Structure searching methods: general discussion

Matthew Addicoat, Claire Adjiman, Mihails Arhangelskis, Gregory Beran, Gerit Brandenburg, Doris Braun, Virginia Burger, Asbjoern Burow, Chris Collins, Andrew Cooper, et al.

▶ To cite this version:

Matthew Addicoat, Claire Adjiman, Mihails Arhangelskis, Gregory Beran, Gerit Brandenburg, et al.. Structure searching methods: general discussion. Faraday Discussions, 2018, 211, pp.133-180. 10.1039/C8FD90030B . hal-02404834

HAL Id: hal-02404834
https://hal.sorbonne-universite.fr/hal-02404834

Submitted on 11 Dec 2019

Faraday Discussions C8FD90030B

We have presented the graphical abstract image and text for your article below. This briefly summarises your work, and will be presented with your article online. It will not appear in the print edition.

DISCUSSIONS

1

Structure searching methods: general discussion

Matthew Addicoat, Claire Adjiman, Mihails Arhangelskis, Gregory Beran, Gerit Brandenburg, Doris Braun, Virginia Burger, Asbjoern Burow, Chris Collins, Andrew Cooper, Graeme Day, Volker Deringer, Matthew S Dyer, Alan Hare, Kim Jelfs, Julian Keupp, Stefanos Konstantinopoulos, Yi Li, Yanming Ma, Noa Marom, David McKay, Caroline Mellot-Draznieks, Sharmarke Mohamed, Marcus Neumann, Sten Nilsson Lill, Jonas Nyman, Artem R. Oganov, Sarah Price, Susan Reutzel-Edens, Michael Ruggiero, German Sastre, Rochus Schmid, Julia Schmidt, Christian Schön, Peter Spackman, Seiji Tsuzuki, Scott Woodley, Shiyue Yang and Qiang Zhu

---

Please check this proof carefully. **Our staff will not read it in detail after you have returned it.**

Proof corrections must be returned as a single set of corrections, approved by all co-authors. No further corrections can be made after you have submitted your proof corrections as we will publish your article online as soon as possible after they are received.

Please ensure that:

- The spelling and format of all author names and affiliations are checked carefully. Names will be indexed and cited as shown on the proof, so these must be correct.
- Any funding bodies have been acknowledged appropriately.
- All of the editor's queries are answered.
- Any necessary attachments, such as updated images or ESI files, are provided.

Translation errors between word-processor files and typesetting systems can occur so the whole proof needs to be read. Please pay particular attention to: tables; equations; numerical data; figures and graphics; and references.

Please send your corrections preferably as a copy of the proof PDF with electronic notes attached or alternatively as a list of corrections – do not change the text within the PDF file or send a revised manuscript. Corrections at this stage should be minor and not involve extensive changes.

Please return your **final** corrections, where possible within **48 hours** of receipt, by e-mail to: faraday@rsc.org. If you require more time, please notify us by email.

Researcher information

If any authors have ORCID or ResearcherID details that are not listed below, please provide these with your proof corrections.

Please check that the ORCID and ResearcherID details listed below have been assigned to the correct author. Authors should have their own unique ORCID iD and should not use another researcher's, as errors will delay publication.

Please also update your account on our online manuscript submission system to add your ORCID details, which will then be automatically included in all future submissions. See here for step-by-step instructions and more information on author identifiers.

First (given) name(s)	Last (family) name(s)	ResearcherID	ORCID
Matthew	Addicoat		
Claire	Adjiman		
Mihails	Arhangelskis		
Gregory	Beran		
Gerit	Brandenburg		
Doris	Braun		
Virginia	Burger		
Asbjoern	Burow		
Chris	Collins		
Andrew	Cooper		
Graeme	Day		
Volker	Deringer		
Matthew S	Dyer		
Alan	Hare		
Kim	Jelfs		
Julian	Keupp		
Stefanos	Konstantinopoulos		
Yi	Li		
Yanming	Ma		
Noa	Marom		
David	McKay		
Caroline	Mellot-Draznieks		
Sharmarke	Mohamed		
Marcus	Neumann		
Sten	Nilsson Lill		
Jonas	Nyman		
Artem R.	Oganov		

Sarah	Price		
Susan	Reutzel-Edens		
Michael	Ruggiero		
German	Sastre		
Rochus	Schmid		
Julia	Schmidt		
Christian	Schön		
Peter	Spackman		
Seiji	Tsuzuki		
Scott	Woodley		
Shiyue	Yang		
Qiang	Zhu		

Queries for the attention of the authors

Journal: **Faraday Discussions**

Paper: **C8FD90030B**

Title: **Structure searching methods: general discussion**

For your information: You can cite this article before you receive notification of the page numbers by using the following format: (authors), Faraday Discuss., (year), DOI: 10.1039/C8FD90030B.

Editor's queries are marked on your proof like this ❶, ❷, etc. and for your convenience line numbers are indicated like this 5, 10, 15,

Please ensure that all queries are answered when returning your proof corrections so that publication of your article is not delayed.

Query Reference	Query	Remarks
1	Please confirm that the spelling and format of your name is correct. Names will be indexed and cited as shown on the proof, so these must be correct. No late corrections can be made. FOR YOUR INFORMATION: You can cite this discussion before the page numbers are assigned with: (authors), Faraday Discuss., DOI: 10.1039/c8fd90030b.	

Faraday Discussions

Cite this: DOI: 10.1039/C8FD90030B

211

DISCUSSIONS

Structure searching methods: general discussion

Matthew Addicoat, Claire Adjiman, Mihails Arhangelskis, Gregory Beran, Gerit Brandenburg, Doris Braun, Virginia Burger, Asbjoern Burow, Chris Collins, Andrew Cooper, Graeme Day, Volker Deringer, Matthew S Dyer, Alan Hare, Kim Jelfs, Julian Keupp, Stefanos Konstantinopoulos, Yi Li, Yanming Ma, Noa Marom, David McKay, Caroline Mellot-Draznieks, Sharmarke Mohamed, Marcus Neumann, Sten Nilsson Lill, Jonas Nyman, Artem R. Oganov, Sarah Price, Susan Reutzel-Edens, Michael Ruggiero, German Sastre, Rochus Schmid, Julia Schmidt, Christian Schön, Peter Spackman, Seiji Tsuzuki, Scott Woodley, Shiyue Yang and Qiang Zhu

DOI: 10.1039/C8FD90030B

(1:[1]1) **Graeme Day** opened the discussion of the introductory lecture by Sarah L. Price: You have nicely outlined a vision for a crystal structure prediction code and what its capabilities should be. In this, you have not said much about what analysis methods this code should have for understanding and interrogating the structure sets that are generated. The crystal structure prediction community, at least in the molecular organic field, has mainly borrowed methods from the analysis of experimental crystal structures. These can be useful, but do we need more sophisticated approaches to get the most information from computed sets of structures? I am interested in hearing your thoughts on what analysis methods need to be part of a general crystal structure prediction code.

Sarah Price answered: This is a difficult question, as it is still a matter of debate how different two structures have to be to count as polymorphs when working with experimental structures, with many of the difficulties relating to polytypes or disorder, where CSP analysis needs further development. Most of the methods used to compare experimental structures are appropriate for CSP generated structures, assuming that duplicates which only differ in the cell choice have already been eliminated. We certainly need to be able to cluster structures into families so that members of the same family could be observed as disorder components of experimental structures or be deemed to be so readily able to transform into the most stable structures of the family somewhere along the nucleation and growth pathway that they could not be observed, or are closely related structures which may appear in the same sample or interconvert. At the moment this is mainly done by experienced human judgement, using tools such as the size of the coordination sphere that can be overlapped, powder pattern similarity, hydrogen bonding motifs, XPAC, *etc.* For complex flexible molecules, deciding how to classify the low energy structures into families is fraught with difficulties, but the families of CSP structures may help interpret the experimental data since the ability of the molecules to rearrange into a more stable structure can depend on the size and quality of the crystallites. The analysis tools we require will be system dependent and closely linked to the prevalence of practically important metastable polymorphs and the accuracy to which their structures and properties are required.

(2:[2]2) **Christian Schön** remarked: Regarding the issue of what we should aim for, I would suggest looking at an often neglected issue: the time scale – equilibration + observation + escape – on which the system equilibrates is observed and (for metastable compounds/structures) leaves the region in its respective configuration space (*e.g.* by transformation into a different structure). If we talk about a relevant crystal structure/polymorph, we must keep in mind the time scale on which we are interested in the structure. People often confuse the stability of a compound with thermodynamic stability – the fact that the potential energy of a structure is low does not mean that it exhibits high stability (*e.g.* there might be a simple transformation route with a low energetic barrier to the global minimum). On the other hand, for a given observational time scale, many structures are kinetically stable even if they are rather high in potential energy. An important step would thus be to classify the structures by their lifetimes (escape times). We already have the computational technology at hand to analyse the kinetic stability of crystalline structures, *e.g.* based on the threshold algorithm[1,2] or transition path sampling.[3,4] A second very important goal should be the development of theoretical methodologies and algorithms that allow us to optimally reach target modifications/structures with a high probability, thus aiding and guiding the experimental synthesis.[5] An important tool could be the application of finite-time thermodynamics concepts to chemical processes.[6,7] Of course, this needs to be done on top of standard crystal structure prediction, and typically requires an order of magnitude more computational effort for the analysis of the barrier structure of the configurational space of the chemical system, and several orders of magnitude more for the analysis and optimization of synthesis routes. Nevertheless, one would envision a multi-level code package that provides such a range of computational capabilities – *e.g.* the G42+ code already does contain such algorithms for both structure prediction and the analysis of barrier structures[8] – even though the predictive modelling of the synthesis route is now only at a stage comparable to that of crystal structure prediction twenty-five years ago.

1 J. C. Schön, *Ber. Bunsenges. Phys. Chem.*, 1996, **100**, 1388–1391.
2 J. C. Schön, H. Putz and M. Jansen, *J. Phys.: Condens. Matter*, 1996, **8**, 143–156.
3 C. Dellago, P. Bolhuis, F. S. Csaijka and D. Chandler, *J. Chem. Phys.*, 1998, **108**, 1964.
4 P. G. Bolhuis, C. Dellago and D. Chandler, *Faraday Discuss.*, 1998, **110**, 421.
5 J. C. Schön, *Adv. Chem. Phys.*, 2015, **157**, 125–134.
6 B. Andresen, *Angew. Chem., Int. Ed.*, 2011, **50**, 2690–2704
7 J. C. Schön, *Z. Anorg. Allg. Chem.*, 2009, **635**, 1794–1806.
8 J. C. Schön, *Proc. Appl. Ceram.*, 2015, **9**, 157–168.

Sarah Price answered: Thank you for clearly emphasising the point about lifetimes of metastable structures. I hope that these approaches will prove suitable for extension to pharmaceutical polymorphs.

(3:[3]3) **David McKay** remarked: On CSP-AIM, how does/should the role of the conversation between CSP people and experimental people fit in? Would this no longer be necessary by writing CSP software that can understand and interpret experimental data?

Sarah Price answered: A CSP_aim code should have encapsulated this vital conversation, so that the code can be used without distinction between computational and experimental scientists. A clear output of the software should be the structures of the possible polymorphs of the system, with the corresponding powder X-ray patterns, solid state NMR spectra, phonon (infrared, Raman, terahertz) spectra *etc.* to allow the polymorphs to be identified even when observed in a poorly crystalline mixed-phase sample. The code should also be capable of providing a suggestion of the experiment most likely to find these polymorphs (*i.e.* specifying pressure, temperature, solvent system or template for sublimation, seeds, *etc.*). There would also be estimates of the relative thermodynamic and kinetic stability, and other properties of interest in the application. We should aim for software that allows seamless comparisons of the experimental data and corresponding data for computer-generated target polymorphs. However, it is essential that there is always a clear distinction between the structures that have been realised in the laboratory, and those which have yet to be verified experimentally.

(30:[4]4) **Alan Hare** opened the discussion of the paper by Yanming Ma:[†] Because the quasi-planar B_{36} species is hexagonal, this can be formulated as $B_6\Sigma_{i=1}^{n}i$, where in this instance, $n = 3$; *i.e.* 6 times the sum of the first n natural numbers. Can it be shown that n is always a local energy minimum? Will $B_{3n(n+1)}$ always be the lowest quasi-planar energy state? Should we expect to see B_{60} too, for instance ($n = 4$)? The structure has a hole in the centre, $i = 0$ (presumably the triangular ring would be under too much strain); now, $B_6\Sigma_{i=0}^{1}i$ gives the visual impression of a nucleus. Does this suggest that if the B_{84} cluster were to be either planar or cage-like it would require multiple nuclei to coalesce? Perhaps this would explain the lower energy of the core-shell. When you first get a glass-like structure – at B_{84}, say – is that "game over" for the traditional crystal? Or is it still possible for the structure to revert to a periodic one at some higher number (such as $n = 5$, B_{90})?

Yanming Ma communicated in response: It cannot be said that n corresponds to the local energy minimum for a planar configuration. For technical reasons, experimental data are only available for the charged clusters. Up to now, all the anionic B_n clusters with $n < 38$ have been experimentally confirmed to be planar, and B_{40} is cage-like. The structures for larger neutral B_n clusters with $n > 40$ are still elusive, with planar, core-shell, bilayered and cage-like configurations

---

[†] The paper by Yanming Ma was not presented at the meeting. Questions were submitted to the author after the meeting.

competing with each other. Whether B_{60} exhibits a planar or core-shell structure should be further checked.

Hexagonal vacancies are crucial for the stability of B clusters. It is reasonable to expect the existence of multiple vacancies in the B_{84} cluster. However the vacancies should be separate from each other (as seen in the alpha-boron sheet[1]). It is generally accepted that B_n clusters with $n > 68$ have core-shell structures, which can be seen as embryos of bulk boron.[2]

1 H. Tang and S. Ismail-Beigi, *Phys. Rev. Lett.*, 2007, **99**, 115501.
2 J. Zhao, L. Wang, F. Li and Z. Chen, *J. Phys. Chem. A*, 2010, **114**, 9969.

(31:[5]5) **Christian Schön** asked: Concerning the "flat" structures shown in Figures 4a and 5c, are they perfectly two-dimensional? Is there another layer of atoms directly behind them, so that they are actually multi-2D-layered structures?

Yanming Ma communicated in reply: No, they are not perfectly two-dimensional, since the plane is curved, but they are single-layered structures without additional layers behind them.

(32:[6]6) **Christian Schön** commented: Considering the first equation, *i.e.* where the energy is written as the sum over "atomic" energies, where these individual energies take some aspects of the local neighborhood of the atoms into account; this procedure is reminiscent of work done in the phase diagram theory community on intermetallic phases, led by Sanchez and DeFontaine since the early 1980s,[1,2] the so-called cluster expansion method. This method takes many (locally optimised) lattice structures, computes their energy (*e.g.* on *ab initio* level), and then derives expressions for the energy as the sum over binary, ternary, *etc.* local clusters of atoms surrounding each atom in the structure. These parameters can then be used to very quickly compute very large arrangements of atoms in a lattice (originally) or in a more general solid structure. There might be some valuable insights for the approach discussed here.

1 J. M. Sanchez and D. DeFontaine, *Phys. Rev. B*, 1980, **21**, 216.
2 J. M. Sanchez, *Phys. Rev. B*, 2010, **81**, 224202.

Yanming Ma communicated in response: Agreed! As far as I know, the cluster expansion method is particularly useful for alloys or systems where the atomic environment is known to some extent. I think the ML method may be more versatile.

(35:[7]7) **Sten Nilsson Lill** asked: On page 8, you discuss an 84-atom boron cluster and mention that there are two candidates for the ground state. What are the spectroscopic fingerprints of these and how do the predicted spectroscopic data compare to the experimental data?

Yanming Ma communicated in reply: To the best of our knowledge, there is no experimental spectroscopic data. In theory, we can calculate the photoelectron spectrum, which will aid future experiments to determination the ground-state structure.

(36:[8]8) **Christian Schön** remarked: It is not quite clear to me whether the experimentally observed structures correspond to the global minima of the potential energy.

Yanming Ma communicated in response: Any local minima on the potential energy surface correspond to mechanically stable configurations, where the gradient vanishes and hence so do all the forces. The experimentally observed structures generally correspond to the global minimum according to the principle of minimum energy, which is essentially a restatement of the second law of thermodynamics.

(38:[9]9) **Virginia Burger** commented: In your paper, it says that "[ML potentials] show accuracy and transferability comparable to those of quantum-mechanical simulations…". I understand that they have similar accuracy, but it is not clear to me that they would have similar transferability, since I thought ML potentials tended to perform much more poorly when extrapololated than on data similar to their training set. I think this sentence would be more useful and clearer if it were supported by citations with examples for accuracy and transferability.

Yanming Ma communicated in reply: Thanks for the good reminder. In the current work, both the training and testing sets consist of random structures generated by CALYPSO and are expected to cover a wide range of the PES. This to some extent demonstrates the high transferability of the ML potential. However, we agree that the current expression is not very accurate, since the transferability of ML potentials generally depends on the quality of the training set.

(56:[10]10) **Volker Deringer** asked: You mention in the Conclusions the possible extension of your potential to the bulk phases of boron. Have you tested the current version, say, on crystalline α-B_{12}? How much information about the bulk can already be "learned" from clusters?

Yanming Ma communicated in response: Currently, we have not tested the potential for bulk boron, since the stress tensor is not included during the training process, but we would like to perform a test according to your comments and provide you with the results as soon as possible.

(806:[11]11) **Julia Schmidt** communicated: Regarding the sentence in your paper: "We expect the RMSEs of the energy and force can be further reduced by increasing the training set size and number of function values of the ACSF descriptor but at the expense of larger computational cost during the training process."; the fact that the RMSE will be reduced if you train the model with more training data, hence with more outliers, makes sense. You currently have a reasonable training : testing ratio of 78% : 22% (12053 : 3571), but if I follow correctly the suggestion above would decrease the test set proportion. Why would increasing the training set size only be beneficial and not lead to overfitting?

Yanming Ma communicated in reply: Since the training and testing sets consist of random structures generated by CALYPSO, they are independent of each other and should cover a wide range of the PES. Therefore, the RMSE of the testing set is a reasonable indicator for the quality of the potential, regardless of the training : testing ratio. We expect the training set not to fully cover the PES. Thus, increasing the number of function values of the ACSF descriptor will lead to a more complete description of the structures, while increasing the training size will include more outliers. This will improve the quality of the potential. However, both the parameters of the ACSF descriptor and the newly included structures should be chosen carefully to avoid overfitting.

(4:[12]12) **Christian Schön** opened the discussion of the paper by Volker L. Deringer: How do the machine learning potentials perform compared with the refined empirical potentials people have been developing and using over a long time – in particular with respect to accuracy, speed, robustness and generalisability?

Volker Deringer replied: These are two conceptually different approaches, so let me start by saying clearly that machine learning (ML) potentials do not aim to replace empirical ones. As a rule of thumb, ML potentials are significantly more computationally demanding than empirical potentials, but are also significantly more accurate—reaching “chemical accuracy” in the kJ mol^{-1} region in many cases. In turn, empirical potentials are normally very robust, because they have physically and chemically motivated forms (so they will correctly describe the repulsion between two atoms that come very close, something that ML models need to “pick up” from the input data). Indeed, robustness is one of the major challenges for making good ML potentials, and it is precisely what we are addressing in this work: the initial atomic configurations are largely randomised, and the fact that our potential can perform progressively better relaxations of these configurations attests to its robustness and flexibility. This is particularly important for systems (such as phosphorus and boron) for which no good general empirical potentials exist.

(5:[13]13) **Michael Ruggiero** asked: Given that your machine learning methods typically generate an interatomic potential, is it possible to determine vibrational dynamics? Can they be taken a step further, to determine properties such as thermal displacement or thermodynamic parameters (*i.e.* due to the dependence on the vibrational partition function)?

Volker Deringer answered: Anything that is based on energies or forces is accessible. If the forces are accurate, any other properties derived from them will be, too. For example, people have computed vibrational (phonon) modes using ML-based interatomic potentials with very good accuracy; once these are there, you could go and obtain vibrational contributions to the free energy, move on to the quasi-harmonic approximation, and so on. I think these are very interesting research directions for the future.

(6:[14]14) **Gerit Brandenburg** remarked: The discussion of the machine learning potentials leads to the comparison of Gaussian kernel regression and artificial neural networks. While comprehensive comparisons are out of reach for this discussion and one might refer to review articles (*e.g.* ref. 1), the computational speed has been mentioned. Independent of the employed descriptor, feed-forward neural networks do not need any matrix operation once they have been optimised. Thus, they can potentially be evaluated extremely quickly on graphic processing units. This holds for increasingly large reference datasets as the evaluation depends on the network geometry and not on the reference set. However, this might not be a major issue for the relatively small datasets used in potential energy interpolations, and the fitting process itself or the smoothness of the resulting potential may be more important. So the best suited machine learning model will strongly depend on the type of application.

1 J. Behler, *J. Chem. Phys.*, 2016, **145**, 170901.

Volker Deringer responded: Without going too deep into aspects of algorithms and implementations, which should be addressed by my co-author Gábor Csányi, I do note that we use a standard sparsification procedure in practice: the fitting algorithm selects a number (say, a few thousand) of the most representative atomic configurations, and there are regression parameters only for these. The cost of a GAP simulation, therefore, scales linearly with the number of “sparse points”, not with the size of the reference database.

After the meeting, Gábor Csányi (a co-author of the paper under discussion, DOI: 10.1039/C8FD00034D) communicated in writing: There is a widely shared misconception that kernel-based methods scale badly with system size. This is due to the pedagogical way these methods are typically introduced in textbooks: a square kernel matrix is formed between every pair of input data points, and is then inverted to obtain the coefficients of the fit, leading to $O(N^3)$ scaling. In practice, one never does this for large data sets. Instead, a rectangular matrix is formed between the N data points and a much smaller set of M data points (typically picked from the N set using various algorithms—this is the “sparsification” Volker Deringer refers to). This leads to an $M \times N$ matrix, for which the operations needed to determine the fit are $O(N)$. After the coefficients are calculated, the predicted properties are $O(M)$, and this does not scale with the database size at all, just like for neural network models. Formally, an equivalence can be drawn between kernel regression and 1-hidden-layer feed-forward neural networks, and in this case M corresponds to the number of neurons in the hidden layer.

In practice, for the large and accurate models published to date, the speed of predicting a force is within the same order of magnitude as our GAPs and Behler’s neural network models. Note that in the GAP case, a significant fraction of time is taken up by computing the representation, before the sum over M is taken, so significant optimisations are still possible in picking better (more quickly converging) basis functions. Typical prediction speeds are 10–100 milliseconds/atom/core, which parallelises extremely well due to the relatively large time spent on computing the force of a single atom. Therefore, applications with > 50 000 atoms are entirely feasible and are currently being undertaken.

(7:[15]15) **Qiang Zhu** enquired: In order to a make a good force field, how many structures do you need? Are there any stop criteria? Can this type of force field be used for MD simulations with large cells?

Volker Deringer replied: A typical value is a couple of thousand reference structures—after which the predicted properties converge to within what one is aiming for. The energy and force error is one possible stopping criterion, but experience has told us that this question is much more intricate. One can fit potentials with a reasonably low numerical error that still fail to capture important physical properties; and one can have potentials with a notable residual error (let's say, for the case of carbon, $\approx$ 1 eV Å^{-1} on the forces for highly disordered liquid and amorphous structures) but which perform exceptionally well for structural and mechanical properties (especially when, as in the liquid, the absolute forces are high anyway).[1] These machine-learned force fields can be, and have been, used for large systems: for example, we recently simulated the deposition of amorphous carbon films using system sizes of > 10 000 atoms,[2] and others have modelled GeTe nanowires to a similar extent,[3] approaching the real physical sizes of devices—and all at close-to-DFT accuracy!

1 Deringer and Csányi, *Phys. Rev. B*, 2017, **95**, 094203.
2 Caro *et al.*, *Phys. Rev. Lett.*, 2018, **120**, 166101.
3 Gabardi *et al.*, *J. Phys. Chem. C*, 2017, **121**, 23827.

(8:[16]16) **Qiang Zhu** remarked: If you did not sample enough configurations (just a couple of thousand), it seems that this would be inadequate to reconstruct the entire potential energy surface. You might get good estimations of the local minima, but the transition state and liquid-like states are unlikely to be described very well.

Volker Deringer responded: This would be true if there were no high-energy structures in the reference database, but the use of RSS here works to our distinct advantage. Perhaps surprisingly, a few thousand configurations are indeed enough to describe the system, including the relaxation trajectories that lead to the local minima. Other databases and potentials have been built for carbon and silicon with comparable numbers of reference structures, and both potentials capture liquid, amorphous, and other higher-energy regions reliably, including excellent agreement with diverse experimental probes for amorphous carbon[1] and amorphous silicon.[2] Interestingly, the liquid structures are often comparably easy to fit using an ML algorithm: they are so highly diffusive that just a handful of snapshots gives you lots of inequivalent atomic environments, and each of these simulation cells contains $3N$ force components, all of which contain valuable information about the potential-energy surface (PES). We also take distinct advantage of the fact that the PES is smooth, and therefore a limited number of sampling points are sufficient to describe it.

1 Caro *et al.*, *Phys. Rev. Lett.*, 2018, **120**, 166101.
2 Deringer *et al.*, *J. Phys. Chem. Lett.*, 2018, **9**, 2879.

(9:[17]17) **Matthew S Dyer** queried: From a practical point of view, it is very simple to use interatomic potentials by reading a small number of parameters from a published table. In practice, how would a researcher use the GAP potential that you have generated for phosphorus?

Volker Deringer replied: They can (and we would encourage them to!) download the potential parameter files, which are provided in XML format.[1] The GAP code, which directly reads these parameter files, is free to download and use for any non-commercial research from the same web address, and my colleagues have interfaced it to the QUIP, ASE, and LAMMPS simulation software. In that, it is arguably much easier and error-proof to use than coding up even a simple interatomic potential model by oneself!

[1] www.libatoms.org

(10:[18]18) **Mihails Arhangelskis** asked: My question concerns the type of DFT calculations used for fitting the interatomic potentials. Is there any advantage in performing DFT geometry optimizations? Presumably geometry optimization would provide more information in the form of atom coordinates and atomic forces around the local minimum of the crystal structure. Is there any advantage to doing that, or are single-point calculations sufficient for deriving the interatomic potentials?

Volker Deringer answered: This is an interesting alternative strategy, which indeed several authors have followed recently.[1,2] We do not do relaxations with DFT, and this a conscious decision—we want to make a potential with the lowest necessary computational effort in terms of reference data, in turn sacrificing some of the advantages (in the quality of early-generation structures) that DFT relaxation would have. Of course, we do perform relaxations too: with the GAP directly, which is progressively improved, as evidenced by the fact that we do find the structure of black P in the fourth generation.

[1] Hajinazar *et al.*, *Phys. Rev. B*, 2017, **95**, 014114.
[2] Kolsbjerg *et al.*, *Phys. Rev. B*, 2018, **97**, 195424.

(12:[19]19) **Artem R. Oganov** remarked: For pure elements, this approach works well. For compounds, how many parameters do you need in the force field? What would you do about long-range interactions?

Volker Deringer responded: For compounds, I think the key step would be fragment-based approaches (as partly seen in our attempts to "find" fibrous phosphorus); trying to find silica polymorphs (or worse, complex silicate compounds) by seeding with isolated atoms would likely waste a lot of resources; initialising the search with [SiO_4] building units should narrow down the search space to a more manageable size. Regarding long-range interactions, see my answer to the previous question—there are existing routes, and it would indeed be extremely interesting to couple these with GAP-RSS and other ML-driven structure-searching techniques.

(13:[20]20) **Virginia Burger** enquired: I noticed in the paper that you deemed the database converged once you had "learned" black phosphorus. Were there other indicators that the space had been sufficiently sampled that might help in making more a general convergence criterion for molecules in the future?

Volker Deringer replied: This is a very interesting point, to which I think the community hasn't found a definitive answer yet. A more general stopping criterion might be the energies of entries in the database (we've done this for boron) or the energy errors (as we've exemplified for Hittorf's phosphorus), but these would still make it necessary to define a numerical convergence criterion, and would hence still be somewhat specific to the system. Another more general approach could mirror what is being done in *ab initio* RSS: identifying minimum structures, and deeming the potential converged only once it finds all the minimum structures repeatedly.

(49:[23]22) **Artem R. Oganov** asked: You mentioned that for you it takes 2000 structure relaxations to get alpha-boron with 12 atoms/cell. Can you get gamma-boron with 28 atoms/cell and beta-boron with > 100 atoms/cell, and how long does this take you?

Volker Deringer answered: We have not tried this yet, but it would be a very interesting next step. Getting the β-rhombohedral boron structure with its ~106 atoms per unit cell out of a free search would certainly be a major accomplishment.

(39:[24]23) **Chris Collins** remarked: When you are fitting your force fields, you get a large number of parameters. Do you have any upper limits on how many parameters your fitting is permitted to use? Do you use any methods to check to see if you are overfitting your training data?

Volker Deringer responded: There are two different types of parameters that one needs to control when fitting GAP models. There are the hyper-parameters, including the expected noise in the input data and the smoothness of the SOAP kernel (that is, how "fuzzy" the comparison between two atomic environments is being made). These need to be controlled by hand but are limited in number, typically fewer than ten, and the sensible ranges are small. There is also the number of regression parameters, normally a couple thousand for SOAP, which can be converged. Generally, overfitting is checked by taking parts of the training database and evaluating the error for those—or by having a separate test set, as was the case in this paper.

(44:[25]24) **Gregory Beran** queried: Given that your machine learning model is based on local geometric descriptors (SOAP), could you comment more on how pressure is included in the model? Is it another input descriptor?

Volker Deringer replied: The (DFT-computed) virial stresses enter the fit, alongside the energies and forces. However, as descriptors we use solely the atomic environments—which are, of course, influenced by the applied pressure. For example, if we fit a GAP-RSS model for phosphorus using applied external pressure during the structural relaxations, it will have a much easier time picking up the six-fold coordinated environments of a simple cubic lattice (see Fig. 5 in the paper).

(45:[26]25) **Gregory Beran** asked: How well does the machine learning model reproduce the DFT enthalpy *versus* pressure curves?

Volker Deringer answered: Once the DFT energy–volume curve can be reproduced with confidence, this curve can be directly translated into an enthalpy–pressure plot of the same confidence level. Figure 5 suggests that data for the high-pressure forms (rhombohedral and simple cubic) can be fitted in principle, but this could of course be improved by including more distorted cells explicitly for this purpose. This is the topic of ongoing work.

(802:[22]26) **Gerit Brandenburg** communicated: The authors cite ref. 1 for using the PBEsol density functional approximation (DFA) in their data generation. As the GAP potential (and actually the whole structure search) builds upon this, more careful validation might be worthwhile. Specifically, it is not clear whether the semi-local DFA is capable of describing the subtle balance of covalent interactions of phosphorus layers and the van der Waals (vdW) interactions between them. To this aim, Table 1 summarizes low-temperature diffraction data for black phosphorus that has been corrected for thermal and zero-point effects, as well as very recent high-level calculations of exfoliation energies. This is compared to results from PBEsol as well as a few other DFAs. While PBE is unsatisfying, PBEsol has errors of 1.9 Å/atom and 40 meV/atom for the unit cell volume and exfoliation energy, respectively, and is reasonably well suited for describing the electronic structure of black phosphorus. However, it should be noted that inclusion of vdW interactions in the DFA framework substantially improves the agreement with the reference data.

1 A. S. Rodin, A. Carvalho and A. H. Castro, *Phys. Rev. Lett.*, 2014, **112**, 176801.
2 A. Brown and S. Rundqvist, *Acta Crystallogr. B*, 1965, **19**, 684.
3 L. Cartz, S. R. Srinivasa, R. J. Riedner, J. D. Jorgensen and T. G. Worlton, *J. Chem. Phys.*, 1979, **71**, 1718.
4 L. Shulenburger, A. Baczewski, Z. Zhu, J. Guan and D. Tománek, *Nano Lett.*, 2015, **15**, 8170.

Table 1 Structural data and exfoliation energy of black P from experient and high-level methodologies compared to density functional approximations

Reference data	E_{exf} [meV/atom]	V [Å^3/atom]	a [Å]	b [Å]	c [Å]
Powder diffr.[2]	—	19.0	3.31	10.48	4.38
Neutron diffr.[3]	—	19.0	3.31	10.47	4.37
Diffr. [QHA corrected][a]		**18.8**	**3.31**	**10.42**	**4.36**
Quantum Monte Carlo[4]	¬80	—			
Ring coupled cluster[5]	**¬92**	—			
Local MP2[6]		18.6	3.27	10.42	4.36
Density functional approximations					
PBE	¬10	21.4	3.31	11.33	4.57
PBEsol	¬52	17.9	3.32	10.29	4.20
PBE-MBD	¬78	19.0	3.31	10.52	4.36
PBE-D3[b]	¬74	19.8	3.31	10.74	4.46
PBE0-D3[b]	¬85	19.4	3.29	10.62	4.44

[a] Thermal and zero-point effects extracted from diffraction data employing the quasi-harmonic approximation at the PBE0-D3(0,ATM) level.[7] [b] London dispersion scheme used in the zero-damping scheme including Axilrod–Teller–Muto-type three-body terms, data from ref. 7.

5 M. Schütz, L. Maschio, A. J. Karttunen and D. Usvyat, *J. Phys. Chem. Lett.*, 2017, **8**, 1290.
6 G. Sansone, L. Maschio, D. Usvyat, M. Schütz and A. Karttunen, *J. Phys. Chem. Lett.*, 2015, **7**, 131.
7 G. Sansone, A. Karttunen, D. Usvyat, M. Schütz, J. G. Brandenburg and L. Maschio, *Chem. Commun.*, 2018, in press.

Volker Deringer communicated in reply: I recall the dispersion interactions in phosphorus allotropes being a generally challenging issue (see also ref. 1); in this light, the performance of state-of-the-art dispersion correction schemes is really encouraging. Our current discussion paper is a proof-of-concept study for GAP-RSS, where one of the milestones has been "discovering" the structure of black phosphorus at all (and at this stage I'm personally very happy with how PBEsol performs); ongoing work beyond this will indeed deal with the inclusion of dispersion interactions in a more general phosphorus potential.

1 F. Bachhuber *et al.*, *Angew. Chem., Int. Ed.*, 2014, **53**, 11629.

(46:[27]27) **Marcus Neumann** enquired: Where do you see the advantages of the random search method? It is surely easy to implement but is generally considered less efficient than methods such as genetic algorithms or Monte Carlo parallel tempering.

Volker Deringer responded: One distinct advantage is precisely the broad selection of initial structures for generating the potential ("broad" here being subject to reasonable constraints): we need an appropriate coverage of the potential-energy surface to start with, and we do want to include high-energy structures. That being said, the general ideas from our paper can be transferred to other structure-searching methods—in fact, the preceding paper coupled GAP fitting to particle-swarm optimisation, and Artem Oganov and Alex Shapeev have combined a novel class of so-called "moment tensor potentials" with evolutionary searching. All these methods have given highly encouraging results so far. I believe the implementation itself is less of an issue: as one example, GAP models can be directly loaded into the popular LAMMPS package, and the latter in turn has been interfaced to many structure-prediction codes.

(14:[28]28) **Peter Spackman** opened the discussion of the paper by Noa Marom: Did you try using different clustering algorithms? How sensitive were the results you found to the choice of clustering algorithm and parameters?

Noa Marom responded: In ref. 55 we compared the performance of affinity propagation to *k*-means clustering.[1] In general, the class of clustering algorithms that can autonomously determine the number of clusters is most suitable for evolutionary niching because the clusters dynamically evolve. In addition, affinity propagation enables defining custom distance metrics. GAtor interfaces with scikit-learn and the user may test any clustering algorithm available therein.

1 X. Li, F. Curtis, T. Rose, C. Schober, A. Vázquez-Mayagoitia and N. Marom, *J. Chem. Phys.*, 2018, **148**, 241701.

(15:[29]29) **Shiyue Yang** asked: The cluster-based fitness is defined by $\overline{f}_{\max}$ as the denominator, and according to the previous definition, this term seems to be

0 all the time. Is it my misunderstanding or is something wrong with the definition?

Noa Marom answered: This is a misunderstanding of the notation. $\bar{f}_{\max}$ is the maximum cluster-based fitness in the population, which is never zero. It does not correspond to the maximal energy in the population. A clarification of the definition of $\bar{f}_{\max}$ has been added to the paper (DOI: 10.1039/C8FD00067K).

(16:[30]30) **Seiji Tsuzuki** enquired: I would like to know the details of the new structure generation procedure in the optimization process using the genetic algorithm. How do you generate new structures with the genetic algorithm? What kind of structure parameters do you change to get new structures?

Noa Marom replied: The crossover and mutation schemes available in GAtor are described in detail in Sections 3.4 and 3.5 of ref. 47.[1]

[1] F. Curtis, X. Li, T. Rose, A. Vázquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli and N. Marom, *J. Chem. Theory Comput.*, 2018, **14**, 2246.

(17:[31]31) **Virginia Burger** remarked: I really liked your niching strategy. I am wondering if you tried tuning the fitness function (equation 4), so that either energy or cluster size would have more weight? For example, since the L descriptor is proportional to the inverse cube root of the unit cell volume, perhaps something like cubing f_i in equation 4 would help reduce the occasional promotion of large unit cell volume structures that is mentioned on page 10 of your paper (DOI: 10.1039/C8FD00067K). Also, I think that the definitions would be clearer if $\bar{f}_{\max}$ were defined in mathematical notation. I had trouble understanding whether $\bar{f}_{\max} = \max_i(\bar{f}_i)$ while reading the paper.

Noa Marom responded: GAtor is an open source code. Its modular structure enables the user to implement and test various fitness functions as they see fit for their own purposes. Clarification has been added to the paper that $\bar{f}_{\max}$ is the maximum cluster-based fitness in the population.

(18:[32]32) **Asbjoern Burow** asked: In how far is the construction of your fitness functions motivated by structural and physical constraints from experimental observations and conditions, if the (free) energy alone is not employed? Can you point out the driving principle behind your fitness constructions?

(20:[33]33) **Artem R. Oganov** enquired: Maybe you have brought more complexity into the algorithm than is necessary. We have been developing evolutionary algorithms for organic and inorganic compounds for many years. For us, genetic drift hasn't been a major problem. It is sufficient to use energy itself as fitness and, as for similar structures, simply kill duplicate structures and among (nearly) identical structures keep one, the structure with the lowest energy.

Noa Marom replied: As explained in detail in Section 3.6.2 of ref. 47,[1] GAtor performs similarity checks twice, before and after relaxation, and duplicate structures are eliminated. As explained here and in ref. 47,[1] evolutionary drift may be the result of the structure of the potential energy landscape with some wide basins that are more likely to be sampled and some narrow funnels that are rarely sampled. The choice of energy method (DFT functional or force field) may also affect the structure of the potential energy landscape and different energy methods may favor particular packing motifs, as we have demonstrated in ref. 47 and 48.[2]

1 F. Curtis, X. Li, T. Rose, A. Vázquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli and N. Marom, *J. Chem. Theory Comput.*, 2018, **14**, 2246.
2 F. Curtis, X. Wang and N. Marom, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**, 562–570.

(21:[34]34) **Artem R. Oganov** asked: Why do you need to divide fitness by the number of structures in the cluster of structures?

Noa Marom responded: As explained in the paper: "With this fitness scheme, structures in less populated clusters with under-represented structural motifs have higher fitness values, and hence a higher probability of being selected for mating. Penalizing the fitness of over-sampled clusters and steering the GA towards under-sampled clusters provides a more uniform sampling of the configuration space. If all the cluster sizes were equal, this fitness scheme would be equivalent to energy-based fitness."

(22:[35]35) **Christian Schön** addressed Noa Marom and Artem R. Oganov: I got the impression that you found many basins containing many local minima separated by very small energy barriers. In that case, even at low temperatures, the system can easily switch between these "structures" on the observational time scale. When looking for free energies, a description including configurational entropy would then be more appropriate than just computing the (harmonic) phonon contribution to the free energy. Do the experiments produce data that would agree with an average over such nearby local minima? Did you try to explore the set of closely related minima, or did you just consider one "representative" minimum?

Noa Marom answered: The GA does not provide any information on the energy barriers between minima. Therefore, we cannot determine whether the system is likely to switch between structures at finite temperature.

Artem R. Oganov noted: This is not what we usually find in our studies of inorganic crystals (but Noa Marom's work is on organic substances). What we usually find is well-defined global minima of the static energy (or enthalpy) at $T = 0$ K, and these persist when you include dynamics. Of course, there are cases similar to what you describe, *e.g.* plastic phases of molecular crystals, where molecules can freely rotate.

(23:[36]36) **Qiang Zhu** enquired: Can you describe the performances of different descriptors? Would it be useful to develop some universal descriptor by using them as the primary descriptors?

Noa Marom responded: The performances of the lattice parameter-based descriptor (L), the radial distribution function (RDF), and the relative coordinate descriptor (RCD) are discussed in detail in the paper. For this system both RDF and RCD perform well, although RCD may perform better for systems that do not have distinct intermolecular interactions like the halogen bonds of Target XIII (for example, molecules comprising only C and H atoms). The L descriptor may be unable to resolve the differences between packing motifs that do not correlate strongly with the unit cell size and shape. GAtor is an open source code and the user may implement and test various descriptors as they see fit for their own purposes.

(24:[37]37) **Virginia Burger** asked: As a follow-up question, why were you not satisfied with the L descriptor? It appears to have the most diverse sampling, and it looks like it came close to finding the experimental structure.

Noa Marom replied: As explained in the conclusion of the paper: "Some differences were observed in the GA behavior when evolutionary niching was performed based on different descriptors. Using the L descriptor resulted in increased sampling of a region characterized by structures with very long *c*-parameters, particularly when the GA was started from the diverse initial pool. This was counterproductive in the case of Target XIII because this region did not contain any important low-energy structures. The L descriptor was additionally found to produce fewer clusters than the RDF and RCD descriptors. This suggests that the sensitivity of the L descriptor may be insufficient to resolve different packing motifs that do not necessarily correlate with the unit cell shape and volume."

(28:[54]39) **Graeme Day** queried: As a general question about genetic algorithms (GAs) for crystal structure prediction, I'm interested in your thoughts about how high up the energy landscape a GA can provide a complete sampling of possible crystal structures. I ask because we are sometimes interested in structures that are in high-energy regions, outside of the typical energy range of polymorphism (7–10 kJ mol^{-1}).

I can see how a GA can be directed to structures with specific properties by including the predicted property in the fitness function. However, this requires that we can quantify a property of interest, which is not always the case; we might not know what property the unknown high-energy structure might have, but still want to predict its existence. This might be the case in a pharmaceutical materials context. So I wanted to know whether we can ensure that we explore the landscape thoroughly, and up to what energy a GA gives all or most of the observable structures when the fitness function is defined by the calculated energy, perhaps along with structural diversity such as in the paper you presented.

As an example, we might be interested in knowing all the polymorphs that are possible because the occurrence of an unanticipated polymorph with associated changes in properties can have important consequences for a pharmaceutical molecule. The changed property could be solubility, crystal habit, stability or a number of other properties. We don't know which of these might change, but a change in any of them is important and needs to be anticipated. So we cannot use one particular property to include in our fitness function. Also, these structures might be quite high in energy, such as structures resulting from dehydration of a hydrate crystal structure, which might be much higher in energy than the usual few kJ mol^{-1} related to polymorphism. Therefore, from a pharmaceutical polymorphism perspective, we need to know what range of lattice energy we are fully exploring with the method we choose for crystal structure prediction.

Noa Marom responded: Genetic algorithms are designed to perform optimization by definition. Furthermore, GAs are designed to converge to the regions of interest with respect to the target property, rather than thoroughly explore other regions of the landscape. To promote broader exploration, one may run the GA several times with different fitness functions. For example, in ref. 60 we tailored GAs to search for structures with low energy, high electron affinity, and low ionization potential.[1] We found that the optimization for target electronic properties rather than energy minimization promoted the exploration of higher energy regions (see Figure 1 therein). In addition to energy and electronic properties, the fitness function may target diversity. If one is completely unable to define target quantities to optimize then perhaps a different approach (*e.g.* random search, Monte Carlo, or MD) would be more suitable than a GA.

1 S. Bhattacharya, B. H. Sonin, C. J. Jumonville, L. M. Ghiringhelli and N. Marom, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **91**, 241115.

(25:[38]42) **Marcus Neumann** asked: It can be quite misleading to test and optimize the performance of a global search algorithm against a single compound. What prevented you from applying the same approach to a larger number of compounds?

Noa Marom answered: As explained in the paper, Target XIII was chosen as a case study for evolutionary niching because “In ref. 47, the experimental structure of Target XIII was rarely generated when using a traditional energy-based fitness function”. Evolutionary niching using the L descriptor was previously applied to Target XXII in ref. 47 and proved beneficial.[1] The evolutionary niching algorithm is by no means optimized for a particular system. Other than the choice of atom pairs for the RDF, which is system specific, the method is completely general. We have emphasized, both here and in ref. 47,[1] that because it is impossible to know *a priori* which GA settings will perform best for a particular system, the recommended best practice is to run GAtor several times with different settings and collect the structures found in all runs, rather than running it with the same settings for many iterations. Furthermore, we recommend using evolutionary niching in at least one run. Due to the short deadline for submitting papers to the *Faraday Discussion* we were unable to run tests for more than one target. We will certainly test GAtor on additional targets in the future.

1 F. Curtis, X. Li, T. Rose, A. Vázquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli and N. Marom, *J. Chem. Theory Comput.*, 2018, **14**, 2246.

(26:[39]43) **Susan Reutzel-Edens** enquired: As a follow-on, how can you tell *a priori* that regions or clusters of structures in the 3D representations outside of where a known experimental structure is found are not important?

Noa Marom replied: It is impossible to tell *a priori* which regions of the PES will be important. This is exactly the reason for using evolutionary niching to promote the exploration of more diverse regions.

(51:[40]44) **Jonas Nyman** remarked: You sound very confident in GAtor finding all the energy minima. You could compare all your predicted structures to those submitted in the blind test to see if your search is really complete. You should get all their structures and then some. Has such a comparison been done?

Noa Marom responded: In general, it is impossible to know whether all the low-lying energy minima have been found in any CSP attempt. Hence, our recommended best practice is to run GAtor several times with different settings and collect the structures found in all runs. If the same low-energy structures consistently appear, it is reasonable to assume that the search was exhaustive (at least within the range of lattice parameters explored). Prior to the sixth blind test, submissions contained only three structures and CIF files of the predicted structures are not available, which makes the comparison difficult for older blind test targets, such as Target XIII. Our experience so far indicates that GAtor's ability to spontaneously generate structures with $Z' > 1$ yields additional low-energy minima, as discussed in ref. 47. For example, for Target II GAtor found a low-energy structure with $Z' = 2$ and a scaffold packing motif, which had not previously been reported by others. Several low-energy structures with $Z' = 2$ were also found for Target XIII.

(11:[41]45) **Sharmarke Mohamed** opened a general discussion of the papers by Yanming Ma, Volker L. Deringer and Noa Marom: We know from various crystal structure prediction (CSP) studies that if the structure generation step does not employ physically realistic intermolecular potentials for sampling the range of forces between the atoms/molecules/ions in question, it will often not be possible to locate the experimental structure in subsequent structure optimization steps, regardless of how accurate the energy model may be. Clearly there are some promising data on the value of using GAPs for accelerating the CSP process. I am curious about the views of both Yanming Ma (paper 11894, DOI: 10.1039/C8FD00055G) and Volker Deringer (paper 19763, DOI: 10.1039/C8FD00034D) on how easy it would be to extend machine learning potentials/algorithms to organic molecular systems beyond the boron (paper 11894, DOI: 10.1039/C8FD00055G) and phosphorus (paper 19763, DOI: 10.1039/C8FD00034D) clusters studied by the authors. My concern is that the combination of the cut-off implicit in the Atom Centered Symmetry Function (ACSF) descriptors, combined with the poor description of long-range effects in DFT methods, may make it challenging to locate global minima corresponding to experimental crystal structures for hydrogen-bonded organic molecular systems (*e.g.* pharmaceuticals) where such long-range interactions are integral to the crystal packing.

Volker Deringer replied: The correct description of long-range interactions is indeed an extremely important goal—both regarding dispersion and electrostatic interactions. Jörg Behler and his co-workers showed how one can generate a separate machine learning (neural network-type) model for the long-range electrostatic interactions and add this to the short-range potential,[1] and later made and applied potentials for water that included van der Waals interactions.[2] There is indeed a recent machine-learned potential for organic molecules, called "ANI-1".[3] Note that even in the absence of an explicit treatment of long-range electrostatics, hydrogen bonding may often be adequately described using ML models with their typical cut-off radii of around 5–6 Å.

1 Artrith *et al.*, *Phys. Rev. B*, 2011, **83**, 153101.
2 Morawietz *et al.*, *Proc. Natl. Acad. Sci. USA*, 2016, **113**, 8368.
3 Smith *et al.*, *Chem. Sci.*, 2017, **8**, 3192.

Yanming Ma replied: I agree. In the case of organic molecular crystals, it is even challenging for DFT to give a very accurate energy order for different crystal packing, and some sophisticated force fields are superior to the ML potentials at the current stage. However, extending the current ML potential to molecular systems is not impossible. Many recent works have shown promising solutions to this problem, *e.g.* expressing the total energy as a sum of the short-range energy, the electrostatic energy and the dispersion correction energy.[1]

1 *J. Phys. Chem. A*, 2013, **117**, 7356.

(33:[42]46) **Scott Woodley** addressed Volker Deringer and Yanming Ma: When I started my research in the field of crystal structure prediction (back in the late 1990s), I was told that it was too computationally expensive to search the energy landscape for the global minimum atomic configuration when the total energy was defined using interatomic potentials (IP) and included an Ewald Summation for the long-range terms. With the increased computer resources that are available today, many groups are now searching the "orders of magnitude more expensive to compute", DFT (or similar electronic structure-based) energy landscapes for low-energy local minima. However, there are still many groups, like mine, who typically first target finding the atomic configurations of the IP local minima on the IP energy landscape, and use these as the initial configurations for electronic structure calculations (including the refinement of the atomic configuration using a standard local optimisation method). My group specialises in developing IP for oxides, for example, where the latter approach is much more efficient. Are there IP models of sufficient quality for boron and phosphorus, and, if so, how efficient is your approach compared to first targeting IP local minima?

Volker Deringer responded: There are no general IP models for elemental boron and phosphorus, which both adopt extremely diverse structures despite being "only" elemental solids. (In fact, there is a single report on an IP model for B that came out just a few years ago, and has been fitted only for the α-rhombohedral form.[1]) This makes these systems particularly suitable targets for ML-driven exploration and potential fitting at the same time—this is why we chose them for GAP-RSS.

1 Pokatashkin *et al.*, *Model. Simul. Mater. Sci. Eng.*, 2015, **23**, 045014.

Yanming Ma responded: Thanks very much for the professional comments. I think both IP and ML potentials have their own advantages. However, to the best of our knowledge, there is no IP model available for isolated boron clusters.

(34:[43]47) **Seiji Tsuzuki** asked: I have a question on the fitting of the potential energy surface using machine learning. I would like to know the number of samples necessary to find the global energy minimum. I suspect that quite a large number of samples are needed to find the global energy minimum. I think that the structure of one of the samples must be close to the global energy minimum, if you want find the global minimum by the fitting of the potential energy surface. I suspect that the number of samples necessary for finding the global minimum is close to the number of local potential minima.

Volker Deringer answered: The number of required samples certainly depends on the system, and on the shape and properties of its potential-energy surface (PES). As a rule of thumb, the reference database for our phosphorus GAP-RSS model contains on the order of 3000 DFT single-point computations; this includes structures that have been relaxed into local minima, but also intermediates and the initial randomised cells.

(40:[44]48) **Michael Ruggiero** addressed Volker Deringer and Yanming Ma: Machine learning is clearly becoming very popular. It is encouraging to see the success of ML in using learned potentials to predict structures, but it seems that armed with the knowledge of the entire potential energy surface there is so much more information available. Can we take this data a step further to predict things like dynamics? From this could we generate accurate thermodynamics? What about temperature dependent structures, anisotropic displacement parameters, phase transitions, *etc.*? Is this a limit of the method, or a limit of the development of tools for analyzing the data? Is there a "hard line" where ML will stop being effective?

Volker Deringer responded: You have stated it already: once the potential energy surface (PES) is known, one can describe all the static and dynamical properties within the Born–Oppenheimer approximation, including the description of thermodynamics. This holds irrespective of whether the description of this PES comes directly from a quantum-mechanical computation or from an (accurate) ML fit. This is what can make ML-based interatomic potentials such powerful tools! Indeed, the search for crystal structures is just one small (but growing) application of these simulation methods—we started GAP-based structure searching around two years ago,[1] but ML-based interatomic potentials have been around much longer. The first application of high-dimensional neural network potentials (in 2008!) dealt with the high-pressure phase transitions in Si;[2] we have discussed the usefulness of ML-based potentials for computing thermal properties in a recent overview paper.[3] Again, anything that you can extract from an atomistic simulation is principally in reach, and that is also where the limit is—we can go up to tens of thousands of atoms with ML potentials, and perhaps to millions in a few years, but even this will correspond to simulation cells sized a few tens of nanometres, not more.

1 Deringer *et al.*, *ChemPhysChem*, 2017, **18**, 873.
2 J. Behler *et al.*, *Phys. Rev. Lett.*, 2008, **100**, 185501.
3 G. C. Sosso *et al.*, *Mol. Simul.*, 2018, **44**, 866.

Yanming Ma responded: The accuracy and transferability of ML potentials are closely related to the training set. With a reasonable choice of reference dataset, I think it is promising to use ML potentials to deal with these problems. However, the training set should be constructed carefully, because the potential may become invalid during extrapolation.

(41:[45]49) **Michael Ruggiero** asked: Given the known structures of polymorphs, how far away are machine learning methods from predicting the relative thermodynamic stabilities of crystalline polymorphs?

Volker Deringer answered: I believe that what is keeping us furthest away is not the ML-*versus*-DFT error but the DFT-*versus*-reality error. In so far as ML potentials give accurate ground-state energies and accurate vibrational free energies, there is no reason not to use them for this purpose.

(42:[46]51) **Scott Woodley** enquired: Sample points on the *ab initio* energy landscape are evaluated in all three papers presented, but, if I remember correctly, only in Noa Marom's paper did I read that additional checks are first made in order to determine whether to evaluate the energy of a newly generated atomic configuration. Surely, particularly for the random sampling approach, you also make initial checks before deciding whether to evaluate a structure?

Volker Deringer replied: For the initial generations, we do not check at all, but instead evaluate all the configurations; we do this on purpose, because we want to sample both low- and (reasonably) high-energy regions. Later on, once the reference database has grown to a reasonable size, we do introduce such checks: we count coordination numbers, with the chemical knowledge in mind that P normally forms three bonds (8–*N* rule). We are currently looking into more general, less element-specific criteria, which I think are very much needed.

Yanming Ma replied: For the CALYPSO method, once a new structure is generated, the bond characterization matrix (BCM) is calculated and applied to examine the similarity/distance of this structure compared with all the previous ones.[1] Similar structures with BCM distances below a certain threshold will be eliminated without energy evaluation by DFT or ML potentials.

[1] *J. Chem. Phys.*, 2012, **137**, 084104.

(43:[47]52) **Scott Woodley** asked: Following on from my previous question and Prof. Christian Schön's point, does the AIRSS software really perform local optimisations on *random* structures during its search of the energy landscape? If I understand correctly, the efficiency of electronic structure codes is much better near local minima, whereas outside local energy basins, or a sufficient distance from the local minimum, the convergence of the electronic structure can be problematic and it is likely that, during a search, more computer resources are spent evaluating regions of the landscape that you are not interested in if you do not avoid assessing such points using an electronic structure code.

Volker Deringer answered: There are reasonable boundary conditions on making the structures, as stated in our paper. Accordingly, the computed energies span a few eV per atom, but do not completely go through the roof. We've seen that happen in early computational experiments with more loose criteria, where we did still converge the electronic-structure calculations but ended up including regions much higher in energy; this definitely didn't help! Generating structures so unrealistic (atoms so close to one another) that the electronic-structure code doesn't converge at all would be another issue; if our GAP-RSS iterations had led to such structures (which they didn't), something would have surely been seriously wrong.

(47:[50]53) **Graeme Day** addressed Volker Deringer and Noa Marom: What do you think is a good measure of the performance of a crystal structure search method? In the most common applications of crystal structure prediction, such as pharmaceutical polymorph screening, we are interested in the entire landscape of possible structures. A common comparison between methods is how quickly they reach the global energy minimum structure. However, the importance of knowing about experimentally realisable higher energy structures gives you a situation where how quickly we get to global minimum is a poor measure of performance. Crystal structure prediction is not simply a global optimisation problem, but relies on global exploration.

Volker Deringer replied: This is an extremely interesting point because it can be linked to the synthesis of (possible) metastable compounds and to the description of amorphous phases. We believe that a broad, stochastic approach such as RSS does help with exactly that—in any case, the aim of our technique is to generate a general interatomic potential for a given chemical composition, not just for its minimum structures.

Noa Marom replied: The measure of performance depends on the objective of the search: If the objective is to find low-energy structures, we would define success as finding all the low-energy structures that could be potential polymorphs, not just the global minimum structure. As explained in Section 3.7 of ref. 47 in our paper:[1] "Because there is no unique way of converging a genetic algorithm, the user specifies simple conditions for when the code should terminate. One option is choosing to terminate the algorithm if a certain number of the best structures in the common population have not changed in a user-defined amount of iterations (*e.g.*, if the top 20 structures have not changed in 50 iterations of the GA). This tracks whether all low-energy structures have been located in a reasonable number of iterations." The GA fitness function can be tailored to optimize for properties other than the total energy, as we did for example in ref. 60.[2] In this case, success would be defined as finding structures with the target properties and/or identifying structure–property relations that provide insight on how to design structures with the target properties. If the objective of the search is to promote diverse sampling through the exploration of high-energy regions, one may construct a GA fitness function that weighs diversity more heavily than energy. In this case, a convergence criterion could be when the algorithm has detected the same number of clusters for a long time, meaning no new packing motifs are being generated so the space has been, theoretically, fully sampled.

[1] F. Curtis, X. Li, T. Rose, A. Vázquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli and N. Marom, *J. Chem. Theory Comput.*, 2018, **14**, 2246.

2 S. Bhattacharya, B. H. Sonin, C. J. Jumonville, L. M. Ghiringhelli and N. Marom, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **91**, 241115.

(48:[51]54) **Artem R. Oganov** remarked: Our inorganic blind test only tested search methods; all teams used identical ranking methods (force fields or DFT functionals and pseudopotentials, with the same force fields or DFT codes), and to prevent leakage of information, avoided experimentally studied systems – all this is unlike the CCDC test for organic crystals. It's about time we repeat this test; we need more teams, more methods. We need to see how different global optimization strategies compare.

Graeme Day responded: I agree that this would be good to repeat. If such a challenge were set up as a test of global optimization strategies, it would be good to see the range of target structures be quite broad, perhaps even including an organic molecular system. Of course, for organic molecular systems, a method's effectiveness for finding the global minimum in lattice energy is only one measure of a method. We know that observed polymorphs often do not correspond to the global minimum and we need to ensure that a good method locates these alternative structures as well.

(50:[52]55) **Marcus Neumann** addressed Artem R. Oganov and Graeme Day: CSP needs to solve two problems: the structure generation problem and the energy ranking problem. In his blind test on inorganic CSP organized several years ago, Artem Oganov focused on the first challenge only by imposing on every participant the use of the same energy calculation method. A constraint of this type may be an interesting option for the next blind test on organic CSP. Especially if all the submitted crystal structures are energy optimized, it would be easy to combine all the submitted structures for each compound into a single crystal energy landscape against which the completeness of each method can then be assessed.

Such a completeness estimate should take into account that:

– Many local minima on the potential energy hypersurface have small attractive basins and are closed to other, deeper local minima. Such minima are not found in every search and are also not relevant for practical CSP. The completeness assessment should only be based on structures that are found by at least two participants.

– CSP methods tend to be biased towards the more stable local minima, which is desirable in terms of performance. So the completeness estimate should be limited to the low-energy region of the potential energy landscape.

Graeme Day replied: I think that this is a good suggestion for a benchmarking exercise of search methods, since it isolates one aspect of the crystal structure prediction problem. However, I am not sure that such an exercise needs to be performed as a blind test. The analysis of these results would be to judge the completeness of the search by comparing all of the low-energy structures generated by different methods. Therefore, I do not see the advantage of this being performed on a blind test molecule (*i.e.* a molecule whose observed crystal structure is not known). The restriction to blind test conditions can really limit the choice of target molecule, which seems unnecessary. Perhaps an exercise of this format could be organised independently of the blind tests.

Artem R. Oganov replied: I completely agree. I hope that the community will continue the inorganic blind tests, and that such rigorous tests of search and – separately – ranking methods will also become part of the organic blind tests.

(52:[56]56) **Sten Nilsson Lill** said: We discussed CCDC blind tests for crystal structure prediction and it was suggested that maybe the blind tests need a new focus, not only who finds the energetically most stable structures. Marcus Neumann suggested combining all the results and seeing who has the best coverage of all the suggested structures. Is it time to also have a property blind test? For example regarding crystalline solubility, *i.e.* including a full thermodynamical cycle of one molecule going from the crystalline phase to the solution phase, or material properties (elasticity, hardness, slip planes, stickiness, *etc.*), or some other key property?

Graeme Day answered: My opinion is that the blind tests of crystal structure prediction should test methods for crystal structure prediction and leave property prediction as a separate test. To combine structure and property prediction in one blind test would add quite a lot of complexity in terms of analysis. Is it better for a method to predict the wrong structure, but get close to the property, or predict the correct structure, while being farther away on the property prediction? I think that we would end up judging the separate parts of the prediction on their own, in the way that we currently try to analyse the crystal structure search and ranking problems separately. In view of this, it is probably more effective to have separate blind challenges for structure and property prediction. There have been similar tests for solubility prediction.[1]

[1] A. Llinàs, R. C. Glen and J. M. Goodman, *J. Chem. Inf. Model.*, 2008, **48**, 1289–1303.

(53:[57]57) **Qiang Zhu** addressed Volker Deringer and Noa Marom: What are the best methods for developing descriptors for crystals? We know that the current descriptors take into account the local chemical environment. Is it possible to develop a universal descriptor?

Volker Deringer responded: Developing descriptors is indeed a very active research topic—one needs mathematical ways to “encode” a crystal structure; this needn’t be in an atom-centred picture, but in practice often is. In the case of interatomic potential fitting, this normally means describing all the neighbours of an atom up to a given cut-off radius (“many-body descriptor”). There is no one best descriptor, but there are several highly successful approaches, including atom-centred symmetry functions (ACSF)[1] and the Smooth Overlap of Atomic Positions (SOAP).[2,3] Both ACSF and SOAP are universal in the sense that they can be applied to any conceivable chemical structure, which to the algorithm is just a list of coordinates and atomic numbers within the cut-off sphere.

Two more comments on that are in order: First, in our current paper, we add non-parametric two-body (2b) and three-body (3b) descriptors to make the fit more robust—which is particularly important for GAP-RSS. The final model is a (properly scaled) linear combination of 2b, 3b, and SOAP terms, following ideas we introduced earlier.[4] Second, when it goes beyond the fitting of interatomic potentials (which is just one of the many current applications of ML methods), there are other directions: as one example, it was recently suggested to represent crystal structures by (simulated) diffraction images, thereby creating two-dimensional “fingerprints”.[5]

1 J. Behler, *J. Chem. Phys.*, 2011, **134**, 074106.
2 A. P. Bartók, *Phys. Rev. B*, 2013, **87**, 184115.
3 A. P. Bartók *et al.*, *Sci. Adv.*, 2017, **3**, e1701816.
4 Deringer and Csányi, *Phys. Rev. B*, 2017, **95**, 094203.
5 Ziletti *et al.*, *Nat. Commun.*, 2018, **9**, 2775.

Noa Marom responded: Many researchers are working on developing descriptors/representations using different strategies. Time will tell whether a universal descriptor will be developed. In analogy, decades of attempts to develop a universal DFT functional have led to a proliferation of different functionals and researchers typically pick a functional that works well for their problem.

(55:[58]58) **Scott Woodley** commented: Artem Oganov earlier mentioned the equivalent inorganic crystal structure blind test that he organised,[1] where an evolutionary and simulated annealing Monte Carlo (MC) algorithm proved much better than locally optimising random structures. In the slide presented at this meeting (see Fig. 1), I showed the output from the simulated annealing approach

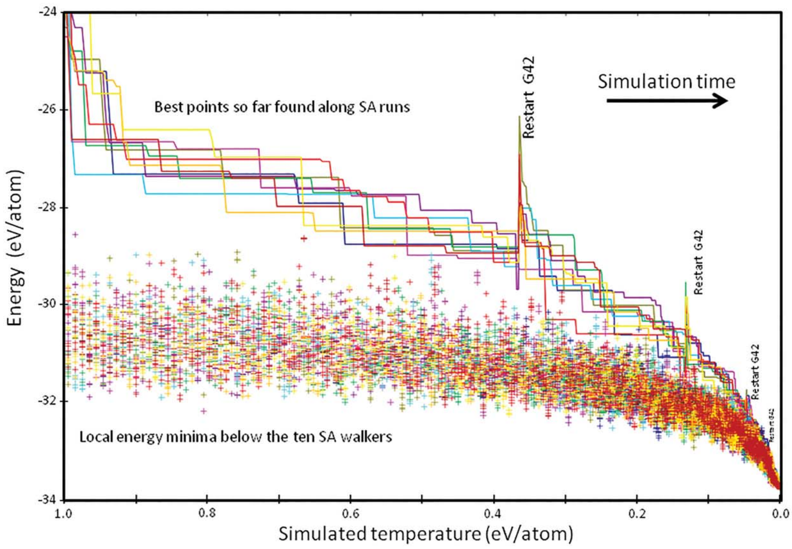


Fig. 1 The trajectories of the best energy found (lines) during ten simulated annealing Monte Carlo runs (each with a unique colour) for a unit cell containing 26 atoms. Each run starts (restarts) with a different random (the last accepted) configuration. For a number of configurations along each trajectory the local minima are reported (crosses). This figure is adapted from ‘From Ergodicity to Extended Phase Diagrams’, Scott M. Woodley and Alexey A. Sokol, *Angew. Chem., Int. Ed.*, 2012, **51**, 2–5.

for one of these test systems, $BaMgAl_4Si_4O_{16}$. It captures two key features of the local landscape of ten independent walkers as they progress across the energy landscape. Each starts from a random point and at a high temperature, which is used in the Metropolis Criterion to decide whether or not to accept a small Monte Carlo (MC) step (random atomic and cell displacements), the walker attempts to make the energy landscape. At each temperature 8000 MC steps are attempted (single-point energy evaluations). The temperature parameter of the Metropolis Criterion is slowly reduced so that the process of annealing is simulated; hence, there is an initial investment of computational resources that the user commits to. The first key feature that is collected by the G42 software[2,3] used to run these simulations (shown in Fig. 1 as a line) is the best energy found during the particular simulation. Given that the G42 software needed to be restarted several times for each walker, due to imposed computer resource time constraints, the first point upon a restart is the actual energy of where the walker was when the simulation was momentarily interrupted. The resulting increase in the lines shown mark temperatures at which the runs were restarted and also give a feeling of how far above the best line the walker typically moves. As temperature is decreased, the walkers gradually move down in energy. The second feature captured by G42 is displayed as crosses in the figure; these are the values of the local energy minima that are accessible to the walker (if only allowed to continue downhill) just before the temperature is dropped. Typically, the number of accessible local minima decreases with temperature. At high temperatures, the crosses represent the energy of the typical configurations that would be quickly found using a random sampling technique coupled with a local optimiser. It can be seen that at a temperature of 0.8 eV/atom there is a very low energy point, however, none of the crosses found at high temperature match the lower energies found from all ten simulated annealing runs. There is an initial computational effort, or cost, required for the simulated annealing runs as it is anticipated that the target structures will be found towards the end of the annealing process. The investment pays off for complex energy landscapes, as the high degree of complexity prevents the likely success of random searching; for less complex systems, the crosses at higher temperatures will be as low as those found at low temperatures.

1 Artem R. Oganov, J. Christian Schön, Martin Jansen, Scott M. Woodley, William W. Tipton and Richard G. Hennig, 'First blind test of inorganic crystal structure prediction methods', in *Modern Methods of Crystal Prediction*, Wiley VCH, 2010, pp. 223–231.
2 J. C. Schön and M. Jansen, *Angew. Chem.*, 1996, **108**, 1358–1377.
3 J. C. Schön and M. Jansen, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 1286–1304.

(57:[59]59) **Christian Schön** enquired: Regarding the ability of random structure search to be competitive with global optimization algorithms such as genetic algorithms or simulated annealing, Artem Oganov organized a structure prediction competition between GA, SimAnn and RSS about ten years ago.[1] Very complicated multinary ionic chemical systems (26, 52 and 60 atoms per variable cell) were searched, with all three codes employing the same empirical potential. RSS was the clear loser, with GA being ahead in two of the problems and SimAnn winning the third search. The conclusion is similar to the experience many people in the combinatorial optimization community have had for decades, *i.e.* random starting point quench searches are not competitive with more refined global optimization searches once the systems become more complex. So, the question is why did you employ the RSS method for your global searches – perhaps you used chemical intuition to select the starting points (although then the name RSS would not really be appropriate!)?

1 A. R. Oganov, J. C. Schön, M. Jansen, S. M. Woodley, W. W. Tipton and R. G. Hennig, in *Modern Methods of Crystal Structure Prediction*, ed. A. R. Oganov, Wiley-VCh, Weinheim, 2011, pp. 223–231.

Volker Deringer responded: We use RSS as a tool for exploring structure space in this paper; I do not intend to compare different structure-searching methods side by side in this discussion. As said above, our paper is not so much about specifically coupling GAP with RSS, and more generally about coupling machine learning interatomic potentials with structure searching.

(58:[60]60) **Noa Marom** asked: Are the initial structures generated by Chris Pickard's algorithm truly random or are physical constraints applied?

Volker Deringer answered: There are physical constraints where they are reasonable. For example, we keep atoms from getting too close to one another by using a "hard-sphere" constraint. Another constraint is given by the density, which we choose to be intermediate between that of red and black phosphorus, and which in the random structures is distributed around that value (this can be seen on the left-hand side of Fig. 2 in the paper, DOI: 10.1039/C8FD00034D).

(100:[100]100) **Matthew S Dyer** opened the discussion of the paper by Julian Keupp: Could this approach be used for other material types, such as covalent organic frameworks or inorganic framework materials? If so, how would this be done in practice?

Julian Keupp and **Rochus Schmid** replied: The reverse topological approach (RTA) – including the blueprint pre-optimization by topoFF – works in principle for any crystalline polymer. We have already used it to predict new hypothetical COFs.[1] In practice, the specific building blocks (organic, inorganic or hybrid) and their connecting atoms need to be defined and the algorithm can start.

1 S. Bureekaew and R. Schmid, *CrystEngComm*, 2013, **15**, 1551–1562.

(101:[101]101) **Marcus Neumann** asked: How many distinct morphologies have been observed for MOFs so far?

Julian Keupp answered: Many known MOFs are observed only in a single morphology. Few examples exist where multiple polymorphs emerge due to different synthetic conditions; one case is MOF-505 and NOT-100 (where the precursors are the same, but the observed PXRD differs slightly). This has been investigated by us using the RTA.[1] This is an example of isotopological polymorphism. I do not know of a MOF featuring topological polymorphs given the exact same precursors, but often slight modifications in the used linker can lead to different topologies.[2–4]

1 S. Bureekaew, V. Balwani, S. Amirjalayer and R. Schmid, Isoreticular isomerism in 4,4-connected paddle-wheel metal–organic frameworks: structural prediction by the reverse topological approach, *CrystEngComm*, 2014, **17**, 344–352.
2 Y.-B. Zhang, H. Furukawa, N. Ko, W. Nie, H. J. Park, S. Okajima, K. E. Cordova, H. Deng, J. Kim and O. M. Yaghi, Introduction of Functionality, Selection of Topology, and Enhancement of Gas Adsorption in Multivariate Metal–Organic Framework-177, *J. Am. Chem. Soc.*, 2015, **137**, 2641–2650.
3 A. Schneemann, R. Rudolf, S. Henke, Y. Takahashi, H. Banh, I. Hante, C. Schneider, S. Noro and R. A. Fischer, Linker functionalisation triggers an alternative 3D-topology for Zn-isophthalate-4,4′-bipyridine frameworks, *Dalton Trans.*, 2017, **46**, 8198–8203.
4 Y. Cai, A. R. Kulkarni, Y.-G. Huang, D. S. Sholl and K. S. Walton, Control of Metal–Organic Framework Crystal Topology by Ligand Functionalization: Functionalized HKUST-1 Derivatives, *Cryst. Growth Des.*, 2014, **14**, 6122–6128.

(102:[102]102) **Asbjoern Burow** enquired: As far as I understood, topologies (of high symmetry) are required as input to predict MOF crystal structure candidates with your approach. Do you see any potential to generalize the rules to less symmetric topologies of H-bonded molecular systems? Is there a way to reverse the procedure to predict topologies? Do you know of any published approaches which attempt this?

Matthew Addicoat replied: I'm pleased to say that ref. 1 + UFF4MOF[2,3] does allow you to form hydrogen-bonded frameworks just as easily as standard covalently-bound frameworks. We published the methodology last year,[4] but essentially specifying the hydrogen bond as a "real" bond with a minimal bond order straightforwardly yields the framework you expect. We've tried both forming the hydrogen bond as the framework forms and embedding the hydrogen bond in the linker, and both give excellent agreement with the reference crystal structure.

1 DOI: 10.1021/jp507643v.
2 DOI: 10.1021/ct400952t.
3 DOI: 10.1021/acs.jctc.6b00664.
4 DOI: 10.1063/1.4985196.

Julian Keupp replied: Our approach is not restricted to topologies (blueprints) of high symmetry. Any blueprint can in principle be used to generate structures. However, we are currently focusing on topologies from the RCSR (and EPINET) database. The reverse of our Reverse Topological Approach (the Topological Deconstruction) is a well established way to deconstruct experimental crystal structures into their underlying topologies, as implemented *e.g.* in the TOPOS code and others. Note that the corresponding TOPOS Database contains a large number of topologies, partly with large unit cells and low symmetry. These could be used in principle, but we have not tried this yet. Furthermore (please consider the corresponding documentation on the RCSR website), many of the topologies have not been observed experimentally but derived from other nets. Another approach to derive new topologies is the EPINET project, where 2D hyperbolic tilings are used as a source to generate topologies.

(103:[103]103) **Andrew Cooper** said: (The following is a draft version of Professor Cooper's question. Please answer as best you can, and we will contact you with the final version. You will then be able to update your answer if necessary.) Some work by Eduardi, MOFs with 12-coordinate linkers and even higher now. Exotic topologies are not represented in known databases. Have you tried to push your method to that sort of level of complexity?

Julian Keupp and **Rochus Schmid** responded: We need a topology as a starting point for the reverse topological approach. Thus, structures with yet unknown topologies cannot be predicted. However, building blocks with a large number of connections and atoms are not a problem for our method.

(104:[104]104) **Andrew Cooper** asked: (The following is a draft version of Professor Cooper's question. Please answer as best you can, and we will contact you with the final version. You will then be able to update your answer if necessary.) Do big systems have too many degrees of freedom?

Julian Keupp and **Rochus Schmid** answered: Of course, the numerical effort scales with the size of the system and the degrees of freedom, but there is no principal limit for our method in general.

(105:[105]105) **Matthew Addicoat** remarked: As we both know there are a number of frameworks that are not well or not completely described by a topology – things like MOF-74/CPO-27 and others that have 1D infinite connectors. How do you deal with those? Do you have a way to decide where to cut the infinite part?

Julian Keupp replied: In the case of MOF-74, there exist different possibilities to use a topology to construct the atomistic model; the topologies *eta* and *etb* both describe one particular realization of the three-fold screw symmetry and can be used with small inorganic cuts of the periodic inorganic structure motif. There also exists the parent net where this ambiguity is not resolved, which is the 5-connected *bnn* topology.[1] By using this topology one could in principle obtain one unit of the inorganic building block, which would in this case have to be periodic itself. For the moment we can only work with the former two topologies, because at the moment we cannot handle periodicity in the building blocks. For the cutting procedure we do not have a clean and off-the-shelf recipe. However, some heuristics we use are:

- Cut as few bonds as possible
- Try to distribute the number of atoms more or less equally
- Try to keep fragments together (*e.g.* don't cut a phenyl in half, …)

These are of course not carved in stone and depending on the application of the current BB, other things need to be considered. It is definitely one of the steps where human input is required.

1 www.mofplus.org/nets/net/bnn

(106:[106]106) **Christian Schön** enquired: Your network topology approach reminds one of the work by Klee,[1,2] a crystallographer from the 1980s, and also, of course, of the work by Friedrichs,[3] LeBail,[4] and Winkler[5] in the 1990s and early 2000s. How does your method compare to their work? A big issue for Klee (and also for Friedrichs, I seem to recall) was the embedding of the network in the three-dimensional space – is this an issue for you? Quite generally, I think the work of the above authors might be helpful when constructing new types of network structures.

1 S. J. Chung, Th. Hahn and W. E. Klee, *Acta Crystallogr. A*, 1984, **40**, 42–50.
2 W. E. Klee, M. Bader and G. Thimm, *Z. Kristallogr.*, 1997, **212**, 553–558.
3 O. D. Friedrichs, A. W. M. Dress, D. H. Huson, J. Klinowski and A. L. Mackay, *Nature*, 1999, **400**, 644–647.
4 *J. Appl. Cryst.*, 2005, **38**, 389–395.
5 B. Winkler, C. J. Pickard, V. Milman and G. Thimm.

Julian Keupp and **Rochus Schmid** replied: In the above works of Klee and also partially in Delgado-Friedrichs' work, their main focus is the generation of topologies in the first place. Herein we use the default embedding (note that we use the words embedding and blueprint synonymously) as obtained from a database (the RCSR) as a starting point and further optimize the cell and vertex positions based on the geometrical needs of the building blocks used. At least in the MOF field, for the embedding there now exists an established methodology to get maximum symmetry embeddings,[1] which restrains the embedding to have edge lengths of 1 and to have a high symmetry space group. This embedding can be computed by Systre (where Delgado-Friedrichs is the main author). Systre is the program used to preprocess topological data for the RCSR database.

The other authors also use topologies as their foundation to generate crystal structures, but for different kinds of systems, and therefore in a different way. Whereas we use larger building blocks that already consist of chemical entities, in other areas smaller units are put onto the vertices of the topology, which can even be single atoms.

1 O. Delgado-Friedrichs, M. D. Foster, M. O'Keeffe, D. M. Proserpio, M. M. J. Treacy, O. M. Yaghi, What do we know about three-periodic nets?, *J. Solid State Chem.*, 2005, **178**, 2533–2554.

(107:[107]107) **Christian Schön** asked: It looks to me that you could draw a straight line (with some scatter) through the correlation plot in Figure 6 (mean AAD *vs.* DE(MOF-FF)). Is this just an accident or is there some real physics behind the "correlation"?

Julian Keupp and **Rochus Schmid** answered: That is indeed the case and we say in the paper that "the relative energies correlate reasonably well with the geometric mean AADs obtained from the topoFF optimization and illustrate its usefulness in order to pre-rank the topologies". The reason here is that the simpler topoFF optimization at the blueprint level corresponds well to the actual atomistic potential energy surface, since the energetics are mainly determined by the deformation of the building blocks, which corresponds well to the summed up average angle deviation (AAD) in topoFF. Only if torsional deformations, for example (not considered in topoFF), start to dominate the energetics, does the topoFF optimization not correlate well anymore with the final atomistic relative energetics.

(108:[108]108) **Caroline Mellot-Draznieks** asked: Since your construction method uses a decoration strategy starting from dimensionless networks such as those enumerated by O'Keeffe, would it be appropriate for the construction or prediction of mixed-linker MOFs possessing, for example, linkers of different sizes and connectivities?

Julian Keupp and **Rochus Schmid** responded: Yes. This is exactly what we have in mind for the next steps. For example, considering the mixed-linker systems from *e.g.* the Shane Telfer group, we could use topoFF to quickly assess which kind of linkers with different lengths or shapes might give rise to stable networks.

(136:[109]109) **Christian Schön** queried: How complete are the sets of networks that you generate? Are there (network-based) structures that you cannot get to by twisting, *etc.*?

Julian Keupp and **Rochus Schmid** replied: We do not generate the networks but use existing databases (like RCSR) as a source, which might not be complete. What we do is use existing networks/topologies to generate "all" potential atomistic structures from these. Note that for molecular fragments – in contrast to zeolites – multiple polymorphs for the same network are possible. Note also that for decreasing symmetry, of course infinite different networks/topologies are possible in principle.

(109:[110]110) **Marcus Neumann** opened the discussion of the paper by German Sastre: You presented a diagram showing a linear increase with time of known zeolite structures. Are these predicted or experimental zeolite structures? How does the increase of distinct predicted structures compare to the increase of known experimental structures?

German Sastre responded: This corresponds to experimental zeolites; the pace is rather constant, about 5–10 new zeolites every year. There are typically millions of predicted zeolites, and in particular several millions are found in two such databases by Treacy-Foster and Deem, respectively. However, how many of these predicted zeolites are feasible? Not all of them, since most of the computer generated zeolites satisfy the criteria of connectivity, as well as some geometrical restraints, which are not enough to guarantee feasibility. Since the conditions of feasibility are not so easy to define, and energy alone is not enough, it becomes difficult to assess how many zeolites are feasible. This is a long-standing and interesting question within the community, to which efforts are being increasingly addressed. A good guide is the known correlation between energy and density, a straight line with a negative slope established by Akporiaye and Price in a remarkable paper in 1989,[1] but this is only for the pure silica (SiO_2) composition. Even for pure silica one wonders how far to the correlation line a zeolite may be in order to disqualify as feasible. But, more importantly, zeolites, initially defined as alumino-silicates, soon expanded into materials where tetrahedral atoms also include P, Ge, B, Mg, Zn, Ga, Be and Ti; and even N and S were found to replace oxygen. With these, and possibly more atoms that can be present, although not of course in any combination, it becomes beyond the current scope to foresee the number of feasible zeolites. Nature is, once more, breathtaking.

[1] Akporiaye and Price, Systematic enumeration of zeolite frameworks, *Zeolites*, 1989, **9**, 23–32.

(110:[111]111) **Sarah Price** said: Your paper concludes that there are 7 new zeolite structures that appear feasible. Zeolites are made by crystallising around a template molecule – have you looked at possible templates for these new structures? Are any experiments planned to make your predicted new zeolites?

German Sastre answered: This is ongoing work, where we are using a newly developed software to search for appropriate templates from a database in order to find one suitable for a zeolite of our choice.

(113:[114]114) **Yi Li** asked: Why do you make a structure with $P1$ symmetry at the beginning and determine the symmetry again in your GA approach?

German Sastre replied: This approach applies to structures that do not qualify as "zeolites" by having a large, but not sufficiently large, fitness. We then try to apply our specific "geometry optimiser", which makes a $P1$ structure, locates oxygen atoms around the Si, and makes a lattice energy minimisation with specifically selected force fields. And, very importantly, this optimisation does allow the volume to change, so that the optimisation is driven from the "non-zeolite" to the "zeolite" realm through the force field without volume restrictions. Then, in a certainly significant number of cases the optimisation is able to generate a new zeolite whose symmetry should be determined since it can be different from the original.

(114:[115]115) **Marcus Neumann** said: You are using a force field for the energy optimization of zeolites. Making a link with earlier presentations, I wonder if a machine learning force field has already been generated for zeolites, and how such a force field would compare to your force field in terms of accuracy and CPU time requirements for energy calculations after parameterization.

German Sastre responded: Machine learning approaches have been used to calculate properties or generate zeolites, relying on geometric features related to local geometry, structure, and porosity. This is opposed to force field approaches which calculate properties explicitly after geometry optimisation, whilst machine learning is focused instead on finding correlations between – in this case – structures and properties. Hence I do not know of any approach which employs machine learning techniques to parameterise a force field for zeolites.

(115:[116]116) **Volker Deringer** commented: Following up on the previous question – I don't know of currently available "machine-learned" force fields specifically for zeolites; the long-range electrostatic interactions are a challenging issue. Note that there are interesting studies on artificial neural network potentials for bulk metal oxides [*e.g.* ref. 1], including fitting neural networks specifically for the long-range part.

1 N. Artrith, T. Morawietz and J. Behler, *Phys. Rev. B*, 2011, **83**, 153101.

(116:[117]117) **Stefanos Konstantinopoulos** asked: A question about force fields: I would like to ask why you use only non-bonded interactions (van der

Waals and electrostatics) to generate zeolite structures. Is there any hindrance to not using bonded interactions?

German Sastre replied: For this particular approach we start with a structure that fails to fully qualify as a zeolite and therefore we need to use force fields capable of as much distortion as possible. “Non-bonding” force fields provide a wider landscape of geometry optimisation. Further, our initial structure will, in many cases, not contain proper bond connectivity, and “bonding” force fields typically preserve the connectivity throughout the geometry optimisation procedure, unlike “non-bonding” force fields.

(119:[118]118) **David McKay** commented: The speaker said bonds, angles and dihedral angles are not included in the force field. I support the speaker in this. By including bonds in the force field, structures found after geometry optimisation are limited to the connectivity of the guess structures, thus limiting the search space explored. In my experience, by allowing the structure to change more during minimisation, fewer structures would later be discarded as invalid and so a more diverse structural landscape would be probed.

(117:[119]119) **Gerit Brandenburg** said: The research group of Jörg Behler has initial work on MOFs, *i.e.* it is indeed possible to generate machine learning potentials for these kind of structures. The corresponding manuscript is in preparation.[1]

[1] Marco Eckhoff and Jörg Behler, Development of a Neural Network Potential for Metal–Organic Frameworks.

German Sastre responded: It is different for MOFs than for zeolites. In MOFs it makes more sense to apply machine learning techniques to parameterise force fields, since MOFs do have much wider chemical variety. Instead, in zeolites specific force fields have been developed for silica, aluminosilicate, silicoaluminophosphate, and silicoaluminogermanate. These four compositions cover a large percentage of the zeolites synthesised and the performance of the existing force fields is sufficiently accurate. Machine learning approaches are more appropriate when we do not know how to link structures and properties. When chemical knowledge is available, an explicit force field is usually more accurate than a machine learning force field. In the near future it is possible that the computer will move towards the human approach and the current trend may become reversed.

(118:[120]120) **Christian Schön** enquired: Concerning the force fields you are using, would it make sense to use/add shell terms to create more bond-like behaviour? Furthermore, regarding the earlier comment/question about symmetries, you might want to employ the KPLOT-code by Rudolf Hundt;[1] the KPLOT code is freely available on the webpage of CRYSTALIMPACT, where a number of algorithms are incorporated that search for symmetries in three-dimensional structures,[2] determine space groups,[3] and compare periodic three-dimensional structures.[4]

1 R. Hundt, *KPLOT Manual – a program for plotting and analysing crystal structures*, ed. J. C. Schön and D. Zagorac, Technicum Scientific Publishing, Stuttgart, 2015.
2 R. Hundt, J. C. Schön, A. Hannemann and M. Jansen, *J. Appl. Cryst.*, 1999, **32**, 413–416.
3 A. Hannemann, R. Hundt, J. C. Schön and M. Jansen, *J. Appl. Cryst.*, 1998, **31**, 922–928.
4 R. Hundt, J. C. Schön and M. Jansen, *J. Appl. Cryst.*, 2006, **39**, 6–16.

German Sastre replied: We have used two core-shell force fields (SLC and SC) but this is only for the oxygen atoms and does not imply bonds between atoms. Using bonds between atoms, either through a core-shell or any other function, would imply a connectivity fixed from the initial structure. Given that the initial structure is not satisfactory, and that is why we want to see if a force field is able to drive it towards a proper zeolite, we need to use force fields capable of as much distortion as possible, and hence "non-bonding" force fields provide a wider landscape of geometry optimisation.

Concerning the algorithms to find symmetry, so far we have used FINDSYM by Harold Stokes,[1] but this does not allow incorporation to analyse big data; and also SPGLIB by Atsushi Togo.[2] None of the approaches used so far are sufficiently robust to work consistently on a high-throughput set where a large diversity of situations can be found. Hence, your suggestion of KPLOT is highly welcome. Thank you.

1 stokes.byu.edu/iso/findsym.php
2 atztogo.github.io/spglib

(801:[121]121) **Alan Hare** communicated: Concerning the small voids we see: can you say why there might exist an exclusion zone forbidding an atom to enter? Years ago, attempts were made – principally by Daniel Kivelson, RIP[1] – to determine void size by the constraint it would impose on the rotation of a nitroxide radical (through its consequential effect on the ESR line-broadening).[2] I am now wondering if an impurity species might nucleate the zeolite, perhaps later vacating it to leave the void.

1 http://www.chem.ucla.edu/news/Kivelson_Obit.html
2 *J. Phys. Chem. B*, 1997, **101**, 8631–8634.

German Sastre communicated in reply: Kivelson's approach gave new insights into how molecules move in liquids, described as translations (in which the molecules move from place to place) and tumbling motions. Templates in zeolites tend to fill the empty voids and this results in significant motion restraints of the template with respect to the liquid phase, not only in translation but also in rotation. The change of the rotational modes of occluded templates in zeolites has recently been explored in our group[1] and sometimes can be used to assess the balance between electrostatic and van der Waals template–zeolite interations. Templates do fill the voids during the synthesis and are only evacuated through calcination, but water molecules may have space in the micropore to partially "solvate" the template and may leave the void under mild conditions. All of this (template + water) contributes to the exclusion zone, but in zeolites most of the exclusion zone (micropore voids) is permanently filled by the organic template.

1 T. Lemishko, J. Simancas, M. Hernández-Rodríguez, M. Jiménez-Ruiz, G. Sastre and F. Rey, *Phys. Chem. Chem. Phys.*, 2016, **18**, 17244–17252.

(120:[122]122) **Doris Braun** opened the discussion of the paper by Christopher Collins: With the FUSE-based probe structure calculations you have shown that you can generate structures with different compositions. Is it possible that different polymorphs are generated within a region of the same composition?

Chris Collins responded: For all of the given probe structure generation runs with FUSE we do retain all of the structures visited (for this work we generated 1 253 298 structures across 42 compositions). So yes, many different polymorphs are generated for a given composition; as for assessing which ones are stable, we would still suffer from all the same issues discussed in the section on crystal structure evaluation.

(123:[124]124) **Volker Deringer** enquired: I'm more generally curious about higher-lying, metastable inorganic polymorphs. There are several others, say for TiO_2, with potentially useful applications (*e.g.* ref. 1) – and at the same time, experimental solid-state chemists are making remarkable progress in developing synthesis routes towards such metastable polymorphs. Aside from the two ICSD entries you show, and brookite, did you observe any other TiO_2 structures? Do you generally see a perspective for using your methodology in that direction?

1 J. Buckeridge *et al.*, *Chem. Mater.*, 2015, **27**, 3844.

Chris Collins answered: We did generate a large number of TiO_2 structures (46 889 that converged), although we did not investigate which ones are synthesisable. This is because our search was set up to find low-energy minima fast and then stop, since our overall objective was to efficiently compute realistic formation energies for the given compositions, not to precisely solve crystal structure(s) and polymorphs. In theory, yes, one could use a method such as FUSE to search for possible polymorphs; to do so we would need to run searches for much longer, although we would still suffer from the same issues as all of the speakers in the discussions on comparing the stability and energy ranking of polymorphs. Currently we have no plans to expand into this area.

(124:[125]125) **Michael Ruggiero** asked: Your method uses individual building blocks to generate crystal structures. How difficult would it be to introduce defects into the models? Can the method be used to find probable vacancy sites, or to determine favourable locations for impurity deposition? Can the method be used for solid solutions?

Chris Collins replied: In principle, yes. Although not mentioned in the paper, as it's a feature we're still working on (for cation vacancies), it is possible to have vacant sites on both the anion and cation lattices. On the anion lattice, this is implemented in the standard sub-module definition, for example one could imagine constructing models of $SrTiO_3$-δ where we would still expect to construct perovskite structures, albeit with oxygens missing from the Ti coordination sphere. For cation vacancies, it is possible to specify a "vacancy" as a species in the initial formula unit; then every time we assemble a sub-module where the cation is labelled as vacancy the site is skipped, thereby creating vacancies in the cation lattice. It is then possible for FUSE to vary the number of vacancies on the lattice on-the-fly.

In practice, I'm not sure that using FUSE would be the most efficient way to study defect chemistry, given that there are a number of groups/software dedicated to looking at the defect chemistry of solids, where the structure you're starting with is pre-defined (examples being SimDope[1] or ChemDASH (currently unpublished)). With FUSE you would be relying on it to first find the crystal structure of interest and locate defects at the same time; I suspect that this would be rather inefficient. Although if you are considering systems where the crystal structure is not known, one could use FUSE first to try to determine the base stoichiometric crystal structure, then follow this up with codes like SimDope to study the potential defect chemistry.

As for broader solid solutions, yes, we can definitely use FUSE for this. Essentially this would be using FUSE as we have in this paper, but in a lower dimensional search space. FUSE would be very well suited for considering solid solutions where the end members are different structures – since there's no way of knowing *a priori* at which point a structure should change between the structure type of each end member, or even if an entirely new structure is formed somewhere in between! An example of this is in our results, if you consider the area of the phase diagram between $(SrTiO_3)_{1-x}$ and $(SrO)_x$, where each end member is a different structure type, with 3 different, stable structures in between, the Ruddlesden–Popper $n = 1$, 2 and 3 phases.

1 L. Enciso-Maldonado, M. S. Dyer, M. D. Jones, M. Li, J. L. Payne, M. J. Pitcher, M. K. Omir, J. B. Claridge, F. Blanc and M. J. Rosseinsky, Computational identification and experimental realisation of lithium vacancy introduction into the olivine $LiMgPO_4$, *Chem. Mater.*, 2015, **27**, 2074–2091 (available from Matthew Dyer, University of Liverpool).

(125:[126]126) **Artem R. Oganov** queried: Key to the efficiency of your method is that you can feed into the input favourable atomic contacts: *e.g.* cations are surrounded by anions and *vice versa*. How useful would this be for intermetallics, for example, where every type of atom can be next to any other type of atom?

Chris Collins responded: The discretisation used in the current version of FUSE is entirely aimed at constructing realistic models for ionic solids, and so it would not be well suited to intermetallics. If we were to look at an intermetallic system (or any other system where anything could neighbour anything else), we would have to design new sub-module motifs and modify the assembly rules to reflect the different chemistry.

(126:[127]127) **Artem R. Oganov** asked: When you put three or four atoms together, how do you choose the angles? Does it matter?

Chris Collins answered: When the sub-modules are initially constructed, the angles on the sub-module are all 90° and with the grid-like placement of ions, the O–M–O angles will all be either 90° or 45°. When the full structure is assembled the unit cell angles are then set according to the lattice type that FUSE is using, which of course will then distort these angles. As to whether or not it matters, we don't know, as this is the only way we've tried assembling the sub-modules/the coordinations of the atoms. My best guess is that it doesn't matter as they will be refined during the geometry optimisation step.

(127:[128]128) **Artem R. Oganov** enquired: What do you do with stoichiometries? Your graph shows that you found many stable stoichiometries, but how do you sample the chemical space?

Chris Collins replied: In most cases, we will design our search pattern based upon two main factors: how many compositions is it reasonable to search with available computing resources and how long is it reasonable to spend doing calculations before going into the lab? For this work, as it was a test case to see how FUSE performs, we initially sampled the $SrO–TiO_2$ line quite densely (20 compositions), and once this worked we then expanded out to include Y_2O_3. We decided that we had time for around 20 additional compositions. So we distributed 20 points evenly across the phase space, and included the two compositions which were not on this evenly distributed grid. This resulted in a total of 42 compositions being sampled in the phase field.

(128:[129]129) **Christian Schön** remarked: Concerning the stacking procedure in FUSE: It looks like you are stacking layers that are electrostatically neutral. But when stacking 3D-blocks like in the AASBU method,[1,2] one can run into the problem that either the blocks or the resulting structures are not charge neutral. Do you ever run into this problem, or how do you avoid that?

1 C. Mellot-Draznieks, J. M. Newsam, A. M. Gorman, C. M. Freeman and G. Ferey, *Angew. Chem., Int. Ed.*, 2000, **112**, 2358.
2 C. Mellot-Draznieks, S. Girard, G. Ferey, J. C. Schön, Z. Cancarevic and M. Jansen, *Chem. Eur. J.*, 2002, **8**, 4103–4113.

Chris Collins responded: The modules that FUSE generates are not required to be electrostatically neutral, it just happens that in the worked example in the paper they were. For example, there would be nothing stopping FUSE generating SrO_2 modules and having non-charge neutral modules does not cause any issues as long as the full structure is charge neutral overall.

When generating structures at random, there can be an issue with structures not being charge neutral. In practice, this is avoided by only choosing the lattice size/shape to correspond to *n* formula units. The code will then create two lists, one of cations and one of anions for *n* formula units, which are then distributed randomly across the sub-modules. The sub-module motifs used are constrained such that the total number of anions used is equal to the number required for *n* formula units.

(129:[130]130) **Christian Schön** asked: What local environments can you actually deal with? For example, if a cation is surrounded by six oxygen atoms, then the cation plus its first coordination sphere by anions is a charged building block, you would need to merge adjacent complex anions to form a real structure. Alternatively, you would have to assign only *e.g.* three of the oxygen atoms to the cation, which makes for strange building blocks. Does this lead to restrictions on the type of building blocks you can employ?

Chris Collins answered: In order to avoid having to merge complexes to form a full 3D structure, as you suggest we work with the sub-modules where only 0–3 anions are assigned to a given cation, which may seem like strange building units to use (I seem to recall that avoiding this very problem was the origin of using the sub-module description used by FUSE). This is (in part) why we have the step to “error check” all of the cation coordination numbers once the full 3D structure is assembled. This also means that the number of building block motifs are restricted, with FUSE generating single atom thick modules. We restricted the sub-modules to be planar, and based upon a grid (as in Figure 1a) this results in only 8 possible motifs.

Error checking of the full 3D structure avoids the sub-modules being assembled in a non-realistic fashion (*e.g.* with all of the anions clustering in one part of the unit cell). However, we currently only check the number of anions in the immediate coordination sphere and do not place any further restrictions on how they are geometrically arranged beyond that imposed by the use of the sub-module motifs.

This method of restricting the cation coordinations doesn’t appear to result in the types of coordination environments in the resulting geometry optimised structures being too restrictive. Since the meeting, I have had the time to analyse the coordination environments of 924 403 of the structures generated in this work, containing 1.445×107 environments. We define our coordination environments by the cation environment’s bond valence sum. We define a pair of ions as bonded when the ion pair contributes over 5% of the formal charge to the cation’s total bond valence sum, as we have done previously.[1] I have included a bar chart, showing the fractional occurrence of the coordination numbers between 0–16 (see Fig. 2). While there are large peaks for each of the elements, all

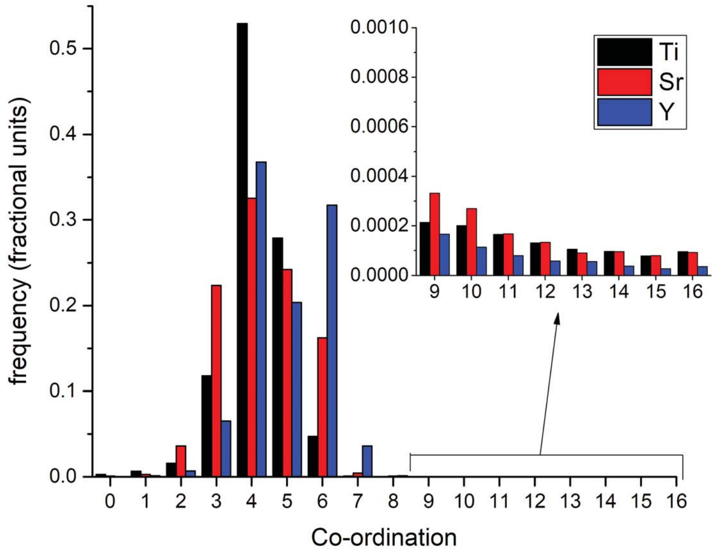


Fig. 2 Fractional occurrence of the coordination numbers between 0–16.

of the coordinations are non-zero (note, though, that some of these structures will be ones which have ended up in non-physical structures).

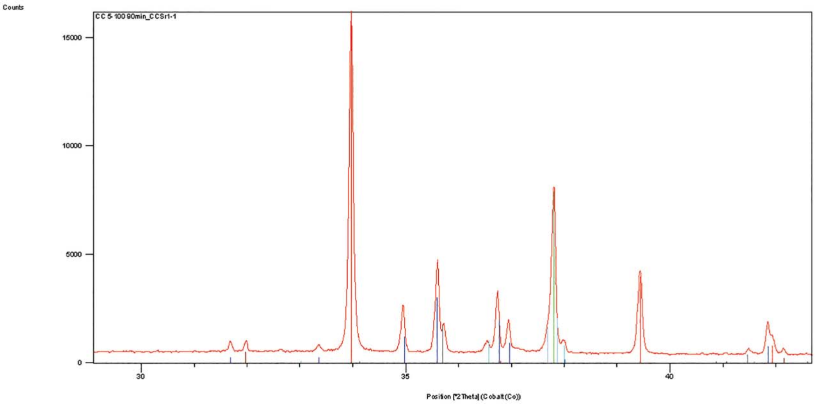
1 C. Collins, M. S. Dyer, M. J. Pitcher, G. F. S. Whitehead, M. Zanella, P. Mandal, J. B. Claridge, G. R. Darling and M. J. Rosseinsky, Accelerated Discovery of Two Crystal Structure Types in a Complex Inorganic Phase Field, *Nature*, 2017, **546**, 280–284 (10.1038/nature22374).

(130:[131]131) **David McKay** queried: On the phase diagram, you said that the white areas, which appear to be below 35 eV/atom, represent stable structures. Why this energy cut-off?

Chris Collins replied: As stated in the paper, this is the threshold we have found previously for locating experimentally realisable compounds. It is quite likely that this is system dependent.

(131:[132]132) **David McKay** asked: How does the result change if this energy cut-off is increased?

Chris Collins responded: If we start to raise the threshold that we consider, the next composition to come under the cut-off is Y_2TiSrO_6 at 48 meV/atom, which is not known to experimentally form a compound. When we tried this in the lab, we



Pattern List:

Compound Name	Chemical Formula
Yttrium Oxide	Y_2O_3
Strontium Yttrium Oxide	SrY_2O_4
Strontium Titanium Oxide	$SrTiO_3$
Yttrium Titanium Oxide	$Y_2Ti_2O_7$
Rutile	TiO_2
Strontium Titanium Oxide	Sr_2TiO_4

Fig. 3 Indexed powder diffraction pattern.

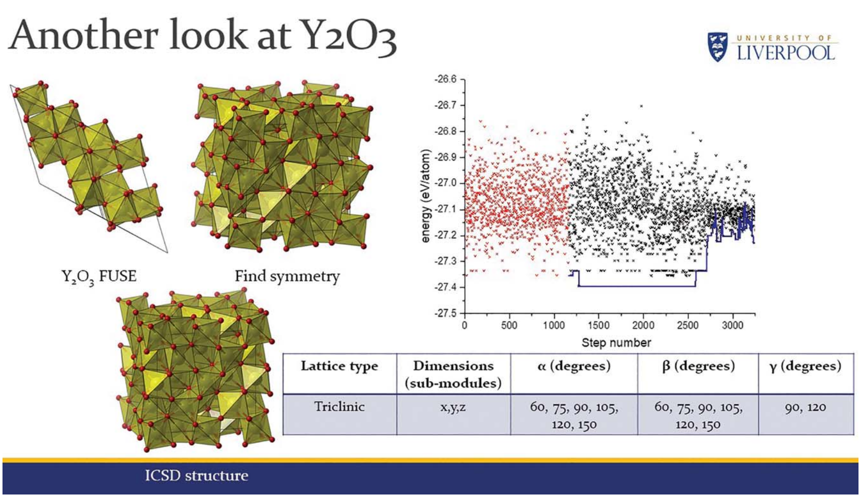
obtained a multi-phase sample containing Y_2O_3, TiO_2, SrY_2O_4, $SrTiO_3$, $Y_2Ti_2O_7$ and Sr_2TiO_4. This confirms that for this phase diagram at least, a cut-off of 35 meV/atom is a reasonable value; however, it is entirely possible that this may vary between phase diagrams. For reference, I have included an indexed powder diffraction pattern (see Fig. 3).

In practice, in the lab we will start synthesis at the lowest energy composition(s) and work up in energy until we obtain samples which only contain mixtures of known phases. So if the phase diagram in the paper were being explored for the first time we would have started with $SrTiO_3$ and finished with Y_2SrTiO_6, experimentally obtaining all of the compounds in the phase diagram.

(132:[133]133) **Sarah Price** remarked: In your paper, you had only found a structure for Y_2O_3 that was 40 meV above the ICSD crystal structure. In your presentation you showed that you had since adapted your procedure so that it now finds the correct structure. Please can you expand on what you needed to change so that this structure was found by your method? Does the updated procedure also find the correct structure for SrY_2O_4?

Chris Collins replied: For reference, I've included the slide from my presentation about Y_2O_3 (see Fig. 4). With the method of constructing unit cells in the paper, it was not possible to construct the correct structure for Y_2O_3 as this would have required 160 ions, 10 above our limit. It was not possible to construct a 40 ion primitive unit cell, so we learnt that we needed to increase the flexibility of the possible unit cell shapes we could create.

The main change we made (we also made a few bug fixes/minor changes to the search routine) was to add in a triclinic lattice approximation. For the triclinic lattice approximation, we allow the *a*, *b* and *c* directions to be of any length, then the α and β angles are randomly chosen from possible values of 60, 75, 90, 105,


Another look at Y2O3

Lattice type	Dimensions (sub-modules)	α (degrees)	β (degrees)	γ (degrees)
Triclinic	x,y,z	60, 75, 90, 105, 120, 150	60, 75, 90, 105, 120, 150	90, 120

Fig. 4 Slide from presentation about Y_2O_3.

120 and 150 degrees. The γ angle is still chosen to be either 90 or 120 degrees. We retained this restriction, because the choice of γ governs what rules are in use to stack modules along *c*, and if we deviated from these values, it is not so easy to automate the stacking.

These changes allowed FUSE to create a 40 ion cell, which was then optimised in GULP into a primitive structure for Y_2O_3. Running the structure through FINDSYM yields the body-centred *Ia*-3 structure. As of yet we have still been unable to obtain the experimental crystal structure of SrY_2O_4 and are currently investigating why. Currently our best result has yielded a structure within 5 meV/atom of the ICSD structure (*vs.* 9 quoted in the paper), and our hypothesis is that there are a lot of low-energy minima in the system (within β10 meV/atom of the ICSD structure), so it is taking far longer to locate the ICSD structure than at the other compositions within our test phase diagram.

(133:[134]134) **Sarah Price** asked: Is this a matter of symmetry that's using a triclinic primitive cell to expand to a larger structure with higher symmetry?

Chris Collins answered: The use of symmetry for tidying up structures after we have completed our search using FUSE is important for arriving at experimental structures. In most cases in our compositional phase diagram, the "raw" structure from FUSE is an odd shape/size in *P*1 symmetry, which FINDSYM then refines to a higher symmetry. Two notable examples of this are for Y_2O_3 and $Y_2Ti_2O_7$: in each case we find a smaller, primitive unit cell, from which FINDSYM then finds the correct, higher symmetry cells. It is also possible, however, to make a larger supercell, which is then reduced to a smaller, higher symmetry cell; this often happens with $SrTiO_3$, for example.

It should be noted, however, that we do not impose any internal symmetry on the structures as they are constructed; if anything, we go out of our way to assemble asymmetric structures by "rattling" all of the atoms off their starting positions, which can then relax into a higher symmetry structure if this is favourable.

(816:[135]135) **Alan Hare** communicated: In growing the lattice – besides doubling the structure along a random axis – would there also be value in creating a supercell other than randomly, from as many units as are needed to retain the shape of the unit cell (so for example in the cubic approximation having dimensions 3×, 5×, *et seq.*)? Could you please elaborate the triclinic approximation (as mentioned prior to the presentation)? What internal symmetry rules are imposed, and can they be tested by treating the monoclinic as a special case?

Chris Collins communicated in reply: In practice there is no reason why we could not allow the unit cell to grow isotopically; as you suggest, it's just something that we have not tried. As for how much value there is to trying this, I simply do not know; I suspect it is something we would have to try and find out! For the triclinic lattice approximation, we allow the *a*, *b* and *c* directions to be of any length, then the α and β angles are randomly chosen from the possible values of 60, 75, 90, 105, 120 and 150 degrees, and the γ angle is still chosen to be either 90 or 120 degrees; this restriction was left in place, as the choice of γ governs what rules are in use to stack modules along *c*.

As for internal symmetry, we do not impose any, so all of the lattice approximations are purely geometric. It was created this way because if we impose internal symmetry (for example by imposing the lowest symmetry space group for each lattice type), then I suspect this would make the construction too restrictive and make it difficult for FUSE to discover crystal structure types with disorder present, for example perovskites with cation disorder.

(134:[136]136) **Kim Jelfs** opened a general discussion of the papers by Julian Keupp, German Sastre and Christopher Collins: Do you have any intention of calculating the properties of the structures you have found? Could you bias your search towards structures with interesting properties? Or investigate how the energy relates to the material's properties?

Chris Collins answered: At the present time, we are not looking to make property predictions for structures we have found, especially since for oxides (our main area of interest), the properties can be very sensitive to the structure. As our research is focused on synthetic chemistry, our focus is to use calculations to lead us to new materials. Once we have obtained them and experimentally verified/solved their crystal structures, we can compute/measure properties. From this point we could then use calculations to predict how to best modify them in order to improve their functionality. In the future, however, as we become more confident in the accuracy of our structural predictions, we may well revise this view! This is just our philosophy; there is no reason why one could not combine a method like FUSE with property prediction, especially if the properties are cheap to compute and one is confident enough in the structural predictions; one could very easily bias a search towards specific properties.

(138:[137]137) **Claire Adjiman** asked: To what extent can one use the desired properties of the MOF, rather than the desired topology, to guide the search?

Julian Keupp answered: Our method (reverse topological approach, RTA, with or without the blueprint optimization with topoFF) aims at a complete generation of all the possible structures by enumerating topologies. The MOF-FF force field energy is used for ranking the stability of the structures. At this point, properties like surface area can also be determined and a systematic search for certain key properties can be performed. However, at the level of structure generation it cannot be used to guide the search.

(135:[138]138) **Christian Schön** addressed Julian Keupp and German Sastre: Can your methods generate interpenetrating networks?

Julian Keupp and **Rochus Schmid** responded: Yes, of course. We can simply interpenetrate (duplicate and shift) the blueprint and start from there, or use the more general approach reported by Baburin.[1] The difficulty is ranking the energies. An interpenetrated system is usually more stable (due to dispersive interactions) than the non-interpenetrated system in the gas phase. However, to predict whether interpenetration is formed, one would need to include solvent effects present in the solvothermal synthesis conditions. This is possible in principle but up to now we have just used simple ideas like a fixed stabilization energy per accessible surface area to approximate this effect.

1 I. A. Baburin, On the group-theoretical approach to the study of interpenetrating nets, *Acta Crystallogr. A*, 2016, **72**, 366–375.

German Sastre responded: So far our methods have not generated interpenetrated structures. Indeed, interpenetrated structures require that not all atoms in the unit cell are part of the same connectivity tree. Our algorithm is sufficiently general to allow this.

(137:[139]139) **Christian Schön** asked: Is there any chemical reason why you would not get interpenetration in zeolites? Some MOFs look like they surely have sufficiently large pores, and large pore zeolites should be feasible and kinetically stable once they have been generated.

German Sastre replied: Although catenated metal–organic frameworks are common, this is not the case for zeolites. The catenation requires, geometrically speaking, a large pore in comparison to the thickness of the framework pillars, and this is not usually found in zeolites. I would expect that zeolites containing rings of at least 16 tetrahedral atoms would be necessary, and catenation also seems to appear more frequently in structures with monodimensional pores. However, zeolites containing extra-large pores do also frequently contain smaller pores in perpendicular directions. From the – currently – 11 zeolites containing such large pores, only two of them (IFO and VFI[1,2]) have a monodimensional channel. These would be my suggestion to try to observe catenation in zeolites; however the structural arguments above explain why this has yet not been observed, to the best of my knowledge.

1 www.iza-structure.org/IZA-SC/framework.php?STC=IFO
2 www.iza-structure.org/IZA-SC/framework.php?STC=VFI

(139:[140]140) **Marcus Neumann** remarked: In organic CSP we struggle to get from the knowledge of the crystal energy landscape to the crystallization of new forms. Do you know of success stories for rational crystallization experiment design in inorganic CSP?

Julian Keupp and **Rochus Schmid** answered: In the context of our topoFF method within RTA and the prediction of “soft” framework materials, things are probably quite different from classic inorganic CSP of hard materials like oxides, *etc.* To some extent, the notion of reticular chemistry[1] essentially means to design new systems in a rational way, based on the knowledge of the shape of the building blocks.

1 O. M. Yaghi, M. O’Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi and J. Kim, Reticular synthesis and the design of new materials, *Nature*, 2003, **423**, 705–714.

Chris Collins answered: I know that others have commented on this and highlighted examples where research groups have been able to identify exact crystal structures of new compounds which have subsequently been synthesised. For our group we deem a “success” to be calculations which lead us to the synthesis of new compounds. As stated in our paper, our aim is primarily to be able to predict new chemical compositions; we do not explicitly aim to predict the crystal structure. As we are looking to synthesise compounds with complex compositions, we do not believe it is currently pragmatic to attempt to perfectly predict structures, hence our use of probe structures.

Providing a compositional target for synthesis sufficiently accelerates the discovery process, since finding the correct compositional regions to study is often an extremely difficult step in large phase spaces. Crystallography of ionic solids is currently good enough that experimental structure determination after computational composition prediction is currently the most pragmatic solution. An example of our philosophy was our computational composition exploration of the Y_3^+–Sr_2^+–Ca_2^+–Ga_3^+–O_2^- phase diagram,[1] where synthesis around a target composition resulted in the synthetic identification of two new compounds in under 40 samples, both with new crystal structures, neither of which we would expect to be able to predict *a priori* at the current time, given both their structural complexity and disorder.

[1] C. Collins, M. S. Dyer, M. J. Pitcher, G. F. S. Whitehead, M. Zanella, P. Mandal, J. B. Claridge, G. R. Darling and M. J. Rosseinsky, Accelerated Discovery of Two Crystal Structure Types in a Complex Inorganic Phase Field, *Nature*, 2017, **546**, 280–284 (10.1038/nature22374).

German Sastre answered: The stabilities of $(ZnO)_x$ nanoclusters could be an example, although they are not crystals but rather 0-D materials. Zeolites are not an example since here the challenge is not to know the crystal structure of SiO_2, unlike organic drugs where an important challenge is identifying the stable crystalline structures of each compound. Boron oxides, with a paper in this *Faraday Discussion*, are a good example of CSP inorganic materials. The structures of perovskites could be another example. Tungsten borides are a recent and elegant example where the structures and stoichiometries that are stable at different P,T conditions have been elegantly predicted in Artem Oganov's group.

(140:[141]141) **Christian Schön** enquired: In the field of inorganic compounds, there are two areas where there have been quite a number of successful syntheses after(!) prediction. One is high-pressure structure prediction: see *e.g.* work by Oganov *et al.* on various high-pressure structures,[1] or work by Cancarevic *et al.* on alkali metal sulfides.[2–9] The other is the synthesis of metastable variants in systems like alkali halides,[10–12] where difficult synthesis routes like the low-temperature atom-beam-deposition method[13] were needed to access these modifications. Nevertheless, the synthesis methods in solid state chemistry are noticeably less developed than those in organic molecular chemistry regarding the targeted synthesis of metastable compounds, which helps to explain why for many years the polymorph richness of bulk inorganic compounds has been gravely underestimated.

[1] A. R. Oganov, C. W. Glass and S. Ono, *Earth Planet. Sci. Lett.*, 2006, **241**, 95–103.
[2] J. C. Schön, M. A. C. Wevers and M. Jansen, *J. Mater. Chem.*, 2001, **11**, 69.
[3] J. C. Schön, Z. Cancarevic and M. Jansen, *J. Chem. Phys.*, 2004, **121**, 2289–2304.
[4] Z. Cancarevic, J. C. Schön and M. Jansen, *Phys. Rev. B*, 2006, **73**, 224114.
[5] A. Grzechnik, A. Vegas, K. Syassen, I. Loa, M. Hanfland and M. Jansen, *J. Solid State Chem.*, 2000, **154**, 603–611.

6 A. Vegas, A. Grzechnik, K. Syassen, I. Loa, M. Hanfland and M. Jansen, *Acta Crystallogr. B*, 2001, **57**, 151–156.
7 A. Vegas, A. Grzechnik, M. Hanfland, C. Mühle and M. Jansen, *Solid State Sci.*, 2002, **4**, 1077–1081.
8 D. Santamaria-Perez, A. Vegas, C. Mühle and M. Jansen, *Acta Crystallogr. B*, 2011, **67**, 109–115.
9 D. Santamaria-Perez, A. Vegas, C. Mühle and M. Jansen, *J. Chem. Phys.*, 2011, **135**, 054511.
10 Z. Cancarevic, J. C. Schön and M. Jansen, *Chem. Asian J.*, 2008, **3**, 561–572.
11 Y. Liebold-Ribeiro, D. Fischer and M. Jansen, *Angew. Chem.*, 2008, **47**, 4428–4431.
12 A. Bach, D. Fischer and M. Jansen, *Z. Anorg. Allg. Chem.*, 2009, **635**, 2406–2409.
13 D. Fischer and M. Jansen, *J. Am. Chem. Soc.*, 2002, **124**, 3488–3489.

Chris Collins responded: Thank you for your comments regarding successes of crystal structure prediction and subsequent synthesis.

(146:[142]142) **Marcus Neumann** addressed Chris Collins, German Sastre and Artem R. Oganov: This question is about the economics and social dynamics of CSP. A previous comment insinuated that there is not much economic/industrial value in inorganic CSP. In organic CSP, the prediction of late-appearing or disappearing polymorph events for pharmaceutical compounds provides a strong economic incentive. At this meeting, inorganic CSP is very strongly represented, which seems to contradict the supposed lesser economic value of inorganic CSP. Where do you see the economic high-value applications of inorganic CSP?

Chris Collins replied: The value of inorganic CSP is in the discovery of new materials. In order to increase the performance of devices, we constantly need to develop new materials from which to construct them. Previously, new materials came from purely synthetic research programs, where a number of revolutionary materials have been discovered, such as the discovery of $LiCoO_2$ (subsequently doped with Mn and Ni), which enabled the wide-scale use of the Li-ion batteries we now take for granted in our everyday lives. Inorganic CSP has the potential to dramatically accelerate the pace at which we are able to discover new materials. So the economic value (from our perspective) is in driving this initial discovery. As for the application, it should be able to contribute to any area of inorganic chemistry where materials discovery is the bottleneck preventing technological progress.

Artem R. Oganov replied: Inorganic CSP is essential for computational materials discovery. This is the future of numerous technologies, where the lack of new materials is the bottleneck. Let me just mention the acute need in new thermoelectric, photovoltaic, thermoresistant, superconducting materials, and the relentless search for novel materials for Li-ion batteries, fuel cells, *etc.* The economic and societal impact is, of course, huge.

German Sastre replied: Inorganic CSP also has an important economic impact in the case of zeolites. They were introduced in the late sixties by Plank and Rosinski, working at Socony/Mobil, in fluid catalytic oil cracking, since they needed catalysts with controlled pores of molecular dimensions in order to obtain higher selectivity and less coking than the previously used amorphous silica alumina. However, knowing the crystal structure is not necessary in order to patent the use of zeolites as catalysts or selective adsorbents. Hence the many industrial applications of zeolites is not a reason to justify the economic impact of

CSP in this family of inorganic materials. However, knowing the structure is key to linking microporosity with properties and ultimately with industrial applications. In the last two decades, computational techniques have had the goal of being able to predict the conditions for the synthesis of new zeolites with specific properties (selected by design). This has only been claimed in a few cases so far,[1] but it is likely that this could be achieved soon in a more systematic and effective way.

[1] T. M. Davis, A. T. Liu, C. M. Lew, D. Xie, A. I. Benin, S. Elomari, S. I. Zones and M. W. Deem, *Chem. Mater.*, 2016, **28**, 708–711, DOI: 10.1021/acs.chemmater.5b04578.

(141:[143]143) **Michael Ruggiero** asked: How important is the role of the solvent for predicting the right structure, *i.e.* structures stabilised by trapped solvent? How likely is your technique to incorrectly predict the experimentally observed structures due to solvent effects or captured gases? Can this be overcome computationally?

German Sastre answered: This is a very good question. The role of the solvent is important but it has usually been neglected in computer simulations regarding the synthesis of zeolites. Usually the solvent is water, with the hydrothermal synthesis being the most employed method of synthesis, but there is also the ionothermal method, using ionic liquids, where the solvent is also the template at the same time. Focusing on water, which is by far the most commonly employed solvent, water molecules alone (without organic template molecules) can certainly stabilise small cavities, although most of these molecules will be evacuated after the usual calcination, whose aim is to remove the template in order to activate the material by making the pores accessible. Water can also be coordinated to template molecules in large cavities. Water can also be responsible for the presence of internal silanol groups forming an internal connectivity defect and clearly affecting the stability of the final material. Water is also responsible, when the synthesis is at $pH > 7$ and when using cationic templates, for the presence of OH^- anions that neutralise the template and become stabilised in – presumably – small cavities. Finally, the commonly employed stabilisation of the material by zeolite–template interactions should also contain another term, usually neglected, for the desolvation energy between the template and the solvent (water molecules). A few of these topics are discussed in a previous publication.[1]

[1] Y. G. Bushuev and G. Sastre, *J. Phys. Chem. C*, 2009, **113**, 10877–10886, DOI: 10.1021/jp9013306.

(143:[144]144) **Sten Nilsson Lill** queried: How flexible are these MOFs and zeolites? Do they breathe? Has this been studied with any molecular dynamics?

Julian Keupp and **Rochus Schmid** responded: In general MOFs can breathe (*e.g.* MIL-53), and this has also been investigated extensively by MD simulations. If you are referring to the particular MOFs we constructed here with topoFF, we have not investigated that in the context of the current paper.

German Sastre responded: Breathing is a kind of long-amplitude vibration that can be attributed to MOFs, but not to zeolites, whose vibrations are typically narrower. This also implies that zeolites show much stronger thermodynamic stability and are better suited for those catalytic processes taking place at temperatures higher than, say, 300 °C. The flexibility of zeolites is, nevertheless, a very important process that often contributes to enhancing their properties. This has been explored since the pioneering work by Deem, Creighton and Newsam in 1992.[1] The flexibility of zeolites contributes to a certain framework deformation when molecules of similar micropore size diffuse, resulting in a minimisation of attractive van der Waals interactions, and enhancing the diffusivity. This is one of the foundations of the concept of enhanced diffusivity, introduced by Eric Derouane, which plays a crucial role in the choice of specific zeolites for selective adsorption processes.

1 M. W. Deem, J. M. Newsam and J. A. Creighton, Fluctuations in Zeolite Aperture Dimensions Simulated by Crystal Dynamics, *J. Am. Chem. Soc.*, 1992, **114**, 7198–7207.

(145:[145]145) **Scott Woodley** remarked: One of the problems with using interatomic potentials is that you may find a suitable set for describing the framework and another set for the template molecules, but not necessarily the set that also includes the interaction between the framework and the template. In your talk, you kindly referenced a number of my publications in *Phys. Chem. Chem. Phys.*[1–3] These papers report how I have adapted a genetic, or evolutionary approach for crystal structure prediction (reference 20 in your paper[4]) and demonstrated how to predict (starting from an initial population of random structures) the framework of microporous structures where the dimensions of the volume within the unit cell that the template molecules occupy are assumed to be known (from experimental measurements) or targeted (you want to predict the framework that has these microporous voids or channels). One of the advantages of any of these approaches is that the interatomic potentials for the template molecules are not required. If a new framework is predicted and a feasible template molecule is then targeted, one could of course then use the code developed by Lewis *et al.*,[5] which uses concepts originally applied in drug design where molecules are "grown" computationally to fit a receptor site. In template design, organic cations are grown to match a target microporous structure. Their software was used to find a template for the microporous aluminophosphate DAF-5. The predicted template succeeded in synthesizing the material rapidly as a pure phase, whereas previous lengthy syntheses had produced impure materials. As a bonus, the predicted site of the template within the host material was verified using microcrystalline single-crystal diffraction techniques.[6]

1 S. M. Woodley, P. D. Battle, J. D. Gale and C. R. A. Catlow, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1815–1822.
2 S. M. Woodley, *Phys. Chem. Chem. Phys.*, 2004, **6**, 1823–1829.
3 S. M. Woodley, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1070–1077.
4 S. M. Woodley, P. D. Battle, J. D. Gale and C. R. A. Catlow, *Phys. Chem. Chem. Phys.*, 1999, **1**, 2535–2542.
5 D. W. Lewis, C. R. A. Catlow, J. M. Thomas, D. J. Willock and G. J. Hutchings, De novo design of structure-directing agents for the synthesis of microporous solids, *Nature*, 1996, **382**, 604–606.
6 G. Sankar *et al.*, Structure of templated microcrystalline DAF-5 ($Co_{0.28}Al_{0.72}PO_4C_{10}H_{20}N_2$) determined by synchrotron-based diffraction methods, *Chem. Commun.*, 1998, **1**, 117–118.

German Sastre replied: The force field necessary to describe the template–zeolite interactions can (and should) be tested by comparing to adsorption energies obtained experimentally. Since this force field is independent of the two other force fields that describe the zeolite framework and the template, it should be possible to select a good choice of three force fields (zeolite, template, and template–zeolite) that describe the system accurately. A good choice of force fields would improve the success of the strategy that you describe in the second part of your comment, where software such as that by Lewis and Willock can be employed in order to "build" the template inside the micropore. Regarding the first part of your comment, your approach based on genetic algorithms is certainly able to predict structures containing predefined voids (mimicking the space filled by the template) without the need for force fields. One advantage is that this is a faster algorithm in terms of computer time and, when big data is at play, as is the case, this is a crucial advantage.

(147:[147]147) **Virginia Burger** asked: Zeolites are often found in nature with contaminants and can take thousands to millions of years to crystallize in nature. In the lab, are contaminants/impurities considered when synthesizing or predicting zeolites?

Julian Keupp and **Rochus Schmid** responded: Within MOFs, similar to zeolites, defects and impurities can dominate the properties (the surface must also be considered as a defect). Theoretical simulations still mostly focus on the ideal system in periodic boundary conditions, but studies of defects[1] and domains[2] (as well as surfaces[3]) have started to appear.

1 J. P. Dürholt, J. Keupp and R. Schmid, The Impact of Mesopores on the Mechanical Stability of HKUST-1: A Multiscale Investigation, *Eur. J. Inorg. Chem.*, 2016, 4517–4523.
2 A. K. Cheetham, T. D. Bennett, F.-X. Coudert and A. L. Goodwin, Defects and disorder in metal–organic frameworks, *Dalton Trans.*, 2016, **45**, 4113–4126.
3 S. Amirjalayer, M. Tafipolsky and R. Schmid, Surface Termination of the Metal–Organic Framework HKUST-1: A Theoretical Investigation, *J. Phys. Chem. Lett.*, 2014, **5**, 3206–3210.

German Sastre responded: We could consider "contaminants" in two senses. In a broad sense, zeolites do allow a very large range of chemical compositions that depend on each particular structure. The zeolites initially defined as "aluminosilicates" are now also called "zeotypes" in order to express other chemical compositions, where tetrahedral atoms also include phosphorus, germanium, titanium, boron, magnesium, zinc, gallium and beryllium. Oxygen can also be replaced by sulfur and nitrogen. The presence of these atoms has an important effect that drives the synthesis towards different structures than when they are absent. In a more strict sense, it has been reported that several synthetic procedures cannot be reproduced when the source of silica or alumina is changed, depending on the manufacturer, or even depending on the batch. It is often not clear in which ways impurities drive the synthesis towards a particular zeolite structure. Finally, an active research topic is "seeding". The presence of certain small crystallites of a zeolite can somehow influence the synthesis towards the appearance of another zeolite, not necessarily the one that has been seeded. Crystallites can dissolve into specific building units that can then be the starting point for a new structure. Finding rules and correlations between the seeding and the final outcome is an active topic of research.

(148:[148]148) **Sarah Price** remarked: I believe that zeolites are very important industrial materials, and that there was patent litigation on whether a zeolite contained Al or not, whose financial implications were very significant relative to those of pharmaceutical polymorphs.

German Sastre replied: ZSM-5 zeolite was patented by researchers at Mobil Oil Co. in 1972[1] as an aluminosilicate zeolite with a range of Al contents and without explicit reference to a pure silica phase that was later patented by researchers at Union Carbide in 1977.[2] A short notice in *Nature* covered the news.[3] Rather than the final legal result, it is important to note how minimal details on the composition can have an enormous impact in the industrial arena. It is important to investigate new ways to synthesise relevant zeolites and circumvent patent protections. An added aspect is that the structure of ZSM-5 was not solved (or published) until 1978 by Kokotailo, Lawton, Olson and Meier,[4] hence the structure is not necessary to patent the material.

1 ExxonMobil Oil Corp., *US Pat.*, US3702886A, 1972.
2 Union Carbide Corp., *US Pat.*, US4061724A, 1977.
3 S. Budiansky, *Nature*, 1982, **300**, 309.
4 G. T. Kokotailo *et al.*, *Nature*, 1978, **272**, 437–438.

Conflicts of interest

Asbjoern Burow and Marcus Neumann work for a company that develops software for organic crystal structure prediction. There are no other conflicts to declare.