Sequence-Dependent Materials

*Published as part of the Accounts of Chemical Research special issue* "*Holy Grails in Chemistry*".

Thomas M. Osborn Popp and Omar M. Yaghi*

Department of Chemistry, Kavli Energy NanoSciences Institute at Berkeley, and Berkeley Global Science Institute, University of California—Berkeley, Berkeley, California 94720, United States

Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

ABSTRACT: Sequence-dependent materials are a class of materials in which a compositionally aperiodic apportionment of functional groups leads to properties where the whole performs better than the sum of the parts. Here, we discuss what defines a sequence-dependent material, and how the concept can be realized in crystals of extended structures such as metal−organic frameworks.
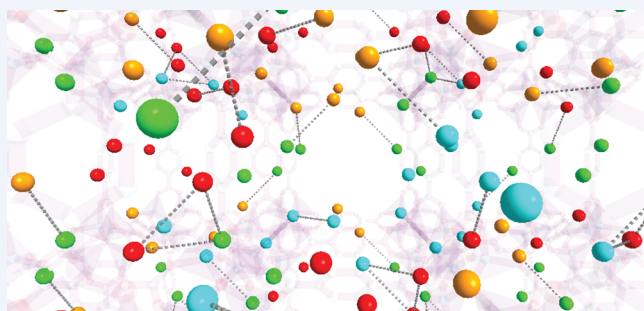
An outstanding objective in chemistry is to synthesize complex solids capable of new and multiple functions beyond the capabilities of current materials. The achievement of this objective, to date, has been impeded by the fact that the introduction of complexity within materials often leads to a loss of synthetic control, and frequently results in materials that are challenging to characterize. We propose materials in which complexity is achieved in the form of rich chemical sequences of varying composition, but without losing the order of the underlying structure, which we call sequence-dependent materials. To achieve this aim, these materials should have an ordered backbone onto which functionalities can be covalently appended, and the spatial arrangement of these functionalities comprises a sequence. In such a construct, the backbone provides a handle on characterization despite the introduction of complexity. The traditional classification of natural and synthetic materials as crystalline or amorphous, depending on their structural periodicity, although useful, is insufficient in the case of such sequence-dependent materials, because it is focused on structure rather than composition. To appreciate sequence-dependent materials, it would be useful to classify materials based in a new way, based on the material's composition and how this composition is distributed through space. We call this the spatial composition of a material. It is the critical factor in the characterization of sequence-dependent materials, and also in how precisely they are designed. In this contribution, we clarify the concept of spatial composition and how it leads to the classification of materials as either compositionally periodic or aperiodic. We then use these terms to discuss sequence-dependent materials and the path toward their development.

The concept of spatial composition may be thought of as a position-dependent molecular formula. In a sequence-dependent material, the local molecular formula varies in space to yield a unique sequence of chemical groups. We can look to

biomolecules such as DNA for inspiration in the design and synthesis sequence-dependent materials. DNA is a one-dimensional material containing four nucleotide units randomly apportioned spatially along its length. The molecular formula of a strand of DNA varies along its length, with each "rung" on the double-helical ladder potentially containing different molecules than the rung above or below it. However, this spatially varying molecular formula contains information: triplets of nucleotides in sequence, known as codons, each coding for one of 21 amino acids. The composition of DNA may thus be thought of as aperiodic. Even before the true biological role of DNA had been fully appreciated, Erwin Schrödinger, in 1944, referred to a solid with spatially varying composition as an "aperiodic crystal," and predicted that such a material would be necessary for the chemistry of life to function.[1]

In order to develop synthetic materials with an aperiodic spatial composition, we need to take a closer look at what makes up biological sequence-dependent materials like DNA. Although the nucleotide base pairs may differ from one part of the double helix to the next, the composition of the sugar−phosphate backbone is conserved along the entire molecule's length. Figure 1 shows the decomposition of DNA into two parts: the first is the compositionally periodic backbone, where the molecular formula repeats, and the second is the compositionally aperiodic nucleotide sequence, where the molecular formula does not necessarily repeat along the length of the double-helix. Compositional periodicity and aperiodicity are concepts that are distinct and decoupled from structural periodicity and aperiodicity. The double-helical backbone of DNA can exist in a twisted and bent state, meaning the strand is not always structurally periodic in a crystallographic sense, yet the molecular
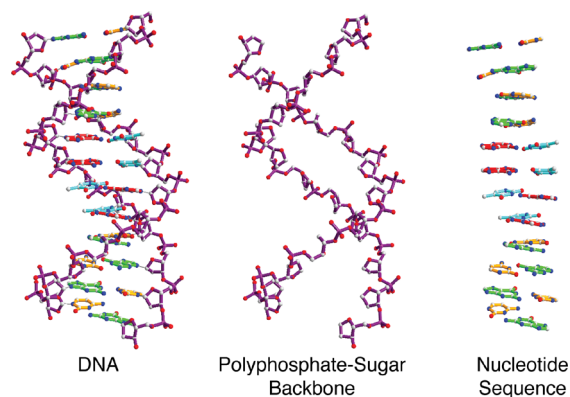
Figure 1. Fragment of DNA decomposed into its compositional parts: the compositionally periodic sugar−phosphate backbone (violet) and compositionally aperiodic sequence of nucleobases (adenine, red; thymine, blue; guanine, green; cytosine, yellow). The sequence of base pairs encodes chemical information, and its aperiodicity is decoupled from the double-helical topology.

formula of the backbone does repeat, making the backbone always compositionally periodic. It is not necessarily true, however, that the base pairs will be repeated periodically in sequence, making this part compositionally aperiodic. Using this construct, where a compositionally periodic backbone serves as the foundation for appending aperiodic chemical sequences, would seem to be critical for making complex materials that are also highly functional. Indeed, the ability for DNA to store information with high fidelity is made possible by the joining of compositionally aperiodic and periodic parts, making an overall compositionally aperiodic material.

Compositional aperiodicity is crucial for the functions of DNA and other biomolecules like proteins. If such biomolecules were compositionally periodic, they would not be capable of the functions they perform. In this context, it seems that even the most useful compositionally periodic synthetic, crystalline materials are truly lacking in their potential. In our discussion, we are focusing on covalently linked aperiodicity onto a backbone, as we would like to be able to design and characterize these sequence-dependent materials with facility. Yet compositional aperiodicity is a powerful property even in materials without a clear distinction between an aperiodic sequence and a backbone, as in the case of doped semiconductors where the

spatial arrangement of dopants is directly related to the performance.

A compositionally periodic backbone is essential to making sequence-dependent materials. It acts to limit the large number of conformational degrees of freedom incurred when appending multiple chemical groups. Though vast, the different types of sequence and functionality possible from the combination of different chemical groups are ultimately limited by the backbone. In designing an ideal backbone, it would be useful to impose a strictly defined size, geometry, and connectivity to the component molecular building units. These constraints enable the formation of a metrically ordered backbone onto which aperiodic chemical functionality may be adorned without perturbing the underlying structure. Additionally, the use of an ordered backbone minimizes structural distortions or variations upon the introduction of a variety of chemical groups. The best structural backbones would use molecular building units restricted to only specific coordination geometry, enabling synthesis of the backbone by design rather than by trial-and-error.

Metal−organic frameworks (MOFs) are an excellent starting point for the synthesis of sequence-dependent materials. They are constructed from organic linkers and metal-oxide clusters, also known as secondary building units (SBUs), to form crystalline, porous solids. Their crystalline structures trace out topological backbones onto which chemical functionalities can be introduced, either as functional groups on the organic linkers, or through the choice of metal cluster. The use of a single type of organic linker or metal SBU in a MOF results in a compositionally periodic material with homogeneous chemical and physical properties. However, one can use multiple types of organic linker to make a compositionally aperiodic MOF with the same topology as a compositionally periodic MOF. This has been shown in the synthesis of multivariate MOFs (MTV-MOFs).[2−4] Like DNA, MTV-MOFs can be thought of as being comprised of a compositionally periodic part and a compositionally aperiodic part (Figure 2), a useful theoretical deconstruction which has been used previously in the analysis of MTV-MOFs.[5] In several cases, MTV-MOFs have successfully been synthesized using two or more types of linker, often showing new bulk properties such as improved selectivity or uptake for the adsorption of gases. While there may be unique sequences in these compositionally aperiodic MTV-MOFs that lead to the improved gas sorption properties, these specific sequences are still unknown, and at best we can say that the compositional aperiodicity in MTV-MOFs
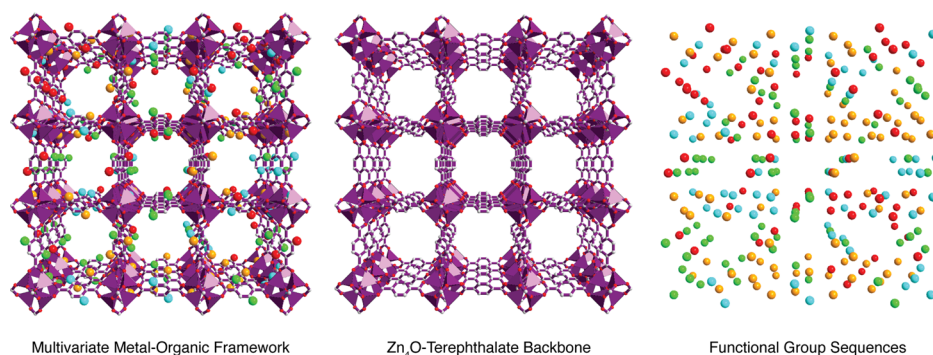
Figure 2. MTV-MOF-5 decomposed into a compositionally and structurally periodic backbone (violet) and an aperiodic decoration of the backbone by four different hypothetical functional groups (red, blue, green, and yellow spheres). The exact specification of a sequence in this material would need to be defined according to the desired application, perhaps as a sequence of groups along one axis of a MOF pore channel, or a sequence of two or more groups on adjacent orthogonal linkers.

may result in synergistic bulk properties. More recently, some MTV-MOFs have been synthesized exhibiting the first glimpse of sequence-dependent properties implemented by design. Postsynthetic covalent reactions were applied to IRMOF-74-III in order to approximate an enzyme active site inside the MOF pore.[6] The final MTV-MOF showed activity for the catalytic process performed by the target enzyme site, while the unmodified parent material showed no such activity. Additionally, other strategies, such as linker insertion and linker exchange, promise to move us closer toward the synthesis of sequence-dependent materials using MTV-MOFs.[7]

Though the prospect of sequence-dependent materials is alluring, successful synthesis of specific sequences is another matter. What defines a chemically relevant sequence depends on the application: a sequence could be defined as a few side chains in a geometry similar to an enzyme active site. However, a sequence could also be chosen as a number of side chains in a pore channel designed to perform multiple functions on a molecule as it passes through. These useful sequences might be found only sparingly throughout an otherwise randomly apportioned material, like discovering a meaningful sentence in Borges' Library of Babel, or perhaps the unique apportionment of functional groups across a single crystal is, in fact, the sequence itself. The designation of sequence and how it is chosen and implemented must be specific to the desired function.

The use of a compositionally periodic backbone to support these sequences should be achievable on the synthesis front-end, but in order to introduce sequences by design and to know what sequences are actually being made, we need appropriate characterization techniques. Sequence-dependent materials raise new questions that cannot be completely answered by the current set of material characterization methods, which puts a bottleneck on the progress toward the synthesis and design of sequence-dependent materials like MTV-MOFs. Preliminary characterizations of MTV-MOFs using solid state nuclear magnetic resonance (NMR) combined with computational modeling have proven effective at characterizing the average distances and apportionments of linkers and linker domains on the nanoscale, a first step toward identifying sequences in these materials.[8−10] In fact, it would seem that the power of solid state NMR in elucidating structure comes down to the idea that it probes the local composition of a material. As such, it is a natural choice for studying the compositional periodicity or aperiodicity of solids. It is also likely that X-ray absorption, microscopy, and computational modeling techniques will play a large role in characterizing sequence-dependent materials like MTV-MOFs. However, we believe that in order to properly characterize these materials, new tools must be developed, either by refining existing techniques (e.g., implementing new NMR pulse sequences) or by developing entirely new methods. Perhaps inspiration for these new methods might be drawn from sequencing methods currently used in biochemistry for peptides and DNA, such as Edman degradation or shotgun sequencing.

The path to sequence-dependent materials poses incredible challenges in both synthesis and characterization. Yet there is vast potential to make materials with functions unlike anything in existence, either natural or synthetic. With control over spatial composition in a MOF, chemical functionalities could trace out paths through the pores for specific adsorbates, creating molecular circuits that could drive logic gates, or count and sort molecules by type. Compartments with different chemical functionalities, either consisting of one pore or multiple, could be linked to perform reactions in series or in tandem. Molecules could be developed to interface with the MOF structure that would encode or read out sequences of functional groups as they diffuse through the pores. These are just a few possible functions that could be made possible through the design of sequence-dependent MOFs and sequence dependent materials.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yaghi@berkeley.edu.

ORCID

Omar M. Yaghi: 0000-0002-5611-3325

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Schrodinger, E. *What is Life?*; Cambridge University Press: Cambridge, 1944.

(2) Deng, H.; Doonan, C. J.; Furukawa, H.; Ferreira, R. B.; Towne, J.; Knobler, C. B.; Wang, B.; Yaghi, O. M. Multiple Functional Groups of Varying Ratios in Metal-Organic Frameworks. *Science* **2010**, *327*, 846−850.

(3) Zhang, Y. B.; Furukawa, H.; Ko, N.; Nie, W.; Park, H. J.; Okajima, S.; Cordova, K. E.; Deng, H.; Kim, J.; Yaghi, O. M. 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.

(4) Thompson, J. A.; Brunelli, N. A.; Lively, R. P.; Johnson, J. R.; Jones, C. W.; Nair, S. Tunable CO2 Adsorbents by Mixed-Linker Synthesis and Postsynthetic Modification of Zeolitic Imidazolate Frameworks. *J. Phys. Chem. C* **2013**, *117*, 8198−8207.

(5) Sue, A. C. H.; Mannige, R. V.; Deng, H.; Cao, D.; Wang, C.; Gándara, F.; Stoddart, J. F.; Whitelam, S.; Yaghi, O. M. Heterogeneity of functional groups in a metal−organic framework displays magic number ratios. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 5591−5596.

(6) Fracaroli, A. M.; Siman, P.; Nagib, D. A.; Suzuki, M.; Furukawa, H.; Toste, F. D.; Yaghi, O. M. Seven Post-synthetic Covalent Reactions in Tandem Leading to Enzyme-like Complexity within Metal−Organic Framework Crystals. *J. Am. Chem. Soc.* **2016**, *138*, 8352−8355.

(7) Yuan, S.; Lu, W.; Chen, Y. P.; Zhang, Q.; Liu, T. F.; Feng, D.; Wang, X.; Qin, J.; Zhou, H. C. Sequential Linker Installation: Precise Placement of Functional Groups in Multivariate Metal−Organic Frameworks. *J. Am. Chem. Soc.* **2015**, *137*, 3177−3180.

(8) Kong, X.; Deng, H.; Yan, F.; Kim, J.; Swisher, J. A.; Smit, B.; Yaghi, O. M.; Reimer, J. A. Mapping of Functional Groups in Metal-Organic Frameworks. *Science* **2013**, *341*, 882−885.

(9) Krajnc, A.; Kos, T.; Zabukovec Logar, N.; Mali, G. A Simple NMR-Based Method for Studying the Spatial Distribution of Linkers within Mixed-Linker Metal−Organic Frameworks. *Angew. Chem., Int. Ed.* **2015**, *54*, 10535−10538.

(10) Jayachandrababu, K. C.; Verploegh, R. J.; Leisen, J.; Nieuwendaal, R. C.; Sholl, D. S.; Nair, S. Structure Elucidation of Mixed-Linker Zeolitic Imidazolate Frameworks by Solid-State ^{1}H CRAMPS NMR Spectroscopy and Computational Modeling. *J. Am. Chem. Soc.* **2016**, *138*, 7325−7336.