Artificial Synthesis of Conjugated Microporous Polymer via Sonogashira-Hagihara Coupling

J.M.H. Thomas,[2] C. Mollart,[1] L. Turner,[1] P. Heasman,[1] P. Fayon,[1] and A. Trewin[1*]

[1] Lancaster University, Department of Chemistry, Bailrigg, Lancaster, UK, LA1 4YB

[2] Institute of Integrative Biology, University of Liverpool, Liverpool, UK, L69 7ZB

Corresponding author email: a.trewin@lancaster.ac.uk

KEYWORDS Nanoporous, CMP, Simulated synthesis, Polymer structure

ABSTRACT: Amorphous network materials are increasingly important with applications including as supercapacitors, battery anodes, and proton conduction membranes. Design of these materials is hampered by the amorphous nature of the structure and sensitivity to synthetic conditions. Here, we show that through artificial synthesis, fully mimicking the catalytic formation cycle and full synthetic conditions, we can generate structural models that can fully describe the physical properties of these amorphous network materials. This opens up pathways for rational design where complex structural influences, such as solvent and catalyst choice, can be taken into account.

1. Introduction

Microporous materials, such as metal-organic frameworks (MOFs)[1-3], covalent organic frameworks (COFs)[4-8], zeolites,[9] and microporous organic polymers (MOPs)[10-12] have important applications. These include gas adsorption, heterogeneous catalysis, chemical separations[13-17], and in energy generation and storage, for example as supercapacitors, ion storage for batteries, and as proton conduction membranes.[18-21] There is a pressing need for the discovery of materials that have high porosity combined with chemical functionality to attain the physical, electronic, and optical properties that are being demanded by these increasingly important applications.

COFs and MOFs can exhibit ultra-high surface areas (over 6000 m^2/g)[22] with good thermal stabilities reported however some studies suggest chemical decomposition in air, for example COF-5 changes colour from dark red to grey in air. In contrast, MOPs have been shown to be very robust with very good physicochemical stabilities. For example, many can be boiled in acid with no loss of porosity. MOPs have a wide synthetic diversity available with a number of different types being available. These include hyper-cross-linked polymers (HCPs),[23] porous aromatic frameworks (PAFs),[24-25] conjugated microporous polymers (CMPs)[26-27] and polymers of intrinsic microporosity (PIMs).[28]

CMPs[26, 29] are hyper-branched polymeric materials linked via π-conjugated bonds.[23] In CMPs the building block nodes are linked together in a manner that allows π-conjugation. There are many different synthetic strategies for forming π-conjugated linkages, resulting in a large range of synthetically and functionality diverse CMP materials. The unique combination of high surface area and extended π-conjugation result in a range of applications unique to CMPs. The interesting electronic structure properties opens up potential for light manipulation, for example light emitting or light harvesting, and electric storage devices, for example as supercapacitors.

Many of the syntheses used to form CMPs include the use of group 10 transition metal catalyzed carbon-carbon coupling reactions. These reactions include palladium catalyzed Heck,[30-31] Sonogashira-Hagihara[32] and Suzuki[27, 33] reactions and the nickel catalyzed Yamamoto[34-35] reaction. These are popular routes to synthesize CMPs due to their versatility and functional group tolerance.[35-36]

The Sonogashira-Hagihara reaction is one of the most commonly used synthetic routes to CMP materials. The mechanism of the cross-coupling follows an oxidative addition reduction elimination pathway. Oxidative addition of a vinyl halide to a Pd(0) metal complex is followed by transmetallation by copper(I)-acetylide. Reductive elimination results in the coupled product and regeneration of the catalyst. The full catalytic mechanism is shown in Figure 1. Because of the requirement for an expensive palladium catalyst, Sonogashira-Hagihara reactions are very expensive and this cost is one of the limiting factors in the search for novel CMP materials.

Here we show that the full catalytic Sonogashira-Hagihara reaction cycle can be simulated so that we can artificially synthesize CMP materials. This greatly

improves our ability to explore new chemical functionality and aid the discovery of novel CMP materials.

2. Computational Methods

2.1 Amorphous structure generation

The generation of representative amorphous network models is challenging as, unlike ordered materials where the crystal structure is known, there is a lack of experimental data to aid construction. A common strategy is to pack polymer chains and use molecular dynamic (MD) simulations to achieve a target density that is obtained from experiment.[37] However, densities can be challenging to obtain experimentally and can oversimplify the inhomogeneous pore structure. This approach can lead to models that match the target experimental density but do not replicate other physical properties well.

A number of computational studies have attempted to follow the polymer synthetic procedure, for example for simple polycondensation reaction mechanisms. This approach has resulted in representative structures with densities that are close to those obtained experimentally that emerge from the generation procedure itself. However, these methods are not able to replicate the catalytic formation mechanism of CMP materials or the desolvation processes that can strongly influence the resulting structures and are limited computationally to small simulation cells.

Here, we use an automated structure generation methodology that exploits GPU hardware for increased speed and size of simulation. This methodology is implemented in the *Ambuild* code. The *Ambuild* code is written in Python and integrates with HOOMD-blue and DL_POLY molecular dynamic simulation codes.[38-40] This enables us to have fine control of the structure building process, conditions and composition of the simulation cell.

2.2 Ambuild Code

The Ambuild code is a Python code written to automate the system generation process. Ambuild seeds an initial simulation cell with stoichiometric quantities of the network building blocks. A number of molecular dynamic (MD) simulation loops are undertaken with structural sampling for potential bond formation (a *Zip* test). This *Zip* step is based on the distance between the respective atom in each building block that will ultimately be within the target bond, defined as an end group atom. Each end group atom has a cap atom. This cap atom defines the reference vector for structural sampling. For each *Zip* step a margin is specified for permitted bonds, for both the bond length and bond angle. Upon bond formation, the structure undergoes a full, rigid-body geometry optimization before returning to the MD loops.

The HOOMD-blue[38-40] GPU-based code is used as the MD and optimization engine throughout, enabling long simulation times and easy integration to the Python code for the reasonably large simulation cell size required. Each MD loop is an NVT (constant number of atoms, constant volume, and constant temperature) MD simulation consisting of 100000 steps with a timestep of 0.5 fs. The Fast Inertial Relaxation Engine (FIRE) rigid-body minimiser is used for optimization of the geometry. Bond and angle parameters and intermolecular forces are taken from the PCFF force field.[41] A harmonic bond potential is used to describe the bonds and dihedrals. Intermolecular forces are described by a Lennard-Jones potential with a cut-off distance of 10 Å. Charges are calculated using the Gasteiger method.[42]

2.3 Computational catalytic mechanism

To replicate the catalytic mechanism, we set specific bonding rules for the CMP network building blocks and the catalyst. The catalyst is described atomistically with the two bonding sites defined by a hydrogen atom respectively, shown in Figure 2. The hydrogen atoms do not have any intermolecular forces associated with them and are there just to define the bonding site. The palladium atom is defined as an end group and the hydrogen atoms defined as the cap atoms. We allow the CMP building blocks to form a bond to the palladium catalyst during a *Zip* test. For the vinyl halide, the halide is defined as the cap atom and the attached carbon as the end group. For the alkyne, we assume that all alkyne groups have reacted to form the copper(I)-acetylide and therefore are all terminated with a copper atom. The copper atom is defined as the cap atom and the attached carbon as the end group. Once a CMP building block is bound to the catalyst, the catalyst and the respective CMP building block are redefined. We allow two redefined CMP building blocks on the same catalyst to bond to each other. We then delete the bonds between the palladium catalyst and the CMP end group and define a new bond between the two carbon end groups of the respective CMP building blocks. The catalyst hydrogens are reinstated, thereby regenerating the catalyst, which is then able to bond to further CMP end groups.

2.4 Computational synthesis process

We mimic the respective experimental procedure as closely as possible. Here, as an example, we describe in detail how we replicate the synthesis of CMP-1 as described by Laybourn *et al.*[43] They describe the synthesis of CMP-1 as follows: 1,3,5-triethynylbenzene (TEB)(1.0 mmol), 1,4-dibromobenzene (DBB)(1.0 mmol), triethylamine (TEA)(7.0 mmol), and DMF (1300 mmol) are mixed under nitrogen and heated to 100 °C. A slurry of 0.04 mmol of the tetrakistriphenyl phosphine palladium(0) catalyst and copper(I) iodide (0.08 mmol) in DMF (1 mL) is added via a wide-bore needle. The reaction is terminated with cold methanol and filtered to leave a solid precipitate. The precipitate is Soxhlet extracted in methanol and dried in a vacuum oven.

To mimic this synthetic procedure, we follow a number of stages; (i) a seeding stage, where the cell is set up, (ii) a network generation stage, (iii) a desolvation stage, where we remove solvent and excess building blocks, (iv) a work up stage that mimics the interaction with methanol (v) a cell size equilibration stage, where we allow the synthesized and desolvated network to adopt its preferred dimension. These stages are described in detail as follows:

(i) Seeding Stage: we seed a simulation cell with 100 molecules of DBB, 100 molecules of TEB, 4 molecules of catalyst, and different amounts of DMF and TEA to replicate the molar ratio of the reaction mixture and different degrees of solvation that may occur in the reaction mixture. The simulation cell length is set by finding the minimum cell size that will accommodate the reaction mixture. For each solvent system, the cell length and the cell contents are identical. The seeding process results in a different starting configuration for each model. The artificial synthetic details are given in Tables S1 and S2. We optimize the seeded structure and run an NVT MD step.

(ii) Network Generation Stage: This stage consists of a loop, the first step of which is to delete any solvent blocks that haven't bonded and run an NVT MD step. This allows the catalyst and any bonded blocks to disperse. We then seed the TEB and DBB blocks back into the cell. A *Zip* test is undertaken to bond any TEB and DBB blocks to the catalyst. A catalyst-bonding step (as described in section 2.3) is undertaken to bond any TEB and DBB that are bonded to the catalyst together. We only allow bonds between DBB and TEB to form during the catalyst-bonding step. This loop is repeated until no more successful *Zip* steps are possible.

(iii) Desolvation stage: Solvent, unreacted building blocks, catalyst and small polymer fragments remain present in the cell after the network generation stage. The polymer undergoes desolvation and so to mimic this process we remove the solvent in a stepwise manner with NVT MD to allow for the cell contents to adapt. It is postulated that the desolvation process also removes the unreacted building blocks, catalyst and small polymer fragments. However, the molecular size that this desolvation removal process is relevant to is not well understood. It is also postulated that alkyne-alkyne homo-coupling can occur during the desolvation process. We therefore have six different approaches to the desolvation strategy that are shown in Figure S9, and are summarized as follows:

Strategy 1: stepwise (slow) removal of unreacted DBB, unreacted TEB, TEA, DMF, catalyst and small oligomers (less than 2 blocks) with NVT MD and an optimization of the structure.

Strategy 2: stepwise removal of unreacted DBB, unreacted TEB, TEA, DMF, catalyst, small oligomers (less than 2 blocks), and large oligomers (less than 3 blocks) with NVT MD and an optimization of the structure.

Strategy 3: stepwise (fast) removal of unreacted DBB, unreacted TEB, TEA, DMF, catalyst and small oligomers (less than 2 blocks) with NVT MD and an optimization of the structure.

Strategy 4: stepwise removal of unreacted DBB, unreacted TEB, TEA, DMF, and catalyst with NVT MD and an optimization of the structure.

Strategy 5: stepwise removal of TEA and DMF solvent with NVT MD and an optimization of the structure.

Strategy 6: stepwise removal of unreacted DBB, unreacted TEB, TEA, DMF, and catalyst with NVT MD and an optimization of the structure and allowing for alkyne-alkyne homo-coupling at each step.

(iv) Workup Stage: The workup with methanol will replace any accessible copper atoms with a hydrogen atom. To replicate the workup, we perform a *Grow* step with a hydrogen molecule. One hydrogen atom is a cap atom and one hydrogen is an end group. The *Grow* step is repeated with an NPT (constant number of atoms, constant pressure, and constant temperature) MD step and optimization until no more *Grow* steps are possible.

A *Grow* step will only be successful if there is enough space for the hydrogen atom to occupy and therefore any copper atoms that are not accessible will remain within the system.

(v) Cell Equilibration: The network has been generated in a solvated environment. Upon desolvation and workup the system will have void space that is potentially unfavourable. In the experiment, the system will contract to fill space efficiently. Here we replicate this by performing an NPT MD step that is repeated until the simulation cell has converged to its optimal volume.

3. Artificial synthesis analysis

3.1 Net generation analysis of CMP-1

Experimentally, Laybourn et al[43] performed a detailed analysis of the CMP-1 network formation mechanism. They rationalized the characterization data taken at sequential timesteps during the experimental synthesis. They concluded that the mechanism proceeds via the formation of soluble oligomers that combine to form insoluble clusters. These clusters then combine to form an insoluble polymer. They describe the cluster formation as occurring via alkyne-bromine cross-coupling as well as formation via alkyne-alkyne homo-coupling.

As the formation mechanism proceeds via the formation of soluble oligomers followed by the formation of insoluble clusters, it is not clear whether the reaction mixture is fully homogenous, where all parts are equally solvated, or whether there is phase separation and some parts contain less solvent. It is conceivable that as the network is forming there are relatively fewer solvent molecules in and around the oligomers than in the bulk reaction mixture. Hence, we have here assessed four different degrees of solvation; Solvation 1- 3 DMF molecules per unit cell, Solvation 2- 13 DMF molecules per unit cell, Solvation 3 – 130 DMF molecules per unit cell, and Solvation 4 – 1300 DMF

molecules per unit cell (the experimental reaction ratio). Tables S1 and S2 show the experimental reaction ratios and the different solvation schemes used here. For each Solvation Scheme, we generate four repeat models to increase the structural sampling and explore the structural diversity of the CMP-1 models. The seeding process results in a different starting configuration of the system and so results in different polymer network configurations being synthesized.

Here, we assess the polymer formation mechanism for Solvation Scheme 3 in detail. In all four models we see the formation of small oligomers initially followed by the formation of small clusters. These clusters then combine to give one large cluster that grows across the cell periodic boundary to form the polymer, shown in Figure S1.

Figure 3 shows the evolution of the size of the largest block present and the number of blocks present in the simulation cell during the network generation phase of Model-1. The number of blocks declines steadily as the generation process steps increases. The number of blocks includes the solvent, catalyst, and starting blocks and so never approaches a singular block. Initially the mass of the biggest block in the system remains relatively steady at a value less than 2000 g mol^{-1}. In this region the system is made up of small oligomers. At step 20 the mass begins to rise as small oligomers are formed. From step 40, the mass remains relatively constant at between 4000 and 6000 g mol^{-1}. In this region larger oligomers form and become intertwined with each other forming a gel-like structure. At step 120 there is a sharp increase in mass that corresponds to these large oligomers bonding to form a network cluster that pervades through the simulation cell. Any further size increase is due to addition of oligomers to the clusters. Figure S2–S3 shows the full network generation process for models 1-4. The same general trend can be seen in each whereby small oligomers form initially, followed by a gel-like phase, and finally polymer formation. This trend is repeated for each solvation model apart from 1300, in which only a very small number of very small oligomers were observed to form after 90 steps. To test if further bond formation is possible in the 1300 system, one model was allowed to continue with further steps. No further bond formation was observed after 130 steps. It is possible that further bond formation could occur over longer time periods but it is not feasible to test this within the scope of this work. Interestingly, as the solvation of the system increases, the gel-like phase is increased from an average of 70 steps for solvation system 3 to an average of 128 steps for solvation system 130. We believe that this is due to the greater number of solvent molecules holding the system further apart and slowing down the rate at which the building blocks can react. Figure S8 shows the largest fragment within the system for each solvation scheme and model.

3.2. Desolvation

Desolvation may result in the removal of solvent only, but it can also result in the removal of building blocks and even of small oligomers. There is no experimental data that describes which moieties are removed from the system upon desolvation. Figure S9 shows the different desolvation strategies that we implemented and the technical details are described in section 2.4(iii). Each strategy is designed to replicate a different real-world scenario. Strategy 5 replicates the removal of solvent only. This desolvation will occur in a system where there are small pores through which only solvent molecules are able to diffuse through. Strategy 4 and Strategy 6 replicate a system in which solvent and building blocks are able to diffuse through the system. In Strategy 6 we also allow homocoupling to occur. In Strategy 1 we remove solvent, building blocks and small oligomers. Strategy 3 is the same but the desolvation occurs at a faster rate, which does not allow for structural reorganisation. Strategy 2 removes solvent, building blocks, small oligomers and larger oligomers. This represents a system which has larger pores and therefore can allow for large molecules to diffuse through the pore structure. Figure S10 shows the resulting structures.

3.4 Workup

The workup stage replaces any freely accessible copper atoms with a hydrogen atom. ESI section 4 shows the resultant weight percent of copper for each system. For the solvation scheme 3 systems the copper weight percentage is 0 wt%Cu. This low weight percent is a result of the high condensation of TEB meaning that very few copper end groups remain and those that do are accessible. For solvent scheme 13 and 130, we see a similar weight percent of mostly between 0 and 5 wt%Cu. For solvent scheme 1300, we see a wide range between 1 and 5 wt%Cu. This reflects the low degree of polymerization in the systems, where there are very few groups remaining after desolvation and so 1 copper atom can result in a high copper weight percentage.

3. 5 Cell equilibration

After desolvation and work up, there are now voids within the framework where solvent, building blocks, or oligomers were located. These may have been enabling the polymer framework to exist in an open form. With their removal, there will now be a force acting upon the polymer framework to compress in order to maximise the intermolecular forces within the polymer framework. The NPT MD simulation allows the polymer structure to respond to the changing forces, the result of which can be observed in the change in the cell size (shown in Tables S3 and S5). For solvent system 3, the resultant systems show no or very little decrease in the cell size with 1.14% being the largest decrease observed. As the solvation becomes larger the decrease in the cell size upon equilibration also becomes larger with a largest decrease of 8.71%,

19.43%, and 55.99% for 13, 130, and 1300 solvent systems respectively.

4. Structure Analysis

Exact atomistic structural analysis of amorphous polymeric materials is not possible as experimental analysis techniques are not able to determine the atomistic locations due to the inherent randomness of the polymer material. We are therefore not able to compare our models to an exact experimentally obtained structure. Instead, we take a holistic approach and compare large-scale structural analysis and simulated properties to experimentally obtained values. Solvent schemes 3, 13, and 130 results in polymer network structures that extend throughout the simulation cell. Figure S8 shows the largest fragment resulting from each net generation growth before desolvation, workup, and cell equilibration. Solvent scheme 1300 did not result in polymer growth with only small oligomers observed to form. Upon desolvation, either the whole system was removed or the small oligomers remained. We have therefore not considered solvent scheme 1300 in our structure analysis discussion.

4.1 Elemental analysis

Elemental analysis is useful for assessing the elemental composition of the polymer, particularly for the percentage of end groups that still remain within the polymer after synthesis. Table 1 shows the average percentage of carbon, hydrogen, and nitrogen in the artificially synthesized CMP-1 models for each desolvation strategy and compares to experimentally determined elemental analysis of CMP-1. In the artificially synthesized models the remaining mass (mass that is not hydrogen, carbon, or nitrogen) is due to phosphorus, copper, palladium, and bromine from the end groups and catalyst that have not been removed during the workup and desolvation steps. Experimentally, the percentages of these elements cannot be determined through elemental analysis and so we compare the percentage remaining after taking account of the carbon, nitrogen and hydrogen. Table S6 shows the averages of the elemental analysis for the models (elemental analysis for all models is given in Table S4). For solvent system 3, we obtain carbon, hydrogen, and other values between 63.20 – 86.61wt%, 2.56 – 4.89wt%, and 10.34 - 34.23wt% respectively depending upon the desolvation strategy. A similar range of values is found for solvent systems 13 and 130.
The experimentally determined elemental analysis is 80.90wt%, 3.48wt%, and 15.58wt% for carbon, hydrogen, and other. This fits within the range of values that we observe for the artificially synthesized systems.

4.2 Density

The bulk density (that is the density of the micropore region of the polymer excluding contributions from mesoporosity) can be estimated from the relationship $W_o = 1/\rho_{app} - 1/\rho_{tr}$, where W_o is the specific pore volume, ρ_{app} is the bulk density, and ρ_{tr} is the absolute density. The bulk density is the more reasonable comparison to our artificially generated CMP-1 systems since we can only simulate the micropore region. For CMP-1, a bulk density of between 0.86 and 1.13 g cm^{-3} is estimated from the absolute densities determined from the pore volumes reported by Dawson *et al*[44-45] and Laybourn *et al*[43] respectively.
Amorphous materials, by definition, do not have a regular extended structure and therefore we would expect that each model structure will be different. It has been postulated that hyper cross-linked polymeric materials collapse upon desolvation similarly to crystalline solvate framework materials that collapse upon desolvation to non-porous materials. For CMP-1 we believe that the local structure and the extent of polymerization throughout the volume and the degree of desolvation determines whether the polymer collapses locally or not. It may be that some volumes of the extended polymer collapse and some areas remain in their original form. The extent of the collapse of the CMP-1 framework that we are able to simulate here may be influenced by the size of the simulation cell. The cell lengths may not be large enough to capture phenomena such as the cluster – cluster framework interpenetration.
Table S5 shows the average density of the artificially synthesized CMP-1 models for each desolvation strategy (densities for all models are given in Table S5). Densities ranging between 0.63 g cm^{-3} and 1.17 g cm^{-3}, 0.71 g cm^{-3} and 1.06 g cm^{-3}, and 0.51 g cm^{-3} and 0.73 g cm^{-3} are obtained for solvent systems 3, 13, and 130, respectively. The experimentally determined bulk density of 1.13 g cm^{-3} is in good agreement with the upper end of these ranges of densities.

4.3 Surface Area

Porosity is a key property of CMP materials. Most CMP materials exhibit reversible microporosity with surface areas obtained from application of the BET equation to the nitrogen uptake isotherm of between 400 – 2000 m^2 g^{-1}. CMP-1 has surface areas that range between 600 – 900 m^2 g^{-1} dependent upon the method of synthesis and specifically the solvent used. Laybourn *et al*, whose synthetic procedure we have based our artificial synthesis procedure upon, obtained a final surface area of 733 m^2 g^{-1}. Furthermore, they were able to estimate the microporosity from the ratio of pore volume calculated at low relative pressures to the pore volume calculated at high relative pressures, $V_{0.1}/V_{Tot}$. A $V_{0.1}/V_{Tot}$ ratio of 0.77 was found indicating a relatively high degree of microporosity. Using this ratio, we can estimate that the contribution to the total surface area from the microporous region is 564 m^2 g^{-1}. Dawson *et al* obtained a surface area of 967 m^2 g^{-1} using the same synthetic conditions. Similarly, we can estimate the contribution to the total surface area from the microporous region to be 755 m^2 g^{-1}. Giving an average

micropore surface area of 655 $m^2 g^{-1}$ for CMP-1. However, as the synthetic method differs to that of Laybourn *et al* we will not use this in our comparison.

There are a number of methods of calculating the geometric surface area of a representative model system. These are based upon the concept of rolling a probe sphere across the interface of the system atoms and calculating the surface area at different points of the interface between the probe and system. These geometric surface areas are useful to visualize and quantify the porosity within the model and can give a good approximation of the experimental surface area. As the models assume full desolvation and are currently limited to small volumes, these geometric surface areas can be considered to represent the idealized maximum possible in the micropore region. Here we use the solvent accessible surface area, which only takes into account volume that is accessible to the sides of the simulation cell, calculated using the Poreblazer program for a nitrogen molecular probe.[46] However, we note that a comparison of geometric surface area values to the experimentally obtained surface area values obtained from application of the BET equation to the nitrogen uptake isotherm is not applicable as they do not directly relate.

Table S5 shows the average solvent accessible surface areas. A wide range of values is determined dependent upon the solvation and the desolvation strategy. For solvent system 3, values range between 0 and 1446 $m^2 g^{-1}$, for solvent system 13 values range between 126 and 1510 $m^2 g^{-1}$, and for solvent system 130 values range between 1561 and 2818 $m^2 g^{-1}$. In general the solvent accessible surface area increases with the level of solvation.

4.4 Micropore volume

Comparing to the micropore volume must take into account the different probes used. Helium has a smaller diameter than nitrogen and therefore will overestimate the micropore volume. We therefore assume that the simulated micropore volume will be higher than the experimental micropore volume for the same system. Using the ratio of the different diameters of the probe, we can estimate that the simulated micropore volume using a helium probe should be approximately 0.3 $cm^3 g^{-1}$ to match the experimental value of 0.23 $cm^3 g^{-1}$ determined using a nitrogen probe. Micropore volume is determined through the nitrogen uptake isotherm. Here, we have calculated the micropore volume using the Poreblazer program for a helium probe. A helium probe has a slightly smaller diameter to molecular nitrogen and therefore a comparison should be comparable but slightly overestimate in comparison to the experimental values.

Table S5 shows the average calculated micropore volumes. For solvent system 3, values range between 0.10 and 0.78 $cm^3 g^{-1}$, for solvent system 13 values range between 0.21 and 0.73 $cm^3 g^{-1}$, and for solvent system 130 values range between 0.69 and 1.19 $cm^3 g^{-1}$. In general the micropore volume increases with the level of solvation.

4.5 Pore size distribution

Laybourn *et al* report CMP-1 with pore size distributions of over 20 Å with a peak at approximately 40 Å calculated from the nitrogen uptake isotherm using non-local density functional theory (NL-DFT). No data below 10 Å was reported due to a lack of pressure points. However, Jiang *et al* report pore size distributions over 5 Å with a peak at approximately 10 Å.

Table S5 shows the average pore limiting diameter (PLD) and the maximum pore diameter (MPD) for the different solvation systems. For solvation scheme 3, the PLD ranges between 1.85 and 4.93 Å and a MPD ranging between 4.94 and 11.89 Å depending upon the solvation scheme. The PLD and MPD increase with the increasing level of solvation for all solvation schemes, indicating that the influence of the degree of solvation is retained even after desolvation occurs. The largest pore sizes are seen for solvation scheme 1300, although as polymerization did not occur for these systems this is a reflection of the emptier cells that remain after desolvation.

Figures S11 to S13 show the plotted pore size distributions for the models generated. No one pore size distribution shows a good match to the experimentally determined pore size distribution determined by Laybourn *et al*. However, peaks below 10 Å are not observed in the experimental pore size distribution due to the technique used. This does not mean that there are no pores with diameters within this region. We therefore look to pore size distributions that have no peaks in the 10-20 Å region, of which there are a number of examples. The pore size distributions calculated here are not able to replicate the larger pore sizes due to the limitations of the cell size.

5. Structure determination

To identify an overall structure of CMP-1, a holistic approach must be taken, considering all the experimental data and the corresponding structural assessments of the models generated.

Firstly we assess the desolvation strategy that would be adopted for each solvation scheme. For solvation scheme 3 and 13, the PLD in each desolvation scenario is such that only small molecules, such as solvent and TEA, would be able to access the pore structure to diffuse out of the material. This means that only desolvation strategy 5 is applicable and the other systems can be discounted. For solvation scheme 130, the PLD in each desolvation scenario is such that solvent, catalyst, and building blocks are able to diffuse through the pore system and out of the material and so desolvation strategies 4 and 6 are applicable. The computational characterization properties of these are summarized in Table 1 in comparison to the experimental characterization data. None of these systems individually match well to the experimental

data. However, it is reasonable to assume that the real system is not homogeneous but rather has areas of different solvation levels.

Laybourn *et al* assessed the formation of the CMP-1 material by taking samples at different points in the synthesis procedure. Samples taken at early stages were composed of spherical particles showing interparticulate mesoporosity. This would fit well with solvation schemes 3 and 13 with only desolvation of solvent as these models have no microporosity. Samples collected later showed a fused particulate morphology with microporosity. This would fit well with the solvation scheme 130 with desolvation of solvent, catalyst, and building blocks. A feasible rationalization would be that phase separation of the catalyst and building blocks occurs to give dense regions that form a denser polymer core of a particulate. Moving from the core to the outer edges of the particulate results in an increase in the solvation and so regions of less dense polymer are able to form at later stages. At some point, perhaps upon precipitation, the particulates are in contact and are able to react via alkyne-bromine cross coupling and alkyne-alkyne homo coupling to fuse the particulates together.

To test this theory, we consider a sphere where the core is composed of polymer formed through solvation scheme 3, a portion that is composed of solvation scheme 13, and an outer core composed of solvation scheme 130. The relative percentage that each contributes can be determined and compared to experimental data to give an overall system for comparison to experimental data. As none of the experimental characterization methodologies give precise and direct measurements, the relative weight that can be attributed to the matching of the computational results to must be considered carefully.

The resulting systems, after taking each characterization method in turn, and producing a system to match that specific value is shown in Tables S7 to S9 and the resulting PSD is shown in Figures S14-16. No one system fits the overall experimental data well, however each does possibly reflect the biases of the characterization technique used and how small structural contributions may dominate results. Fitting to the micropore and surface area results appears to give the best overall fit, giving a good match for density as well as the PSD. A summary of this model is given in Figure 4. In this model, a dense core is formed in a phase that has limited solvation (solvent system 3) from which only solvent is able to be removed upon desolvation (desolvation strategy 5) with no shrinkage of the polymer. This accounts for 40% of the spherical particulate volume. Surrounding this core is a region formed from a higher solvated phase (solvation scheme 13) from which only solvent is able to be removed upon desolvation (desolvation strategy 5) with no shrinkage of the polymer. This accounts for 36% of the spherical particulate volume. Surrounding this core is a region formed from a higher solvated phase (solvent system 130). When the spherical particulates meet, they are able to fuse together through alkyne-alkyne homocoupling or alkyne-bromine cross coupling. From this region we can remove solvent and the building blocks and allow homocoupling (desolvation scheme 4 and 6). In these regions, the volume can shrink by up to 75%. This accounts for 24% of the spherical particulate volume. Overall, the system contraction is 4% upon desolvation. As we have only modeled four different levels of solvation, models that take account of other solvation schemes may result in systems that better match the overall real-world system.

6. Summary

In summary, we have developed a methodology for generating structural models of amorphous conjugated microporous polymers through full imitation of the synthetic process. The structural systems produced are consistent with experimentally synthesized materials and observed properties and give information about the structure and formation mechanism.

Furthermore, this artificial synthesis approach can be applied to other systems with different building blocks and synthesis conditions. We believe that we can therefore apply this approach to systems that have yet to be explored and this will open new pathways to rational design strategies for the discovery of new materials.

SUPPORTING INFORMATION

The Supporting Information contains synthetic details, artificial synthetic details, polymer growth analysis, net generation results summary, desolvation strategy description and example, resultant model properties, model pore size distributions, and system fitting results.

ACKNOWLEDGEMENT

REFERENCES

1. Eddaoudi, M.; Moler, D. B.; Li, H.; Chen, B.; Reineke, T. M.; O'Keefe, M.; Yaghi, O. M., Modular chemistry: secondary building units as a basis for the design of highly porous and robust metal-organic carboxylate frameworks. *Acc. Chem. Res.* **2001,** *34*, 319-330.
2. Kitigawa, S.; Kitaura, R.; Noro, S., *Angew. Chem., Int. Ed.* **2004,** *43*, 2334-2375.
3. Han, S. S.; Goddard, W. A., Lithium-Doped Metal-Organic Frameworks for Reversible H2 Storage at Ambient Temperature. *Journal of the American Chemical Society* **2007,** *129* (27), 8422-8423.
4. Côté, A. P.; Benin, A. I.; Ockwig, N. W.; O'Keeffe, M.; Matzger, A. J.; Yaghi, O. M., Porous, crystalline, covalent organic frameworks. *Science* **2005,** *310* (5751), 1166-1170.
5. Hunt, J. R.; Doonan, C. J.; LeVangie, J. D.; Cote, A. P.; Yaghi, O. M., Reticular synthesis of covalent organic borosilicate frameworks,. *J. Am. Chem. Soc.* **2008**, 11872-11873.
6. Cote, A. P.; El-Kaderi, H. M.; Furukawa, H.; Hunt, J. R.; Yaghi, O. M., *J. Am. Chem. Soc.* **2007,** *129*, 12914 - 12915.
7. El-Kaderi, H. M.; Hunt, J. R.; Mendoza-Cortes, J. L.; Cote, A. P.; Taylor, R. E.; O'Keefe, M.; Yaghi, O. M., Designed synthesis of 3D covalent organic frameworks, . *Science* **2007,** *316*, 268-272.

8. Wan, S.; Guo, J.; Kim, J.; Ihee, H.; Jiang, D., A Belt-Shaped, Blue Luminescent, and Semiconducting Covalent Organic Framework, . *Angew. Chem. Int. Ed.* **2008,** *47*, 1-5.
9. Soldatov, D. V.; Ripmeester, J. A., Organic zeolites. In *Nanoporous Materials Iv*, 2005; Vol. 156, pp 37-54.
10. Dawson, R.; Stöckel, E.; Holst, J. R.; Adams, D. J.; Cooper, A. I., Microporous organic polymers for carbon dioxide capture. *Energy & Environmental Science* **2011,** *4* (10), 4239-4245.
11. Wood, C. D.; Tan, B.; Trewin, A.; Su, F.; Rosseinsky, M. J.; Bradshaw, D.; Sun, Y.; Zhou, L.; Cooper, A. I., Microporous Organic Polymers for Methane Storage. *Advanced Materials* **2008,** *20* (10), 1916-1921.
12. Xu, S.; Luo, Y.; Tan, B., Recent Development of Hypercrosslinked Microporous Organic Polymers. *Macromolecular Rapid Communications* **2013,** *34* (6), 471-484.
13. Germain, J.; Fréchet, J. M. J.; Svec, F., Nanoporous Polymers for Hydrogen Storage. *Small* **2009,** *5* (10), 1098-1111.
14. Dawson, R.; Cooper, A. I.; Adams, D. J., Nanoporous organic polymer networks. *Progress in Polymer Science* **2012,** *37* (4), 530-563.
15. Barbour, L. J., Crystal porosity and the burden of proof. *Chemical Communications* **2006,** (11), 1163-1168.
16. Cheetham, A. K.; Ferey, G.; Loiseau, T., Open-Framework Inorganic Materials. *Angew. Chem. Int. Ed.* **1999,** (38), 3268 - 3292.
17. Jones, J. T. A.; Hasell, T.; Wu, X.; Bacsa, J.; Jelfs, K. E.; Schmidtmann, M.; Chong, S. Y.; Adams, D. J.; Trewin, A.; Schiffman, F.; Cora, F.; Slater, B.; Steiner, A.; Day, G. M.; Cooper, A. I., Modular and predictable assembly of porous organic molecular crystals. *Nature* **2011,** *474* (7351), 367-371.
18. Fayon, P.; Thomas, J. M. H.; Trewin, A., Structure and Properties of a Nanoporous Supercapacitor. *The Journal of Physical Chemistry C* **2016,** *120* (45), 25880-25891.
19. Kou, Y.; Xu, Y.; Guo, Z.; Jiang, D., Supercapacitive Energy Storage and Electric Power Supply Using an Aza-Fused π-Conjugated Microporous Framework. *Angewandte Chemie International Edition* **2011,** *50* (37), 8753-8757.
20. Zou, J.; Trewin, A.; Ben, T.; Qiu, S., High Uptake and Fast Transportation of LiPF6 in a Porous Aromatic Framework for Solid-State Li-Ion Batteries. *Angewandte Chemie International Edition* **2020,** *59* (2), 769-774.
21. Heasman, P.; Trewin, A., Uptake and Diffusion of Ions in Organically Synthesized Porous Carbon for Battery Anode Applications. *The Journal of Physical Chemistry C* **2019,** *123* (42), 25603-25610.
22. Furukawa, H.; Ko, N.; Go, Y. B.; Aratani, N.; Choi, S. B.; Choi, E.; Yazaydin, A. O.; Snurr, R. Q.; O'Keeffe, M.; Kim, J.; Yaghi, O. M., Ultrahigh Porosity in Metal-Organic Frameworks. *Science* **2010,** *329* (5990), 424-428.
23. Wood, C. D.; Tan, B.; Trewin, A.; Niu, H.; Bradshaw, D.; Rosseinsky, M. J.; Khimyak, Y. Z.; Campbell, N. L.; Kirk, R.; Stoeckel, E.; Cooper, A. I., Hydrogen Storage in Microporous Hypercrosslinked Organic Polymer Networks,. *Chem. Mat.* **2007,** *19*, 2034-2048.
24. Ben, T.; Ren, H.; Ma, S.; Cao, D.; Lan, J.; Jing, X.; Wang, W.; Xu, J.; Deng, F.; Simmons, J.; Qiu, S.; Zhu, G., Targeted Synthesis of a Porous Aromatic Framework with High Stability and Exceptionally High Surface Area. *Angewandte Chemie International Edition* **2009,** *48* (50), 9457-9460.
25. Trewin, A.; Cooper, A. I., Porous Organic Polymers: Distinction from Disorder? *Angewandte Chemie International Edition* **2010,** *49* (9), 1533-1535.
26. Jiang, J.; Su, F.; Trewin, A.; Wood, C. D.; Campbell, N. L.; Niu, H.; Dickinson, C.; Ganin, A. Y.; Rosseinsky, M. J.; Khimyak, Y. Z.; Cooper, A. I., *Angew. Chem. Int. Ed.* **2007,** *46*, 1 - 5.
27. Weber, J.; Thomas, A., Toward Stable Interfaces in Conjugated Polymers: Microporous Poly(p-phenylene) and Poly(phenyleneethynylene) Based on a Spirobifluorene Building Block. *Journal of the American Chemical Society* **2008,** *130* (20), 6334-6335.
28. McKeown, N. B.; Budd, P. M., Polymers of intrinsic microporosity (PIMs): organic materials for membrane separations, heterogenous catalysis and hydrogen storage. *Chem. Soc. Rev.* **2006,** *35*, 675-683.
29. Jiang, J.-X.; Su, F.; Trewin, A.; Wood, C. D.; Niu, H.; Jones, J. T. A.; Khimyak, Y. Z.; Cooper, A. I., Synthetic Control of the Pore Dimension and Surface Area in Conjugated Microporous Polymer and Copolymer Networks. *Journal of the American Chemical Society* **2008,** *130* (24), 7710-7720.
30. Wang, D.; Xue, L.; Li, L.; Deng, B.; Feng, S.; Liu, H.; Zhao, X., Rational Design and Synthesis of Hybrid Porous Polymers Derived from Polyhedral Oligomeric Silsesquioxanes via Heck Coupling Reactions. *Macromolecular Rapid Communications* **2013,** *34* (10), 861-866.
31. Wang, D.; Yang, W.; Feng, S.; Liu, H., Constructing hybrid porous polymers from cubic octavinylsilsequioxane and planar halogenated benzene. *Polymer Chemistry* **2014,** *5* (11), 3634-3642.
32. Jiang, J. X.; Su, F.; Trewin, A.; Wood, C. D.; Campbell, N. L.; Niu, H.; Dickinson, C.; Ganin, A. Y.; Rosseinsky, M. J.; Khimyak, Y. Z.; Cooper, A. I., Conjugated microporous poly (aryleneethynylene) networks. *Angewandte Chemie-International Edition* **2007,** *46* (45), 8574-8578.
33. Rose, M.; Klein, N.; Bohlmann, W.; Bohringer, B.; Fichtner, S.; Kaskel, S., New element organic frameworks viaSuzuki coupling with high adsorption capacity for hydrophobic molecules. *Soft Matter* **2010,** *6* (16), 3918-3923.
34. Schmidt, J.; Werner, M.; Thomas, A., Conjugated Microporous Polymer Networks via Yamamoto Polymerization. *Macromolecules* **2009,** *42* (13), 4426-4429.
35. Ben, T.; Ren, H.; Ma, S. Q.; Cao, D. P.; Lan, J. H.; Jing, X. F.; Wang, W. C.; Xu, J.; Deng, F.; Simmons, J. M.; Qiu, S. L.; Zhu, G. S., Targeted Synthesis of a Porous Aromatic Framework with High Stability and Exceptionally High Surface Area. *Angewandte Chemie-International Edition* **2009,** *48* (50), 9457-9460.
36. Lu, W.; Yuan, D.; Sculley, J.; Zhao, D.; Krishna, R.; Zhou, H.-C., Sulfonate-Grafted Porous Polymer Networks for Preferential CO_2 Adsorption at Low Pressure. *Journal of the American Chemical Society* **2011,** *133* (45), 18126-18129.
37. Abbott L J, H. K. E.; Colina, C. M., Polymatic: a generalized simulated polymerization algorithm for amorphous polymers. *Theor. Chem. Acc.* **2013,** *132*, 1.
38. http://codeblue.umich.edu/hoomd-blue, HOOMD-blue web page: .
39. Anderson, J. A.; Lorenz, C. D.; Travesset, A., General purpose molecular dynamics simulations fully implemented on graphics processing units. *Journal of Computational Physics* **2008,** *227* (10), 5342-5359.
40. Nguyen, T. D.; Phillips, C. L.; Anderson, J. A.; Glotzer, S. C., Rigid body constraints realized in massively-parallel molecular dynamics on graphics processing units. *Computer Physics Communications* **2011,** *182* (11), 2307-2313.
41. Sun, H., Ab initio calculations and force field development for computer simulation of polysilanes. *Macromolecules* **1995,** *28* (3), 701-712.
42. Gasteiger, J.; Marsill, M., *Tetrahedron* **1990,** *36*, 3219-3228.
43. Laybourn, A.; Dawson, R.; Clowes, R.; Hasell, T.; Cooper, A. I.; Khimyak, Y. Z.; Adams, D. J., Network formation mechanisms in conjugated microporous polymers. *Polymer Chemistry* **2014,** *5* (21), 6325-6333.
44. Dawson, R.; Laybourn, A.; Clowes, R.; Khimyak, Y. Z.; Adams, D. J.; Cooper, A. I., Functionalized Conjugated Microporous Polymers. *Macromolecules* **2009,** *42* (22), 8809-8816.
45. Dawson, R.; Laybourn, A.; Khimyak, Y. Z.; Adams, D. J.; Cooper, A. I., High Surface Area Conjugated Microporous Polymers: The Importance of Reaction Solvent Choice. *Macromolecules* **2010,** *43* (20), 8524-8530.
46. Sarkisov, L.; Harrison, A., Computational structure characterisation tools in application to ordered and disordered porous materials. *Molecular Simulation* **2011,** *37* (15), 1248-1257.

ASSOCIATED CONTENT

Supporting Information. Synthetic and artificial synthetic information, structural results, desolvation process, property results, structure fitting results.

AUTHOR INFORMATION

Corresponding Author

* a.trewin@lancaster.ac.uk

Author Contributions

The manuscript was written through contributions of all authors.

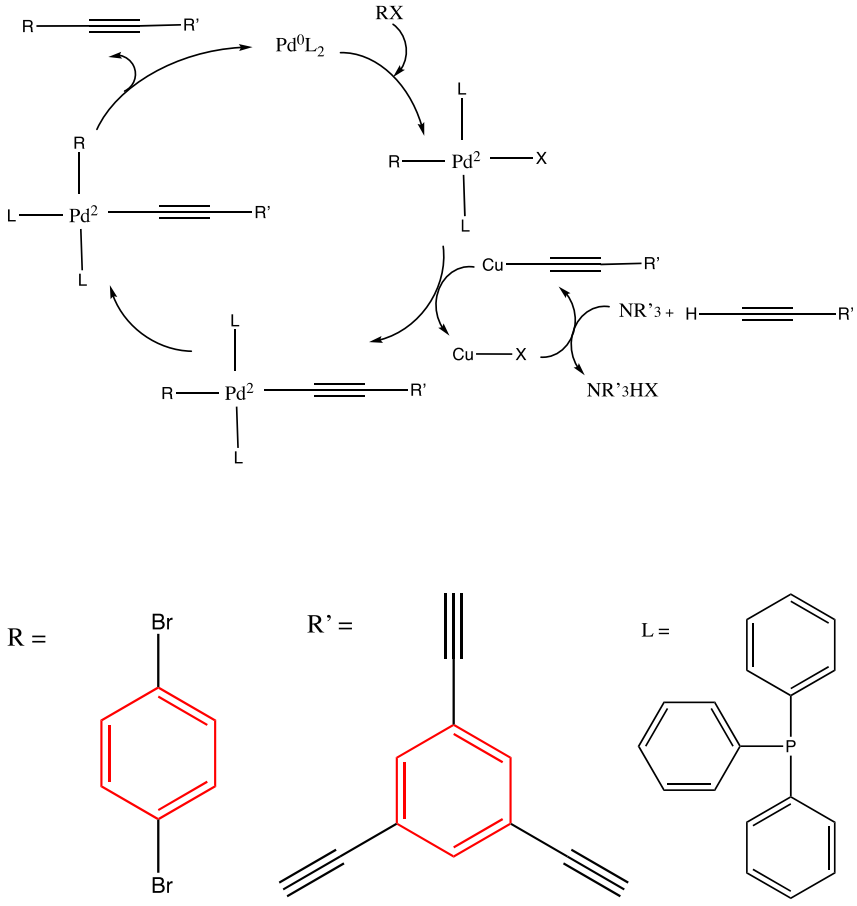

Figure 1. The catalytic cycle for the Sonogashira-Hagihara carbon coupling reaction to produce CMP-1.

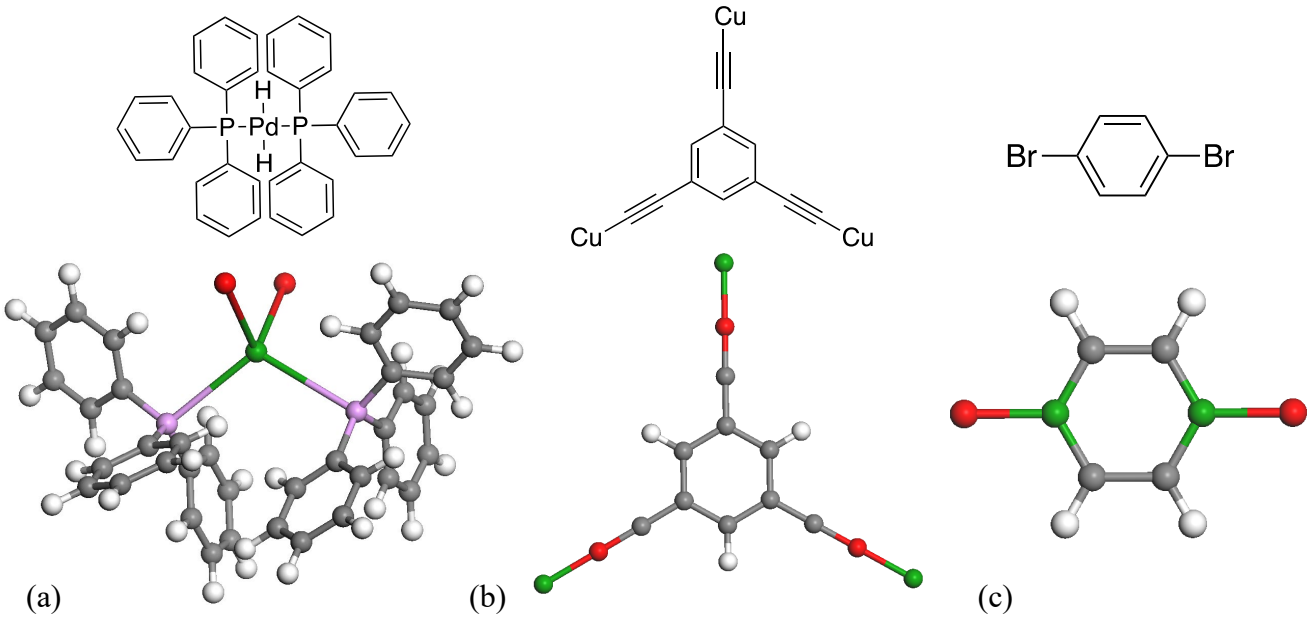

Figure 2. The blocks used in the Ambuild code (a) catalyst, (b) the CMP-1 building block copper(I) triethynylbenzene, and (c) the CMP-1 building block 1,4-dibromobenzene. In each case the cap atom is colored green and the end group is colored red. Bonds are formed between end groups and the cap atoms are deleted upon bond formation.

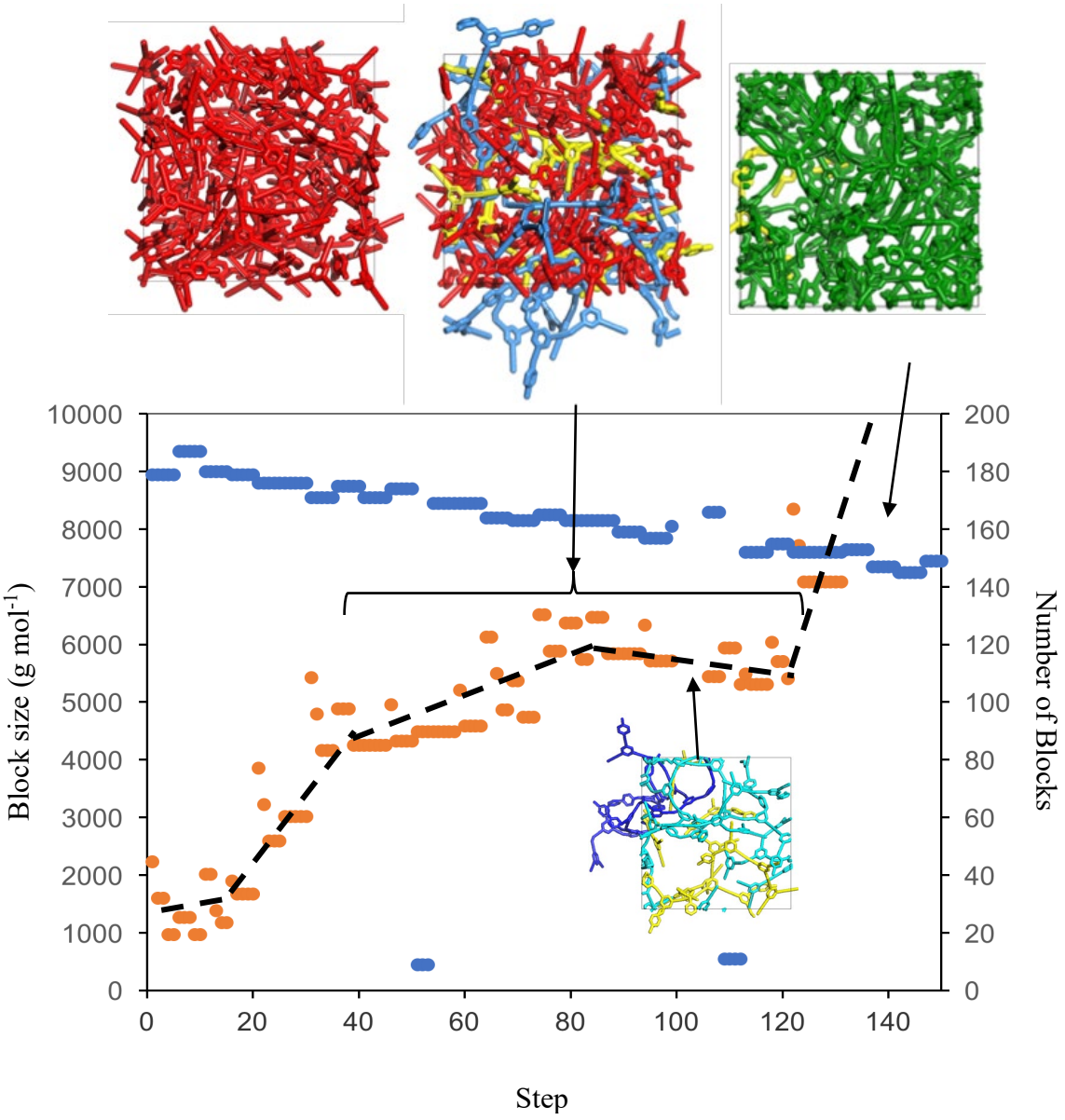

Figure 3. The orange dots show the molecular weight of the biggest block in the system. The blue dots show the number of blocks in the system for Model-1 with a solvation of 3. The network growth mechanism. The building blocks (red) coalesce to give small oligomers (yellow) that then give larger oligomers (blue) that become intertwined and gel-like. These are then able to very quickly form the polymer network (green). Snapshots of the network growth at steps 1, 120, 220, and 320. Catalyst and solvent not shown. Inset shows the gelation of two larger oligomers (blue) and smaller oligomers (yellow).

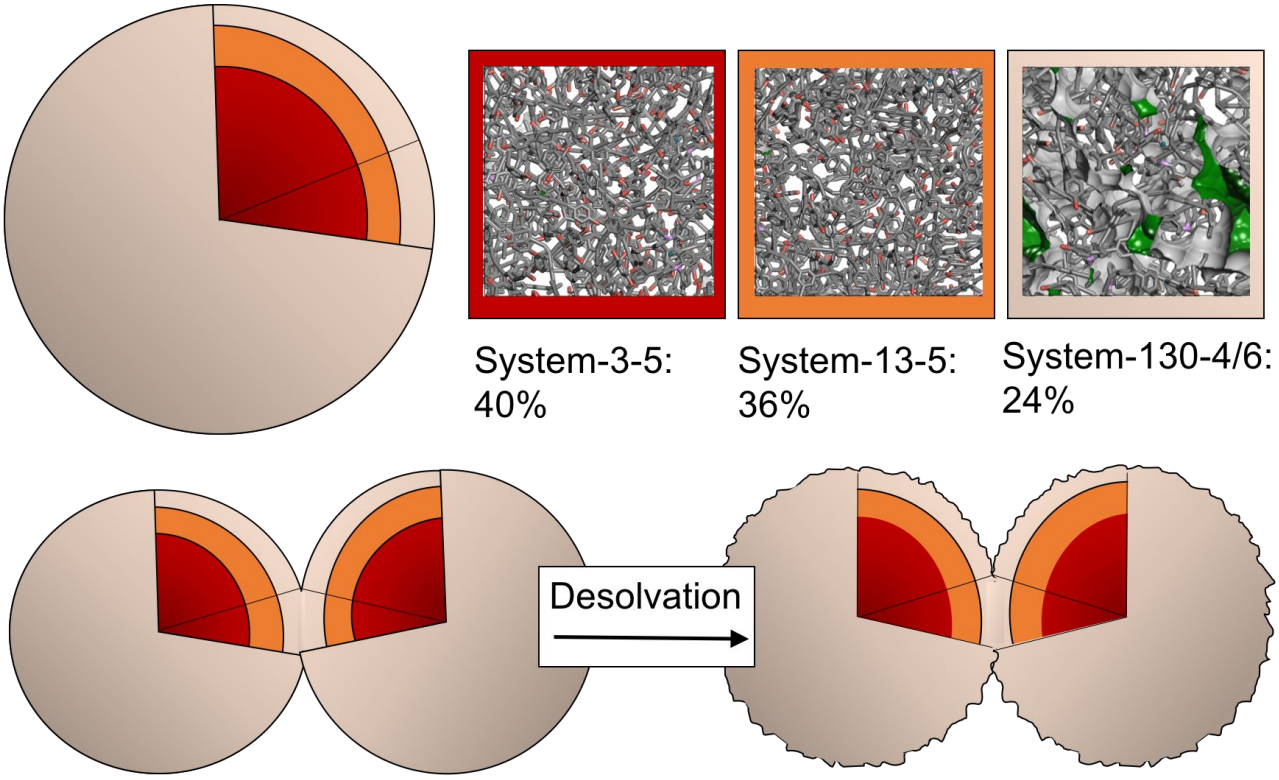

Figure 4. The proposed resulting structure of CMP-1. Spherical particulates are formed that have a dense core (40% of the particulate volume) formed in a phase with limited solvation of the polymer building blocks. Surrounding the core is a region that formed in higher solvation regions (36% and 24% of the particulate volume). When the spherical particulates meet, they are able to fuse together through alkyne-alkyne homocoupling or alkyne-bromine cross coupling. Upon desolvation, only solvent is removed from the dense core regions leaving behind a non-porous area that does not show any decrease in volume. Solvent, catalyst, and polymer building blocks are removed from the outer regions leaving behind a porous structure that has reduced in volume by 4%. System 3-5 is synthesized with 3 solvent molecules per unit cell and is desolvated using strategy 5 (only solvent molecules are removed) and is represented in the spheres by the red volume. System 13-5 is synthesized with 13 solvent molecules per unit cell and is desolvated using strategy 5 (only solvent molecules are removed) and is represented in the spheres by the orange volume. System 130-4/6 is synthesized with 130 solvent molecules per unit cell and is desolvated using either strategy 4 (solvent and building block molecules are removed) or strategy 6 (solvent and building block molecules are removed and homo coupling is allowed) and is represented in the spheres by the beige volume.

Solvation (number of DMF)	Desolv. strategy	Average He MP volume $cm^3\ g^{-1}$	Average Density $g\ cm^3$	Average AccSA $m^2\ g^{-1}$	Average PLD Å	Average MPD Å	Elemental Analysis		
							wt% C	wt% H	wt% Remaining
3	5	0.10	1.17	0	1.85	4.94	63.20	2.56	34.23
13	5	0.21	1.06	126	2.74	7.19	62.78	2.63	34.60
130	4	0.86	0.54	2541	7.67	13.95	81.89	3.05	15.06
130	6	0.74	0.68	1561	6.00	11.50	78.34	2.73	19.19
Experiment*		0.23	1.13	564	Peak at 15-20 Å		80.90	3.48	15.58

Table 1. Density and surface area properties of the simulated CMP-1 models with different desolvation strategies. Geometric surface areas calculated using Poreblazer for artificially synthesized CMP-1 averaged for Models 1-4. *Experiment is from Layborn *et al.*[43] Surface area is micropore surface area only.

Table of Contents Image