Moving beyond the constraints of chemistry via crystal structure discovery with isotropic multiwell pair potentials

Julia Dshemuchadse[a,1], Pablo F. Damasceno[b,2], Carolyn L. Phillips[c], Michael Engel[a,d], and Sharon C. Glotzer[a,b,e,3]

[a]Department of Chemical Engineering, University of Michigan, Ann Arbor, MI 48109; [b]Applied Physics Program, University of Michigan, Ann Arbor, MI 48109; [c]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439; [d]Institute for Multiscale Simulation, Interdisciplinary Center for Nanostructured Films, Friedrich-Alexander University Erlangen-Nürnberg, 91058 Erlangen, Germany; and [e]Biointerfaces Institute, University of Michigan, Ann Arbor, MI 48109

The rigid constraints of chemistry—dictated by quantum mechanics and the discrete nature of the atom—limit the set of observable atomic crystal structures. What structures are possible in the absence of these constraints? Here, we systematically crystallize one-component systems of particles interacting with isotropic multiwell pair potentials. We investigate two tunable families of pairwise interaction potentials. Our simulations self-assemble a multitude of crystal structures ranging from basic lattices to complex networks. Sixteen of the structures have natural analogs spanning all coordination numbers found in inorganic chemistry. Fifteen more are hitherto unknown and occupy the space between covalent and metallic coordination environments. The discovered crystal structures constitute targets for self-assembly and expand our understanding of what a crystal structure can look like.

self-assembly | crystal structures | isotropic pair potentials

Do we know all conceivable crystal structures? This question appears naïve at first, because crystallography is a mature field, but the list of reported inorganic crystal structures is not necessarily representative of all kinds of order that are possible on other scales. Atomic crystal structures are affected by the discreteness of the periodic table and the resulting constraints on chemical bonding (1). Molecular crystals (2), metal organic frameworks (3), nanoparticle superlattices (4), and other soft-matter assemblies (5) are free from these chemical constraints and can exhibit entirely new types of crystallographic order distinct from those found with atoms. A universal list of all plausible crystal structures in systems of particles ranging from the angstrom to the micrometer scale would benefit the search for—and design of—new materials.

Crystal structures observed on the atomic scale are subject to the laws of quantum mechanics and to the discrete nature of the atom. The constraints of the chemical bond limit the ways in which atoms can be arranged; in particular, inorganic compounds display geometries that are specific to different kinds of bonding, for example, tetrahedral coordinations in the case of covalent bonding with sp^3-hybridized orbitals, or high coordination numbers (CNs) in the case of metallic bonding. For example, in the case of water, angular information is usually directly encoded into the computational model (6, 7); however, structures with similar local arrangements have recently been observed in simulations with isotropic multiwell potentials (8, 9). In addition to known crystal structures, these tunable pair potentials can also be parametrized to model particles "in between" the discontinuous types of behaviors that are possible on the atomic scale, due to the quantized nature of the realm of electronic interactions. As a result, these interaction potentials can model systems on the mesoscopic soft-matter length scale, where particle properties and shapes are highly variable. How do these crystal structures differ from those observed on the atomic scale?

Here we show that molecular dynamics simulations of single-component systems interacting via simple isotropic oscillating pair potentials (OPPs) can produce—via crystallization from disordered initial conditions—the majority of reported one-component as well as several multicomponent inorganic crystal structures (10). Our findings are obtained with pair potentials that encode coordination geometry solely via the shape of the radial function of the interaction, and contain no angular terms. We characterize simulation outcomes semiautomatically with the help of computational crystal structure identification techniques. Among the crystal structures we observe are many previously unknown structures, some resembling known ones and some surprisingly complex. The previously unknown crystal structures cluster near each other in parameter space in islands of complexity, demonstrating that certain pair potential features promote specific coordination environments, and that certain coordination environments are inherently prone to crystallographic diversity. Because the functional form of the potential is generic, chemical constraints do not play a role in our simulations, and the

Significance

Which crystal structures are possible if the restrictions of the quantum realm are lifted? Our knowledge of ordered particle geometries was previously restricted to the kinds of structures observable in hard condensed matter—on the atomic scale. Here, we use freely tunable computational models to represent particles with variable properties, and we determine the crystal structures into which they self-assemble. The resulting arrangements often correspond to structures known from atomic-scale materials; however, we discover a comparable number of previously unknown crystal structures with different local coordination motifs, incompatible with the limitations of the chemical bond. Our results can be used to engineer soft condensed matter with unprecedented, ordered geometries, paving the way toward materials with potentially novel properties.

APPLIED PHYSICAL SCIENCES

www.manaraa.com

particles explore feasible geometries more freely. Such geometries should be accessible, if not to atomic crystals, then to colloidal crystals of nanoparticles where valence is easily tunable.

Materials and Methods

We chose interaction potentials that are as simple as possible (i.e., have the fewest and simplest features) while at the same time flexible enough to allow all possible coordination environments. The OPP and the Lennard-Jones Gauss potential (LJGP), illustrated schematically in Fig. 1, fulfill these conditions. Both potentials are short ranged, with up to two attractive wells located at relative depths and positions tunable by two independent parameters. In the following, we identify and categorize all three-dimensional crystals that self-assemble with these potentials within their two-dimensional parameter spaces.

The OPP and LJGP model different classes of materials. The OPP is representative of functional forms that mimic the behavior of intermetallic systems based on electronic interactions (11). Whereas many pair potential shapes for the study of intermetallics are empirically adapted to specific compositions, the present study varies the potential shapes continuously and systematically. A first glimpse of the structures accessible within this family of potentials revealed a variety of structures, including a few simple ones, a Frank–Kasper phase, clathrates, and a series of icosahedral quasicrystals (9), and more recently, the phase diagram of the OPP has been explored via inverse materials design (12). The LJGP, on the other hand, has been used to model particle interactions in soft-matter systems (13, 14).

The functional form of the OPP that we use is based on the Mihalkovič–Henley interaction potential (11). Its shape is closely related to the potential form used for the self-assembly of complex structures in monatomic systems (15). The OPP is given by

$$V_{OPP}(r) = \frac{1}{r^{15}} + \frac{\cos(k(r-1) + \phi)}{r^3}. \qquad [1]$$
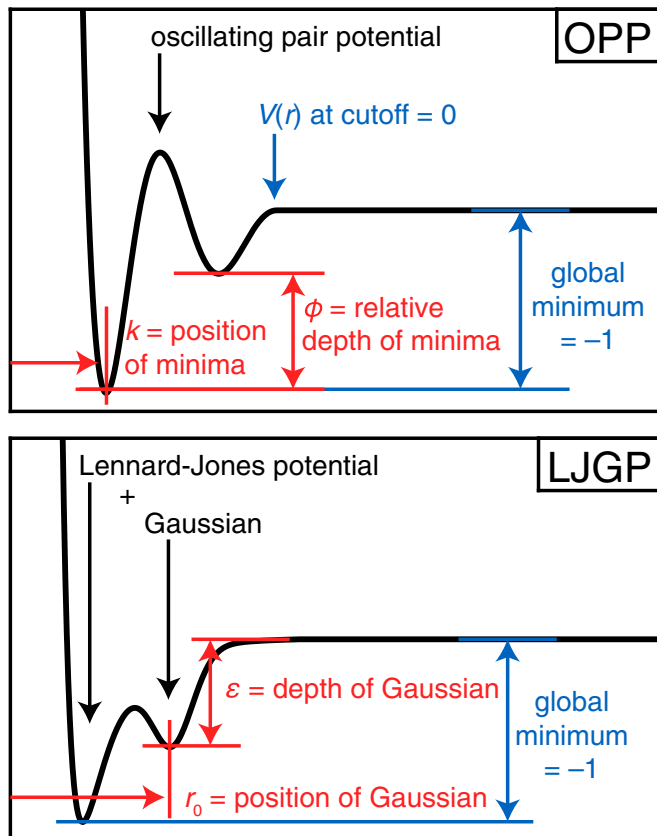
Fig. 1. Schematic illustration of the OPP (*Top*) and the LJGP (*Bottom*). The OPP is characterized by a wavenumber k and a phase shift ϕ defining the relative well positions and their relative depths. The LJGP is the sum of a Lennard–Jones potential and a negative Gaussian potential. Parameters r_0 and ϵ determine the position and depth of the Gaussian.

Parameters k and ϕ describe the wavenumber and phase shift of a damped oscillation, respectively. The oscillation is truncated and shifted to zero at the local maximum following the second attractive well.

The LJGP is constructed by adding a negative Gaussian to the Lennard-Jones pair potential (13),

$$V_{LJGP}(r) = \frac{1}{r^{12}} - \frac{2}{r^6} - \varepsilon \exp\left(-\frac{(r - r_0)^2}{2\sigma_G^2}\right), \qquad [2]$$

where we fixed the width of the Gaussian to $\sigma_G^2 = 0.02$ (and σ_{LJ} in the Lennard-Jones potential is set to $2^{-1/6}$ and $\varepsilon_{LG} = 1$). The two remaining parameters are the position r_0 and depth ε of the Gaussian. Because the potential converges rapidly, we truncated it at $r = 2.5$. The parameters k and ϕ for the OPP and the parameters ε and r_0 for the LJGP define the two parameter spaces to be explored.
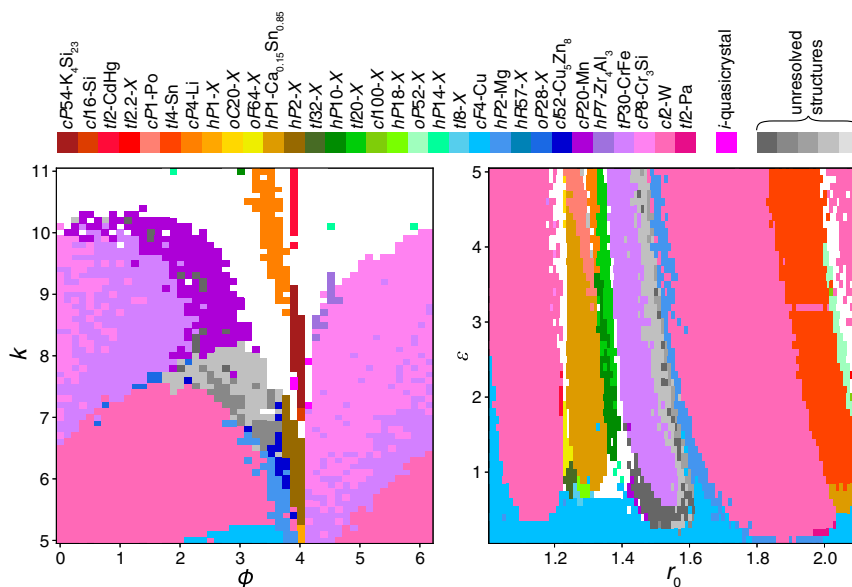
We performed all molecular dynamics simulations with the open-source simulation package HOOMD-blue (16, 17). Independent simulations were run at state points $\{k, \phi\}$ with $5 \leq k \leq 11$, $\Delta k = 0.1$ and $0 \leq \phi < 2\pi$, $\Delta \phi = \pi/25$ for the OPP (3,050 simulations in total) and at state points $\{r_0, \varepsilon\}$ with $1 < r_0 \leq 2.1$, $\Delta r_0 = 0.01$ and $0 < \varepsilon \leq 5$, $\Delta \varepsilon = 0.1$ for the LJGP (5,500 simulations). Systems of $N = 4,096$ (OPP) and $N = 2,744$ (LJGP) particles were initialized in random arrangements at low density, effectively corresponding to zero pressure. Simulations were started at high temperature and gradually cooled in the isochoric-isothermal (*NVT*) ensemble over at least 15×10^6 time steps (with $\delta t = 0.01$) to promote the formation of single-crystalline assemblies. Additional, slower simulations were carried out in parameter regions where no ordering had been observed with faster cooling, as detailed in *SI Appendix*. Because all potentials are attractive and the systems were initialized at low densities, our simulations generally resulted in a compact cluster of particles that does not span the simulation box across periodic boundaries. Thus, mechanical stresses and structural defects often present in bulk simulations relax in our setup.

We analyzed the final frames of all simulation trajectories to identify self-assembled structures in crystallographic detail. Phase diagrams were obtained, first, automatically with global bond-orientational order parameters and, later, manually via analysis of bond-orientational order diagrams (BODs) and visual inspection of the three-dimensional structures. To extract unit cells, we isolated a single-crystalline grain from the final frame and then—via the real-space particle positions as well as the BOD—aligned the crystal with the relevant high-symmetry directions, depending on the crystal system. Periodicity along these directions was analyzed via a cylindrical BOD that identifies the most frequent interparticle distances along specified directions. The determined unit cell vectors were then used to project all particle positions onto one unit cell, which, if correct, clearly shows that the projected particle positions cluster. The list of averaged cluster positions constitutes the extracted crystal structure. Space groups were determined manually or using Spglib (18). To check whether the identified structures are already known, we consulted the Inorganic Crystal Structure Database (ICSD) (19).

The diverse structures observed in this study are characterized by CNs. The CN of a particle is its number of nearest neighbors. Particles are considered to be nearest neighbors if their spacing falls within the first peak of the radial distribution function (RDF). The average CN, $\langle CN \rangle$, is the average of CN over all particles, since a given structure can have multiple distinct particle environments. CN and $\langle CN \rangle$ can be determined irrespective of periodicity and the degree of order; however, in this study, we specify the CN values according to the different Wyckoff positions in the determined unit cells and based on the overall RDF of the assembly. The $\langle CN \rangle$ is then determined by weighing the different CN values with the multiplicity associated with each Wyckoff position in the unit cell of the respective crystal.

Results

Phase Diagrams from Simulation Data. Simulation outcomes as a function of the two potential parameters are summarized in phase diagrams. The phase diagram of the OPP (Fig. 2, *Left*) is periodic in the ϕ parameter and consists of a few large domains of crystal structures separated by a narrow stripe of high structural diversity in the range $3 < \phi < 4$. In this range, an icosahedral quasicrystal was discovered (9, 20). At large k, the OPP oscillates rapidly, and simulations do not exhibit clear long-range order. The phase diagram of the LJGP (Fig. 2, *Right*) consists of several large domains that extend in the ε parameter direction. High structural diversity appears in the range $1.2 < r_0 < 1.6$. With few exceptions, the LJGP always leads to crystalline order. In both phase diagrams, regions of unusually complex order are marked

Dshemuchadse et al.
Moving beyond the constraints of chemistry via crystal structure discovery with isotropic multiwell pair

Fig. 2. Phase diagrams of OPPs (*Left*) and LJGPs (*Right*) indicating self-assembled crystal structures. Each colored rectangle corresponds to one simulation. Colors denote different crystal structures consistently across both diagrams as specified in the color bar (top), together with Pearson symbols and representative compounds. White regions indicate simulations that assemble structures with no discernible long-range order.

by shades of gray. The following discussion focuses on the 31 crystal structures indicated by colors in Fig. 2.

All 31 crystal structures are illustrated in Fig. 3 and listed in Table 1. Crystal structures are identified by their Pearson symbol, which consists of a lowercase letter representing the crystal system (c, cubic; t, tetragonal; o, orthorhombic; h, hexagonal/trigonal; m, monoclinic; and a, triclinic), an uppercase letter representing the lattice centering (P, primitive; I, body centered; F, face centered; $A/B/C/S$, base centered; and R, rhombohedrally centered), and the number of atoms per unit cell, as well as the representative compound, usually the first one that was found to exhibit the crystal structure. While 16 of the crystal structures

observed in these datasets correspond to known compounds, some of which have been reported in simulation before (9, 15, 21), we also discovered 15 previously unknown crystal structures. We indicate these structures without atomic equivalents by replacing the representative compound with an X.

Analysis of Coordination Numbers for Known Structures. An atom's CN reflects the nature of its chemical bonding. We analyzed our dataset of self-assembled crystal structures by determining CN of individual particles as well as $\langle CN \rangle$ of all structures in Table 1. The CN values present in the known crystal structures among our data cover almost the full range of CNs represented by
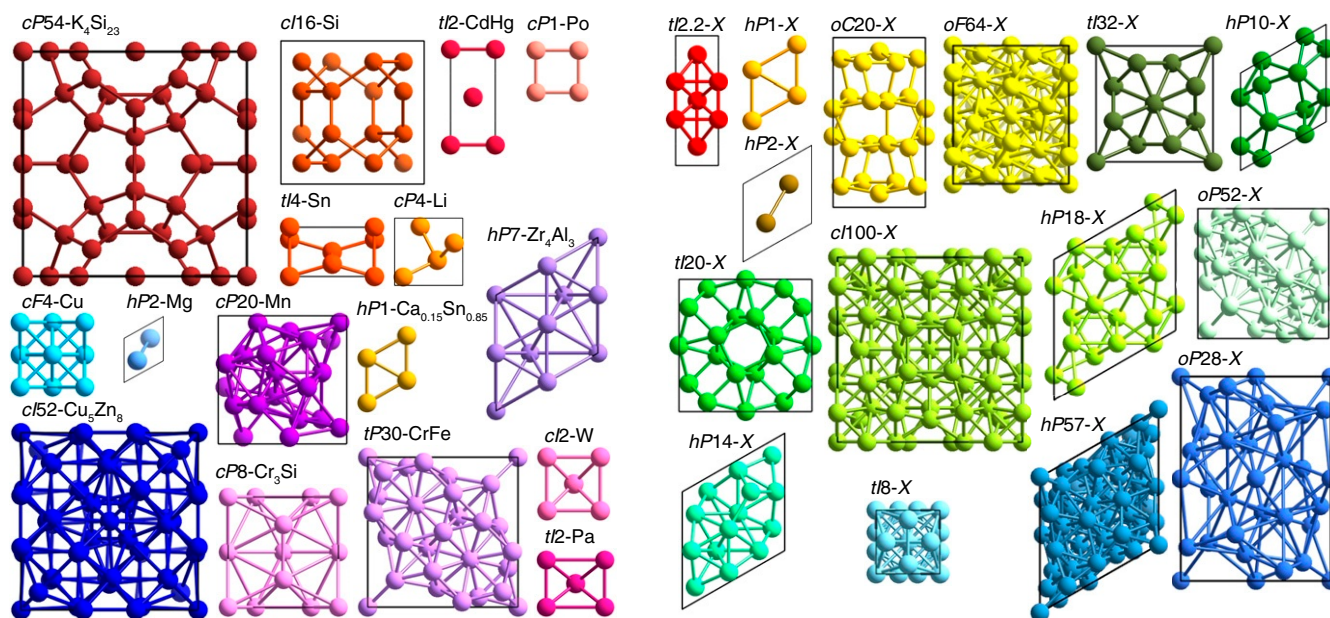
Fig. 3. Unit cells of the crystal structures reported in this study. Unit cells of all 16 known crystal structures (*Left*) and all 15 previously unknown crystal structures (*Right*) are shown. The particles are colored according to the color bar in Fig. 2; unit cell outlines are shown in black.

www.manaraa.com

Table 1. List of crystal structures divided into known crystal structures and previously unknown crystal structures

Pearson symbol		Space group	CN	⟨CN⟩	OPP	LJGP	Colloquial name or comments
Known crystal structures							
$cP54$-K_4Si_{23}	▪	$Pm\bar{3}n$	0/4	3.4	+		"clathrate-I"
$cI16$-Si	▪	$Ia\bar{3}$	4	4	+		high-pressure silicon
$tI2$-CdHg	▪	$I4/mmm$	4	4	+	+	—
$cP1$-Po	▪	$Pm\bar{3}m$	6	6		+	"sc"
$tI4$-Sn	▪	$I4_1/amd$	6	6		+	"β-tin"
$cP4$-Li	▪	$P4_132$	6	6	+	+	structure prediction only
$hP1$-$Ca_{0.15}Sn_{0.85}$	▪	$P6/mmm$	8	8		+	—
$cF4$-Cu	▪	$Fm\bar{3}m$	12	12	+	+	ccp/"fcc"
$hP2$-Mg	▪	$P6_3/mmc$	12	12	+	+	hcp
$cI52$-Cu_5Zn_8	▪	$I\bar{4}3m$	12/13	12.7	+		"γ-brass"
$cP20$-Mn	▪	$P4_132$	12/14	12.8	+	+	"β-manganese"
$hP7$-Zr_4Al_3	▪	$P6/mmm$	12/14/15	13.4	+		Frank-Kasper (FK) phase
$tP30$-CrFe	▪	$P4_2/mnm$	12/14/15	13.5	+	+	"σ-phase"/FK phase
$cP8$-Cr_3Si	▪	$Pm\bar{3}n$	12/14	13.5	+	+	"A15"/FK phase
$cI2$-W	▪	$Im\bar{3}m$	14	14	+	+	"bcc"
$tI2$-Pa	▪	$I4/mmm$	14	14		+	distorted "bcc"
–	▪				+		icosahedral quasicrystal
Previously unknown crystal structures							
$tI2.2$-X	▪	$I4/mmm$	4 to 5	4.8		+	—
$hP1$-X	▪	$P6/mmm$	6	6	+		stretched $hP1$-$Ca_{0.15}Sn_{0.85}$
$oC20$-X	▪	$C222_1$	6	6		+	—
$oF64$-X	▪	$Fddd$	6 to 8	7.0		+	—
$hP2$-X	▪	$P6_3/mmc$	8	8	+		compressed hcp/$hP2$-Mg
$tI32$-X	▪	$I4/mcm$	8 to 11	8.8		+	modulated/disordered
$hP10$-X	▪	$P6_3/mcm$	8/10	9.2	+	+	stacked octahedra + columns
$tI20$-X	▪	$I4/mcm$	9/10	9.2		+	related to incommensurate $tI19.3$-Rb
$cI100$-X	▪	$Im\bar{3}m$	8 to 12	9.3		+	—
$hP18$-X	▪	$P\bar{3}$	10	10		+	stacked octahedra
$oP52$-X	▪	$Pbcm$	9 to 12	10.3		+	—
$hP14$-X	▪	$P6_3mc$	11/12	11.1	+	+	octahedra + tetrahedra
$tI8$-X	▪	$I4/mmm$	11/12	11.3		+	modulated/disordered
$hR57$-X	▪	$R\bar{3}$	12	12		+	—
$oP28$-X	▪	$P2_12_12$	12 to 16	12.6	+		FK phase
–	▪▪▪▪▪				+	+	unresolved crystal structures

Given are the Pearson symbol, space group, CN of the first-shell environments, ⟨CN⟩, occurrence in the OPP and LJGP phase diagrams ("+" = yes), and colloquial name or comments. All particles are equivalent in the simplest observed structures (e.g., $cF4$-Cu, $hP2$-Mg, $cI2$-W), but several structures have multiple different—albeit similar—CNs.

one-component structures in the ICSD—ranging from 4 to 15. (We do not observe the molecular-type structures with ⟨CN⟩ = 1 to 3.) CN values are concentrated at the low end with CN ≤ 6, representative of covalent bonding, and at the high end with CN ≥ 12, representative of metallic bonding. We now discuss the known structures in order of increasing CN.
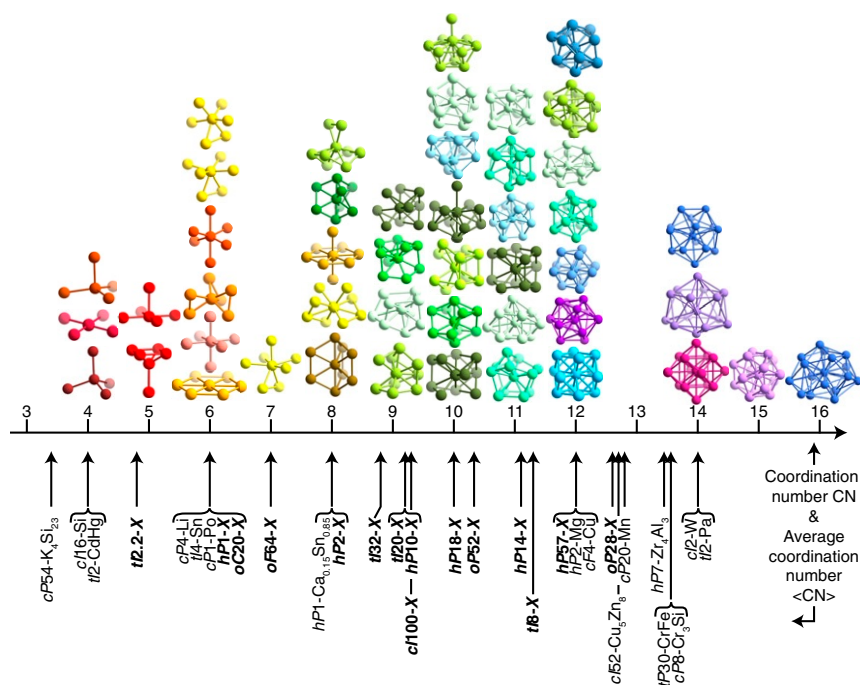
Atoms with CN = 4 correspond to sp^3-hybridized, covalently bonded silicon- and carbon-based frameworks. A representative structure that we observe is clathrate-I $cP54$-K_4Si_{23}. Most particles in clathrate-I build up a tetrahedral framework of fullerene-like cages with 20 or 24 vertices. The remaining particles occupy cage centers and do not have nearest neighbors at the shortest interparticle distance. Another structure on the low end of the CN range is the high-pressure silicon structure $cI16$-Si. The structure $tI2$-CdHg is crystallographically equivalent to $tI2$-Pa, but, due to its different unit cell dimensions ($c/a_{tI2\text{-CdHg}} = 2.2$ vs. $c/a_{tI2\text{-Pa}} = 0.8$; with $c/a_{cI2\text{-W}} = 1$), their topologies and thus their structures differ. The $tI2$-CdHg consists of flat square lattice layers, where each neighboring layer is offset by $1/2 \times (1, 1, 1)$.

Atoms with CN = 6 include covalent environments in chemistry, for example, for aluminum atoms, but also occur in salts like NaCl. Known one-component crystal structures with CN = 6 are the simple cubic packing $cP1$-Po, the β-tin phase $tI4$-Sn, and a chiral, high-pressure lithium phase $cP4$-Li. Ionic compounds can have higher values, such as CN = 8 in the case of CsCl. An exam-

ple of a one-component structure with CN = 8 is the stacking of the two-dimensional triangle tiling in $hP1$-$Ca_{0.15}Sn_{0.85}$.

Crystal structures with CN = 8 to 11 are rare in the ICSD. Among the elements, these CNs are known in high-pressure phases only. No one-component atomic structures exist with CN values of exclusively 5 or 7. The only atomic crystal structures that have local instances of CN = 5 or 7 are complex phases of boron. (In our data, but also in the ICSD, odd-numbered CN values are less common than even-numbered ones.)

At the high end of the CN spectrum with CN ≥ 12 lie metallically bonded compounds, a group that is most prominently represented by sphere packings such as the cubic close-packed (ccp, often also termed fcc after its face-centered cubic Bravais lattice; here $cF4$-Cu) and hexagonal close-packed (hcp; here $hP2$-Mg) structures. Another common structure with high CN is the body-centered cubic structure (often termed bcc; here $cI2$-W) with CN = 14. The $tI2$-Pa represents a distorted version of $cI2$-W. The more complex β-manganese structure $cP20$-Mn has two Wyckoff positions with CN = 12 and CN = 14 that correspond to distorted Frank–Kasper polyhedra. Other prominent examples of complex high-CN structures that we observe are γ-brass $cI52$-Cu_5Zn_8 and three Frank–Kasper phases: $hP7$-Zr_4Al_3, the "A15"-phase $cP8$-Cr_3Si, and the σ-phase $tP30$-CrFe. All of these phases represent multicomponent sphere packings having local environments with CN = 12 to 16.

Dshemuchadse et al.
Moving beyond the constraints of chemistry via crystal structure discovery with isotropic multiwell pair

Fig. 4. Coordination polyhedra in the 31 crystal structures self-assembled with OPPs and LJGPs (displayed in colors representing their structures in Figs. 2 and 3). Coordination polyhedra are grouped by CN in the range CN = 4 to 16. Reported structures are listed by Pearson symbols below the CN axis and indicated by arrows pointing to their ⟨CN⟩ (the largest average CN observed is ⟨CN⟩ = 14). Pearson symbols of previously unknown structures are labeled in bold font.

Analysis of Coordination Numbers for Previously Unknown Crystal Structures. Two previously unknown structures we observe in our data have the same Pearson symbol, space group, and Wyckoff position as known structures—but different c/a ratios. The $hP1$-X is a stretched variant of $hP1$-$Ca_{0.15}Sn_{0.85}$. Both structures are periodic stackings of triangular lattice layers directly on top of one another, corresponding to an AA stacking (distinct from the common ABCABC and ABAB stackings found in $cF4$-Cu and $hP2$-Mg, respectively, which correspond to dense packings due to the alternating layer positions). Particles in $hP1$-$Ca_{0.15}Sn_{0.85}$ have a hexagonal bipyramid for a coordination polyhedron (CN = 8) in contrast to the flat hexagon coordination (CN = 6) in $hP1$-X. The distance between hexagonal layers in $hP1$-X is so large that the layers become partially decorrelated, and stacks of two to three A layers followed by a stacking fault are common. The other structure with a familiar-looking Pearson symbol—$hP2$-X—also differs from $hP2$-Mg in its larger c/a ratio. The $hP2$-X has a bicapped triangular prism (CN = 8) as coordination polyhedron, while $hP2$-Mg has an anticuboctahedron (CN = 12).

Four of the previously unknown structures are best understood as columnar arrangements. The stackings in the direction of the columns can be periodic or exhibit nontrivial order. In $hP18$-X, particles form columns of face-sharing octahedra. In $hP10$-X, columns of face-sharing octahedra coexist with columns of individual particles. The $tI20$-X is related to the high-pressure phase $tI19.3$-Rb (22). Both have columns of face-sharing square antiprisms and intermediate columns of individual particles, but, while $tI20$-X is periodic, the column periodicities are incommensurate in $tI19.3$-Rb, requiring a four-dimensional superspace description (23, 24). Finally, $tI32$-X consists of columns of edge-sharing tetrahedra on a square grid together with columns of individual particles. While the columns of tetrahedra are fully ordered, the individual particles occur in three sites along the column direction corresponding to three equivalent local environments. Averaged over the simulation volume, the three sites occur with equal probability. Whether the decoration of the sites

is decorrelated or locally correlated without global order (25, 26) cannot be decided based on our data. Longer and larger simulations are necessary to fully characterize the presence of incommensurate modulations and correlated disorder.

We also observe partial disorder in two other previously unknown structures. The $tI8$-X consists of alternating buckled and flat layers. The buckled square lattice layers are fully ordered. The flat triangle–square tiling layers (corresponding to the semiregular snub square tiling, $3^2.4.3.4$) adopt one of two orientations. Relative positions of the two orientations are at least partially decorrelated across the buckled layer. The $tI2.2$-X consists of a $tI2$-CdHg–like underlying structure: flat square lattice layers alternating in their positions with an offset of $1/2 \times (1, 1, 1)$. Additional interstitial particles are positioned just above or below the flat layers and are offset with respect to the in-layer squares, effectively forming a flat square pyramid. The high global symmetry of this particle position combined with its low occupancy leads to CN = 5 for 80% of the particles in this structure, resulting in ⟨CN⟩ = 4.8.

The seven remaining structures cannot be understood as stackings of layers or packings of columns. They combine some of the geometric building blocks described above in a fully three-dimensional fashion. The $hP14$-X is a network of face-sharing octahedra as well as vertex- and face-sharing tetrahedra. The $oP52$-X is a framework of square antiprisms, octahedra, and tetrahedra, among other polyhedra. The $hR57$-X exclusively exhibits environments with CN = 12, only approximately a third of which are icosahedral in shape. The remaining coordination environments are too distorted to be assigned regular polyhedral shapes. The $oP28$-X is a Frank–Kasper phase previously postulated (27) but not yet reported. It consists of all four, partially distorted, topologically close-packed Frank–Kasper polyhedra with CN ∈ {12, 14, 15, 16}. Particles in $oC20$-X have CN = 6, with 40% of them surrounded by slightly distorted, and another 40% by strongly distorted, octahedral environments, and the remaining 20% surrounded by triangular prismatic environments. The $oF64$-X features particle environments with CN = 6

APPLIED PHYSICAL SCIENCES

to 8. Particles with CN = 7, 8 have irregularly shaped environments, and 25% of particles with CN = 6 are surrounded by distorted octahedra. The even more broadly distributed coordination environments in $cI100$-X comprise cuboctahedra with CN = 12 but also more irregular shapes for CN = 8 to 10.

Besides these 15 previously unknown structures described here, we observed a few additional unusually complex structures. Their location is indicated by shades of gray in the phase diagrams in Fig. 2. Preliminary analysis indicates that these crystal structures are so complicated that they cannot be resolved with the methods of this study. They will be discussed in a separate publication with more advanced structure analysis tools. Crystallographic details for all described structures, as well as visualizations, are given in *SI Appendix*.

Discussion and Conclusion

Altogether, we documented the self-assembly of 16 known crystal structures and 15 previously unknown crystal structures without atomic equivalents using multiwell isotropic pair potentials in a one-component system using molecular dynamics simulations. The structures reported have unit cell sizes containing up to 100 particles and also include aperiodic and partially disordered structures. Particles in these structures have CNs between 4 and 16. The structures with atomic equivalents exhibit ⟨CN⟩s of 4, 6, 8, and 12 to 14. While the low and medium ⟨CN⟩ values occur mostly discretely—that is, with all particles in these structures having the same CN—high ⟨CN⟩s seem to be more tunable. The presence of a variety of slightly differing noninteger ⟨CN⟩ values is correlated with the higher structural complexity in this group, which is a prerequisite for multiple symmetrically inequivalent particle positions and environments with different CNs = 12 to 16 to coexist in the same crystal structure.

We find that the previously unknown crystal structures are predominantly clustered in the intermediate range with ⟨CN⟩ = 7 to 11 (Table 1). This range is characterized by a particularly high number of distinct coordination polyhedra (shown in Fig. 4). At the same time, one-component atomic structures almost never have these CN values: The restrictions dictated by the periodic table of elements and the nature of chemical bonds select for the crystal structures currently known and available to materials scientists. Apparently, the list of all geometrically possible crystal structures that can be stable is much larger and extends beyond the constraints of chemistry. The previously unknown crystal structures we report are candidates for soft-matter and nanoscale materials.

Our results complement inverse methods for materials design (8, 28–30) that target specific geometries and fuel both experimental and computer-driven materials searches across the scales where coarse-grained interactions can be used to explore particle assembly. Particles in our simulations readily assemble many simple structures with only a few particles in the unit cell, but also highly complex ones with up to 100 particles in the unit cell. The simplicity of the models (isotropic interactions with few features in the pair potential) and our focus on one-component systems does not preclude structural complexity. While the here-discussed simulations are conducted at effectively zero pressure, additional structural variety can likely be accessed by observing the self-assembly behavior of these systems as a function of pressure (e.g., ref. 29). The kinetic pathways permitting this complexity, and whether similar behavior is also present in binary mixtures or compounds with even more components, will be subjects of future work.

1. L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry* (Cornell University Press, 1960).
2. G. R. Desiraju, Crystal engineering: From molecule to crystal. *J. Am. Chem. Soc.* **135**, 9952–9967 (2013).
3. H. Furukawa, K. E. Cordova, M. O'Keeffe, O. M. Yaghi, The chemistry and applications of metal-organic frameworks. *Science* **341**, 1230444 (2013).
4. M. A. Boles, M. Engel, D. V. Talapin, Self-assembly of colloidal nanocrystals: From intricate structures to functional materials. *Chem. Rev.* **116**, 11220–11289 (2016).
5. S. C. Glotzer, M. J. Solomon, Anisotropy of building blocks and their assembly into complex structures. *Nat. Mater.* **6**, 557–562 (2007).
6. V. Molinero, E. B. Moore, Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B* **113**, 4008–4016 (2009).
7. G. A. Cisneros et al., Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chem. Rev.* **116**, 7501–7528 (2016).
8. A. Jain, J. R. Errington, T. M. Truskett, Dimensionality and design of isotropic interactions that stabilize honeycomb, square, simple cubic, and diamond lattices. *Phys. Rev. X* **4**, 031049 (2014).
9. M. Engel, P. F. Damasceno, C. L. Phillips, S. C. Glotzer, Computational self-assembly of a one-component icosahedral quasicrystal. *Nat. Mater.* **14**, 109–116 (2015).
10. W. Steurer, J. Dshemuchadse, *Intermetallics – Structures, Properties, and Statistics* (Oxford University Press, 2016).
11. M. Mihalkovič, C. L. Henley, Empirical oscillating potentials for alloys from ab initio fits and the prediction of quasicrystal-related structures in the Al-Cu-Sc system. *Phys. Rev. B* **85**, 092102 (2012).
12. P. Zhou, J. C. Proctor, G. Van Anders, S. C. Glotzer, Alchemical molecular dynamics for inverse design. *Mol. Phys.* **17**, 3968–3980 (2019).
13. M. Engel, H.-R. Trebin, Self-assembly of monatomic complex crystals and quasicrystals with a double-well interaction potential. *Phys. Rev. Lett.* **98**, 225505 (2007).
14. M. Engel, H.-R. Trebin, Structural complexity in monodisperse systems of isotropic particles. *Z. Kristallogr.* **223**, 721–725 (2008).
15. M. Elenius, F. H. M. Zetterling, M. Dzugutov, D. C. Fredrickson, S. Lidin. Structural model for octagonal quasicrystals derived from octagonal symmetry elements arising in β-Mn crystallization of a simple monatomic liquid. *Phys. Rev. B* **79**, 144201 (2009).
16. J. A. Anderson, C. D. Lorenz, A. Travesset, General purpose molecular dynamics simulations fully implemented on graphics processing units. *J. Comput. Phys.* **227**, 5342–5359 (2008).
17. J. Glaser et al., Strong scaling of general-purpose molecular dynamics simulations on GPUs. *Comput. Phys. Commun.* **192**, 97–107 (2015).
18. A. Togo, I. Tanaka, Spglib: A software library for crystal symmetry search. arXiv [Preprint](2018). https://arxiv.org/abs/1808.01590 (Accessed 3 December 2018).
19. F. I. Z. Karlsruhe, Inorganic crystal structure Database – ICSD. https://icsd.fiz-karlsruhe.de. Accessed 5 November 2018.
20. M. Spellings, S. C. Glotzer, Machine learning for crystal identification and discovery. *AIChE J.* **64**, 2198–2206 (2018).
21. F. H. M. Zetterling, M. Dzugutov, S. Lidin, γ-Brass crystallization in a simple monotomic liquid. *MRS Proc.* **643**, K9.5 (2000).
22. M. I. McMahon, S. Rekhi, R. J. Nelmes, Pressure dependent incommensuration in Rb-IV. *Phys. Rev. Lett.* **87**, 55501 (2001).
23. T. Janssen, G. Chapuis, M. De Boissieu, *Aperiodic Crystals – Fom Modulated Phases to Quasicrystals* (,Oxford University Press, 2007), vol. 20.
24. S. Van Smaalen, *Incommensurate Crystallography* (, Oxford University Press, 2007), vol. 21.
25. D. A. Keen, A. L. Goodwin, The crystallography of correlated disorder. *Nature* **521**, 303–309 (2015).
26. T. R. Welberry, *Diffuse X-ray Scattering and Models of Disorder* (, Oxford University Press, 2010), vol. 16.
27. M. D. Sikirić, O. Delgado-Friedrichs, M. Deza, Space fullerenes: A computer search for new Frank–Kasper structures. *Acta Crystallogr. A* **66**, 602–615 (2010).
28. B. A. Lindquist, R. B. Jadrich, T. M. Truskett, Communication: Inverse design for self-assembly via on-the-fly optimization. *J. Chem. Phys.* **145**, 111101 (2016).
29. P. F. Damasceno, S. C. Glotzer, M. Engel, Non-close-packed three-dimensional quasicrystals. *J. Phys. Condens. Matter* **29**, 234005 (2017).
30. C. S. Adorf, J. Antonaglia, J. Dshemuchadse, S. C. Glotzer, Inverse design of simple pair potentials for the self-assembly of complex structures. *J. Chem. Phys.* **149**, 204102 (2018).

6 of 6 | **PNAS**
https://doi.org/10.1073/pnas.2024034118

Dshemuchadse et al.
Moving beyond the constraints of chemistry via crystal structure discovery with isotropic multiwell pair