

GASP: software for geometric simulations of flexibility in polyhedral and molecular framework structures

Stephen A. Wells & Asel Sartbaeva

To cite this article: Stephen A. Wells & Asel Sartbaeva (2015) GASP: software for geometric simulations of flexibility in polyhedral and molecular framework structures, *Molecular Simulation*, 41:16-17, 1409-1421, DOI: [10.1080/08927022.2015.1032277](https://doi.org/10.1080/08927022.2015.1032277)

To link to this article: <https://doi.org/10.1080/08927022.2015.1032277>



© 2015 The Author(s). Published by Taylor & Francis.



Published online: 13 May 2015.



Submit your article to this journal [↗](#)



Article views: 2158



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 9 View citing articles [↗](#)

MOLECULAR SIMULATION OF FRAMEWORK MATERIALS

GASP: software for geometric simulations of flexibility in polyhedral and molecular framework structures

Stephen A. Wells* and Asel Sartbaeva

Department of Chemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK

(Received 23 January 2015; final version received 17 March 2015)

Template-based geometric simulation is a specialised method for modelling flexible framework structures made up of rigid units using a simplified, localised physical model. The strengths of the method are its ability to handle large all-atom structural models rapidly and at minimal computational expense, and to provide insights into the links between local bonding and steric geometry and global flexibility. We review the implementation of geometric simulation in the ‘GASP’ software, and its application to the study of materials including zeolites, perovskites and metal–organic frameworks. The latest version (5) of GASP has significant improvements and extensions, in particular an improved algorithm for relaxation of atomic positions, and the capacity to handle both polyhedral and molecular structural units. GASP is freely available to researchers.

Keywords: geometric simulation; polyhedral rotation; zeolite; perovskite; metal–organic framework

1. Introduction

Template-based geometric simulation is a specialised method for the study of flexible frameworks. Such frameworks are composed of relatively rigid subunits (clusters) connected by relatively flexible linkages; that is, the energetic penalty for distortions of the clusters is an order of magnitude greater than the energetic penalty for flexion in the linkage. In the case of a 3D framework silicate, such as quartz or a siliceous zeolite, the SiO_4 tetrahedra are the relatively rigid clusters, while the Si–O–Si bridge is a more flexible linkage, as indicated by the wide range of Si–O–Si angles observed in different framework silicate structures.[1,2] In a metal–organic framework (MOF), both the organic linker molecules and the coordination polyhedra around metal centres may be treated as rigid clusters, with the potential for flexibility at the interface between them.

The central concept of geometric simulation is to represent the bonding within rigid units, not by a collection of two-, three- and four-body empirical potentials, but by a template or ‘ghost’ representing the ideal bonding geometry of the cluster, either polyhedral or molecular. Harmonic constraints linking each atom to the corresponding vertex of a template then penalise any deviation from the ideal cluster geometry.[3,4] In the course of a ‘geometric relaxation’, the positions of atoms and templates are mutually updated to minimise both deviations from cluster geometry and steric overlap of atomic spheres. Thus the method implements a simple physical model which includes the strongest, most local forces in the system – covalent bonding and steric

exclusion – while neglecting all longer ranged interactions. The method is computationally robust and inexpensive, with results typically being generated in seconds or minutes on a single processor, e.g. on a laptop or desktop computer. This makes the method particularly suitable for the investigation of large system sizes and for the rapid exploration of hypothetical conditions of strain and/or extra framework contents in porous frameworks, on its own or as an adjunct to simulations with conventional methods. The process of fitting cluster geometries to groups of atoms can also be used as an analysis tool to investigate structural models produced by other methods, particularly reverse Monte Carlo (RMC) modelling based on total scattering data.[5–7]

The geometric simulation approach operates at an entirely different level of theory from *ab initio* electronic structure and all-atom empirical-potential methods; it does not generate a detailed energy landscape, but rather explores framework geometries that satisfy local steric and bonding geometric requirements. This simplification can in fact be a powerful generator of insight, revealing the significance or otherwise of the long-range (charge, polarity) effects that are deliberately excluded from the geometric simulation model.

1.1 Scope of our review

Essential mathematical details in the implementation of geometric simulation are discussed (Section 2) for the benefit of current and future users and for those wishing to

*Corresponding author. Email: s.a.wells@bath.ac.uk

make use of geometric simulation concepts in their own software. We will take a brief overview (Section 3) of some of the previous applications of geometric simulation to polyhedral mineral structures – such as aluminosilicates, including zeolites and perovskites – and to proteins, viewed as molecular frameworks. These topics are discussed in additional detail in our recent review [8]. In this work we focus on recent developments, especially (i) improvements and extensions in current geometric simulation software, (ii) the concepts of the intrinsic and extrinsic flexibility windows in a zeolite (Section 4) and (iii) the application of geometric simulation to MOFs (Section 5).

1.2 Software implementation

Geometric simulation for periodic framework structures is implemented in a piece of software titled ‘GASP’, for ‘Geometric Analysis of Structural Polyhedra’. The name ‘GASP’ is intended as a respectful pun on the widely used and extremely comprehensive ‘GULP’ simulation package.[9] Version 1 of GASP was written in Fortran 90; subsequent versions have been written in C++. In 2014 GASP was rewritten from scratch (in C++) to improve program workflow, implement an improved algorithm for the update of atomic positions, remove unwanted legacy features, and provide a more legible and consistent input and control format. This current version, GASP v.5, consists of the main GASP code and a small suite of associated utility programs; it has been successfully

compiled and run under Windows, Linux and Mac OSX environments. In its current form, GASP requires an input structure in XTL format, and a control input file containing parameters and commands. GASP is freely available to researchers. Prospective users should email the corresponding author to receive a copy of the code, a set of example input files and a comprehensive manual.

2. Mathematical notes: bivectors, rotors and fitting

The heart of the geometric simulation method is the fitting of a template to a cluster of atoms. Fundamentally this is a matter of finding a rotation which matches the orientation of the template and cluster. In geometric simulations the rotation is described using the bivector mathematics of geometric (Clifford) algebra.[10] Since this approach is not commonly encountered in the physical sciences literature, we will briefly go over the essential concepts. The process of fitting a template to a group of atoms, and of relaxing the atoms to fit the templates, is illustrated schematically in Figure 1.

2.1 Geometric algebra in two dimensions

In conventional vector calculus, the mathematics of a 2D plane would be described in terms of scalar quantities (based on the unit scalar, 1) and vector quantities (based on two perpendicular unit vectors e_1 and e_2). To these, Clifford algebra adds an additional entity: a directed area element, the bivector $B = e_1e_2$. For clarity we shall

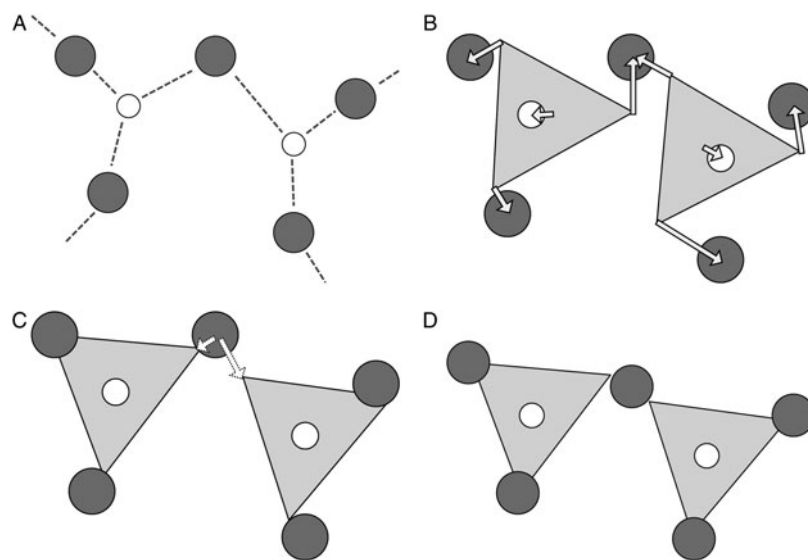


Figure 1. Schematic illustration in 2D of the geometric simulation process. (a) Two approximately trigonal, vertex-sharing clusters in a 2D framework. (b) Initial construction of a template with ideal geometry over each cluster; mismatches are identified between each atom and the corresponding vertex. (c) Rotational fitting of templates by minimisation of mismatches; this generally leaves some residual mismatch, as shown for the shared vertex atom. (d) Relaxation of atomic positions to reduce mismatches. Geometric analysis constitutes steps (a)–(c). Geometric simulation proceeds by iterative repetition of steps (c) and (d).

capitalise variables representing bivectors and higher grad objects. Vectors in Clifford algebra can be multiplied directly; the product of perpendicular vectors is a bivector while that of parallel vectors is a scalar, e.g. $e_1e_1 = 1 = e_2e_2$. The bivector is directed in the sense that $B = e_1e_2 = -e_2e_1$. The bivector can act to rotate the basis vectors by multiplication either from the left, rotating 90° clockwise ($Be_1 = e_1e_2e_1 = -e_2$; $Be_2 = e_1e_2e_2 = e_1$) or from the right, rotating anticlockwise ($e_1e_1e_2 = e_2$; $e_2e_1e_2 = -e_1$).

The 2D bivector also squares to -1 ($B^2 = e_1e_2 \times e_1e_2 = -e_1e_1e_2e_2 = -1$) and so it exponentiates exactly like the conventional scalar unit imaginary i . We can write the identity $\exp(B\theta) = \cos \theta + B\sin \theta$. This $\exp(B\theta)$ object can act to rotate a vector through θ radians. It is convenient to define a rotor operator $R = \exp(-B\theta/2)$ and its ‘reverse’ $\bar{R} = \exp(+B\theta/2)$, so that rotation of an arbitrary vector v anticlockwise by θ radians is achieved by the two-sided operation $Rv\bar{R}$. The rotor R is an example of a ‘multivector’ object, as it contains scalar and bivector components.

2.2 Geometric algebra and rotations in three dimensions

The general utility of this approach becomes clearer once we go from two to three dimensions. In this case we develop a geometric algebra of eight entities; from the scalar 1 and unit vectors e_1, e_2, e_3 we develop three bivectors $B_3 = e_1e_2, B_1 = e_2e_3, B_2 = e_3e_1$, and a directed volume element, the pseudoscalar $J = e_1e_2e_3$. The pseudoscalar can convert between a bivector and its Hodge dual (perpendicular vector), e.g. $JB_3 = e_1e_2e_3 \times e_1e_2 = -e_3$ etc. A bivector rotates vectors in its own plane and generates the pseudoscalar with a perpendicular vector. Our 2D rotor R containing the bivector e_1e_2 , if applied to a general vector, will act to rotate the in-plane (e_1 and e_2) components of the vector, but will leave the e_3 component quite unaffected – all components of $e_1e_2e_3$ cancel when $Rv\bar{R}$ is evaluated.

The even-grade objects (scalar and bivectors) in 3D have an algebra very similar to Hamilton’s quaternions: $B_i^2 = -1, B_iB_j = -B_k$. A general plane can be represented by a unit bivector \hat{B} with components $c_1B_1 + c_2B_2 + c_3B_3$; by the Hodge dual this is the plane normal to the vector $c_1e_1 + c_2e_2 + c_3e_3$. Thus a general rotation can be carried out by a rotor $R = \exp(-\hat{B}\theta/2)$.

Within the geometric analysis software GASP, a rotation is parameterised by the three components of a bivector $B_\theta = 2\hat{B} \sin(\theta/2)$; effectively a ‘vector’ b_θ whose direction is the rotation axis and whose magnitude is very close to the rotation magnitude in radians.

2.3 Rotations of polyhedra and clusters

The fitting operation carried out by GASP for each cluster in a framework – e.g. an SiO_4 polyhedron – involves

minimising the mismatch between the positions of a set of vertices in a template, and the matching set of vertices based on the current positions of the atoms making up the cluster. It is trivial to make the centres of the template and cluster coincident, so we only need to find a suitable rotation. For each vertex q , we construct a vector \overline{pq} from the (Cartesian) position of the centre to the position of the vertex in the template. It is convenient to construct \overline{pq} as a function of the rotation ‘vector’ b . We can then consider the vector mismatch ε_q between $\overline{pq}(b)$ and the matching vector \overline{pq}' based on the current positions of the atoms in the cluster: $\varepsilon_q = \overline{pq}(b) - \overline{pq}'$. Thus for example the x component of ε_q is calculated as

$$\begin{aligned} \varepsilon_x = & \overline{pq}_x \left(1 - \frac{1}{2} (b_y^2 + b_z^2) \right) + \overline{pq}_y \left(-Cb_z + \frac{1}{2} b_x b_y \right) \\ & + \overline{pq}_z \left(Cb_y + \frac{1}{2} b_x b_z \right) - \overline{pq}'_x, \end{aligned} \quad (1)$$

where C is the scalar quantity

$$C = \sqrt{\left(1.0 - \frac{1}{4} (b_x^2 + b_y^2 + b_z^2) \right)} \quad (2)$$

equal to $\cos(\theta/2)$. Expressions for $\varepsilon_{y,z}$ can be obtained by cyclic permutation of components of b , and are fully tabulated in reference [3].

A rotation is found by minimising a penalty function $E = \sum_q \varepsilon_q^2$ as a function of the rotation parameters b for this cluster. The gradient of E with respect to the components of b is easily constructed analytically, and the full expression is tabulated in reference [3]. GASP uses the secant method to minimise E by finding the zero of the gradient, thus identifying the rotation parameters b which best fit the template to the cluster. The equations for ε_q can now be used, setting $\overline{pq}' = 0$, to generate a rotated version of the template, superimposed on the cluster. This process of fitting is illustrated in the method schematic, Figure 1(b,c).

The residual differences between template and cluster positions can be used to quantify the distortion of the cluster from the ideal geometry defined by the template (see Figure 1(c)). It is a useful feature that this distortion is a distance measure (that is, distortion can be described by the mean-squared or RMS vertex–atom mismatch) which is easier to grasp than a collection of bond length and angle measurements. Indeed the first application of GASP was as a form of geometric analysis, to assess large structural models of polyhedral framework materials produced by RMC modelling based on total scattering data (see Section 3.1).

During the initial development of geometric simulation, applied to regular polyhedral structural units, it was

typical for the centre–vertex vectors \overline{pq} to correspond to interatomic bonds such as Si–O, with the centre position of the cluster coinciding with the position of the central Si atom. However, there is no requirement in the mathematics of the fitting process for this to be the case. The method can equally well be applied to a group of atoms representing a molecular fragment, for example a peptide unit in a protein backbone. In this case, the centre of the cluster is simply the geometric centre of the positions of member atoms, and need not coincide with the position of any individual atom. The centre–vertex vectors \overline{pq} then do not lie along interatomic bonds, but simply describe the positions of all the atoms in the cluster.

A further process – geometric relaxation – can be carried out by seeking to minimise these residual differences as a function of the positions of both atoms and templates. If this is to be done, we must consider how to relax the position of an individual atom, which may belong to multiple clusters and thus have multiple ‘ideal’ positions, and which may furthermore be in steric contact with other atoms.

2.4 Relaxation of atomic positions

We now consider an atom A which belongs to n clusters and which is in steric overlap with m other atoms (clearly if n, m are both zero then nothing need be done). For each cluster i there is a vertex position which the atom should move towards, while each contact j is an object which the atom should move away from. We note that the most significant radius for zeolite frameworks is that of the tetrahedral oxygen atoms, for which we use a standard radius of 1.35 Å [11]; steric radii for all elements present in the input structure can be defined by the user in the program control input file. The neighbour search for contacts is short range and highly localised, and can therefore be undertaken using a coarse-gridding approach which scales linearly with system size. Atoms are assigned to cells in a coarse grid with spacing of order 4Å. Contacts need only be searched for in the same cell and its immediately adjacent neighbouring cells. Our goal is to generate a displacement vector for the atom so that we can update its position. Previous versions of GASP constructed this displacement by averaging a set of vectors representing (i) movement of the atom to a vertex position and (ii) movement directly away from a steric contact until just out of contact distance. The most recent version of GASP, however, uses an improved approach which has not previously been reported in detail, and which we will now derive.

We place the atom A at the origin of a temporary Cartesian coordinate system, so that it currently has a position $r_0 = (0, 0, 0)$. During the fitting it will develop a position $r = (x, y, z)$ which is the displacement vector we require. Each template vertex to which the atom belongs

has a position r_i , and each contact atom j has a position r_j . The r_i, r_j do not change during this part of the process, that is, r_i, r_j are parameters while r is variable. The vector from a vertex to the atom A is thus $\Delta r_i = r - r_i$, and the vector from a contact atom to the atom A is $\Delta r_j = r - r_j$. The ‘penalty’ for the offset between the atom and a vertex position is $|\Delta r_i|^2 = (r - r_i)^2$; a harmonic term going to zero when the atom A is coincident with the vertex i .

For steric contacts the situation is a little more complicated, as we must consider the current distance between the atoms ($|\Delta r_j|$) and their ideal contact distance d_j , which is the sum of their steric radii. The distance by which the atoms overlap is $d_j - |\Delta r_j|$ and an appropriate harmonic penalty is this distance squared, which is $d_j^2 - 2d_j|\Delta r_j| + |\Delta r_j|^2$.

The net penalty for our atom A is thus

$$P = \sum_i |\Delta r_i|^2 + \sum_j \left(d_j^2 - 2d_j|\Delta r_j| + |\Delta r_j|^2 \right) \quad (3)$$

and we seek a position r for the atom that will minimise P by making $\nabla P = 0$. Considering for example the x component of this gradient, and noting that $\partial_x |\Delta r_k|^2 = 2(x - x_k)$ while $\partial_x |\Delta r_k| = (x - x_k)/|\Delta r_k|$, we find that $\partial_x P = 0$ when

$$\sum_i (x - x_i) + \sum_j (x - x_j) \left(1 - \frac{d_j}{|\Delta r_j|} \right) = 0. \quad (4)$$

We can now seek a suitable value of x (and y, z) by iterating towards a self-consistent solution, using an ‘old’ value of x and $|\Delta r_j|$ to generate a ‘new’ value of x . Beginning with our initial position at the origin, and the corresponding values of $|\Delta r_j|$, we use this rearrangement of the preceding equation:

$$x_{\text{NEW}} = \frac{\sum_i x_i + \sum_j x_j (1 - (d_j/|\Delta r_{j,\text{OLD}}|))}{\sum_i 1 + \sum_j (1 - (d_j/|\Delta r_{j,\text{OLD}}|))} \quad (5)$$

and iterate until x_{NEW} and x_{OLD} are effectively identical. Of course y, z must likewise be updated at the same time, as $|\Delta r_j|$ depends on x, y, z .

A few noteworthy features of this approach are as follows:

- (i) If an atom has no steric contacts, it is moved in a single step to the mean position of all the template vertices to which it belongs, as desired;
- (ii) As constructed the steric exclusion constraint is holonomic, so the function also penalises the separation of atom A from a contact j by more than the ideal contact distance d_j . This does not offer a difficulty in practice, since if the atoms do move out of contact, they will not be found in contact at the start of the next cycle of relaxation,

- and so no further penalty will apply;
- (iii) The derivation above gives equal weight to the vertex (i) terms and the steric overlap (j) terms. In practice the terms can be weighted by a user-defined damping parameter multiplying the sums over j ; this term within GASP is set to 0.5 by default and can be adjusted;
 - (iv) When atoms are forced into a very severe steric overlap (where the overlap distance is a substantial fraction of the sum of their radii), the fitting routine may attempt to move them into exact superposition (!) as the gradient of their overlap penalty is zero here. This is recognisable as GASP reports an overlap equal to the ideal contact distance;
 - (v) Since the space of rotations is non-Euclidean and cyclic, all coordinate or parameter systems describing rotations will display a coordinate singularity or degeneracy at some point. Using the rotor parameterisation described here, this problem arises as the magnitude of a rotation approaches 180° . In consequence, the rotor fitting process can become numerically unstable as the magnitude of the rotation increases. Within the GASP code, large rotations are handled by a stabilising procedure; when b becomes large, the initial vectors pq are redefined using the current value of the rotation and fitting then continues from this new origin. A ‘running total’ is also kept when this occurs so that the correct final total value of b can be reported.

2.5 Iterative geometric relaxation

The full geometric relaxation of a structure proceeds iteratively. Each fitting cycle has two stages; in the first stage all templates are fitted over clusters, and in the second stage all atoms are relaxed based on template vertex positions and steric contacts (see Figure 1(c,d)). In our experience, the fitting process is generally stable once begun and the least squares fits will eventually converge to a minimum. Typically in each cycle, the magnitude of the displacements of the atoms decreases. The process is complete, typically after some tens or hundreds of cycles, when either (i) the greatest displacement of any atom in the previous step is below a user-specified very small threshold value, e.g. 10^{-7} Å, indicating that the process is effectively jammed, or (ii) the worst mismatch between any atom and template vertex has dropped below a user-specified threshold, typically 10^{-3} Å,[12] indicating that the match to template geometry is very good. If the framework has substantial internal flexibility (redundant degrees of freedom), the resulting configuration may not be uniquely determined by the cell parameters. Rather, a range of configurations are

compatible with the bonding constraints, and GASP reports the first such configuration it encounters. Further exploration, if desired, can be carried out by performing a small random perturbation of the atomic positions and then a further geometric relaxation.

A typical application of GASP is to study the behaviour of a given framework as its cell parameters are varied. Earlier versions of GASP used the cell parameter provided for the input structure; as a result a new input file had to be prepared for each cell parameter to be investigated. Version 5 now includes a ‘new cell’ feature allowing a new set of cell parameters ($a, b, c, \alpha, \beta, \gamma$) to be specified as an option. Hence a single input structure can be used for a wide ranging investigation of cell parameters. This feature is particularly important for applications to MOFs (Section 5) where the geometry of molecular clusters is defined by the input structure.

3. Overview of applications of geometric simulation

In this section we briefly review applications of geometric analysis and simulation to the study of tetrahedral mineral frameworks (3.1), especially dynamic disorder in dense silicates, ‘strain screening’ by polyhedral rotations and the pressure behaviour and framework folding mechanisms of zeolites; to octahedral and mixed frameworks (3.2), particularly perovskites and zirconium tungstate; and to protein flexibility in structural biology (3.3). Further sections are devoted to the ‘flexibility window’ phenomenon in zeolites (4), and its extension taking account of extra framework content; and to geometric simulations for MOFs (5), a new area of application for the method.

3.1 Tetrahedral frameworks

The first application of geometric analysis using GASP was in the analysis of large structural models generated by RMC modelling based on total neutron scattering data [5,6,7]. These models are based on both sharp (Bragg) and diffuse scattering, and so they capture dynamic disorder as well as the average (crystal) structure. GASP was applied both to analyse individual snapshots, quantifying the range of distortions compared to ideal tetrahedral geometry and to compare independently generated snapshots with the same topology. This comparison decomposes the differences between the models into components of distortion and polyhedral rotation, and thus revealed a dominant component of rigid-unit motion [13,14] in the dynamic disorder of quartz.[4] The hexagonally symmetric β phase of quartz is revealed as a dynamic average; the local instantaneous structure is more similar to the trigonally symmetric structure of α quartz.

The flexibility which allows this large-amplitude rotational dynamic disorder also allows a framework to accommodate static distortions locally by collective

rotations of nearby polyhedra [15]. This leads to the phenomenon of ‘strain screening’, where geometric analysis shows polyhedral distortion decreasing rapidly with increasing distance from a defect site such as an Al/Si substitution [16]. Ionic and molecular motion through polyhedral frameworks is likewise strongly influenced by flexibility. In frameworks such as quartz, O–O distances defining a channel radius or pore aperture can vary by amounts on the order of an Ångström with little energetic penalty. This permits the framework to adapt locally to the presence of mobile ions such as Li^+ , reducing the activation energy for diffusion. Geometric simulations have been applied to elucidate this phenomenon in systems such as quartz, aiding in the interpretation of experimental data.[17–19]

A useful application of geometric simulation is to study the response of a framework structure to an applied strain, for example the compression induced experimentally in a diamond-anvil cell. In zeolites such as edingtonite (EDI framework) [20] and levyne (LEV framework) [21] geometric simulations have revealed subtle and counter-intuitive connections between strain in the unit cell and the local changes arising from collective rotational motion of the framework tetrahedra. The question of strain response in zeolites will be developed further in Section 4 in relation to the ‘flexibility window’.

3.2 Octahedral and mixed frameworks

Geometric simulation is of course not limited to tetrahedral frameworks. The octahedral clusters encountered in minerals such as the perovskites are likewise suitable for geometric analysis and simulation. For fully coordinated octahedral systems such as the perovskites SrSnO_3 [22] and SrTiO_3 , [23] rigid-unit motion is found to be generally more restricted than it is in the framework silicates – a natural consequence of the framework topology, with each octahedron being constrained by links to six neighbouring clusters rather than four. However, geometric analysis still provides useful information on octahedral tilting in the framework.

Simulations using geometric clusters as constraints have also been very informative in understanding manganite perovskite frameworks, in which the MnO_6 octahedra can be regular or Jahn–Teller (JT) distorted depending on the oxidation state of Mn. Modelling using GASP showed that ‘stripe’ patterns in LaMnO_3 , conventionally attributed to variations in charge or oxidation states, can also arise from ordered patterns of JT distortion.[24] Further RMC modelling using geometric cluster constraints led to a new understanding of the high-temperature phase of LaMnO_3 . This system shows an apparent discrepancy between the long-range average crystal structure, with regular octahedra, and local structural probes indicating persistence of JT distortion.

The geometric modelling accounts for this using a quadrupolar order parameter which does not entirely vanish in the high-temperature phase.[25]

A particularly interesting system is the mixed octahedral/tetrahedral framework on zirconium tungstate, which displays isotropic negative thermal expansion (NTE) in a cubic structure. Geometric analysis of structural models confirms the significance of collective polyhedral rotations in generating NTE.[26] Subsequent materials were developed to maximise the scope for such motion and display NTE on a colossal scale.[27]

3.3 Protein flexibility in structural biology

Geometric simulation has also been applied in biophysics as a method for simulating flexible motion in proteins.[28] The ‘FRODA’ approach (framework rigidity optimised dynamic algorithm) makes use of templates to represent the geometry of portions of a protein structure. Rigid groups are identified using the ‘FIRST’ rigidity analysis software,[29] within which FRODA is implemented, and vary in size from single methyl groups to clusters spanning multiple secondary structure units or entire protein domains, depending on the distribution of covalent and non-covalent constraints in the protein. Geometric simulation can then be used to explore the flexible motion intrinsic to the structure. Template-based geometric simulation derived from FRODA is also used in the FRODAN [30] method for the generation of pathways of conformational transition, and in NMsim,[31] a method combining elements of geometric simulation, elastic network modelling and empirical-potential molecular mechanics.

A particularly effective application for FRODA is the rapid simulation of large amplitudes of flexible motion, by exploring low-frequency normal modes identified by coarse-grained elastic network modelling.[32] This allows FRODA to rapidly traverse the landscape of conformations that are compatible with the bonding constraints of the input structure. In several recent papers this approach has been used to investigate target flexibility in protein folding simulations [33]; side chain crosslinking in calmodulin caused by the anticancer drug cisplatin [34]; large-scale motion in a multidomain protein, ERp27 [35] (reconciling an apparent disagreement between crystal and solution structures); cooperative and independent functional motions in dimeric citrate synthase structures from across the tree of life [36]; and flexibility in native and mutant structures of calexcitin.[37]

The scale of the motions that are easily accessible using geometric simulation is illustrated in Figure 2. Panel a shows an ‘open’ crystal structure of the dimeric enzyme citrate synthase, in which two binding clefts are both in an open state. Panel b shows this structure during geometric simulation of motion along a symmetric closing mode,

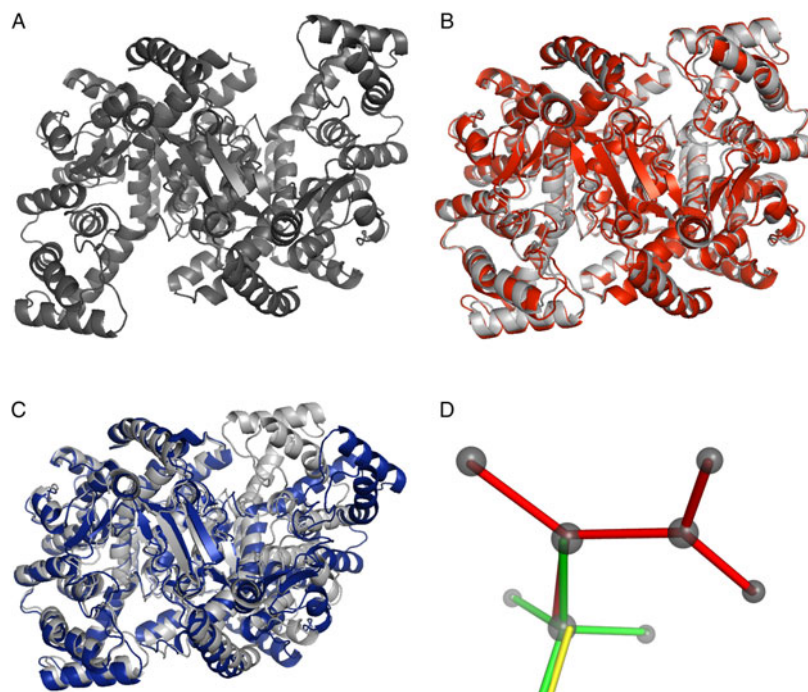


Figure 2. (Colour online) Illustrative example of geometric simulations for modelling of protein flexibility. Panel (a) crystal structure of dimeric citrate synthase with binding clefts open, shown as backbone cartoon. Binding sites lie in the clefts between the main body of the dimer and the small domains at left and right; (b) overlay of crystal structure of citrate synthase with binding clefts closed (light grey) and a simulation of symmetric closing motion; (c) overlay of crystal structure of citrate synthase with binding clefts closed (light grey) and a simulation of antisymmetric motion; (d) methods schematic showing the overlap of molecular cluster templates along rotatable dihedral angles.

overlaid on a known crystal structure with the two clefts in a closed state; and panel c shows the structure during geometric simulation of motion along an antisymmetric mode, which closes one cleft and opens the other. The latter motion is not directly evidenced by crystallography but is revealed by simulation as a natural motion of the protein.[36]

A feature of the FRODA implementation of geometric simulation is that the ‘ghost’ templates, defining rigid clusters with the geometry of molecular fragments, generally overlap along bonds with variable dihedral angles; this is in contrast to the original implementation in GASP for periodic mineral frameworks, where polyhedral clusters meet at a shared vertex but do not overlap. Thus when dihedral angles are the variables, adjacent clusters typically have two atoms in common rather than one, as illustrated in Figure 2(d). We may consider a four-atom ‘molecule’ $A-B-C-D$, in which the bond lengths AB , BC , CD and the angles ABC , BCD are fixed but the dihedral $ABCD$ is variable, permitting rotation around the $B-C$ bond. In geometric simulation we will identify two rigid clusters, $A'-B'-C'$ and $B''-C''-D''$. Constraints link the atom B to the vertices B' and B'' , and atom C to vertices C' and C'' . As a result the cluster bonds $B'-C'$ and $B''-C''$ will be kept parallel, but rotation around the $B-C$ bond is permitted. Another difference from GASP is that

the bonding geometry in FRODA is defined by the input structure rather than being generated mathematically as regular polyhedra.

The latest version of GASP has a more general logic for the identification and handling of clusters, and so is capable of modelling frameworks containing both polyhedral and molecular (overlapping) rigid clusters. It is this extension, making use of concepts developed in FRODA, that makes GASP v.5 suitable for the study of MOFs.

4. The flexibility window in zeolites

Geometric simulation has been applied to investigate compression mechanisms in zeolites (Section 3.1), tetrahedral distortion in real and hypothetical zeolites and AlPOs,[38] and to the detection of unfeasible hypothetical zeolite frameworks [39] generated through symmetry-constrained inter site bonding search.[40,41] This focus on zeolites led to the following question: in a zeolite framework, is it possible in principle for the tetrahedral geometry to be made ideal and perfect? Or are the small distortions typically seen in crystal structures in fact inevitable, given the framework topology and geometry?

Ideal tetrahedral geometry signifies that all $Si-O$ distances are equal to the ideal bond length, e.g. 1.61 Å,

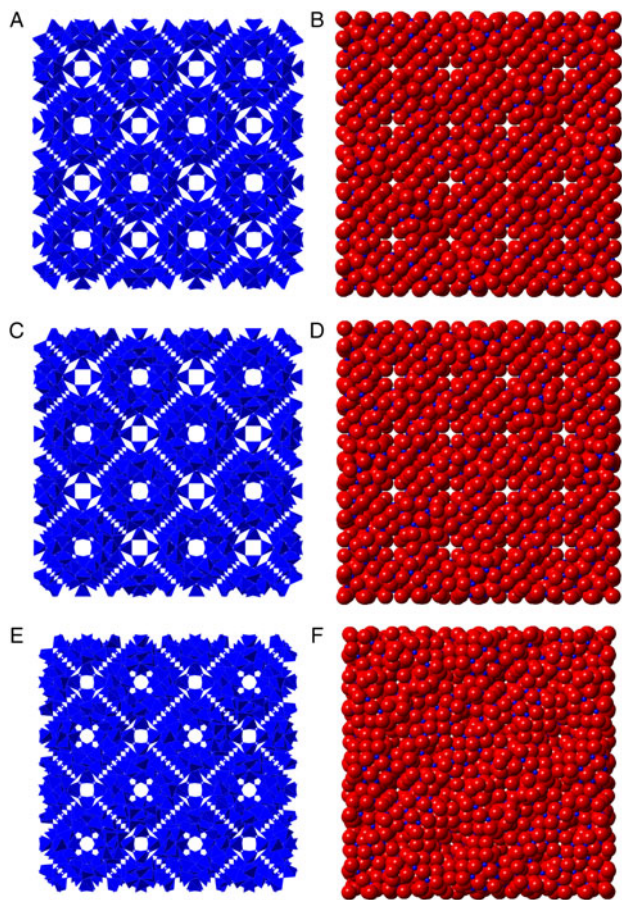


Figure 3. Application of geometric simulation to zeolite framework LTN. Panels show a single unit cell of the structure in tetrahedral (left) and space-filling (right) views. (a,b): structure at maximum geometric expansion; (c,d): structure with cell parameter found experimentally at ambient conditions; (e,f): structure at maximum geometric compression. Note the similarity of the ambient to the maximally expanded case.

and that all O—Si—O angles are equal to the tetrahedral angle $\arccos(-1/3)$, to within some defined precision. In a geometric simulation, this is equivalent to a requirement that the mismatch between any atom and its corresponding vertex on a tetrahedral template be less than a defined threshold, which we set at 0.001\AA . The bridging Si—O—Si angles are left unconstrained.

The results of this investigation were unexpected and striking.[12,42,43] It emerges that, first, actually existing real (natural or synthetic) zeolites can indeed attain ideal geometry when modelled as silica in geometric simulations. Second, this can be done over a range of densities; limits arise on expansion from extension of Si—O bonds, and in compression from steric collisions of oxygen atoms on adjacent tetrahedra ('codimeric' oxygens). This range we dub the 'flexibility window'. Third, zeolites under ambient conditions are found to lie towards the *low*-

density edge of the window. From this point of view, zeolites can be viewed as a form of 'expanded condensed matter' – the frameworks seek to be expanded as far as the framework topology will allow.

Figure 3 illustrates the use of geometric simulation to explore the flexibility window of the LTN zeolite framework. This structure is notable for its extremely large cubic unit cell ($a = 35.62\text{\AA}$) [11] containing 768 tetrahedra; however, GASP rapidly relaxes the structure, with 2304 independently mobile atoms. The investigation to determine the range of the flexibility window as illustrated in Figure 3 took only a few minutes on a single processor. The initial input is an all-atom representation of the structure in XTL format, essentially a unit cell and list of coordinates. Such all-atom inputs are easily generated from the fully symmetric structure in CIF format by export of coordinates in a structure viewer such as CrystalMaker [44]. An initial geometric relaxation using the cell parameter at ambient conditions rapidly idealises the tetrahedral geometry and confirms that the structure lies within its flexibility window. The relaxed structure under these conditions is shown in Figure 3(c,d). Progressively increasing the cubic a parameter, we soon find an upper limit to expansion, beyond which distortions are inevitable as bonds begin to be over-extended. For LTN this occurs when the unit cell volume has increased by 2.5%. This condition is illustrated in Figure 3(a,b); the similarity of the ambient and maximally expanded cases is visually evident. Decreasing the a parameter, we find a much greater scope for compression, with the framework folding freely until arrested by collisions of codimeric oxygen atoms. Here this occurs when the unit cell volume has decreased by 14% and is illustrated in Figure 3(e,f).

The majority of hypothetical zeolites, by contrast, are found to lack the 'flexibility window' property [45] even if the framework is assigned a low energy by interatomic potential calculations. This suggests that the flexibility window is a key criterion in the selection of hypothetical frameworks as candidates for synthesis [42]; structures lacking this property cannot be formed without causing strain in the building units, and so will be disfavoured relative to structures with a flexibility window. In recent years, GASP has been applied to assess new frameworks submitted to the Zeolite Structure Commission (Prof. M.M.J. Treacy, personal communication by email, 2015).

The flexibility window concept also provides valuable insights into interpreting experimental data on zeolites under compression. In zeolites of the analcime group, displaying phase transitions from high to low symmetry forms when subjected to pressures of 1–2 GPa, the transition appears to occur when the structure is compressed to the limit of the flexibility window. [46–48] An intriguing finding is that structures displaying pressure-induced amorphisation at low pressures do so while the framework lies within its flexibility window.

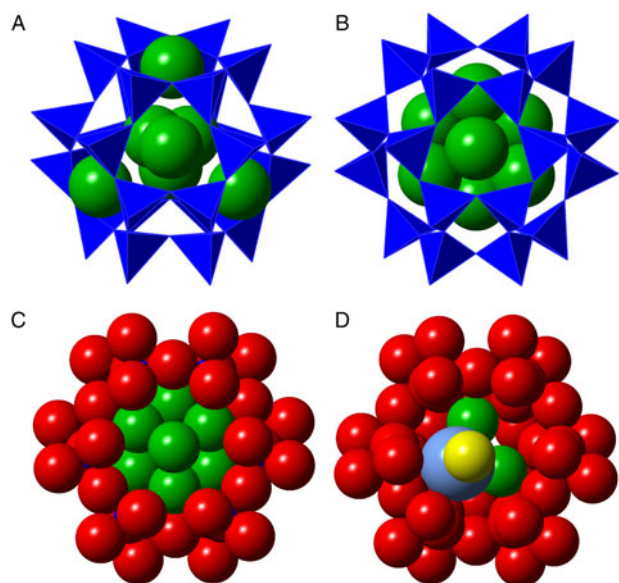


Figure 4. (Colour online) Geometric simulations of framework flexibility with extra framework contents. Panel (a) framework tetrahedra and extra framework ‘water’ spheres in a beta cage of siliceous faujasite from a refined crystal structure [51]; (b) geometrically relaxed framework and contents, showing substantial rearrangements to resolve steric clashes; (c) geometrically relaxed framework and contents, with framework oxygens in space-filling view. One of the framework six-ring apertures has been removed from the view to allow the contents to be seen; (d) cage occupation by a combination of water and methanol molecules. Figures are from the authors’ work in reference [50].

If the compression conditions move the structure out of its flexibility window, e.g. through pore occupation by penetrating pressure media,[49] the structure then resists amorphisation. From this point of view, the amorphisation of the framework is a process *permitted* by framework flexibility.

4.1 Extrinsic and intrinsic flexibility window

In a recent publication [50] we have further developed the concept of the flexibility window by considering the effect of explicitly present extra framework content. The flexibility window as originally defined may be termed ‘intrinsic’, being a property of the framework alone. The ‘extrinsic’ window, in contrast, is affected by interactions between the framework and the extra framework content. This study was conducted based on a structural refinement of siliceous faujasite by Colligan et al. [51], in which it is noticeable that extra framework content nominally refined as ‘water oxygens’ lies in positions that are not in fact sterically possible, due to clashes with the framework and/or other extra framework sites (see Figure 4(a,b)).

The compression experiments were carried out using a methanol/ethanol/water pressure medium; we, therefore, considered cage occupation by various combinations of water and methanol molecules. For this study, water molecules were represented as spheres while methanol was represented as two spheres representing the methyl and hydroxy groups. Framework and content geometries generated during these simulations are shown in Figure 4 (b–d). Geometric relaxation of the structure with water occupying all refined extra-framework sites, with both framework and content atoms mobile, shifts the extra-framework water from the refined positions to a different, approximately close-packed, arrangement, as shown in Figure 4(b,c). The experimental extra framework electron density may be better accounted for if the cages are occupied by both water and methanol molecules in a disordered arrangement, as shown in Figure 4(d).

The limits of the flexibility window in compression behave more or less as expected. At zero or low loadings, we observe the framework-controlled limit of the intrinsic window. At higher loadings, the extrinsic window takes over, as framework–content interactions begin to limit the degree of compression that the framework can achieve. An unexpected feature, however, is the limitation of the extrinsic window on *expansion*. When the beta cages are heavily loaded (containing both methanol and water molecules), the maximum expansion the framework can achieve is less than the limit of the intrinsic window! The cages require a degree of internal flexibility to accommodate their bulky contents, and this internal flexibility is lacking at the expanded edge of the intrinsic window.

5. Geometric simulation and flexibility in MOFs

There is considerable current research interest in questions of rigidity and flexibility in MOFs.[52] If materials of this type are to attain their potential for industrial applications, design principles for control of their stability, rigidity and flexibility are a prerequisite. For this reason, the capacity to handle MOFs, ZIFs (zeolitic imidazolate frameworks) and similar frameworks with non-polyhedral components is now included in version 5 of GASP. We will briefly review the features needed to make MOFs tractable in GASP (Section 5.1) and an illustrative case study (Section 5.2).

5.1 Geometry and rigidity of non-polyhedral framework units

The geometry of polyhedral units is communicated to GASP by specifying a polyhedral shape, e.g. ‘tet’ for tetrahedron, the identity of the centre and vertex species, e.g. ‘Si O’, and an ideal bond length, e.g. ‘1.61’ (Å). Non-polyhedral molecular units are now specified by a series of

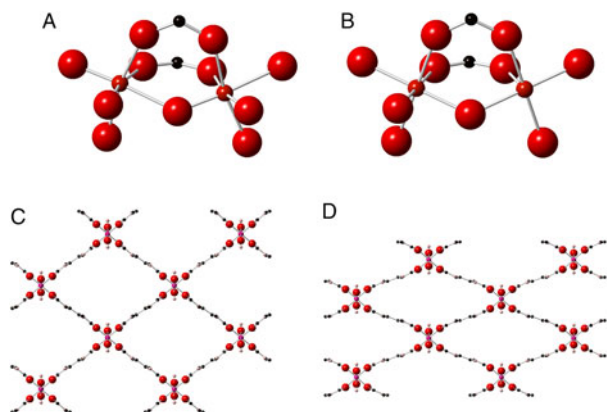


Figure 5. (Colour online) Investigation of rigidity and flexibility in MOFs. Panel (a) adjacent VO_6 octahedra in MIL-47 linked by a bridging oxygen vertex and two carbonyl groups, showing the apparent distortion from octahedral geometry in the crystal structure; (b) the same octahedra after geometric relaxation of the framework, showing almost regular geometry; (c) view down the channels of MIL-53; (d) simulation of ‘wine rack’ motion in MIL-53.

lines, each identifying a ‘central’ atom species, e.g. ‘C’, followed by all species to which it may bond. Terminal (singly bonded) species need not be specified explicitly. Thus for example the bonding in a typical small organic molecule might be specified by the lines ‘C C H N O’, ‘N C H’ and ‘O C H’. With this information, GASP will identify a series of overlapping bonded groups each consisting of a ‘central’ atom and its bonded neighbours. The geometry of each such cluster, as found in the input structure, will then be maintained in subsequent geometric simulations. An atom can simultaneously be a vertex species in a polyhedral group and part of a molecular group – for example the metal-coordinating carboxylate oxygens found in many MOFs.

If no further processing took place, all dihedral angles in molecular framework components would be variable. However, in general, the linkers in MOFs show considerable rigidity due to delocalised bonding extending over adjacent sp^2 hybridised atoms. GASP, therefore, carries out a further phase of cluster unification if desired. sp^2 hybridised carbon and nitrogen atoms are recognised by their trigonal bonding, and adjacent cases are unified into a single cluster, thus rigidifying the molecule appropriately. Cluster rigidity can be fully controlled by the user on a bond-by-bond basis when required: the bonding topology can be exported from GASP in a simple text format and edited to label specific bonds as rotatable or locked. Non- sp^2 rigidity arising from over constraint, as for example in adamantanes, arises naturally from the overlap of rigid clusters in GASP.

Sensible steric radii must also be assigned to the atoms present in the structure. This is to some extent a question of

judgement, and detailed results – especially on the limits in compression of the flexibility window – will depend on the values chosen. The radii assigned should therefore be reported explicitly when the results are affected by steric contacts within the framework. In the study below (5.2), the results depend only on the bonding geometry; for completeness we note that the radii assigned were $C = 1.70$, $H = 1.0$, $\text{Cr} = 0.6$, $O = 1.35 \text{ \AA}$.

Flexibility in the structure can then be explored by specifying new cell parameters for the system; GASP will impose these new parameters but maintain the bonding geometry of clusters specified in the original input. Geometric relaxation will determine how the framework responds to the strain, within the geometric simulation model, and whether the molecular clusters can maintain their shape or must be distorted. In the general case, a flexibility window is a 6D hypershape, as the cell parameters $a, b, c, \alpha, \beta, \gamma$ can all be varied independently. An exploration within a particular crystal system is generally more constrained; a cubic system has only a single relevant parameter, a hexagonal system has two.

With these new capabilities in GASP version 5, the concept of the ‘flexibility window’ can now be generalised from zeolites to MOFs, ZIFs and flexible frameworks in general.

5.2 Case study: rigidity and strain in MIL-47 and MIL-53

The MIL-47/MIL53 system provides an interesting first test case. These materials are metal dicarboxylate MOFs with very similar structures, distinguished by the oxidation state of the metal ion and the consequent presence or absence of a bridging hydroxy group in the framework. In this case, two frameworks with essentially identical topology display strikingly different flexibility behaviour: rigidity in one case,[53] flexible compliance in the other, with the structure folding by a ‘wine-rack’ mechanism. [54] We have therefore examined a vanadium MIL-47 and a chromium MIL-53 framework using GASP version 5. The goal of this pilot study is, first, to see what insights geometric analysis and simulation can produce into the rigidity of the framework, and second, to determine whether geometric simulation can produce the wine-rack mechanism as a stress-free motion of the framework.

Considering first the crystal structure of MIL-47,[55] we note that the coordination of the metal (vanadium) centre by oxygens is sixfold but shows substantial deviations from regular octahedral geometry. Indeed, the metal ion appears to lie off the plane formed by four coordinating carboxylate oxygens, being much closer to one of the bridging oxygens than to the other, as shown in Figure 5(a), due to the formation of the $[\text{VO}]^{2+}$ vanadyl unit. This immediately raises the question of whether regular octahedral geometry around the metal centre is in

fact possible. Investigating with GASP using a polyhedral specification of ‘oct v o 1.97’ – the bond length being chosen from the average of the bonds observed in the crystal structure – we find that at the cell parameters given, the distortion of octahedra cannot be reduced below 0.060 Å (for comparison, the input crystal structure displays a distortion of 0.174 Å). The geometry of the octahedra after relaxation is shown in Figure 5(b).

Consideration of the connections between any two adjacent metal centres in the *a*. direction makes clear the origin of the distortion. Adjacent metal ions are linked through one bridging oxygen and through two carboxylate groups, defining a very restrictive bonding geometry. An incompatibility between the bond lengths in the MO₆ octahedra, the O–C–O carboxylate group and the spacing between adjacent metal centres (set by the *a* cell parameter) leads to distortions. The distortions could be reduced further by adjustments of the *a* parameter and the V–O bond length; for example, with the given octahedral geometry, a slight expansion of the *a* parameter from 13.64 to 14.05 Å reduces the bonding distortion in the framework to a mere 0.0025 Å.

In the MIL-53 framework, the bridging oxygen between adjacent metal centres is converted to a bridging hydroxyl, reflecting the different oxidation states of the V and Cr ions in the frameworks. The Cr metal centre is found in square planar coordination by four carboxylate oxygens, with the hydroxy groups making up an ‘octahedron’; the Cr–OH bond differs from the Cr–O_{carboxylate} bond, at 1.90 versus 1.97 Å. This can be represented in GASP by specifying a square-planar Cr:O_{carboxylate,4} geometry and a Cr:OH ‘bar’ geometry. Thus the bond lengths are all constrained, but the OH–Cr–O_{carboxylate} angle is permitted to vary. Analysis of the input crystal structure using these settings reports a small distortion in the square planar units of 0.032 Å; geometric relaxation reduces this distortion to less than 0.001 Å, effectively zero. Thus if we take the view that the change from a bridging oxygen to a bridging hydroxyl group acts to reduce constraints on the bonding geometry around the metal centre, we find that the MIL-47 framework contains intrinsic stresses but the MIL-53 framework can be geometrically relaxed.

To investigate flexibility and the capacity for large-scale motion in MIL-53, we make use of the ‘new cell ...’ feature to vary the unit cell of the framework and observe the response of the framework. GASP seeks to maintain the geometry around the metal centres (square planar Cr–O + bars Cr–OH) and the molecular geometry of the dicarboxylate linkers. To investigate ‘wine rack’ motion, we consider variations in the *b* and *c* parameters. An initial test shows that neither parameter can be varied on its own without the immediate onset of unresolvable distortions in the framework due to incompatible bonding geometry. However, a variation in which *b* is increased

and *c* reduced in an appropriate ratio does generate a stress-free motion in which the framework folds as expected. This is illustrated in Figure 5(c), showing the crystal structure of MIL-53 with pores open, and in Figure 5(d), in which the *b*, *c* parameters have been substantially altered from their initial values (*b* = 16.73 and *c* = 13.04 Å) to *b* = 18.40, *c* = 10.00 Å. Here *c* was set arbitrarily and we searched for a value of *b* that allows the framework geometry to remain undistorted.

The results reported here are simply those of a pilot study demonstrating the applicability of GASP to the study of MOFs. Detailed results will depend on subtle interactions between the local bonding geometry around metal sites (polyhedral shapes, bond lengths), the steric geometry (atomic radii) assigned to the atoms and the global geometry of the framework (unit cell parameters). Elucidating these interactions will be highly productive of insights for the design and understanding of MOFs. We anticipate that the investigation of rigidity and flexibility of MOFs using GASP will be a rich seam for future study. Of particular interest will be the comparison of results from the real-space approach of GASP and reciprocal-space approaches, such as those recently applied to MOFs by Rimmer et al. [56]; and the comparison to other framework flexibility approaches such as the bar and hinge approach of Sarkisov et al. [57].

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

AS thanks the Royal Society for fellowship funding. SAW acknowledges funding from EPSRC project grant EP/K004956/1.

References

- [1] Wragg DS, Morris RS, Burton W. Pure silica zeolite-type frameworks: a structural analysis. *Chem Mater.* 2008;20(4): 1561–1570. doi:10.1021/cm071824j.
- [2] Baur WH. Straight Si–O–Si bridging bonds do exist in silicates and silicon dioxide polymorphs. *Acta Crystallogr B.* 1980;36(10): 2198–2202. doi:10.1107/S0567740880008382.
- [3] Wells SA, Dove MT, Tucker G. Finding best-fit polyhedral rotations with geometric algebra. *J Phys Condens Matter.* 2002;14(17): 4567–4584. doi:10.1088/0953-8984/14/17/327.
- [4] Wells SA, Dove MT, Tucker MG, Trachenko K. Real-space rigid-unit-mode analysis of dynamic disorder in quartz, cristobalite and amorphous silica. *J Phys Condens Matter.* 2002;14(18):4645–4657. doi:10.1088/0953-8984/14/18/302.
- [5] Dove MT, Tucker MG, Keen A. Neutron total scattering method: simultaneous determination of long-range and short-range order in disordered materials. *Eur J Mineral.* 2002;14(2):331–348. doi:10.1127/0935-1221/2002/0014-0331.
- [6] Wells SA, Dove MT, Tucker MG. Reverse Monte Carlo with geometric analysis – RMC + GA. *J Appl Crystallogr.* 2004;37(4): 536–544. doi:10.1107/S0021889804008957.

- [7] Keen DA, Tucker MG, Dove MT. Reverse Monte Carlo modelling of crystalline disorder. *J Phys Condens Matter*. 2005;17(5): S15–S22. doi:10.1088/0953-8984/17/5/002.
- [8] Wells SA, Sartbaeva A. Template-based geometric simulation of flexible frameworks. *Materials*. 2012;5(12):415–431. doi:10.3390/ma5030415.
- [9] Gale JD. GULP: a computer program for the symmetry-adapted simulation of solids. *J Chem Soc Faraday Trans*. 1995;93:629–637.
- [10] Gull S, Lasenby A, Doran C. Imaginary numbers are not real—The geometric algebra of spacetime. *Found Phys*. 1993;23(9):1175–1201. doi:10.1007/BF01883676.
- [11] Baerlocher C, Meier W, Olson D. Atlas of Zeolite framework types. Amsterdam: Elsevier; 2001.
- [12] Sartbaeva A, Wells SA, Treacy MMJ, Thorpe F. The flexibility window in zeolites. *Nat Mater*. 2006;5(12):962–965. doi:10.1038/nmat1784.
- [13] Dove MT, Heine V, Hammonds KD. Rigid unit modes in framework silicates. *Mineral Mag*. 1995;59(397):629–639. doi:10.1180/minmag.1995.059.397.07.
- [14] Dove MT, Pryde AKA, Heine V, Hammonds KD. Exotic distributions of rigid unit modes in the reciprocal spaces of framework aluminosilicates. *J Phys Condens Matter*. 2007;19(27): 275209. doi:10.1088/0953-8984/19/27/275209.
- [15] Goodwin AL. The crystallography of flexibility: local structure and dynamics in framework materials. *Z Kristallogr*. 2009;30: 1–11.
- [16] Goodwin AL, Wells SA, Dove MT. Cation substitution and strain screening in framework structures: the role of rigid unit modes. *Chem Geol*. 2006;225(3–4):213–221. doi:10.1016/j.chemgeo.2005.08.015.
- [17] Sartbaeva A, Redfern SAT, Lee T. A neutron diffraction and Rietveld analysis of cooperative Li motion in beta-eucryptite. *J Phys Condens Matter*. 2004;16(29):5267–5278. doi:10.1088/0953-8984/16/29/018.
- [18] Sartbaeva A, Wells SA, Redfern SAT. Li + ion motion in quartz and β -eucryptite studied by dielectric spectroscopy and atomistic simulations. *J Phys Condens Matter*. 2004;16(46):8173–8189. doi: 10.1088/0953-8984/16/46/005.
- [19] Sartbaeva A, Wells SA, Redfern SAT, Hinton RW, Reed JB. Ionic diffusion in quartz studied by transport measurements, SIMS and atomistic simulations. *J Phys Condens Matter*. 2005;17(7): 1099–1112. doi:10.1088/0953-8984/17/7/004.
- [20] Gatta GD, Wells SA. Rigid unit modes at high pressure: an explorative study of a fibrous zeolite-like framework with EDI topology. *Phys Chem Miner*. 2004;31(8):465–474. doi:10.1007/s00269-004-0413-z.
- [21] Gatta GD, Wells SA. Structural evolution of zeolite lewyne under hydrostatic and non-hydrostatic pressure: geometric modelling. *Phys Chem Miner*. 2006;33(4):243–255. doi:10.1007/s00269-006-0077-y.
- [22] Goodwin AL, Redfern SAT, Dove MT, Keen DA, Tucker MG. Ferroelectric nanoscale domains and the 905 k phase transition in SrSnO_3 : a neutron total-scattering study. *Physical Review B*. 2007; 76(17):174114. doi:10.1103/PhysRevB.76.174114.
- [23] Hui Q, Tucker MG, Dove MT, Wells SA, Keen DA. Total scattering and reverse Monte Carlo study of the 105 k displacive phase transition in strontium titanate. *J Phys Condens Matter*. 2005;17(5): S111–S124. doi:10.1088/0953-8984/17/5/012.
- [24] Sartbaeva A, Wells SA, Thorpe MF, Božin ES, Billinge SJL. Geometric simulation of Perovskite frameworks with Jahn-Teller distortions: applications to the cubic manganites. *Phys Rev Lett*. 2006;97(6):065501. doi:10.1103/PhysRevLett.97.065501.
- [25] Sartbaeva A, Wells SA, Thorpe MF, Božin ES, Billinge SJL. Quadrupolar ordering in LaMnO_3 revealed from scattering data and geometric modeling. *Phys Rev Lett*. 2007;99(15):155503. doi:10.1103/PhysRevLett.99.155503.
- [26] Tucker MG, Goodwin AL, Dove MT, Keen DA, Wells SA, Evans JSO. Negative thermal expansion in ZrW_2O_8 : mechanisms, rigid unit modes, and neutron total scattering. *Phys Rev Lett*. 2005;95 (25):255501. doi:10.1103/PhysRevLett.95.255501.
- [27] Conterio MJ, Goodwin AL, Tucker MG, Keen DA, Dove MT, Peters L, Evans SO. Local structure in $\text{Ag}_3[\text{Co}(\text{CN})_6]$: colossal thermal expansion, rigid unit modes and argentophilic interactions. *J Phys Condens Matter*. 2008;20(25):255225. doi:10.1088/0953-8984/20/25/255225.
- [28] Wells SA, Menor S, Hespeneide BM, Thorpe MF. Constrained geometric simulation of diffusive motion in proteins. *Phys Biol*. 2005;2(4):S127–S136. doi:10.1088/1478-3975/2/4/S07.
- [29] Jacobs DJ, Rader AJ, Kuhn LA, Thorpe F. Protein flexibility predictions using graph theory. *Proteins Struct Funct Bioinform*. 2001;44(2):150–165. doi:10.1002/prot.1081.
- [30] Farrell DW, Speranskiy K, Thorpe MF. Generating stereochemically acceptable protein pathways. *Proteins Struct Funct Bioinform*. 2010;78(14):2908–2921. doi:10.1002/prot.22810.
- [31] Ahmed A, Rippmann F, Barnickel G, Gohlke H. A normal mode-based geometric simulation approach for exploring biologically relevant conformational transitions in proteins. *J Chem Inform Modell*. 2011;51(7):1604–1622. doi:10.1021/ci100461k.
- [32] Jimenez-Roldan JE, Freedman RB, Römer RA, Wells SA. Rapid simulation of protein motion: merging flexibility, rigidity and normal mode analyses. *Phys Biol*. 2012;9(1):016008. doi:10.1088/1478-3975/9/1/016008.
- [33] Burkoff NS, Várnai C, Wells SA, Wild DL. Exploring the energy landscapes of protein folding simulations with Bayesian computation. *Biophys J*. 2012;102(4):878–886. doi:10.1016/j.bpj.2011.12.053.
- [34] Li H, Wells SA, Jimenez-Roldan JE, Römer RA, Zhao Y, Sadler PJ, O'Connor PB. Protein flexibility is key to cisplatin crosslinking in calmodulin. *Protein Sci*. 2012;21(9):1269–1279. doi:10.1002/pro.2111.
- [35] Amin NT, Wallis AK, Wells SA, Rowe ML, Williamson RA, Howard MJ, Freedman RB. High-resolution NMR studies of structure and dynamics of human ERp27 indicate extensive inter-domain flexibility. *Biochem J*. 2013;450(2):321–332. doi:10.1042/BJ20121635.
- [36] Wells SA, Crennell SJ, Danson MJ. Structures of mesophilic and extremophilic citrate synthases reveal rigidity and flexibility for function. *Proteins Struct Funct Bioinform*. 2014;82(10):2657–2670. doi:10.1002/prot.24630.
- [37] Erskine P, Fokas A, Muriithi C, Rehman H, Yates L, Bowyer A, Findlow S, Hagan R, Werner J, Miles A, Wells SA, Wood S, Cooper J. X-ray, spectroscopic and normal-mode dynamics of calxectin: structure-function studies of a neuronal calcium-signalling protein. *Acta Crystallogr D*. 2015;71(Pt 3):615–631. doi:10.1107/S1399004714026704.
- [38] Zwiijnenburg MA, Simperler A, Wells SA, Bell RG. Tetrahedral distortion and energetic packing penalty in 'zeolite' frameworks: linked phenomena? *J Phys Chem B*. 2005;109(31):14783–14785. doi:10.1021/jp0531309.
- [39] Wells SA, Foster MD, Treacy MMJ. A simple geometric structure optimizer for accelerated detection of infeasible zeolite graphs. *Microporous Mesoporous Mater*. 2006;93(1–3):151–157. doi:10.1016/j.micromeso.2006.02.010.
- [40] Treacy MMJ, Randall KH, Rao S, Perry JA, Chadi DJ. Enumeration of periodic tetrahedral frameworks. *Z Kristallogr*. 1997;212: 768–791.
- [41] Treacy MMJ, Rivin I, Balkovsky E, Randall KH, Foster MD. Enumeration of periodic tetrahedral frameworks. II. Polynodal graphs. *Microporous Mesoporous Mater*. 2004;74(1–3):121–132. doi:10.1016/j.micromeso.2004.06.013.
- [42] Rivin I. Geometric simulations: a lesson from virtual zeolites. *Nat Mater*. 2006;5(12):931–932. doi:10.1038/nmat1792.
- [43] Kapko V, Dawson C, Treacy MMJ, Thorpe MF. Flexibility of ideal zeolite frameworks. *Phys Chem Chem Phys*. 2010;12(30): 8531–8541. doi:10.1039/c003977b.
- [44] *CrystalMaker*; software available from CrystalMaker Software Limited at <http://www.crystallmaker.com/index.html>
- [45] Kapko V, Dawson C, Rivin I, Treacy MMJ. Density of mechanisms within the flexibility window of zeolites. *Phys Rev Lett*. 2011; 107(16):164304. doi:10.1103/PhysRevLett.107.164304.
- [46] Sartbaeva A, Gatta GD, Wells SA. Flexibility window controls pressure-induced phase transition in analcime. *Europhys Lett*. 2008; 83(2):26002. doi:10.1209/0295-5075/83/26002.
- [47] Gatta GD, Sartbaeva A, Wells SA. Compression behaviour and flexibility window of the analcime-like feldspathoids: experimental

- and theoretical findings. *Eur J Mineral*. 2009;21(3):571–580. doi:10.1127/0935-1221/2009/0021-1923.
- [48] Wells SA, Sartbaeva A, Gatta GD. Flexibility windows and phase transitions of ordered and disordered ANA framework zeolites. *Europhys Lett*. 2011;94(5):56001. doi:10.1209/0295-5075/94/56001.
- [49] Sartbaeva A, Haines J, Cambon O, Santoro M, Gorelli F, Levelut C, Garbarino G, Wells SA. Flexibility windows and compression of monoclinic and orthorhombic silicalites. *Phys Rev B*. 2012;85(6):064109. doi:10.1103/PhysRevB.85.064109.
- [50] Wells SA, Leung KM, Edwards PP, Sartbaeva A. Flexibility windows in faujasite with explicit water and methanol extra-framework content. *Dalton Trans*. 2015;44(13):5978–5984. doi:10.1039/C4DT03150D.
- [51] Colligan M, Forster PM, Cheetham AK, Lee Y, Vogt T, Hriljac JA. Synchrotron X-ray powder diffraction and computational investigation of purely siliceous zeolite Y under pressure. *J Am Chem Soc*. 2004;126(38):12015–12022. doi:10.1021/ja048685g.
- [52] Li W, Henke S, Cheetham AK. Research update: mechanical properties of metal-organic frameworks – influence of structure and chemical bonding. *APL Mater*. 2014;2(12):123902. doi:10.1063/1.4904966.
- [53] Kolokolov DI, Jobic H, Stepanov AG, Plazanet M, Zbiri M, Ollivier J, Guillermin V, Devic T, Serre C, Férey G. Comparison of the dynamics of MIL-53(Cr) and MIL-47(V) frameworks using neutron scattering and DFT methods. *Eur Phys J. Special Topics* 2010;189:263–271.
- [54] Chen L, Mowat JPS, Fairen-Jimenez D, Morrison CA, Thompson SP, Wright PA, Düren T. Elucidating the breathing of the metalorganic framework MIL-53(Sc) with ab initio molecular dynamics simulations and *in situ* X-ray powder diffraction experiments. *J Am Chem Soc*. 2013;135(42):15763–15773. doi:10.1021/ja403453g.
- [55] Barthelet K, Marrot J, Riou D, Férey G. A breathing hybrid organic-inorganic solid with very large pores and high magnetic characteristics. *Angew Chem*. 2002;114(2):291–294. doi:10.1002/1521-3757(20020118)114:2<291:AID-ANGE291>3.0.CO;2-I.
- [56] Rimmer LHN, Dove MT, Goodwin AL, Palmer DC. Acoustic phonons and negative thermal expansion in MOF-5. *Phys Chem Chem Phys*. 2014;16(39):21144–21152. doi:10.1039/C4CP01701C.
- [57] Sarkisov L, Martin RL, Haranczyk M, Smit B. On the flexibility of metal-organic frameworks. *J Am Chem Soc*. 2014;136(6):2228–2231. doi:10.1021/ja411673b.