEW), 4-trifluoromethylphenyl azide (XAZPAS), and the dithiine guest (GARLAQ), all of which have  $\Delta E_{conf}=0$  kJ/mol, are reasonably rigid molecules. Three further structures containing vanillin (XAZNUK) and (1R)-(-)-menthyl acetate (VUKYUX and VUKYOR) adopt the global minimum conformer in their experimental CS crystal structures; however, these have small, non-zero (up to 0.6 kJ/mol) calculated  $\Delta E_{conf}$ , which is considered negligible once a visual comparison is made with the corresponding global minimum energy conformer. These small energetic discrepancies provide a measure of the numerical noise in the molec-



Table 1 Calculated strain, conformational, and total relative energies of the observed molecular geometries and RMSD of atomic positions of the guest with its optimised isolated conformer. Structures are numbered according to the chosen configuration. Where multiple guests are present, these are labelled A to D.

MOF type	CSD Refcode	Structure <sup>a</sup>	Guest	$\Delta E_{strain}$ (kJ/mol)	RMSD (Å)	$\Delta E_{conf}$ (kJ/mol)	$\Delta E_{total}$ (kJ/mol)	$\Delta E_{MOF}^b$ (kJ/mol)
ZnI <sub>2</sub>	LABMOT <sup>3</sup>	1		-	-	-	-	16.06
	XAZPEW <sup>14</sup>	1		3.99	0.042	0.00	3.99	4.53
	XAZNUK <sup>14</sup>	1	A	9.87	0.060	4.24	14.10	19.83
			B	3.31	0.071	0.01 <sup>c</sup>	3.33	
			C	2.41	0.034	0.01 <sup>c</sup>	2.42	
			D	7.81	0.124	0.01 <sup>c</sup>	7.81	
	ZOQSUV <sup>12</sup>	1	A	4.09	0.127	0.17	4.26	8.79
			B	3.37	0.129	0.16	3.53	
			C	6.73	0.101	0.18	6.90	
	XAZPAS <sup>14</sup>	1		2.50	0.094	0.00	2.50	7.19
	CENBOP <sup>26</sup>	2		17.33	0.465	8.74	26.07	10.43
	GARLAQ <sup>27</sup>	1		18.16	0.185	0.00	18.16	16.82
	LIRLIL <sup>25</sup>	2	A	23.39	0.577	5.13	28.52	9.98
B			5.98	0.102	22.54	28.53		
LIRLEH <sup>25</sup>	1	A	19.93	0.423	3.81	23.76	9.10	
		B	5.23	0.085	21.05	26.28		
ZnCl <sub>2</sub>	CENBUI <sup>26</sup>	4		13.30	0.371	1.83	15.13	101.83
	RECJAN <sup>26</sup>	1		25.92	0.651	6.73	32.65	16.89
	EZEWUE <sup>28</sup>	3		10.99	0.385	0.57	11.56	14.26
	COQBIW <sup>29</sup>	1	A	12.56	0.404	28.45	41.01	17.49
			B	16.24	0.481	19.98	36.21	
	EZEXAL <sup>28</sup>	6	A	12.41	0.432	6.87	19.28	18.07
			B	18.39	0.965	5.03	23.42	
	EZEVUD <sup>28</sup>	1		6.82	0.112	1.72	8.54	13.02
	EZEWOY <sup>28</sup>	2		40.80*	0.609*	79.75*	120.55*	13.28*
VUKYUX <sup>11</sup>	1	A	4.66	0.195	0.45 <sup>c</sup>	5.11	20.70	
		B	4.61	0.143	0.02 <sup>c</sup>	4.63		
ZnBr <sub>2</sub>	VUKYOR <sup>11</sup>	1	A	3.08	0.060	0.63 <sup>c</sup>	3.71	3.41
			B	3.41	0.110	0.05 <sup>c</sup>	3.46	

<sup>a</sup> For disordered structures where multiple configurations have been considered, the lowest energy configuration is reported. For the complete set of results see the Supporting Information.

<sup>b</sup> The value of  $\Delta E_{MOF}$  is quoted per minimum framework unit, i.e.  $(ZnX_2)_3(tpt)_2$ , or 81 atoms, rather than per unit cell as quoted in the calculation output, to make the values comparable.

<sup>c</sup> It is important to note here that although 1*R*-(-)-methyl acetate and vanillin have a non-zero  $\Delta E_{conf}$ , geometric comparison with the set of generated conformers shows that the conformer in the experimental structure matches the global minimum energy conformer, and therefore the non-zero energy values are indicative of calculation errors resulting from the convergence tolerances.

\* These data values are included to show that calculations on this structure have been carried out, however they will not be discussed as the calculated energies are unusually high and may have resulted from the use of restraints/constraints on the disordered molecule.

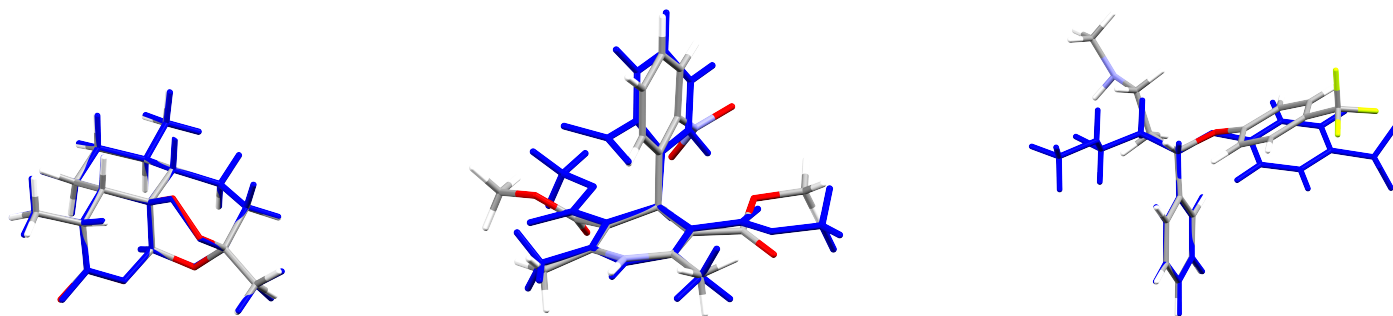


Fig. 4 Overlays of the crystalline geometry of (a) (+)-artemisinin (XAZPEW), (b) nifedipine (RECJAN) and (c) fluoxetine (CENBOP) in the optimised CS structure (element colours) with its gas phase (global energy minimum) conformer (blue).

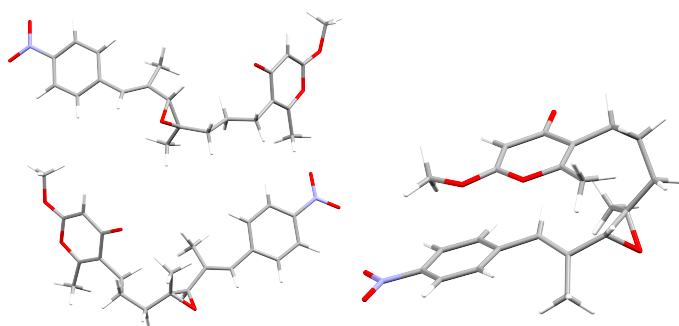


Fig. 5 (a), (b): the two unique crystalline molecular geometries of tenebrathin in the pores of the optimised CS structure (CSD Refcode CO-QBIW) and (c): the gas phase (global energy minimum) structure of tenebrathin.

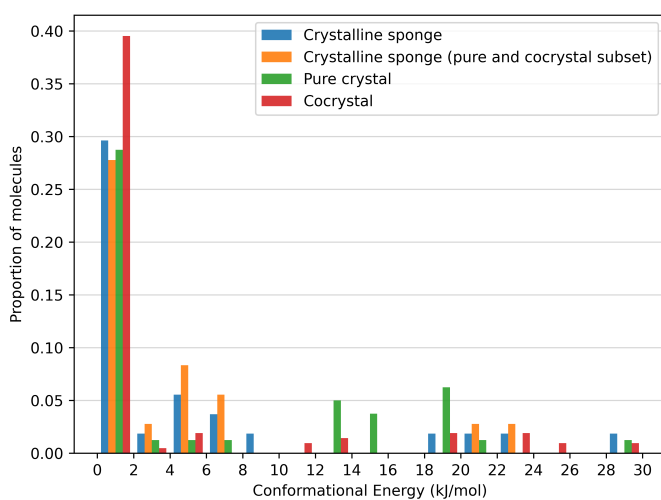


Fig. 6 Histogram detailing the distribution of conformational energies of the guest molecular geometries in all CS structures (blue), subset of CS guests with pure and cocrystal structures (orange), pure crystal (green) and cocrystal (red) structures studied. Histogram bin width = 2 kJ/mol.

ular optimisations applied in this work.

Although it is rare for molecules to adopt their lowest energy conformer in CS structures, for over half of the unique molecules (59%, 16 out of 27 molecules in all structures) the crystalline geometry remains close to the global minimum (below 2.0 kJ/mol). In 23 out of 27 unique guest molecules, the conformational energy lies below 10 kJ/mol: molecules usually adopt low energy conformers in CS structures.

For the remaining molecules, their geometry in the CS structures corresponds to a conformer that is energetically far above the gas phase global energy minimum. The structures LIRLIL and LIRLEH each contain a biaryl guest with high  $\Delta E_{conf}$ , at 22.5 and 21.0 kJ/mol, respectively; both CS structures also contain a second guest with much lower  $\Delta E_{conf}$  (see Table 1). In the gas phase, the global minimum conformer for both BBA-8,12-OMe and BBA-8,10,12-OMe adopts a geometry stabilised by intramolecular hydrogen bonding. The biaryl guests which adopt a local conformation with no intramolecular hydrogen bonds have a high  $\Delta E_{conf}$ , due to the high energetic barrier for geometric distortion/destabilisation of the molecule. The biaryl guests with low  $\Delta E_{conf}$  adopt geometries with intramolecular hydrogen bonding intact after relaxation to their local conformation (see Fig. 7). Both guest molecules of tenebrathin (COQBIW), as mentioned previously in discussion of  $\Delta E_{total}$ , also adopt relatively high energy conformers in their CS structure, with  $\Delta E_{conf}$  = 19.98 and 28.45 kJ/mol.

Previous work by Thompson<sup>30</sup> showed that conformers with large surface areas are favoured in the crystal structures of flexible molecules, as it provides greater available surface area for stabilising intermolecular interactions, while lower energy conformers typically display relatively compact geometries to enable intramolecular stabilisation. This is what we observe for tenebrathin, which adopts extended conformers in its CS structure, compared to the folded global energy minimum conformer (Fig. 5), that could be stabilised by general van der Waals interactions with solvent and the framework. This is not the case for high  $\Delta E_{conf}$  values in LIRLIL and LIRLEH however, where the driving force for these molecules comes from breaking intramolecular hydrogen bonding stabilised by strong, specific interactions with the framework (Fig. 8).

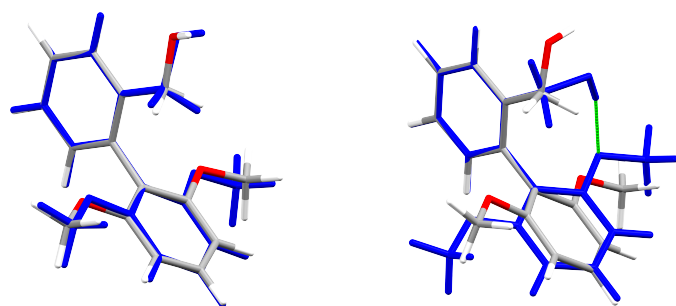


Fig. 7 Overlays of the crystalline molecular geometries (element colours) of the BBA-8,12-OMe guests in the CS (LIRLEH) with their respective local conformers (blue). Intramolecular hydrogen bonding is indicated with a green line.

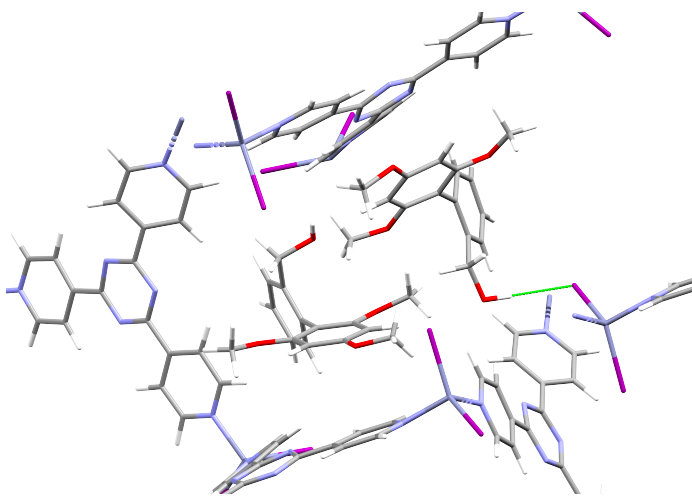


Fig. 8 Structure of two unique BBA-8,10-12-OMe guest molecules in the CS pores (LIRLIL). The highlighted guest-framework interaction relates to the guest with high strain energy, a result of distortion of the local conformer away from intramolecular hydrogen bonding.

### 3.3 Intramolecular Strain Energy

The distortion of any molecule in a crystal structure away from the relaxed isolated molecular geometry involves an increase in the intramolecular energy,  $\Delta E_{strain}$ . This energy is calculated as the difference in energy of the molecule in the optimised crystal structure from its local energy minimum on the conformational energy surface (Fig. 1).

The distribution of  $\Delta E_{strain}$  is detailed in Fig. 9a for the 26 distinct guest geometries in the 16 optimised experimental CS structures outlined in Fig. 2. The distribution includes the strain energies of each independent molecule found within each CS structure configuration. Calculated CS strain energies range up to 26 kJ/mol and are concentrated in the lower end of the distribution: over a third (10 out of 26 guest geometries) have calculated  $\Delta E_{strain} < 5$  kJ/mol and 59% of guest molecules (16 out of 26 guest geometries) have  $\Delta E_{strain} < 10$  kJ/mol. These molecules adopt geometries energetically near to a gas phase optimised conformer in the CS structures.

There is a correlation between the intramolecular strain energy  $\Delta E_{strain}$  of the guest within the CS pores and the RMSD in atomic positions between the crystalline guest molecular geometry and the optimised isolated guest (see Fig. 9b), confirming that large distortions in the molecular geometry are usually associated with a relatively large energetic cost. There are a few cases where significant geometry changes have a relatively low energetic cost, such as risperidone (CSD Refcode: EZEXAL) which has an extended structure; or where a small change in molecular distortion leads to a high strain energy, such as the dithiine guest (GARLAQ). The latter example contains two highly-strained 4-membered rings in the molecular structure, for which there is a large energetic penalty for making relatively small geometric changes.

There are several instances of surprisingly high strain energies above 20 kJ/mol. The highest values calculated from the molecules studied here are 23.39 kJ/mol in BBA-8,10,12-OMe (LIRLIL) and 25.9 kJ/mol in nifedipine (RECJAN). High strain energies are not restricted to the most flexible molecules; two molecules with relatively few flexible degrees of freedom both show a relatively large strain energy in their CS structures: GARLAQ ( $\Delta E_{strain} = 18.16$  kJ/mol, 2 rotatable bonds) and CENBUV structure 1 ( $\Delta E_{strain} = 15.50$  kJ/mol, 4 rotatable bonds).

### 3.4 Pure and cocrystal structures

To understand how the CS pore environment differs from typical crystal packing, the conformational energetics of molecules in their pure and cocrystal structures have been compared to those calculated for the CS structures. The CS method is used primarily for structure elucidation of substances that do not crystallise or are produced in minute quantities, so several guest molecules in this study do not have any non-CS crystal structures which can be studied. Of those that do have crystal structures, there are 10 molecules with pure (single-component) structures and 7 molecules with cocrystal structures, including 6 which have both pure and cocrystal structures. The complete set included in this study is detailed in **Tables S3 and S4**.



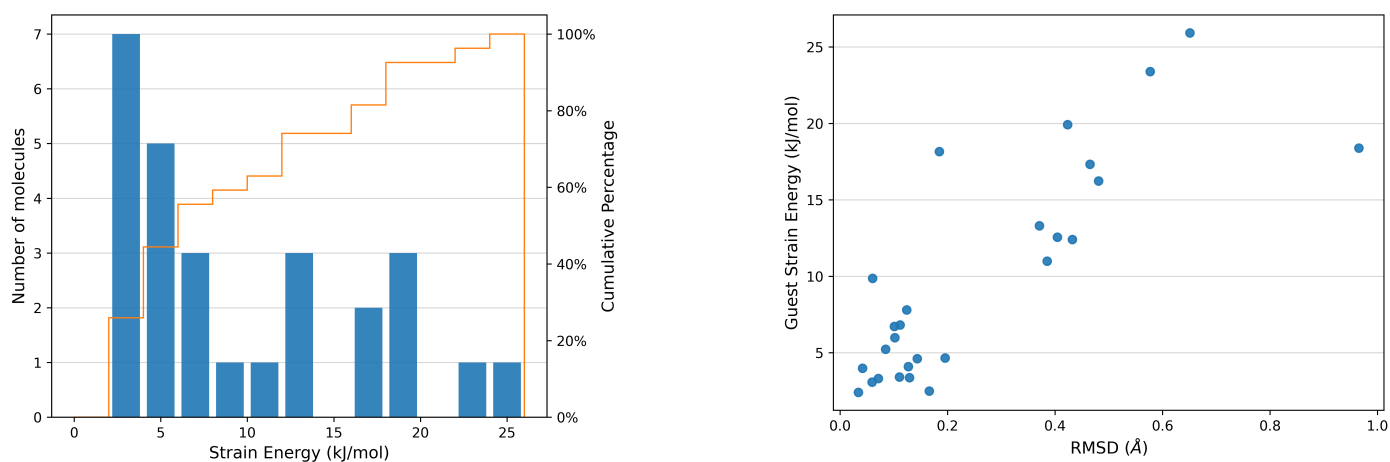


Fig. 9 (a) Histogram detailing the distribution of intramolecular strain energies of the CS guest molecules studied, with a cumulative percentage of the distribution of molecules as a function of strain energy (orange line). (b) Plot of  $\Delta E_{strain}$  against RMSD of atomic positions between the optimised crystalline molecular geometry and the conformation of the optimised isolated molecule of the CS guest molecules studied. Histogram bin width = 2 kJ/mol.

The distributions of  $\Delta E_{total}$  in conformers present in CS structures are compared to those in pure and cocrystal structures in Fig. 10. Molecules typically have a lower total relative energy in the CS compared to both their pure and cocrystal structure counterparts. The  $\Delta E_{total}$  distribution for molecules in their pure and cocrystal structures ranges up to 45 kJ/mol, higher than those observed in the CS at 34 kJ/mol. For pure and cocrystal structures the distributions peak between 12-14 and 10-12 kJ/mol, respectively, whereas for CS structures this is observed between 2-4 kJ/mol. For the subset of CS structures where the guest has pure and/or cocrystal structures, 47% of unique molecules have a calculated  $\Delta E_{total}$  below 10 kJ/mol. In comparison, the proportion of unique molecules with  $\Delta E_{total}$  below 10 kJ/mol is much lower in pure and cocrystal structures: 12.5% (5 out of 40) of molecules in pure crystal structures and 20% (23 out of 114) of molecules in cocrystal structures.

The shift to higher  $\Delta E_{total}$  for cocrystals could likely be a consequence of the strong interactions between coformers, which will have an impact on molecular geometry and  $\Delta E_{total}$ . The distribution also has a larger  $\Delta E_{total}$  range, suggesting that the intermolecular interactions with the pore environment of the CS are weaker than in typical crystal structures. The  $\Delta E_{total}$  distribution in Fig. 10b is also skewed by unequal proportions of cocrystal structures: some molecules have many more structures than others and if the number of structures per molecule of interest were equal, the distribution would be flatter. Fig. S2 shows the  $\Delta E_{total}$  distribution for cocrystal structures, broken down by the molecule of interest: there are significantly more cocrystal structures than pure crystal and CS structures, providing greater sampling of cocrystal structures and resulting in an inconsistent comparison.

As previously described, a large proportion of calculated  $\Delta E_{conf}$  is limited to <2 kJ/mol and consequently  $\Delta E_{strain}$  has the biggest contribution to  $\Delta E_{total}$  of guest molecules. The distribution of  $\Delta E_{strain}$  for both pure and cocrystal structures is detailed in Fig.

S3b. Calculated strain energies range up to 28 kJ/mol and peak between 10-12 kJ/mol. Previous conformational studies of single-component organic crystal structures has shown that strain energies range up to 20 kJ/mol and the majority (75%) of strain energies lie below 10 kJ/mol<sup>30</sup>. Comparison to the  $\Delta E_{strain}$  distribution of the subset of 10 CS guest molecules that have pure and/or cocrystal shows that the distribution peaks in the 2-4 kJ/mol bin (Fig. S3a) and only 61% have a strain energy below 10 kJ/mol (11 out of 18 unique molecules). Similarly to CS structures, there is a positive correlation between  $\Delta E_{strain}$  and RMS deviation of atomic positions between the experimental and gas phase local conformer molecular geometries, for both pure and cocrystal structures (Fig. S4).

The increased probability of a higher  $\Delta E_{strain}$  in pure and cocrystal structures suggests that the close-packed nature of typical crystal packing provides greater stability of molecular distortion than interactions in the pores of the CS framework. It should be noted that Thompson's study<sup>30</sup> excluded molecules with potential for intramolecular hydrogen bonding, while the molecules in this study have greater potential for hydrogen bonding. These strong interactions could also lead to a higher  $\Delta E_{strain}$ .

### 3.5 Framework Strain

Quantifying the strain of the CS framework gives insight into the breathing behaviour of the interpenetrated structure and how it is affected upon uptake and exchange of guest molecules. All guest and solvent molecules were removed from the DFT optimised CS structures before running single-point energy and geometry optimisation calculations to find the intramolecular strain energy of the framework,  $\Delta E_{MOF}$ , as described in Section 2.3. Calculated  $\Delta E_{MOF}$  values are reported per minimum framework unit ((ZnX<sub>2</sub>)<sub>3</sub>(tpt)<sub>2</sub>, 81 atoms) and can be found in Table 1. Structure information and cell volume changes are found in Table S5.

A CS structure containing only solvent was included to under-

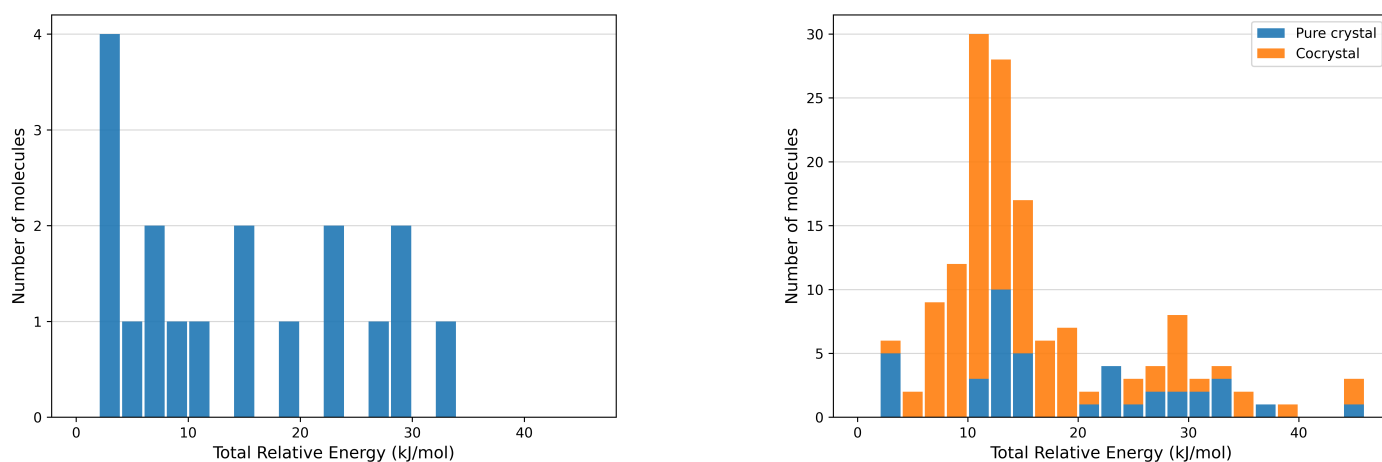


Fig. 10 (a) Histogram detailing the distribution of total relative energies  $\Delta E_{total}$  of CS guests with pure and cocrystal structures and (b) Histogram detailing the distribution of the total relative energies  $\Delta E_{total}$  of the pure and cocrystal structures studied. Histogram bin width = 2 kJ/mol.

stand the impact of a non-interacting guest on the framework. The  $ZnI_2$  structure (CSD Refcode: LABMOT) was chosen due to the inclusion of cyclohexane, present in many structures in this study. The structure has a calculated  $\Delta E_{MOF}$  of 16.06 kJ/mol with only a slight increase in cell volume of 0.62% after all solvent molecules are removed, demonstrating that non-interacting solvent in the pores accounts for some of the induced framework strain, but do not significantly distort the structure.

The CS frameworks have a calculated  $\Delta E_{MOF}$  of up to 21 kJ/mol per minimum unit, and only 7 out of 16 structures studied have a framework strain higher than the solvent-only structure.  $\Delta E_{MOF}$  is on the same order of magnitude as for single molecules, indicating that breathing of the framework does not come at a large energetic penalty. Furthermore, guest exchange does not cause significant amounts of distortion in the framework: the change in cell volume between the optimised experimental structure and the optimised empty framework is typically less than 3%. This exhibits the breathing capabilities of the framework that occurs during the thermodynamic guest exchange process, and does not appear to correlate with framework strain (Fig. S6).

The system including dapsone in the  $ZnI_2$  framework (CSD Refcode: CENBUIV) is an outlier in the subset. The framework has a calculated  $\Delta E_{MOF}$  = 101.83 kJ/mol, an order of magnitude greater than typical values. In addition, the cell volume contracts by 15.76%, much greater than all other structures. The guest molecule in the experimental structure is clearly defined, with no constraints or restraints applied, and there is no mention of unusual observations relating to the framework by the authors.<sup>28</sup> The framework geometry is unusual: a typical CS structure has a unit cell volume of 15,000-16,000  $\text{\AA}^3$  whereas CENBUIV has a volume of 12,000  $\text{\AA}^3$ . The high framework strain could result from a combination of factors. Dapsone and other molecules with similar N-containing nucleophilic functional groups are known for interacting with and breaking down the CS framework<sup>26</sup>; the nucleophilic guest attacking the coordination network could therefore cause the framework to become unstable and collapse. In addition,

the high occupancy of the guest molecule, combined with the presence of water molecules and n-hexane solvent, could stabilise the high framework strain via strong interactions.

Insight into the effect of changing the halide ion within the framework can be gained from comparison of  $ZnCl_2$  and  $ZnBr_2$  analogues containing the same guest molecule. The experimental structures VUKYUX ( $ZnCl_2$ ) and VUKYOR ( $ZnBr_2$ ) containing 1R(-)-menthyl acetate are included in this study. The 1R(-)-menthyl acetate guest molecules are present in the same exchange sites in both structures, with additional presence of one and two  $CHCl_3$  solvent molecules, respectively. Calculated framework strain values show a large difference between each system:  $\Delta E_{MOF}^{av}$  = 3.41 kJ/mol for  $ZnBr_2$ , with an overall cell volume increase of 0.05%, and  $\Delta E_{MOF}^{av}$  = 20.7 kJ/mol for  $ZnCl_2$ , with an overall cell volume reduction of 6.27%. The  $ZnCl_2$  structure also has an overall smaller experimental unit cell size than the  $ZnBr_2$  structure (15,720.4  $\text{\AA}^3$  versus 16,176.6  $\text{\AA}^3$ ). It is clear that guest exchange causes greater distortion of the  $ZnCl_2$  framework due to smaller overall pore volumes, resulting in higher framework strain.

The average framework strain for the  $ZnCl_2$  systems ( $\Delta E_{MOF}^{av}$  = 16.74 kJ/mol) is larger than for the  $ZnI_2$  systems ( $\Delta E_{MOF}^{av}$  = 11.41 kJ/mol).<sup>\*</sup> There are several factors which could have caused this: the size of the guest molecules in the pores, and the overall unit cell and therefore pore size. In this study, the average molecular weight of guest molecules in  $ZnCl_2$  structures (296.8 g/mol) is larger than in  $ZnI_2$  structures (245.0 g/mol). The average unit cell sizes of  $ZnCl_2$  systems is also larger than  $ZnI_2$  systems, at 14,335.1  $\text{\AA}^3$  and 15,793.9  $\text{\AA}^3$ , respectively. While the increased strain in the  $ZnCl_2$  frameworks could simply be due to the larger molecules present in the pores, accounting for the average unit cell sizes of the  $ZnCl_2$  and  $ZnI_2$  structures indicates that the higher  $\Delta E_{MOF}^{av}$  of  $ZnCl_2$  systems is affected by a combination of factors. The reduced pore size of the  $ZnCl_2$  system results in greater distortion of the framework to accommodate the overall larger guest molecule size, ultimately leading to higher

framework strain.

### 3.6 What is the effect of multiple guests in the pore?

It is not uncommon for guest molecules to soak into several exchange sites within the pores of the CS. There are 8 CS structures in this study which feature multiple guests in the asymmetric unit of the unit cell. Energetic differences between guest molecules in the same CS structure are observed, resulting from molecules experiencing different interactions in each exchange site. For all but one CS structure with multiple guests,  $\Delta\Delta E_{total}$  is below 5 kJ/mol (Table 1), but is much larger,  $\Delta\Delta E_{total} = 11.68$  kJ/mol, in the CS structure of vanillin (XAZNUK).

The intermolecular interactions experienced by the molecule, resulting from differences in pore environment, can have a profound effect on  $\Delta E_{strain}$  and  $\Delta E_{conf}$ . To understand how molecular geometry is affected by multiple guests present in the pore, we can look at the structures containing BBA-8,10,12-OMe (LIRLIL) and BBA-8,12-OMe (LIRLEH). Each structure has two guest molecules: guest A with high  $\Delta E_{strain}$  and low  $\Delta E_{conf}$ , and guest B with low  $\Delta E_{strain}$  and high  $\Delta E_{conf}$  (see Table 1). The high  $\Delta E_{strain}$  of guest A and high  $\Delta E_{conf}$  of guest B results from distortion that breaks the intramolecular hydrogen bond within the molecule; in the case of guest A, it is simply intramolecular strain/distortion, whereas for guest B, the complete change in geometry to a conformation with no intramolecular hydrogen bonding is energetically costly (see Fig. 7). The guest with high  $\Delta E_{strain}$  in each structure relaxes to a local conformational minimum stabilised by intramolecular hydrogen bonding. The number and type of interactions present in these two structures were recently discussed in a systematic study.<sup>25</sup> In LIRLIL, interactions between BBA-8,10,12-OMe Guest A and the framework seem to be dominated by polar O—Ar interactions, while Guest B experiences mostly aromatic Ar—Ar, in addition to the long-range dispersion interactions present in the pore environment. Guest-guest interactions between A and B are also observed, which will contribute to the overall molecular strain and conformational changes from the lowest energy conformer.

The experimental CS structure of risperidone in the ZnI<sub>2</sub> framework (EZEXAL) contains two guest molecules in the asymmetric unit and enables a quantification of the effect of multiple guests in the pore. Guest B has an increased induced intramolecular strain compared to guest A ( $\Delta E_{strain} = 12.35$  kJ/mol and 18.43 kJ/mol, respectively), which is most likely due to experiencing stronger interactions in its exchange site. Configurations where only one molecule of risperidone is present have also been considered to investigate the interactions between guests. The calculated  $\Delta E_{conf}$  of guest A and B is the same in all configurations, regardless of the total number of guests present. When considering only Guest B in the pore, the strain energy is unaffected: the calculated  $\Delta E_{strain}$  of 18.13 kJ/mol is only 0.30 kJ/mol less than configurations where both guests A and B are present (Table S2), corresponding to a small RMS deviation in atomic positions of

0.505 Å. Guest A, however, when considered as the only guest in the pore, has a calculated  $\Delta E_{strain}$  of 8.26 kJ/mol: compared to configurations where both guests are present, the RMS deviation in atomic positions is small (0.166 Å), but causes an increase in  $\Delta E_{strain}$  of nearly 4 kJ/mol. While there are no observed direct guest-guest interactions in the experimental and optimised structure, the increase in intramolecular strain energy of guest A when guest B is present demonstrates the influence of dispersive forces between molecules in the pores even when no direct atom-atom interactions are observed.

It is clear from these examples that the number of interactions a guest experiences can have a pronounced effect on the intramolecular strain energy of the guest molecules when there are multiple guests in the pore.

## 4 Conclusions

This computational study provides important insights into the influence of the crystalline sponge (CS) environment on guest molecule conformation, which should inform the interpretation and use of molecular geometries that are obtained from CS structures.

Guest molecules in CS structures, particularly larger molecules with several rotatable bonds, almost always experience some distortion away from their most stable gas phase geometries; total relative energies,  $\Delta E_{total}$ , calculated as the difference between the geometry in the CS structure and the global energy minimum of the isolated molecule, are frequently small (below 5 kJ/mol in a third of structures studied here), but range up to 41 kJ/mol. Partitioning these conformational energetics into the conformer energy of the nearest gas phase conformer to the geometry in the CS structure, and strain energy of molecule (distortion away from the nearest local minimum), provides additional insight into the conformational effects of the CS environment on guest molecules.

The largest energetic effects (highest  $\Delta E_{total}$ ) usually correspond to breaking of strong non-bonded intramolecular interactions, such as intramolecular hydrogen bonds or stabilising interactions from folding of long, flexible molecules, that are present in the most stable geometry of the isolated molecule. In the absence of such effects, molecules in CS structures usually adopt a geometry that is energetically close to one of their lowest energy conformers, although this is often not the lowest possible conformer of the isolated molecule.

By comparing the conformations of guest molecules in CS structures with those in pure, close-packed crystal structures and co-crystals of the same molecules, we find that the molecular geometries from CS structures have a lower total relative energy, and lower molecular strain, than those found in conventional crystal structures. These results suggest that CS structures represent a more "neutral" environment for studying the geometry of organic molecules than conventional crystal structures, providing a closer representation of molecules' intrinsically preferred geometries.

## Author Contributions

E. M. Soper: methodology, investigation, analysis, writing. S. J. Coles: supervision, analysis, writing. G. M. Day: conceptualisation, supervision, methodology, analysis, writing.

\* We have excluded the structures CENBUB and EZEWOY from averaging values due to their unusually high respective framework and guest energies.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank R.C. Carroll for his help with interpreting experimental CS structures. We thank the EPSRC for funding. The authors acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work. Via our membership of the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/R029431), this work also used the ARCHER2 UK National Supercomputing Service (<https://www.archer2.ac.uk>).

## References

- 1 Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen and M. Fujita, *Nature*, 2013, **495**, 461–466.
- 2 N. Zigon, V. Duplan, N. Wada and M. Fujita, *Angewandte Chemie International Edition*, 2021, **60**, 25204–25222.
- 3 M. Hoshino, A. Khutia, H. Xing, Y. Inokuma and M. Fujita, *IUCrJ*, 2016, **3**, 139–151.
- 4 J. P. Metherall, R. C. Carroll, S. J. Coles, M. J. Hall and M. R. Probert, *Chem. Soc. Rev.*, 2023, **52**, 1995–2010.
- 5 M. Ernst, J. D. Evans and G. Gryn'ova, *Chemical Physics Reviews*, 2023, **4**, 041303.
- 6 F.-X. Coudert and A. H. Fuchs, *Coordination Chemistry Reviews*, 2016, **307**, 211–236.
- 7 J.-R. Li, R. J. Kuppler and H.-C. Zhou, *Chemical Society Reviews*, 2009, **38**, 1477.
- 8 M. Ernst and G. Gryn'ova, *ChemPhysChem*, 2022, **23**, e202200098.
- 9 M. Ernst and G. Gryn'ova, *Helvetica Chimica Acta*, 2023, **106**, e202300013.
- 10 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallographica Section B*, 2016, **72**, 171–179.
- 11 T. R. Ramadhar, S.-L. Zheng, Y.-S. Chen and J. Clardy, *Chemical Communications*, 2015, **51**, 11252–11255.
- 12 T. R. Ramadhar, S.-L. Zheng, Y.-S. Chen and J. Clardy, *Acta Crystallographica Section A Foundations and Advances*, 2015, **71**, 46–58.
- 13 L. M. Hayes, N. J. Press, D. A. Tocher and C. J. Carmalt, *Crystal Growth and Design*, 2017, **17**, 858–863.
- 14 T. R. Ramadhar, S.-L. Zheng, Y.-S. Chen and J. Clardy, *CrystEngComm*, 2017, **19**, 4528–4534.
- 15 G. W. Waldhart, N. P. Mankad and B. D. Santarsiero, *Organic Letters*, 2016, **18**, 6112–6115.
- 16 F. Habib, D. A. Tocher and C. J. Carmalt, *Materials Today: Proceedings*, 2022, **56**, 3766–3773.
- 17 A. D. Cardenal and T. R. Ramadhar, *CrystEngComm*, 2021, **23**, 7570–7575.
- 18 R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro and B. Kirtman, *WIREs Computational Molecular Science*, 2018, **8**, e1360.
- 19 J. Laun, D. V. Oliveira and T. Bredow, *Journal of Computational Chemistry*, 2018, **39**, 1285–1290.
- 20 M. F. Peintinger, D. V. Oliveira and T. Bredow, *Journal of Computational Chemistry*, 2013, **34**, 451–459.
- 21 B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano and P. Ugliengo, *CrystEngComm*, 2008, **10**, 405–410.
- 22 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *The Journal of Chemical Physics*, 2010, **132**, 154104.
- 23 J. G. Brandenburg, M. Alessio, B. Civalleri, M. F. Peintinger, T. Bredow and S. Grimme, *The Journal of Physical Chemistry A*, 2013, **117**, 9282–9292.
- 24 P. Pracht, F. Bohle and S. Grimme, *Physical Chemistry Chemical Physics*, 2020, **22**, 7169–7192.
- 25 R. C. Carroll, D. C. Harrowven, J. E. Pearce and S. J. Coles, *IUCrJ*, 2023, **10**, 497–508.
- 26 F. Sakurai, A. Khutia, T. Kikuchi and M. Fujita, *Chemistry – A European Journal*, 2017, **23**, 15035–15040.
- 27 P. Chauhan, S. Javed, D. C. Levendis and M. Fernandes, *Journal of Molecular Structure*, 2022, **1251**, 132054.
- 28 L. Rosenberger, C. von Essen, A. Khutia, C. Kühn, K. Georgi, A. K. Hirsch, R. W. Hartmann and L. Badolo, *European Journal of Pharmaceutical Sciences*, 2021, **164**, 105884.
- 29 S. Hoshino, T. Mitsuhashi, T. Kikuchi, C. P. Wong, H. Morita, T. Awakawa, M. Fujita and I. Abe, *Organic Letters*, 2019, **21**, 6519–6522.
- 30 H. P. G. Thompson and G. M. Day, *Chem. Sci.*, 2014, **5**, 3173–3182.