
Artificial Synthesis and Characterisation of Amorphous Microporous Framework Materials

Catherine Mollart

Chemistry

Lancaster
University



This thesis is submitted for the degree of Doctor of Philosophy in the
Chemistry department at Lancaster University.

August 2024

Catherine Mollart

Artificial Synthesis and Characterisation of Amorphous Microporous Framework Materials

Abstract

In this thesis, four microporous polymer materials are simulated using an artificial synthesis protocol. The resulting models are then compared to experiment to rationalise the structure and properties.

The first material discussed is CMP-1, used to rationalise the influence of reaction solvent choice on the porosity of conjugated microporous polymers (CMPs). It was established that the polarity of the solvent relative to the monomer building blocks is crucial to the formation of the polymer framework and resulting pore structure.

The second material is a hypercrosslinked polymer (HCP), which, when loaded with azobenzene, shows differing porosities and gas uptakes depending on the presence and isomer of azobenzene. These differences were rationalised due to changes in the micropore region of the pore size distribution, and the ability of *cis*-azobenzene to interact with carbon dioxide *via* dipole-quadrupole interactions, rather than the ability of the carbon dioxide to diffuse throughout the material.

The third material is organically synthesised porous carbon (OSPC)-1, composed of sp^3 hybridised carbon nodes connected by sp hybridised carbon linkers. The solid-state nuclear magnetic resonance (NMR) spectrum of this framework appears to show an alternative structure, and it was rationalised that the framework is composed of dense, interpenetrated, and non-porous polymer chains, surrounded by a thin shell of open, porous OSPC-1, explaining the experimental porosity and NMR.

Finally, CTF-1, a covalent triazine framework, is simulated to rationalise the differences in the Fourier-transform infrared spectra of CTF-1 synthesised using high-temperature ionothermal conditions, and the same material synthesised at room temperature, named P1. It was established that the amorphous P1 structure is able to incorporate additional structural diversity within the system, composed of neutral intermediates and alternative ring features formed during the kinetically controlled reaction. The remaining peaks in the spectrum were rationalised as absorbed guest molecules that interact favourably with the framework.

Acknowledgements

Firstly, I would like to thank my supervisor, Professor Abbie Trewin, without whom this PhD would not have been possible. Thank you, Abbie, for your guidance, support, and Microsoft Teams whiteboards throughout my academic career to date! It has been a pleasure to work for you, and I am very grateful for all of the opportunities you have given me throughout my PhD (including nominating me for the FST Dean's Award, we got there in the end!).

I would also like to thank my parents, David and Anita Mollart, for their constant support, even whilst I was "copying and pasting Coldplay album covers" (aka generating structures of amorphous microporous polymers following the full synthetic protocol and catalytic mechanism). A special thanks must also go to Pauline Pattinson and Caroline Nurse for their positivity and faith in my abilities.

Thank you to Kurtis Stanistreet-Welsh, Amy Crisp, and Ellena Sherrett, for the everyday rants when computers suffered power cuts or simulations didn't work first time round, the therapy was much appreciated!

Additionally, I would like to thank all of the collaborators I have had the privilege to work with throughout my PhD, in particular Dr Sam Lau and Dr Aotian Liu, and the summer and final-year undergraduate project students I have mentored, especially Jess Steele, Maria Ozog, and Sarah Holcroft.

Thank you to Dr Luís Pinho for his help with preparing for my viva and for his many questions regarding my work; Harry Brough for acting as chief proof-reader for my theory chapter, and Naseem Aktar for her positivity and Strictly Come Dancing expertise.

Thank you to Dr Michael Peach for educating me on the difference between 'blue' and 'teal', my colour palette has been transformed since. Thank you also for your expertise with Gaussian, LaTeX, and mini Colin the caterpillar cakes.

And finally, last but absolutely not least, I would like to thank the Chemistry department at Lancaster University for funding my PhD; and the Armourers and Brasiers' Company, Royal Society of Chemistry, and Lancaster University Graduate College for providing me with travel grants to attend UK-based and international conferences, which have really helped to broaden my scientific knowledge.

Declaration

I declare that the thesis has been composed by myself, and that the work has not be submitted for any other degree or professional qualification. I confirm that the work submitted is my own, with the exception of collaborative work and publications where specifically indicated. I also declare that the length of this thesis (67,778 words) does not exceed the maximum of 80,000 words.

The work within Chapter 4 would not have been possible without the experimental data provided by Dr Aotian Liu, Dr Cher Hon Lau and Professor Xianfeng Fan. I gratefully acknowledge the research provided by these collaborators.

Unless otherwise stated, all figures, schemes and tables were produced by the author.

Publications

1. Conjugated microporous polymer frameworks for sustainable energy materials – elucidating the influence of solvents on the porosity properties for future design principles
C. Mollart and A. Trewin*, *J. Mater. Chem. A*, 2024, **12**, 4159–4168.
Contribution: conceptualisation, data collection, analysis and manuscript preparation. This work forms the basis of **Chapter 3**.
2. Anhydrous Proton Conduction Through a Chemically Robust Electrolyte Enabling a High-Temperature Non-Precious Metal Catalyzed Fuel Cell
J. Zou, Y. Zhao, C. Mollart, M. J. G. Peach, P. Fayon, P. Heasman, P. A. T. J. Fletcher, J. Xu, W. Liang, A. Trewin* and T. Ben*, *Small*, 2024, **20**, 2308677.
Contribution: data collection and analysis.
3. Mesoscale artificial synthesis of conjugated microporous polymers
C. Mollart*, B. Ciborowski, and A. Trewin*, *Mol. Syst. Des. Eng.*, 2023, **8**, 1456–1461.
Contribution: conceptualisation, data collection, analysis and manuscript preparation. This work is briefly mentioned in **Chapter 3**.
4. Photo-Modulating CO₂ Uptake of Hypercross-linked Polymers Upcycled from Polystyrene Waste
A. Liu, C. Mollart, A. Trewin, X. Fan, and C. H. Lau*, *Chem. Sus. Chem.*, 2023, **16**, e202300019.
Contribution: data collection, analysis and contribution to the manuscript. This work forms the basis of **Chapter 4**.
5. Artificial synthesis of covalent triazine frameworks for local structure and property determination
C. Mollart, S. Holcroft, M. J. G. Peach, A. Rowling and A. Trewin*, *Phys. Chem. Chem. Phys.*, 2022, **24**, 20025–20029.
Contribution: conceptualisation, data collection and analysis. This work forms the basis of **Chapter 6**.
6. Rationalising the influence of solvent choice on the porosity of conjugated microporous polymers
C. Mollart and A. Trewin*, *Phys. Chem. Chem. Phys.*, 2020, **22**, 21642–21645.
Contribution: conceptualisation, data collection, analysis and manuscript preparation. This work contributes to **Chapter 3**.

7. Artificial Synthesis of Conjugated Microporous Polymers via Sonogashira-Hagihara Coupling

J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon, and A. Trewin*, *J. Phys. Chem. B*, 2020, **124**, 7318–7326.

Contribution: data collection and analysis. This work contributes to **Chapter 3**.

8. A new combined computational and experimental approach to characterise photoactive conjugated polymers

C. Mollart, P. Heasman, P. A. T. J. Fletcher, P. Fayon, J. M. H. Thomas, V. Franckevičius, M. J. G. Peach* and A. Trewin*, submitted to *Adv. Mater.* in July 2024.

Contribution: data collection and analysis.

In preparation

1. OSPC-1: Savvy use of computational NMR to characterise the trickster anode material

C. Mollart, A. Rowling, E. Sherrett, M. J. G. Peach and A. Trewin*, target submission to *J. Am. Chem. Soc.* in Summer 2024).

Contribution: data collection and analysis. This work forms the basis of **Chapter 5**.

2. Influence of Catalyst Choice on the Preparation of CMP networks

E. Sherrett, C. Mollart, A. Rowling and A. Trewin*, target submission Summer 2024.

Contribution: data collection and analysis.

3. Artificial synthesis of TCNQ-CTF

C. Mollart, E. Sherrett, A. Rowling and A. Trewin*, target submission Summer 2024.

Contribution: conceptualisation, data collection, analysis and manuscript preparation.

Contents

1	Introduction	1
2	Theory and Methodology	5
2.1	Summary	5
2.2	Computational theory	5
2.2.1	Introduction	5
2.2.2	Determining the energy of a chemical system	5
2.2.3	Geometry optimisation and the potential energy surface	23
2.2.4	Molecular dynamics	25
2.2.5	Determining charge distribution	30
2.2.6	Summary	30
2.3	Experimental theory	30
2.3.1	Porosity	30
2.3.2	Solubility	35
2.3.3	X-ray Photoelectron Spectroscopy	37
2.3.4	Nuclear Magnetic Resonance Spectroscopy	38
2.3.5	Fourier-Transform Infrared Spectroscopy	40
2.3.6	X-ray Diffraction	41
2.3.7	Microscopy	42
2.4	Methodology	42
2.4.1	Ambuild	42
2.4.2	HOOMD-blue	51
2.4.3	Poreblazer	53
2.4.4	Materials Studio	54
2.4.5	DL_POLY	54
2.4.6	Visual Molecular Dynamics	55
2.4.7	Gaussian	55
2.4.8	CASTEP	56
3	Artificial synthesis of a conjugated microporous polymer as a case study for rationalising the porosity properties of CMP materials	57
3.1	Summary	57

3.2	Introduction	58
3.3	Methodology	84
3.4	Results and Discussion	89
3.4.1	Degree of Solvation Systems	89
3.4.2	Phase Separated Systems	93
3.4.3	Solvent/Monomer Phase Interface Systems	97
3.5	Conclusions	112
4	Artificial synthesis of a hypercrosslinked polymer as a case study for rationalising the gas uptake properties of HCP materials	114
4.1	Summary	114
4.2	Introduction	115
4.3	Methodology	126
4.4	Results and Discussion	139
4.4.1	Comparison of pore size distribution to experiment	139
4.4.2	CO ₂ diffusion and uptake	147
4.5	Conclusions	152
5	Artificial synthesis of organically synthesised porous carbon as a case study for rationalising the solid-state NMR spectra of OSPC materials	154
5.1	Summary	154
5.2	Introduction	155
5.3	Methodology	166
5.3.1	OSPC-1 fragment models	166
5.3.2	OSPC-1 periodic models	168
5.4	Results and Discussion	170
5.4.1	OSPC-1 fragment models	170
5.4.2	OSPC-1 periodic models - crystalline topologies	174
5.4.3	OSPC-1 periodic models - chain structures	177
5.4.4	Energy profiles	182
5.4.5	Full NMR spectrum modelling	189
5.4.6	Structure rationalisation	195
5.5	Conclusions	198
6	Artificial synthesis of a covalent triazine framework as a case study for rationalising the FT-IR spectra of amorphous CTF materials	199
6.1	Summary	199
6.2	Introduction	200
6.3	Methodology	208
6.4	Results and Discussion	213

6.4.1 Network growth	213
6.4.2 Structure analysis and comparison to experiment	216
6.5 Conclusions	223
7 Conclusions	225
7.1 Future work	227
Appendices	228
A	228
B	242
C	245
D	255
Bibliography	278

Chapter 1

Introduction

Chemistry begins in the stars.

Peter Atkins

Porous materials, which contain open and interconnected, and/or closed and occluded voids throughout (Figure 1.1a),^{1,2} have been a known concept for a number of years. The first known porous materials were natural frameworks such as honeycomb,^{3,4} sponges^{5,6} and zeolites.⁷ Inspired by these, synthetic zeolites⁸ and zeo-types (such as zeolitic imidazolate frameworks),⁹ metal-organic frameworks (MOFs),¹⁰ covalent organic frameworks (COFs),¹¹ and microporous organic polymers (MOPs)¹²⁻¹⁷ were later introduced to the field (Figure 1.1b).

Synthetic zeolites, MOFs and COFs are crystalline materials, often exhibiting high porosity due to the clearly defined pore structure within the network,⁵ for example, MOF-210 exhibits an ‘ultra-high’ Brunauer-Emmett-Teller (BET) surface area of 6240 m² g⁻¹.¹⁰ In contrast, MOPs are typically, with a few exceptions in the case of covalent triazine frameworks (CTFs),¹⁴ fully amorphous materials with no defined unit cell.⁵ There are a number of subclasses that fall under the MOP categorisation, shown in Figure 1.2, with some materials falling into more than one subclass.¹⁸ The common feature arising in each class is the microporous nature of each polymer, indicative of pores within the micropore classification defined by the International Union of Pure and Applied Chemistry (IUPAC) of less than 2 nm in diameter.^{1,2} MOPs are typically more chemically and thermally stable than MOFs and COFs due to their hypercrosslinked polymer skeleton,¹⁹ however, the amorphous nature of these materials does typically cause a reduction in the porosity,²⁰ with one exception to this being PAF-1, a porous aromatic framework with a BET surface area above 5600 m² g⁻¹.¹⁵ Where relevant, the subclass of MOP is described in detail within the respective results chapter in this thesis.

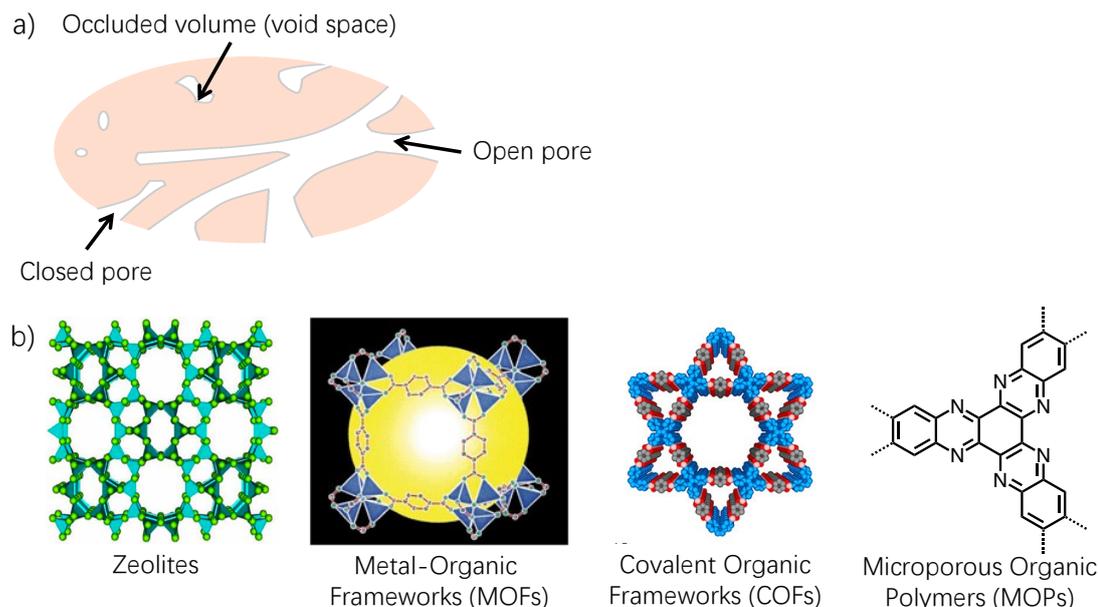


Figure 1.1: Illustration of porous materials. a) Cartoon representation of a porous material indicating the different types of pore. b) Examples of a zeolite, MOF, COF, and MOP material. The zeolite figure is reproduced and adapted with permission from reference 7, licensed under CC-BY. Copyright 2017 Scientific & Academic Publishing. The MOF figure is reproduced with permission from reference 21 with permission from the Royal Society of Chemistry. The COF figure is reprinted (adapted) with permission from K. Geng, T. He, R. Liu, S. Dalapati, K. T. Tan, Z. Li, S. Tao, Y. Gong, Q. Jiang and D. Jiang, *Chem. Rev.*, 2020, **120**, 8814–8933. Copyright 2020 American Chemical Society.

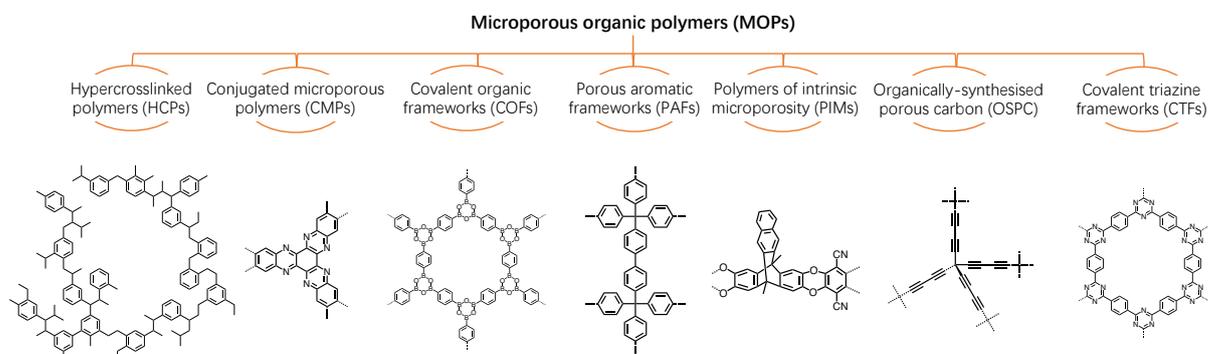


Figure 1.2: The different types of MOP material, with illustrations of each structure type underneath. Structures left to right: Styrofoam-based HCP, aza-CMP, COF-1, PAF-1, triptycene-based PIM, OSPC-1, CTF-1.

Porous materials have a large range of applications, including energy storage,^{16,22,23} gas uptake and capture,^{24–29} molecular separations and sieves,^{7,30} heterogeneous catalysis,^{31,32} proton conduction,³³ medical applications,³⁴ and wastewater treatment.²⁴ These are fuelled by the plethora of subclasses within the porous materials field, with particular types of porous material being more suited to certain applications than others, and this has led to many reports being published within this field.

The atomic structure of crystalline porous materials such as MOFs and COFs can be easily rationalised using x-ray diffraction (XRD) due to the clearly defined unit cell. However, amorphous MOPs, which are formed under kinetic control and therefore have no clearly defined unit cell, are very challenging to characterise experimentally.²⁰ This means that a characterisation toolkit must be utilised, with each item in the toolkit revealing a different piece of the atomic structure puzzle.²⁰ This holistic approach includes Fourier-transform infrared spectroscopy (FT-IR), x-ray photoelectron spectroscopy (XPS) and elemental analysis, which dictate the functional groups present in the material, including those arising from residual monomers.²⁰ Additionally, solid-state nuclear magnetic resonance (ssNMR) can establish the extent of bonding within the material and, in the case of ¹³C ssNMR, the carbon environments within the structure.²⁰ Finally, the porosity of the materials can be characterised using a pore size distribution (PSD) function to show the range of pore sizes and BET and/or Langmuir analysis to give the surface area and adsorption isotherm of the framework.^{20,35-37} More details about each characterisation technique will be covered in Chapter 2 section 2.3.

However, whilst using this plethora of experimental techniques within the characterisation toolkit does help to understand the structure of amorphous MOPs, simulation is essential to fully rationalise the structure on the atomic scale due to the many different structural configurations that can yield the same experimental characterisation results.^{20,38} Simulation must therefore be included within the characterisation toolkit with the caveat that it is necessary to ensure that the simulated structure is representative of the experimental material that it is being compared to.³⁸ For this reason, the Ambuild simulation code was used to model the polymer structures simulated as part of this thesis as it allows fine control over the synthetic protocol, and is able to mimic each step in the catalytic mechanism in an approach termed ‘artificial synthesis’.¹⁹ More details about Ambuild are given in Chapter 2 section 2.4.1. Other simulation codes that can be used to model these structures, which do not have the same control over the artificial synthesis approach, include Poly-matic³⁹ and Zebedde,⁴⁰ which have been previously used to model PIMs⁴¹ and hypercrosslinked polymers,⁴² and CTF fragments,^{43,44} respectively.

This thesis demonstrates the use of simulation to generate structural models of amorphous MOPs *via* artificial synthesis, following the full synthetic mechanism and catalytic pathway to best compare to the experimental material. When these models were subsequently characterised and analysed, each simulated material gave insight into the experimental structure and was able to rationalise the experimental properties and/or spectrum in each case (Figure 1.3). Each results chapter discusses a different MOP material and the rationalisation of a different experimental property or spectrum. It is antici-

pated that in the future, the approaches described within this thesis can be exploited to pre-screen materials before synthesis to establish which are the best contenders for a particular application.

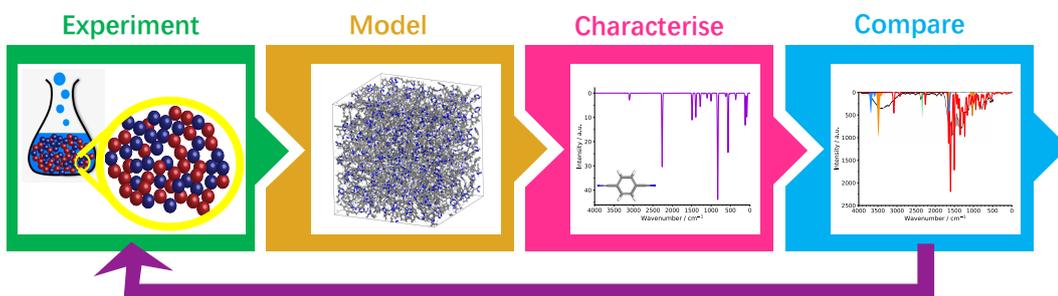


Figure 1.3: The general approach taken within this thesis, where a representative structural model of the experimental material is artificially synthesised. Fragments of the simulated models are then analysed using a variety of computational characterisation techniques. Finally, the simulated structure, properties and spectra are compared back to the experimental, showing comparable properties and/or spectra in each case. Structural model and FT-IR spectra were reproduced and adapted from reference 45 with permission from the PCCP Owner Societies.

Chapter 2

Theory and Methodology

Nobody understands quantum mechanics.

RP Feynman

2.1 Summary

This thesis involves the simulation of amorphous microporous frameworks to rationalise the experimental properties of each material studied. Each method utilised to simulate the materials discussed here is described within the ‘Computational theory’ section of this chapter. Following this, the ‘Experimental theory’ section details the background knowledge required for each technique employed to compare to the simulated properties. Finally, the ‘Methodology’ section details the numerous software codes used within the work described in this thesis.

2.2 Computational theory

2.2.1 Introduction

This section describes the background of some common approaches taken to simulate chemical systems. The approaches utilised within this thesis are then covered in greater depth.

2.2.2 Determining the energy of a chemical system

The Schrödinger equation

The time-dependent Schrödinger equation,⁴⁶⁻⁵¹ where i is the imaginary unit defined by $i^2 = -1$, \hbar is the reduced Planck’s constant, Φ is the wavefunction,

R are the nuclear positions, t is the time, and \hat{H} is the Hamiltonian operator, is given in Equation 2.1.⁴⁶⁻⁵²

$$i\hbar \frac{d\Phi(R, t)}{dt} = \hat{H}\Phi(R, t) \quad (2.1)$$

The time-dependent Schrödinger equation describes how an atomic, ionic, or molecular system evolves over time, based on its wavefunction.^{46-51,53} The wavefunction gives information about the positions of each particle in the system. The positions of the particles of an N particle system are described by the coordinates $R = (R_{1x}, R_{1y}, R_{1z}, R_{2x}, \dots, R_{Nz})$, where, for example, R_{1x} is the x -coordinate of nucleus 1. The probability that the particles are located in these positions is defined by the square of the absolute value of the wavefunction, which gives a probability density function at each timestep.⁵² \hat{H} contains terms that describe the potential and kinetic energy of each particle in the system. When operating on the wavefunction, the Hamiltonian operator informs how the wavefunction varies in time.^{53,54}

This indicates that if the wavefunction of the system is known, it is possible to calculate how it will change in time by solving the right-hand side of the Equation 2.1.⁵³ For example, if the wavefunction of the initial system configuration at time $t = 0$ is known, it can be introduced into the time-dependent Schrödinger equation (Equation 2.1) along with the compiled Hamiltonian operator for the system, and the right-hand side of Equation 2.1 solved.⁴⁶⁻⁵² This would indicate the rate of change of the wavefunction over time, δt , which could then be used to calculate the wavefunction at time $= 0 + \delta t$, repeated to assess how the system evolves throughout a larger timescale.⁵³

However, as the time-dependent Schrödinger equation⁴⁶⁻⁵¹ (Equation 2.1) is dependent on calculating the wavefunction of the system using the coordinates of every nucleus and electron present, at each time evaluated, the calculation becomes incredibly complex very quickly. This means that it is not currently possible to solve in a reasonable time for even small molecules.⁵³

The first approximation made is that the dynamical contribution to the equation can be considered independently from the remaining terms, as many systems can be described sufficiently using a time-independent wavefunction.⁵³ This means that the wavefunction of a particular system type and configuration is only dependent on the positions of the particles, R , regardless of the time.⁵³

Secondly, the Born–Oppenheimer approximation can be applied,⁵⁵ which states that the electronic wavefunction $\Psi(r)$, where r describes the electronic coordinates, can be evaluated at a fixed nuclear configuration. This would correspond to a nuclear kinetic energy of zero.⁵³ The total wavefunction is then

represented as the product of the electronic $\Psi(r)$ and nuclear $\Theta(R_N)$ wavefunction components. This approximation can be made due to the larger mass ($> 1800\times$) of nuclei in comparison to electrons. This approximation reduces the number of spatial coordinates that the wavefunction depends on, making the calculation easier to compute.^{53,55} Importantly, the interactions between electrons and nuclei are not substantially affected due to the ability of the wavefunction to be separated into electronic and nuclear components, where in the electronic solution, the nuclei are considered as an external potential.⁵³

As a result of these two approximations, the electronic time-independent Schrödinger equation (Equation 2.2)⁴⁶⁻⁵¹

$$\hat{H}_{\text{elec}}\Psi(r) = E\Psi(r) \quad (2.2)$$

was established, which relates the total electronic energy of the system, E , at state Ψ to the Hamiltonian operator.⁴⁶⁻⁵¹ For any system, there will be multiple electronic configurations, leading to multiple wavefunctions and corresponding system energies. The minimum-energy structure is known as the ground state, indicating the most favourable system configuration, and all other solutions to the equation are referred to as excited states.⁵³

The Hamiltonian operator can then be expanded as given, in atomic units, in Equation 2.3,

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^P \sum_{i=1}^N \frac{Z_A}{r_{iA}} + \sum_{j>i}^N \sum_{i=1}^N \frac{1}{r_{ij}} + \sum_{B>A}^P \sum_{A=1}^P \frac{Z_A Z_B}{r_{AB}} \quad (2.3)$$

where the first term defines the kinetic energy of the electrons, \hat{T} , the second term is the electron-nuclear attraction, $\hat{V}_{\text{N-e}}$, the third term is the electron-electron repulsion, $\hat{V}_{\text{e-e}}$, and the final term is the nuclear-nuclear repulsion, $\hat{V}_{\text{N-N}}$. Within the Born-Oppenheimer approximation,⁵⁵ the nuclear-nuclear repulsion energy is constant.

Whilst the time-independent Schrödinger equation is much more applicable to calculate the properties of chemical systems due to its reliance on a smaller number of terms in comparison to the time-dependent Schrödinger equation, there still comes a point where the systems are too large to be able to calculate the exact solutions for in a reasonable timeframe.⁴⁶⁻⁵¹ This means that throughout computational chemistry, a variety of approximations are applied to solve the time-dependent Schrödinger equation for a particular system.⁴⁶⁻⁵¹ Depending on the degree of accuracy required in the calculated properties and the system size, different approximations can be realistically employed.

In attempting to solve the time-dependent Schrödinger equation,⁴⁶⁻⁵¹ wavefunction-based methods such as Hartree–Fock (HF)⁵⁶⁻⁶⁰ use Slater determinants to describe the system.^{53,61-63} The separability of the wavefunction, which is the basis of the Born–Oppenheimer approximation,⁵⁵ also means that a many-electron wavefunction can be approximated as a product of many one-electron wavefunctions. However, the wavefunction must be anti-symmetrised to satisfy both the Pauli principle⁶⁴ and the indistinguishability of electrons from one another. Slater determinants^{62,63} are an efficient way to represent an anti-symmetrised many-electron wavefunction. A generic Slater determinant^{62,63} for an N -electron system is given below. The $\frac{1}{\sqrt{N!}}$ pre-factor is used as a normalisation factor to ensure that the wavefunction contains N electrons and is correctly interpretable as a probability density. ψ_i describes the wavefunctions of each electron, and \mathbf{x}_i describe the positions of each electron.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix}$$

Hartree–Fock theory uses a single Slater determinant to describe the system.^{56-60,62,63} In contrast to HF, correlated wavefunction methods, such as configuration interaction and coupled-cluster calculations, linearly combine multiple Slater determinants to describe the system. Each Slater determinant represents excited electron configurations to describe electron correlation effects. Møller–Plesset perturbation theory similarly uses information about excited configurations to include electron correlation *via* perturbative corrections. These correlated wavefunction methods offer an improvement over the standard Hartree–Fock approach, typically leading to more accurate ground-state properties.^{53,56-63} However, these can only be computed within a reasonable timescale for very small systems containing tens of atoms,⁶⁵ so are not relevant for the systems discussed here.

Each approximate solution to the time-independent Schrödinger equation utilised within this thesis is discussed below. Figure 2.1 shows the comparison of the relative system sizes that can be computed with each method, the timescale, and length scale.

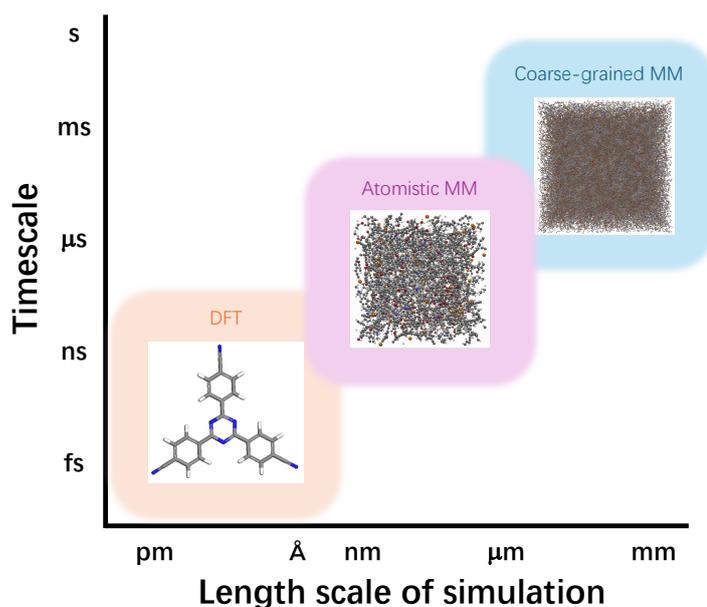


Figure 2.1: Plot detailing some of the common approximations to the time-independent Schrödinger equation⁴⁶⁻⁵¹ used within computational chemistry, as a function of system size, timescale and length scale. MM - molecular mechanics, DFT - density functional theory.

Density functional theory

A second approach is density functional theory (DFT),⁶⁶⁻⁶⁹ used to calculate an approximate solution to the time-independent Schrödinger equation⁴⁶⁻⁵¹ for small systems consisting of approximately hundreds of atoms.⁶⁵ Instead of relying on the wavefunction, which requires many variables, to calculate the energy of the system, and therefore identify the ground state structure, it was proposed that the energy, $E[\rho(r)]$, could be calculated as a functional of the electron density, $\rho(r)$. The electron density can be described using only the x , y and z coordinates in three-dimensional space, making it a much simpler quantity to obtain.⁵³ The smaller number of required variables in comparison to wavefunction methods also makes DFT a less computationally expensive solution to the time-independent Schrödinger equation.⁴⁶⁻⁵¹ The electron density can be used to generate a probability distribution function of finding each electron throughout the system. The term ‘functional’ arises as the electron density is a function of the electronic coordinates, r , and the energy is a function of the electron density.⁵³ This approximation⁶⁶⁻⁶⁹ can be applied as the electron density provides sufficient information about the number of electrons and positions and charges of nuclei to define the Hamiltonian and the system energy. This satisfies all of the requirements of the Schrödinger equation.^{46-51,70}

The first exploration into DFT in the 1920s involved the approximation that the system is composed of homogeneous electrons spread with uniform den-

sity across an infinite space, known as jellium, or the uniform electron gas.⁵³ These electrons are able to interact in the presence of an external potential including both exchange and correlation, unlike HF, which does not consider electron correlation. Within the Born–Oppenheimer approximation,⁵⁵ where the nuclei are fixed and the nuclear–nuclear repulsion term is set to zero, $E[\rho(r)]$ is derived from the sum of the Coulombic repulsion between electrons, $J[\rho(r)]$, the kinetic energy, $T[\rho(r)]$, the Coulombic attraction between the nuclei and electrons, $E_{\text{N-e}}[\rho(r)]$, and the exchange–correlation (XC) functional, $E_{\text{XC}}[\rho(r)]$ (Equation 2.4).⁵³

$$E[\rho(r)] = J[\rho(r)] + T[\rho(r)] - E_{\text{N-e}}[\rho(r)] + E_{\text{XC}}[\rho(r)] \quad (2.4)$$

This is defined as the Thomas–Fermi–Dirac model.^{71–73}

The electron–electron repulsion can be defined in terms of the density. The classical part, $J[\rho(r)]$, is written as a functional in terms of the interaction between the densities at point \mathbf{r} and point \mathbf{r}' , given in Equation 2.5,

$$J[\rho(r)] = \int \frac{\rho(\mathbf{r}) \cdot \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r} \mathbf{d}\mathbf{r}'. \quad (2.5)$$

The kinetic energy, $T[\rho(r)]$, can be calculated exactly for the uniform electron gas using Equation 2.6,

$$T[\rho(r)] = C_{\text{TF}} \int \rho^{\frac{5}{3}}(\mathbf{r}) \mathbf{d}\mathbf{r}, \quad (2.6)$$

where C_{TF} is the Thomas–Fermi constant $(\frac{3}{10}(3\pi^2)^{\frac{2}{3}})$.^{72,73} Whilst the form of $T[\rho]$ for the uniform electron gas is known exactly, the form is not more generally known. However, an expression for $T_{\text{s}}[\phi]$, the non-interacting kinetic energy, which depends implicitly on the density *via* the Kohn–Sham orbitals ϕ , provides a valuable approximation to T .

The nuclear–electron attraction, $E_{\text{N-e}}[\rho(r)]$ is given in Equation 2.7,

$$E_{\text{N-e}}[\rho(r)] = \int \rho(\mathbf{r})\nu(\mathbf{r}) \mathbf{d}\mathbf{r}. \quad (2.7)$$

$\nu(r)$ is the external potential experienced by the electrons as a result of the positions and charges of the nuclei in the system, Equation 2.8,

$$\nu(\mathbf{r}) = - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|}. \quad (2.8)$$

Finally, the exchange–correlation energy functional, $E_{\text{XC}}[\rho(r)]$, is calculated as the differences in the exact and non-interacting kinetic energies, T and T_{s} , respectively, and the total and classical electron–electron repulsions, E_{ee} and

J , respectively, given in Equation 2.9

$$E_{\text{XC}}[\rho(r)] = (T - T_s) - (E_{ee} - J). \quad (2.9)$$

However, despite the uniform electron gas theory having sound promise, the results obtained from applying the approximation that a uniform electron gas model could be used to describe a real chemical system were not particularly accurate, and further development of this theory was not undertaken for another 30 years.^{53,66}

In 1964, Hohenberg and Kohn expanded this theory by proving that if the exact exchange–correlation functional were known, and the exact system density used as an input, the exact energy would be obtained.⁶⁶ Importantly, any approximate density would yield a higher energy. Hohenberg and Kohn proposed a variational principle, which states that when comparing the energies obtained by multiple density inputs, the density which corresponds to the lowest energy is the most accurate. Hartree–Fock,^{56–60} which does not rely on an exchange–correlation functional, is variational and considered a bound of the real energy, where the exact energy is always lower than or equal to the HF energy.⁶⁶

However, whilst the general form of Equation 2.4 has been established for a number of years and the electron–electron repulsion and nuclear–electron attraction can be calculated, the challenge remained of how to evaluate the kinetic energy and XC functional terms.⁵³ Based on this, Kohn and Sham developed the current standard approach where the system density of a non-interacting n electron system can be expressed as the sum of densities across a set of occupied one-electron Kohn–Sham molecular orbitals, ψ_i (Equation 2.10).^{53,67–69}

$$\rho(r) = \sum_i^n |\psi_i(r)|^2 \quad (2.10)$$

These molecular orbitals can be used to calculate the non-interacting kinetic energy of the system.⁵³ Kohn and Sham proposed that the non-interacting system could be used to represent the density of a fully-interacting system with the same number of electrons by applying an external potential to generate a density equal to the ground state density of the fully interacting system.^{67–69} This approximation to the kinetic energy term means that only the XC functional remains to be fully resolved. If the exact XC functional is found, the DFT calculation will yield an exact solution to the Schrödinger equation.^{46–51,53} Therefore, there is a demand for the development of new XC functionals that are best able to approximate this term, giving a better quality of calculation that leads to more accurate structures and corresponding properties.⁵³

Currently, there are a variety of different classes of XC functionals that can be utilised as the $E_{\text{XC}}[\rho(r)]$ term within Equation 2.4, with each class occupying a different rung on ‘Jacob’s Ladder of Chemical Accuracy’ (Figure 2.2).⁷⁴ Moving up the ladder, more detailed information regarding the density throughout the system is available, leading to a better description and a more accurate XC functional. However, the computational expense also increases moving up the ladder, meaning it is important to decide what level of accuracy is required to describe the system of interest.^{53,74}

The lowest quality XC functionals, and the least expensive to compute, are the local XC functionals, such as the local density approximation (LDA),⁷⁵⁻⁷⁷ which require only the electron density at a given set of coordinates. For LDA,⁷⁷ the XC functional term (Equation 2.11)^{70,78-80}

$$E_{\text{XC}}^{\text{LDA}}[\rho(r)] = \int \rho(r) \cdot \epsilon_{\text{XC}}[\rho(r)] \, \text{d}r \quad (2.11)$$

is calculated by integrating over the entire volume of the system and evaluating the energy term associated with the electron density, $\epsilon_{\text{XC}}[\rho(r)]$, in each infinitesimal segment of the total volume. The energy term of each segment is obtained from summing the exchange, $E_{\text{X}}^{\text{Slater}}[\rho(r)]$ and correlation, $E_{\text{C}}^{\text{VWN}}[\rho(r)]$, components to the energy (Equation 2.12).⁷⁸⁻⁸¹

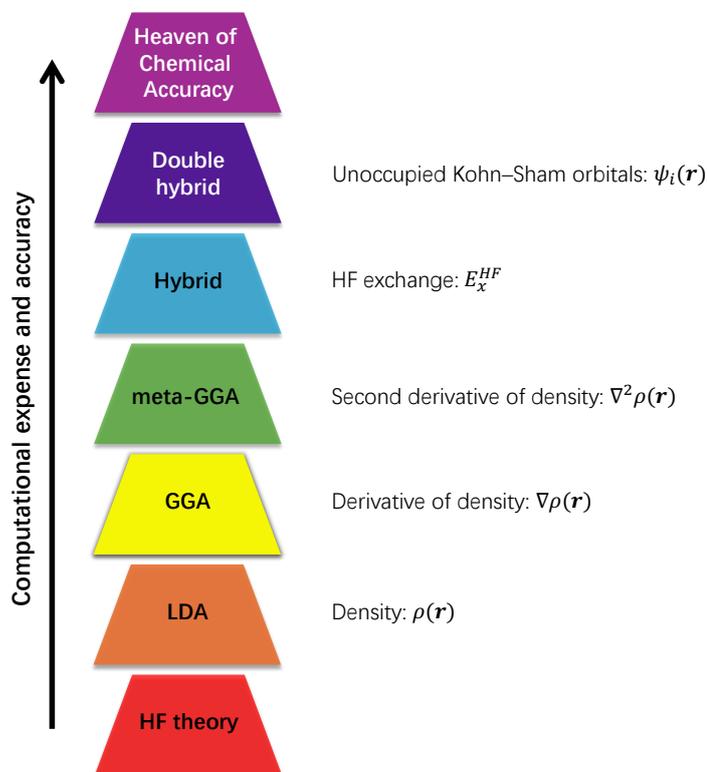


Figure 2.2: Jacob’s Ladder of Chemical Accuracy of exchange–correlation functionals. Figure reproduced and adapted from reference 82 with permission from The Materials Research Society, Springer Nature.

$$E_{\text{XC}}^{\text{LDA}}[\rho(r)] = E_{\text{X}}^{\text{Slater}}[\rho(r)] + E_{\text{C}}^{\text{VWN}}[\rho(r)] \quad (2.12)$$

The exchange term of each component is calculated by introducing the density of that segment into the equation proposed by Dirac and Slater for the exchange of the uniform electron gas (Equation 2.13),⁷⁸

$$E_{\text{X}}^{\text{Slater}}[\rho(r)] = -C_x \int \rho^{\frac{4}{3}}(r) \, \text{d}r \quad (2.13)$$

making the approximation that this can be used to describe the exchange of a non-uniform system. The correlation term is obtained using the method proposed by Vosko, Wilk and Nusair (VWN), where an approximate correlation is obtained for a given density through parametrisation of the correlation energies from high level calculations of a range of uniform densities.^{79,81} However, whilst this approach works reasonably well for periodic systems such as silicon, which are comparable with the uniform electron gas model, they do not describe chemical bonding well. This is because the uniform electron gas describes a system consisting of uniform electron density throughout. It is able to describe systems with slowly varying density, such as periodic systems, well as the density is more uniform than for systems containing localised electron density as a result of chemical bonding. This means that typically, higher quality XC functionals are preferred.⁵³

The next two rungs up ‘Jacob’s Ladder of Chemical Accuracy’⁷⁴ house the generalised gradient approximation (GGA) and meta-GGA XC functionals, such as PBE, BLYP and BP86 (GGA XC functionals),⁸³⁻⁸⁶ and the Minnesota and TPSS XC functionals (meta-GGAs).^{87,88} In addition to the electron density in each infinitesimal volume segment in the system, these also consider the first derivative of the electron density through x (the gradient, Equation 2.14),^{70,80,84-86}

$$E_{\text{XC}}^{\text{GGA}}[\rho(r)] = \int \rho(r) \cdot \epsilon_{\text{XC}}[\rho(r), \nabla\rho(r)] \, \text{d}r \quad (2.14)$$

and, in the case of meta-GGAs, the second derivative of the electron density (Equation 2.15),^{70,80,87,88} which is related to the kinetic energy density.^{53,80}

$$E_{\text{XC}}^{\text{meta-GGA}}[\rho(r)] = \int \rho(r) \cdot \epsilon_{\text{XC}}[\rho(r), \nabla\rho(r), \nabla^2\rho(r)] \, \text{d}r \quad (2.15)$$

The final rungs on ‘Jacob’s Ladder of Chemical Accuracy’⁷⁴ before reaching the ‘heaven of chemical accuracy’ contain hybrid XC functionals such as PBE0 and B3LYP,^{79,85,89-92} which add in some exact HF exchange, calculated using the occupied Kohn-Sham orbitals, into a GGA functional in order to benefit from the dynamic correlation obtained from the DFT approach but also to

reduce the quantity of non-dynamic correlation occurring within the exchange term of the GGA XC functional, as this is often overestimated. The XC term for a hybrid functional is generally written as in Equation 2.16

$$E_{XC}^{\text{hybrid}} = \alpha E_X^{\text{HF}} + (1 - \alpha) E_X^{\text{GGA}} + E_C^{\text{GGA}} \quad (2.16)$$

where some HF exchange energy is incorporated with a corresponding reduction in the exchange energy of the GGA. The quantity of HF exchange incorporated into the XC functional is often determined by calibration to near-exact calculations. However, the quantity of HF exchange incorporated within some XC functionals, such as PBE0, is determined non-empirically.^{80,89} The typically unmodified GGA correlation term is still included to incorporate dynamical correlation.^{53,80} Finally, double-hybrids, which rely on virtual Kohn-Sham orbitals, are the final, and most computationally expensive rung on ‘Jacob’s Ladder of Chemical Accuracy’.⁷⁴ These are very similar to ‘standard’ hybrid XC functionals, so are not discussed further.

A selection of XC functionals have been used within the results chapters of this thesis. They include B3LYP, a hybrid functional composed of 20% HF exchange⁵⁶⁻⁶⁰ and 80% DFT exchange.^{79,85,90,91} This functional was utilised to model the Fourier-transform infrared (FT-IR) spectra of covalent triazine frameworks in Chapter 6 section 6.4.2. This functional, in combination with the 6-31G* basis set, is well-established for reporting high quality vibrational frequencies.^{45,79,85,90,91,93,94} B3LYP was also used in Chapter 5 section 5.4.1 to calculate the nuclear magnetic resonance (NMR) spectra of molecular fragments of organically-synthesised porous carbon (OSPC)-1 due to the wide applicability of this functional.^{79,85,90,91,95} Perdew–Burke–Ernzerhof (PBE),⁸⁶ a GGA functional, was utilised to model the periodic OSPC-1 models in Chapter 5 sections 5.4.2 and 5.4.3.⁹⁵ PBE is a standard functional to model periodic systems due to the increased expense of these calculations compared to molecular DFT calculations requiring a less expensive XC functional. Finally, CAM-B3LYP⁹⁶ was utilised within Chapter 4 sections 4.3 and 4.4.2 to calculate the isomerisation barrier of azobenzene and the binding energies of carbon dioxide to *cis*- and *trans*-azobenzene, respectively.²⁶ CAM-B3LYP is a range-separated hybrid functional, where the ratios of HF and DFT exchange vary with the distance between electrons.

The XC functional, system coordinates, basis set, and a set of initial guess orbitals are specified as inputs to the DFT calculation, where the basis set provides a flexible description of the electron density. In the case of atom-centred basis sets, the basis set is composed of a set of functions representing the atomic orbitals within the system.⁵³ The Kohn–Sham equations are then solved for that geometry.⁶⁷⁻⁶⁹

At a given structure, the orbitals can be converged to a minimum-energy using self-consistent field (SCF) methods. The initial energy of the system, at a given structure, is calculated using a series of guess orbitals. The orbitals are then converged to self-consistency, yielding the lowest energy electron configuration for that method, within the chosen basis set. This is the process undertaken in, for example, single-point energy, frequency, and NMR calculations. However, as the energy is a functional of both the orbitals and the nuclear positions, the total electronic energy can be minimised by allowing the nuclear positions to change. This process is called geometry optimisation and works by allowing the orbitals to converge *via* the SCF procedure as above at a particular system configuration, re-converging the orbitals, and evaluating a new total electronic energy, to search for the lowest possible energy molecular configuration.

Whilst molecular DFT becomes challenging when the system size exceeds hundreds of atoms, there is the option to model larger solid-state systems using periodic DFT, which takes advantage of periodicity, or translational symmetry, within the structures. Bloch's Theorem states that for a periodic system, the potential energy, V , that a nucleus at position \mathbf{R} applies to its nearby electrons is identical to that at periodic displacements \mathbf{L} from \mathbf{R} (Equation 2.17).^{53,97,98}

$$V(\mathbf{R}) = V(\mathbf{R} + \mathbf{L}) \quad (2.17)$$

This means that the electron density, $\rho(\mathbf{r})$, will also be periodic throughout the system (Equation 2.18),^{53,97,98}

$$\rho(\mathbf{r}) = \rho(\mathbf{r} + \mathbf{L}). \quad (2.18)$$

As the electron density and square of the wavefunction both yield a probability distribution function of finding each electron throughout the system, the square of the wavefunction is also periodic.^{53,97,98} For this condition to be met, each wavefunction is written as the product of a complex number known as the plane-wave, $e^{i\mathbf{k}\mathbf{r}}$, and a spatial term, $u_n(\mathbf{r})$, which is equal to $u_n(\mathbf{r} + \mathbf{L})$ (Equation 2.19).^{53,97,98}

$$\psi_n(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} \cdot u_n(\mathbf{r}) \quad (2.19)$$

\mathbf{k} is the crystal momentum, a wavevector that takes a range of values to describe how ψ can vary, with each \mathbf{k} value corresponding to a different orbital in the system. In comparison with a molecular DFT calculation generating a set of molecular orbitals, the periodic calculation yields a set of bands. Each band contains a set of orbitals with a unique $u_n(\mathbf{r})$ value per band. The different or-

bitals making up the band each have a different \mathbf{k} value and energy.^{53,97,98}

When running DFT on a system, whether periodic or molecular, a basis set must also be specified to describe how the electron distribution should be represented throughout the system. This can either be done using localised atom-centred basis sets, such as Slater or Gaussian-type basis sets,^{62,63} or plane-waves.⁹⁹ Localised atom-centred basis sets use atom-centred basis functions, each containing angular and radial components, which are coupled together to describe the overall behaviour of the system. The angular component of the basis functions uses the spherical harmonics of the hydrogenic one-electron wavefunctions to describe the angular behaviour of the system, where including a larger number of angular momentum functions leads to increased flexibility within the basis set. Using a Gaussian-type basis set as an example, the radial component utilises Gaussian functions to describe the radial behaviour of the system. Whilst each individual Gaussian function cannot describe the electron behaviour exactly, they are very efficient to compute. Therefore, a large number of Gaussian functions can be utilised to realistically describe the system with reasonable computational expense.

In contrast to atom-centred basis sets, plane-waves are naturally delocalised, and hence suited to extended, periodic systems.⁹⁹ In practice, a series of plane-waves all oscillating at different frequencies are used to represent the electron distribution. These plane-waves can be combined in different contributions to describe the system, through interfering constructively and destructively to describe inhomogeneous electron distributions. An appealing feature of plane-wave basis sets is that their number (and hence accuracy and expense) can be systematically increased or reduced by changing the cutoff energy (Equation 2.20),⁹⁹

$$E_{\text{cutoff}} = \frac{1}{2}G_{\text{max}}^2 \quad (2.20)$$

which modifies the number of plane-waves included within the basis set.⁹⁹ This in turn leads to different flexibilities of interference, allowing finer or coarser localisation of electron density in different areas of the system. Combining more plane-waves by increasing the energy cutoff allows for a finer description of the electron density distribution, leading to a realistic representation.

Whilst Equation 2.19 describes how the wavefunction of a system can be written as a function of plane-waves in 'real' space, it is also important to understand how the wavefunction can be expanded into a basis set made of plane-waves in reciprocal space, or k-space, $e^{i\mathbf{G}\mathbf{r}}$. The plane-waves in k-space have

vectors that are reciprocal to the crystal lattice, \mathbf{G} (Equation 2.21)^{53,100}

$$\psi_n(\mathbf{r}) = \sum_G C_G \cdot e^{i\mathbf{G}\mathbf{r}}. \quad (2.21)$$

Reciprocal space occurs *via* the Fourier-transform of real space, and *vice versa*. The reciprocal space function, $F(\mathbf{Q})$ (Equation 2.22),¹⁰¹

$$F(\mathbf{Q}) = \int \rho(r) \cdot \exp(i\mathbf{Q} \cdot r) \, dr, \quad (2.22)$$

defines how the density of a system in real space, $\rho(\mathbf{r})$ can be described using waves of periodicity Q .¹⁰¹

Exploiting the periodicity within the system reduces the calculation size from that of an infinite material to that of the unit cell, making it possible to use DFT to calculate the properties of the infinite system in a reasonable timescale. During the calculation, the Brillouin zone, which is the minimum unit cell of the reciprocal lattice containing all points between the origin and the first Bragg plane,¹⁰² is sampled by taking a range of k -points which are spread throughout reciprocal space. Whilst increasing the number of k -points and cutoff energy increase the accuracy and quality of the calculation, this also increases the associated computational expense.^{53,98,99}

Atomistic molecular mechanics

Whilst the solutions to the time-independent Schrödinger equation⁴⁶⁻⁵¹ described so far involved using high-level quantum-mechanical approaches to calculate the potential energy of a molecule in a given geometry, these approaches quickly become challenging and computationally expensive as the system size increases beyond approximately a few hundred atoms. Additionally, depending on the choice of quantum mechanical approach and model chemistry (the choice of XC functional and basis set) utilised, the result of the calculation may not describe the system well, despite the computational expense required to calculate the system energy.⁵³

Therefore, an alternative approach, known as molecular mechanics, can instead be utilised to study systems containing approximately 100,000 atoms, typically studied within periodic unit cells to explore bulk effects.^{53,65} Here, each atom is treated as a spherical centre of mass with a charge and radius, typically the van der Waals radius, connected by chemical bonds, considered as springs.^{53,103} This increase in system size means that solvent molecules, which are typically present in excess in chemical synthesis, can be incorporated explicitly into the cell as atomistic building blocks.¹⁰⁴

The Born–Oppenheimer approximation is once again applied,⁵⁵ this time dis-

counting the electronic behaviour of the system and instead describing the potential energy of the system as a function of the nuclear coordinates, R .^{103,105} This means that whilst the electronic structure cannot be assessed using MM, larger-scale bulk effects such as gelation and phase separation can be studied with this approach.^{53,103}

When employing MM, the system is described using a forcefield, which contains the equilibrium bond lengths, angles, dihedrals, and intermolecular forces for each unique atom type within the structure. For example, sp^3 , sp^2 and sp hybridised carbon atoms are all distinct atom types, and even within each atomic hybridisation, the forcefield type will vary depending on what the atom is bonded to. The force constants associated with the forces interacting on each atom type, which define the energy penalty required to distort from the equilibrium values, are also specified within the forcefield.⁵³ These values are obtained semi-empirically from high-level calculations of small molecules or from experiment, and are then generalised for similar materials. The forcefield variables are then incorporated into equations to describe the energy of the system using classical mechanics, where the total potential energy $V(R)_{\text{total}}$ is given as the sum of each of the interaction types within the system (Equation 2.23).¹⁰³

$$V(R)_{\text{total}} = V(R)_{\text{bond}} + V(R)_{\text{angle}} + V(R)_{\text{torsion}} + V(R)_{\text{cross}} + V(R)_{\text{electrostatic}} + V(R)_{\text{vdW}} \quad (2.23)$$

The total potential energy is made up of the short-range bonding and the longer-range non-bonding terms. The bonding interactions are composed of bond stretching, $V(R)_{\text{bond}}$, angle bending, $V(R)_{\text{angle}}$, and torsions, where the dihedral angles within the molecule vary, $V(R)_{\text{torsion}}$ (Figure 2.3).¹⁰⁶⁻¹⁰⁹ Cross-terms, $V(R)_{\text{cross}}$, are also included within the bonding interactions in high-level forcefields due to additional interactions as a result of the other bonding interactions, for example, the influence of changing the bond stretching force constant on the angle bend.⁵³

The non-bonding interactions between pairs of atoms comprise the electrostatic fixed charge interactions of the molecule, $V(R)_{\text{electrostatic}}$, arising from Coulombic interactions, and the van der Waals intermolecular forces, $V(R)_{\text{vdW}}$.¹⁰³ Whilst the number of bonded terms scales linearly with the number of atoms in the system, N , the non-bonded terms scales as a multiple of N^2 . As the non-bonded terms tend towards zero energy at infinite intermolecular separations, a cutoff term is generally applied to reduce computational expense by only considering the non-bonded terms within the cutoff distance.¹⁰³

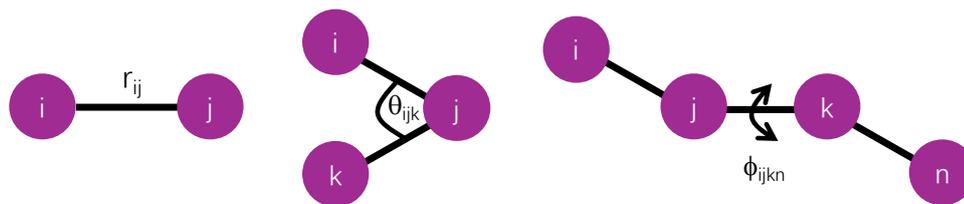


Figure 2.3: Schematic of the bond, angle, and dihedral terms (left to right) included within the bonding interactions of a molecular mechanics forcefield. Figure inspired by references 106–109.

There are a number of forms that can be used to describe the potential energy of each term, for example, the harmonic, Morse, and quartic potentials can be used to describe the bonds within the system,¹⁰⁷ whilst the harmonic, cosine harmonic, and quartic potentials (Figure 2.4)^{108,109} can be used to describe the angles, among others.^{108,109} The dihedral angles can be described using the harmonic, cosine harmonic, periodic, and cosine potentials, among others.¹⁰⁶ The potential forms used to describe the system should be chosen to best replicate the potential energy surface. The forms of the bond, angle, dihedral, and intermolecular interactions utilised within this thesis are covered in more detail in the ‘Ambuild’ section of the methodology (section 2.4.1).

Additionally, there are a number of forcefields that are currently available for use in MM simulations, including general forcefields that are applicable to a large range of systems such as the universal forcefield (UFF),¹¹⁰ the family of Assisted Model Building with Energy Refinement (AMBER) forcefields, with the most generic being the Generalised Amber Forcefield (GAFF),¹¹¹ and the Chemistry at HARvard Molecular Mechanics (CHARMM) forcefields, including the CGenFF forcefield.^{112,113} Other forcefields are more tailored towards specific systems, such as the COMPASS forcefield,¹¹⁴ designed to model alkanes and benzene derivatives; and the Polymer Consistent Forcefield (PCFF),¹¹⁵ designed to model interactions within polymer systems.

Each forcefield will typically yield slightly different parameters to describe the same interaction as a result of the functions used to define the forcefield, and the type of systems compared to when fitting the model, for example, polymers, proteins, *etc.* Table 2.1 shows as an example how a selection of different forcefields describe the equilibrium bond length and force constant of a carbon-carbon triple bond.

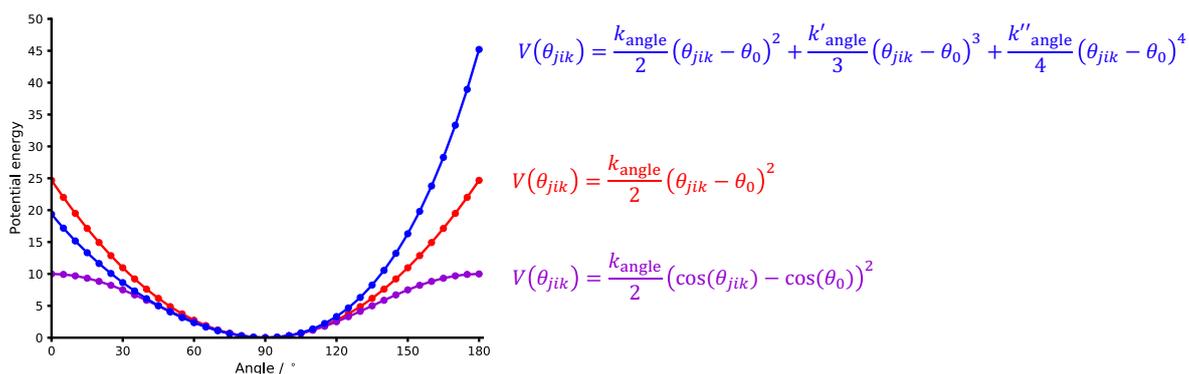


Figure 2.4: A selection of potential forms that may be used to describe the angles within a system described using a molecular mechanics forcefield as a function of the equilibrium angle, θ_0 , force constant(s), k_{angle} , k'_{angle} and k''_{angle} , and angle of interest, θ_{jik} . The equilibrium angle used in each potential is 90° . Harmonic potential – red, cosine harmonic potential – purple, quartic potential – blue. Equations replicated from references 108 and 109.

Table 2.1: Comparison of the bond stretching interaction parameters for a carbon-carbon triple bond amongst different forcefields. The equilibrium bond lengths and force constant values, r_{ij} and k_{bond} , respectively, were obtained from references 111, 113, 115. The values used within the Ambuild code, adapted from the PCFF,¹¹⁵ are included for comparison. As discussed later in Chapter 2 section 2.4.1, the models generated within Ambuild are optimised using a rigid-body approach, meaning only new interactions formed between building blocks are optimised. This means that generally, the force constant used in Ambuild is approximately double that of the PCFF to describe the rigid nature of the building blocks.

Forcefield	Description of interaction type	$r_{ij} / \text{\AA}$	$k_{\text{bond}} / \text{kcal mol}^{-1}$
PCFF	<i>sp</i> carbon involved in a triple bond	1.204	800.0
CGenFF	Internal alkyne R-C \equiv C	1.220	960.0
GAFF	Carbon-carbon bond	1.526	2086.0
Ambuild parameter as used	<i>sp</i> carbon involved in a triple bond	1.181	1691.6

Forcefields can be grouped within one of three classes, Class I, Class II and Class III.¹¹⁶ Of these, Class I, which includes the AMBER and CHARMM forcefields,¹¹⁷ calculates the total potential energy, $V(R)_{\text{ClassI}}$, as a function of bond stretching, angle bending, torsional, and non-bonding electrostatic and van der Waals interactions (Equation 2.24),^{116,117}

$$V(R)_{\text{ClassI}} = V(R)_{\text{bond}} + V(R)_{\text{angle}} + V(R)_{\text{torsion}} + V(R)_{\text{electrostatic}} + V(R)_{\text{vdW}}.$$

In contrast, Class II forcefields such as the PCFF also incorporate bond-bond and bond-angle coupling terms (Equation 2.24),¹¹⁶

$$V(R)_{\text{ClassII}} = V(R)_{\text{ClassI}} + V(R)_{\text{bond-bond}} + V(R)_{\text{bond-angle}}, \quad (2.24)$$

which can improve the overall quality of the model due to the additional inter-

actions. Class III forcefields again increase in complexity from Classes I and II due to the incorporation of charge polarisation, in contrast to fixed charges in Classes I and II.¹¹⁶

The PCFF is utilised by a number of research groups working within the area of amorphous porous materials simulation,^{19,38,41,104,115} making it an appropriate choice to model polymer frameworks such as those described in this thesis. It is important to ensure for any system of interest that the forcefield chosen is appropriate to replicate the chemistry of the system, yielding simulation results that are comparable to experiment.⁵³

So far, all of the forcefields described have been unreactive and therefore unable to explicitly describe the making and breaking of chemical bonds in the simulation cell, and the resulting changes in the forcefield type of each atom in doing so.¹¹⁸ As a result of this, one approach taken to model chemical reactions within this thesis and described within Chapter 6 section 6.3 is to define the forcefield with each atom being assigned the forcefield type that would occur within the reacted system, for example, that an *sp* carbon within the nitrile group of a monomer could be described using the forcefield type of an *sp*² carbon within the imine group of the resulting polymer.⁴⁵ In contrast to this approach, reactive forcefields such as ReaxFF^{119,120} do not require all interactions to be defined explicitly at the beginning of the simulation, allowing for the simulation of chemical reactions. Instead, these forcefields are trained using high-level quantum mechanical calculations, enabling the reaction to occur using molecular dynamics, and benefitting from the larger system sizes resulting from that, whilst still retaining high levels of accuracy from the quantum-mechanical training of the dataset.¹¹⁹ However, reactive forcefields are typically trained for very specific reactions, meaning that any unexpected side reactions or intermediates, as observed in the study discussed in Chapter 6 section 6.4.1,⁴⁵ may not be identified using these methods.^{118,119}

Zhang and co-workers recently reported a reactive forcefield known as a machine learning interatomic potential (MLIP), parametrised for C, H, N and O and named ANI-1xnr.^{118,121} This forcefield was tested on five case study systems, including methane combustion and carbon solid-phase nucleation, that were independent of the training set utilised to refine the model. These test cases showed energies and forces that were of comparable accuracy to those derived using quantum mechanical methods, but with a calculation speed that was up to seven orders of magnitude faster.¹²¹ The aim of this MLIP was to be a general reactive forcefield to model a range of systems composed of C, H, N and O in the condensed-phase (*i.e.*, periodic systems), reducing the requirement for significant chemical intuition and MLIP expertise to train the

model for a specific scenario.^{118,121}

ANI-1xnr is based on the ANI MLIP, a neural network-based model trained to describe a large dataset of organic systems.^{122,123} Active learning was utilised to establish which new findings from the simulations should be added to the dataset, reducing human bias and allowing for unexpected side reactions and intermediates to occur.¹²⁴ The reaction pathways were sampled using the *ab initio* nanoreactor (NR), first reported by Wang and co-workers.^{125,126} This utilises an external bias to enhance the likelihood of chemical reactions occurring from the high-velocity collisions of small molecules. The NR approach was used to successfully model graphene ring formation from ethyne, and glycine formation in the so-called 'Miller experiment'.^{125,126}

From the combination of the ANI model, active learning, and the *ab initio* nanoreactor, ANI-1xnr was established as a general reactive MLIP.^{118,121} This showed results comparable to experiment or high-level computation across the five case studies, with, crucially, no requirement to re-train the model between test systems. Additionally, this model explored significant additional regions beyond the scope of a non-reactive, near-equilibrium active learning dataset, ANI-1x, suggesting an improvement over the non-reactive precursor.^{118,121}

Similarly to a traditional non-reactive forcefield, ANI-1xnr was designed to only compute intermolecular interactions, such as van der Waals forces and Coulombic interactions, within the cutoff radius. The main challenge with the ANI-1xnr MLIP at present is the limitation to the number of elements parametrised within the model, which renders it at present unfeasible for use to describe elements discussed within the systems in this thesis such as bromine, copper, phosphorus, and palladium. However, this was addressed as a future prospect by the authors of the original ANI-1xnr publication.^{118,121}

The ANI-1xnr MLIP^{118,121} appears to be a vast improvement on conventional reactive forcefields, which are trained for very specific chemical reactions and struggle to identify unexpected intermediates and side reactions, yet, only a small number of elemental types are currently considered within this model. Whilst it is an intriguing prospect for the future study of the systems described within this thesis on the proviso that all elements required have been parametrised within this forcefield, it is not currently feasible for use. Instead, the PCFF was utilised throughout this work as an established forcefield within the area of amorphous porous materials simulation.^{19,38,41,104,115}

Coarse-grained molecular mechanics

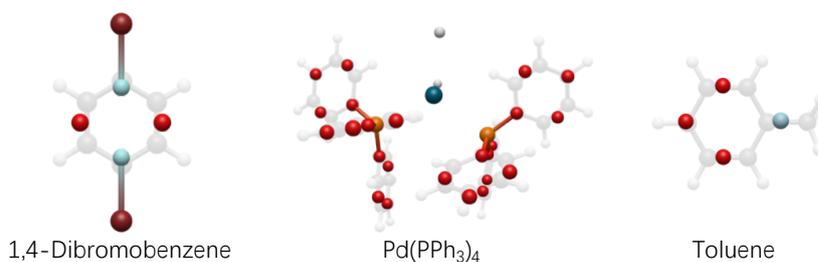


Figure 2.5: Superimposed all-atom (faded) and coarse-grained (bold) building blocks for 1,4-dibromobenzene, Pd(PPh₃)₄ and toluene solvent. In the case of 1,4-dibromobenzene and Pd(PPh₃)₄, the building blocks are designed as hybrid coarse-grain to maintain the bonding sites as fully atomistic. Figure reproduced and adapted from reference 104 with permission from Institution of Chemical Engineers (IChemE) and the Royal Society of Chemistry.

It is also possible to coarse grain a chemical system, which reduces the number of atoms present by using a single grain to represent a group of atoms. The grains are designed to replicate the shape and chemistry of the atoms contained within them.¹²⁷⁻¹²⁹ Whilst there is an inherent error by reducing the number of atoms within the system, the resulting increase in the system size that can be modelled allows for larger-scale effects to be studied.¹⁰⁴ As for atomistic MM simulations, specialised forcefields have been designed for coarse-grained systems. For example, a common application of coarse graining a chemical system is to replace each methyl group with a single grain, which is well described by United Atom (UA) forcefields.^{130,131} Another popular forcefield designed for coarse-grained systems is the MARTINI forcefield, often utilised for biological systems such as proteins and peptides.¹²⁹ It is also possible to coarse grain solvent molecules, for example, DeVane and co-workers reported a coarse-grained toluene molecule, represented by four grains instead of the fifteen atoms contained within the fully atomistic molecule.¹²⁷ Hybrid coarse-grain models can also be utilised, which allow for the bonding sites to be retained as fully atomistic whilst representing the remainder of the molecule using coarse grains (Figure 2.5).¹⁰⁴

2.2.3 Geometry optimisation and the potential energy surface

The approaches detailed above show how the energy of a system in a given arrangement can be described using quantum chemical or forcefield-based techniques, however, the energy of a chemical system is highly dependent on the molecular configuration. For an isolated N -atom system, there are $3N - 6$ possible degrees of freedom, some of which will give a different potential energy. This generates a potential energy surface for each system, with the energy changing as a result of the atomic positions.⁵³ The ground state structure, obtained by exactly solving the time-independent Schrödinger equation

for a given system,⁴⁶⁻⁵¹ is defined as the minimum energy configuration. This corresponds to the nearest local minimum of the surface, with no imaginary vibrational modes observed on undertaking a frequency calculation of the structure. Sampling any other configuration within the system leads to a resulting increase in energy. Transition state structures, which are referred to as ‘first-order saddle points’, are structures that are energy minima in $N - 1$ dimensions, and an energy maximum in the remaining dimension. These structures correspond to a gradient with respect to the change in energy as a function of the atomic coordinates of zero, and have one imaginary vibrational mode in a frequency calculation.⁵³ Local minima are also present on the potential energy surface, which appear higher in energy than the ground state structure, but lower in energy than their surrounding environment (Figure 2.6).¹³²

Whilst in an ideal world, the global minimum could be found by simply calculating the geometry of every possible configuration, and then selecting the one with the lowest energy, this is not possible for any meaningful chemical system composing more than a handful of atoms due to the number of calculations required to model this. Therefore, the process of geometry optimisation is designed to determine the molecular configuration leading to the lowest possible energy, the nearest local minimum.⁵³ There are a number of ways to do this.

The first, and the simplest method, involves giving an input structure at a particular conformation, R_0 , a ‘guess’ structure. The potential energy, $V(R_0)$ is then calculated for this structure. Following this, a random change is made to the atomic coordinates, ΔR , and the potential energy of the new conformation calculated, $V(R_0 + \Delta R)$. If the energy after the change is more favourable than before the change, *i.e.*, $V(R_0 + \Delta R) < V(R_0)$, the change is accepted. If not, a different random change is applied to R_0 . As the change in energy begins to converge approaching the minimum, smaller changes should be made to the coordinates to ensure that the minimum can be located.⁵³

A second, and more efficient way to optimise the geometry, is known as the steepest descent method. This involves calculating the potential energy, $V(R_0)$, and gradient of the energy, $\frac{\delta V(R_0)}{\delta R}$, for conformation R_0 . Moving from conformation R_0 along the direction defined by the gradient, referred to as a ‘line search’, the lowest energy structure at conformation R_1 is located. The potential energy, $V(R_1)$, and gradient of the energy, $\frac{\delta V(R_1)}{\delta R}$ are then calculated for conformation R_1 , and a further line search obtained, until the energy fits within the convergence criteria, yielding a minimum energy structure.⁵³

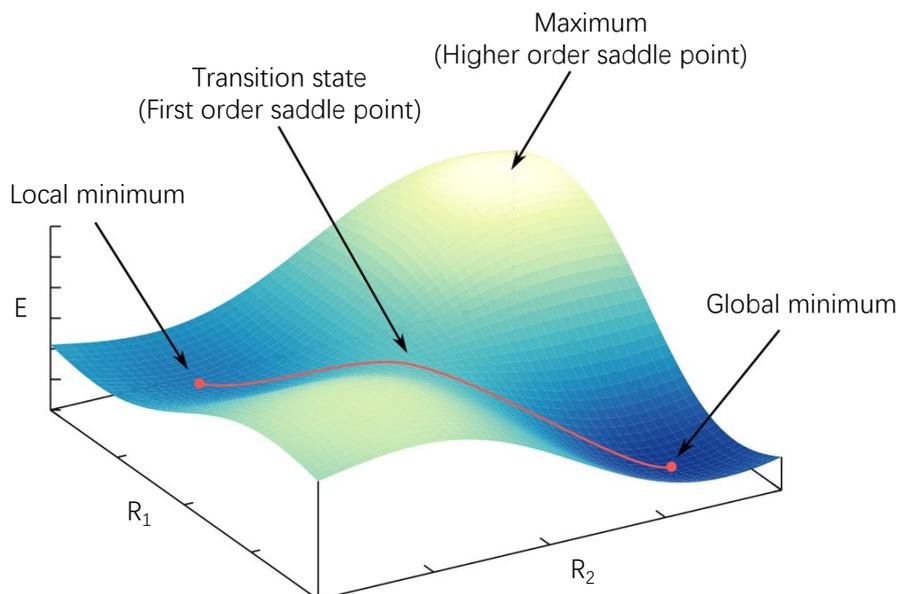


Figure 2.6: Example of a potential energy surface, indicating the global and local minima, saddle points and maximum. The red line shows the path from the local minimum to the global minimum. Figure adapted with permission from reference 132, licensed under CC-BY-NC-ND 4.0. Copyright 2021 American Chemical Society.

2.2.4 Molecular dynamics

Once the forcefield has been chosen to describe the system in a given configuration, the dynamical motion of the system can then be established using molecular dynamics (MD). There are two main classifications of MD: classical MD, which is the approach used within this thesis, using classical Newtonian mechanics to simulate how the system moves as a function of time upon application of energy into the system, and *ab initio* MD, which utilises electronic structure theory to calculate forces on-the-fly during the MD simulation.^{53,133,134}

Using classical MD, the dynamical motion of the system is calculated from Newton's law of motion, where the force acting on each atom, F , at time t , is equal to the atomic mass, m , multiplied by its acceleration, a . The force also relates directly to the potential energy of the system, where the force is equal to the negative first derivative (gradient) of the energy (Equation 2.25).^{53,135}

$$F = m \cdot a = -\frac{\delta V(R)_{\text{total}}}{\delta R} \quad (2.25)$$

From the forces and masses, the acceleration at time t can be calculated (Equation 2.25). Following this, the velocity, v , at time t can be calculated from the acceleration (Equation 2.26).^{53,135,136}

$$a(t) = \frac{\delta v(t)}{\delta t} \quad (2.26)$$

The position of the atoms, R , at time t can be calculated from the velocity (Equation 2.27).^{53,135,136}

$$v(t) = \frac{\delta R(t)}{\delta t} \quad (2.27)$$

From the initial position and velocity of each atom in the system, the velocities and positions of the atoms at time $t + \Delta t$, where Δt corresponds to the integration timestep, can then be calculated using the Euler equations (Equations 2.28–2.29).⁵³

$$v(t + \Delta t) = v(t) + a(t) \cdot \Delta t \quad (2.28)$$

$$R(t + \Delta t) = R(t) + v(t) \cdot \Delta t \quad (2.29)$$

A summary of the full MD simulation protocol is given in Figure 2.7.^{53,135} Whilst these equations appear simple, the velocity and acceleration of the atoms change over the course of the integration timestep, meaning that the newly obtained positions and velocities may not be accurate. A more accurate way to obtain the position of each atom at time $t + \Delta t$ is approximated using a Taylor expansion at time t , resulting in a modification of the Euler position equation (Equation 2.30).^{53,135}

$$R(t + \Delta t) = R(t) + v(t) \cdot \Delta t + \frac{1}{2} \Delta t^2 a(t) + \dots \quad (2.30)$$

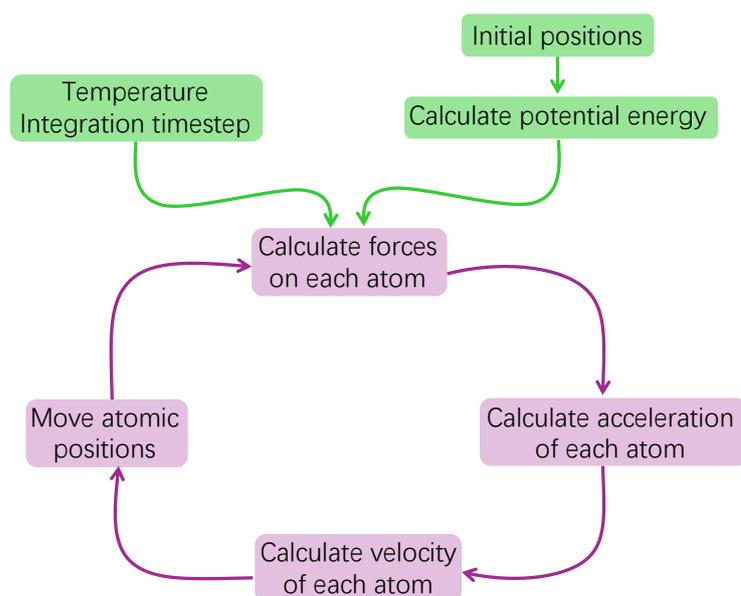


Figure 2.7: Flow chart depicting the full MD simulation protocol. Figure inspired by references 53 and 135.

The kinetic energy of the system is increased or decreased by modifying the atomic velocities. *In silico*, this is done by increasing or decreasing the temperature utilised within the simulation. This is because the kinetic energy is proportional to both the square of the velocity, v , and the temperature, T , as shown in Equation 2.31,

$$\text{Kinetic energy} = \frac{1}{2}mv^2 = \frac{3N}{2}RT, \quad (2.31)$$

where m is the mass, N is the number of particles, and R is the universal gas constant. The relationship between kinetic energy and temperature is defined in the kinetic molecular theory of gases.¹³⁷

The error occurring from assuming that the velocity and acceleration will remain constant over the course of the timestep can be minimised by using a smaller timestep. However, as the total simulated time is calculated from the product of the number of steps and the integration timestep, decreasing the integration timestep increases the computational expense as a larger number of steps are required to simulate the molecular motion over the same timescale.⁵³ It is therefore important to ensure that the integration timestep employed is appropriate by considering the level of accuracy required from the simulation.¹³⁵

There are a number of algorithms that are used in practice to obtain accurate positions and velocities of atoms throughout an MD simulation. The first of these is the Verlet algorithm, named after the French physicist Loup Verlet.¹³⁸ This algorithm utilises the position and acceleration of the system at current time t and the position at previous time $t - \Delta t$ to establish the positions of each atom at time $t + \Delta t$, by summing together the terms from times t and $t - \Delta t$, with the process repeated for each step considered (Equations 2.32–2.33).^{53,135,138}

$$R(t - \Delta t) = R(t) - v(t) \cdot \Delta t + \frac{1}{2}\Delta t^2 a(t) + \dots \quad (2.32)$$

$$R(t + \Delta t) = 2R(t) + \Delta t^2 a(t) - R(t - \Delta t) + \dots \quad (2.33)$$

However, this algorithm does have some downsides. Firstly, as the approach relies on the positions at the previous time, $t - \Delta t$, an alternative method must be used to begin the algorithm. Secondly, the velocities themselves are not included as terms within the Verlet algorithm (Equation 2.33), meaning that alternative approximations such as the mean value theorem must be employed to calculate these, leading to additional errors in the simulation.¹³⁵

With these issues in mind, the Velocity Verlet algorithm was later developed.¹³⁹ The Velocity Verlet algorithm is one of the most common algorithms used in

MD simulations, and relies on the positions, velocity and acceleration at time t to establish the positions at time $t + \Delta t$. This is done by firstly defining an initial set of positions and the corresponding velocities for time t . From these positions, the acceleration can be found using Equation 2.1. The positions at time $t + \Delta t$ are then obtained using the Taylor expansion of the Euler position equation (Equation 2.30) at time t to update the atomic coordinates for time $t + \Delta t$. From the updated positions and the subsequent force acting on each atom (Equation 2.25), the acceleration at time $t + \Delta t$ can be calculated. Finally, the velocity at time $t + \Delta t$ can be evaluated using a Taylor expansion of the Euler velocity equation (Equation 2.28) to give Equation 2.34.^{53,135,139}

$$v(t + \Delta t) = v(t) + \frac{1}{2}\Delta t[a(t) + a(t + \Delta t)] \quad (2.34)$$

The Leapfrog algorithm is a common alternative to the Velocity Verlet.¹⁴⁰ This uses the position at time t and the velocity at time $t + \frac{1}{2}\Delta t$ (Equation 2.35) to establish the positions and velocities of the system at subsequent timesteps.¹⁴⁰

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + a(t) \cdot \Delta t \quad (2.35)$$

This is done by firstly calculating the velocities of the system at time $t + \frac{1}{2}\Delta t$ (Equation 2.35). From the velocities at time $t + \frac{1}{2}\Delta t$ and positions at time t , the positions at time $t + \Delta t$ can then be calculated (Equation 2.36).

$$r(t + \Delta t) = r(t) + v\left(t + \frac{1}{2}\Delta t\right) \cdot \Delta t \quad (2.36)$$

Then, the process repeats, obtaining the velocities at time $t + \frac{3}{2}\Delta t$ and using these to determine the positions at time $t + 2\Delta t$. As the positions and velocities are calculated at different times from one another and alternate which is calculated at a larger time, this is considered analogous to frogs ‘leaping’ over one another, giving the algorithm its name.¹⁴⁰

It should also be considered that whilst the acceleration at time t is required to obtain the velocity at time $t + \frac{1}{2}\Delta t$, this can be easily calculated from Newton’s law of motion (Equation 2.25) from the position at time t . Unlike the Velocity Verlet algorithm, the Leapfrog algorithm does not require the acceleration to be calculated for time $t + \frac{1}{2}\Delta t$ to obtain the subsequent velocity, making the Leapfrog algorithm less computationally expensive compared to the Velocity Verlet.^{135,139,140} As the two algorithms calculate the velocity at different times throughout the simulation (Velocity Verlet = $t, t + \Delta t, t + 2\Delta t$ etc.; Leapfrog = $t + \frac{1}{2}\Delta t, t + \frac{3}{2}\Delta t, t + \frac{5}{2}\Delta t$ etc.), it is not possible to change between algorithms during the simulation.¹³⁵

As well as the algorithm employed during the MD simulation, the thermodynamic properties of the system can be varied by changing the ensemble. Each ensemble holds a trio of conditions constant, given in parentheses alongside the name, whilst allowing all of the others to change. The reaction conditions considered are: the number of molecules (expressed in moles), N , simulation cell volume, V , system energy, E , temperature, T , pressure, P , and chemical potential, μ .¹⁴¹

Common ensembles include the microcanonical (NVE), canonical (NVT), isothermal-isobaric (NPT) and grand-canonical (μ VT). Of these, arguably the two most commonly used within MD simulations are the NVT and NPT. The NVT ensemble allows the cell pressure to change whilst scaling the velocities of the system to ensure that the kinetic energy and temperature remain constant, replicating a thermostat, where a higher temperature corresponds to increased kinetic energy in the system and therefore larger velocities.^{141,142} This is done by applying a scaling factor, λ , to the atomic velocities in the system to change the kinetic energy of each atom and the resulting system temperature (Equation 2.37).¹⁴² The scaling factor is chosen to maintain the assigned temperature throughout, with values > 1 and < 1 increasing and decreasing the system temperature, respectively.¹⁴² Common methodologies that can be used to retain the system temperature include the Berendsen,¹⁴³ Andersen,¹⁴⁴ Nosé–Hoover,¹⁴⁵ and Bussi–Donadio–Parrinello thermostats.¹⁴⁶

$$v_i^{\text{new}} = v_i^{\text{old}} \quad (2.37)$$

The NPT ensemble, contrastingly, allows the cell volume to change whilst keeping the pressure constant by modifying the size of the simulation cell, replicating a barostat,¹⁴⁷ where a smaller simulation cell corresponds to an increased pressure. This is done in a similar way to the thermostat algorithm, by applying a scaling factor, λ , to the atomic coordinates in the system to change the cell volume (Equation 2.38). Here, the scaling factor is chosen to maintain the assigned pressure throughout, mimicking placing the sample in a ‘pressure bath’.¹⁴⁷ λ is raised to the exponent of $\frac{1}{3}$ as this corresponds to multiplying the cell volume of system i by λ .¹⁴⁷ Common methodologies that can be used to scale the simulation cell size to control the pressure include the Berendsen,¹⁴³ Andersen,¹⁴⁴ and Parrinello–Rahman barostats.¹⁴⁸⁻¹⁵⁰

$$r_i^{\text{new}} = r_i^{\text{old}} \cdot \lambda^{\frac{1}{3}} \quad (2.38)$$

Typically, an MD simulation will use a combination of NVT and NPT ensembles to equilibrate the system and ensure that the model chosen is representative of the real-world material, however, the ensemble is chosen to replicate the chemistry of the system in each case.¹⁴¹

2.2.5 Determining charge distribution

Charge distribution across a molecule, such as the dipole moment, can be calculated using the Mulliken population analysis technique. In the linear combination of atomic orbitals, molecular orbitals (MO) can be calculated as the sum of the normalised atomic orbitals composing the molecular orbital.^{151,152} The square of the wavefunction of the MO can then be used to give the probability distribution of finding an electron within the MO. The Mulliken electron population of each atom is given by the density and overlap matrices, and the Mulliken atomic charge is then calculated from subtracting the Mulliken electron population from the atomic number of that atom.^{151,152}

2.2.6 Summary

This computational theory section has addressed a number of common methodologies used for simulation of chemical systems of varying sizes. It is important to consider the trade-off between system size and accuracy and carefully choose the approach that will yield the most relevant information for the system of interest in order to design the best new materials and rationalise those already existing.

2.3 Experimental theory

It is also necessary to understand the range of methodologies utilised to characterise the materials experimentally. This section gives the background to the origin of data used to compare to the simulated case studies.

2.3.1 Porosity

Porosity is the ratio of space within a material occupied by void spaces, known as pores, which are categorised by the International Union of Pure and Applied Chemistry (IUPAC) based on their dimensions. Following IUPAC guidelines, micropores, mesopores, and macropores are pore voids with a diameter of < 2 nm, 2-50 nm, and > 50 nm, respectively.²

There are a variety of ways to measure the porosity of a material. Firstly, and perhaps the most quantitative, is to calculate the pore volume, which is the volume of gas that can be adsorbed into the pores of the material.⁵ This can be evaluated from an adsorption/desorption isotherm, where gas molecules are added into the porous material at constant temperature and increasing pressure. The isotherm is plotted as the volume of adsorbed gas, typically N₂, as a function of relative pressure $\frac{p}{p_0}$, where p is the equilibrium pressure and p_0 is the saturation vapour pressure of the adsorbate. As the isotherm is plotted in terms of relative pressure, the range is given as 0-1. Adsorption-desorption

isotherms fall into one of six types (Figure 2.8), which can give information about the relative pore sizes within the material.¹⁵³ For example, a type I isotherm is composed of a monolayer of adsorbed gas, showing a plateau in the quantity of adsorbed gas upon completion of a monolayer, whilst a type VI isotherm is composed of multilayers, adsorbed one-by-one. In contrast, type IV isotherms, typically indicative of materials containing micro- and mesopores, show a difference in the adsorption and desorption isotherms, known as hysteresis, occurring due to the difference in pathways occurring during adsorption and desorption.^{153,154}

Experimentally, isotherms can be calculated using a number of different models, each of which utilising a different model equation and adsorption mechanism to calculate the adsorption. The possible adsorption mechanisms are: chemisorption, the formation of a chemical bond between the adsorbate and adsorbent, leading to formation of a monolayer of adsorbed gas; physisorption, van der Waals interactions between the adsorbent and adsorbate, leading to multilayer formation, and ion exchange (Figure 2.9).¹⁵⁵

Adsorption-desorption isotherms can be simulated using packages such as Towhee,¹⁵⁶ Music,¹⁵⁷ and Materials Studio.¹⁵⁸ These methods, which collect isotherms at fixed points within the simulation, do not allow for inclusion of dynamic porosity within the system, where pores can open and close throughout the isotherm collection process as the materials move and/or swell. This means that isotherm simulation of microporous polymer materials such as the ones studied within this thesis are not altogether representative of the system, and so they have not been computed within this thesis.

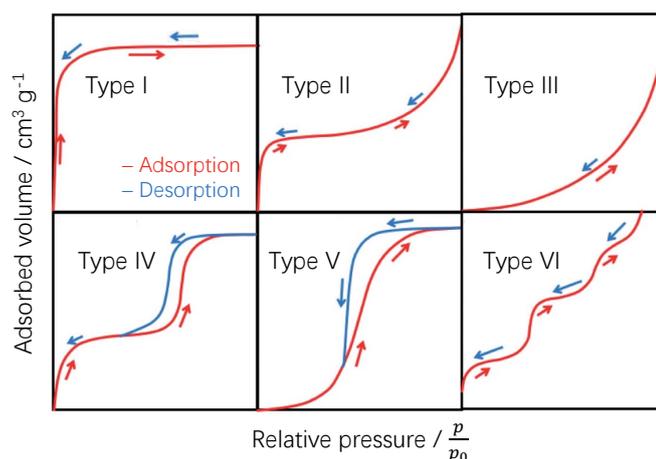


Figure 2.8: The six types of adsorption-desorption isotherm. Key: adsorption – red, desorption – blue, indicated by the arrows. Figure reproduced and adapted from reference 153 with permission from the Royal Society of Chemistry.

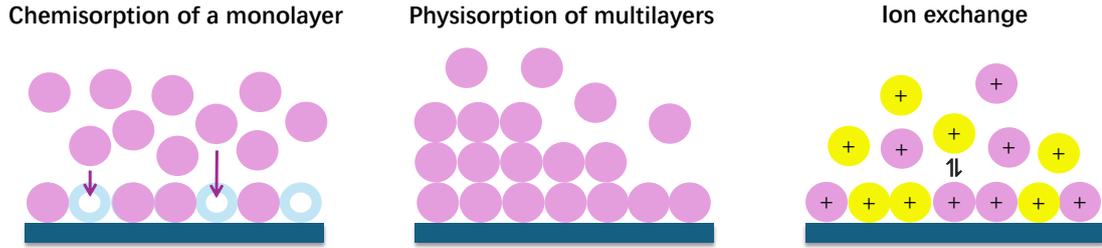


Figure 2.9: The three adsorption mechanisms of an adsorbate gas onto an adsorbent material. Key: adsorbent surface – dark blue rectangle, available adsorption sites – pale blue open circles, adsorbates – purple and yellow circles (the different colours signify different chemical species). Figure reproduced and adapted from reference 155 with permission from Elsevier.

A common example isotherm model is the Langmuir model (Equation 2.39),¹⁵⁹

$$\frac{P}{q} = \frac{1}{q_m} \cdot P + \frac{1}{K \cdot q_m} \quad (2.39)$$

which considers adsorption of a monolayer *via* chemisorption.^{159,160} Here, P is the pressure, q is the quantity of adsorbed gas, q_m is the capacity of a monolayer (obtained from Equation 2.40),³⁶

$$q_m = \frac{1}{\text{intercept} + \text{gradient}} \quad (2.40)$$

and K is the Langmuir rate constant. The isotherm is then plotted, with an y -intercept of $\frac{1}{K \cdot q_m}$ and a gradient of $\frac{1}{q_m}$.¹⁶⁰

This model does not account for multilayer formation and assumes that each site on the surface is energetically identical, with no possibility for interactions to occur between adsorbed molecules.¹⁶¹ However, it is known that not every material will only form a monolayer of adsorbed gas, leading to the Brunauer-Emmett-Teller (BET) model (Equation 2.41),³⁶

$$\frac{\frac{p}{p_0}}{q \left(1 - \frac{p}{p_0}\right)} = \frac{C - 1}{C \cdot q_m} \cdot \frac{p}{p_0} + \frac{1}{C \cdot q_m} \quad (2.41)$$

an application of the Langmuir model which also considers adsorption of multilayers *via* physisorption.^{35,36} Here, $\frac{p}{p_0}$ is the relative pressure, and C is the BET constant. The isotherm is then plotted, with an y -intercept of $\frac{1}{C \cdot q_m}$ and a gradient of $\frac{C-1}{C \cdot q_m}$.³⁶ The BET theory fixes two key issues with the assumptions of the Langmuir model: firstly, it allows for multilayer formation, yet retains that each layer be described using the Langmuir model. Secondly, interactions are allowed to occur between adsorbates on neighbouring layers, which is an advantage given that the system is unlikely to be fully inert.³⁵

A second way to calculate the porosity of a material is to deduce its surface area, which is a measure of the area per mass, of the surface of the material.

Experimentally, the surface area is calculated directly from the isotherm by fitting a linear trendline to the points in the pressure range of $0.05 < \frac{p}{p_0} < 0.30$, and then solving the resulting equation to calculate the surface area (Equation 2.42).³⁶

$$SA_{\text{Total}} = \frac{q_m \cdot N_A \cdot A_{\text{ads}}}{M_{\text{ads}}} \quad (2.42)$$

Here, N_A is Avogadro's number, A_{ads} is the cross-sectional area of the adsorbed gas molecule, typically nitrogen, and M_{ads} is the mass of an adsorbed gas molecule.³⁶ The specific surface area is then calculated by dividing the surface area obtained in Equation 2.42 by the sample mass.

However, whilst this is a very commonly employed method to calculate the surface area of a material, the requirement to fit a trendline leads to very large errors in the calculated surface areas. This means that even when calculating the surface area of the same sample using the same instrument multiple times, a range of surface areas will be obtained. This is particularly prudent for microporous materials, which contain a smaller number of data-points at low pressure values due to high uptakes at low pressures, leading to an increased challenge of fitting an appropriate trendline to the dataset.³⁶

Computationally, a range of software tools, such as Poreblazer,^{162,163} Materials Studio,¹⁵⁸ and Zeo++,¹⁶⁴ can be used to simulate the surface area of a material. This is done by rolling a probe molecule (typically a nitrogen molecule, to match experiment, setting the probe radius as 1.82 Å, the kinetic radius of N₂) across the surface of the material, then calculating the total area covered by the probe. However, depending on the software package utilised, there are a number of different types of surface area that can be calculated, each subtly different in how the probe interacts with the surface of the sample (Figure 2.10).^{37,44,162}

The Connolly¹⁶⁵ surface area is calculated as the surface area covered by the probe from the point of contact of the probe onto the surface, whilst the accessible surface area calculates the surface area from the centre of the probe.⁴⁴ It has been reported by several groups that the accessible surface area is a better representation of the experimental material than the Connolly surface area, giving a far superior match to reported experimental surface areas, despite the popularity of the Connolly within biological systems.^{37,44,165} This is due to the ability of the accessible surface area to describe interactions occurring between the adsorbate and adsorbent, along with considering the entropy of adsorbed guests. As a result, the accessible surface area gives the upper limit of the expected surface area of a material, with any large deviations from this being rationalised as due to defects, incomplete desolvation

or network interpenetration of the experimental sample.³⁷ Additionally, surface areas can be calculated as the network-accessible (or solvent-accessible) area, or as a total surface area.¹⁶² The network-accessible surface area is the complete surface area accessible to the probe *via* a diffusive pathway and is comparable to the reported experimental surface area. The total surface area is the network-accessible surface area plus the regions accessible to the probe that cannot be accessed *via* a diffusive pathway, *e.g.*, the closed pores.¹⁶² Calculation of the simulated pore volume is also subject to a variety of types following the same general idea.¹⁶⁶

Another way to characterise the porosity of a material is using a pore size distribution (PSD), which, as the name suggests, gives an indication of the range of pore sizes within the sample. Pore size distributions of experimental samples are often calculated using non-local density functional theory (NLDFT).¹⁶⁷ This works by specifying as input a model that most appropriately describes the system.¹⁶⁷ For example, a one-dimensional infinite slit model, where the assumption is made that pores can be described as infinite slits with ideal walls modelled on graphite, is often utilised for porous carbon.¹⁶⁸

NLDFT is then used along with this input to calculate the fluid density of adsorbent within the pores of the material.¹⁶⁷ A series of theoretical isotherms are generated, known as a kernel. The isotherms within the kernel demonstrate how the amount of adsorbed gas changes as a function of pore size and pressure. The theoretical isotherms are then compared by the gas sorption instrument to the experimental isotherm, giving the pore size distribution of the material.¹⁶⁷

A key advantage of using NLDFT to calculate the PSD of a material is its ability to calculate the distribution across the micro-, meso-, and macropore range.¹⁶⁷ This is in contrast to approaches such as the Horváth-Kawazoe and Barrett-Joyner-Halenda, which are used to compute the PSD of the micropore, and meso- and macropore regions, respectively.^{169,170} Additionally, the act of fitting the theoretical isotherms to the experimental is undertaken by the software, making this a facile method to compute the PSD. However, as the PSD is based on an approximation that the pores can be described using one of the available models, the accuracy of the computed PSD is dependent on the fit to the model.¹⁶⁷

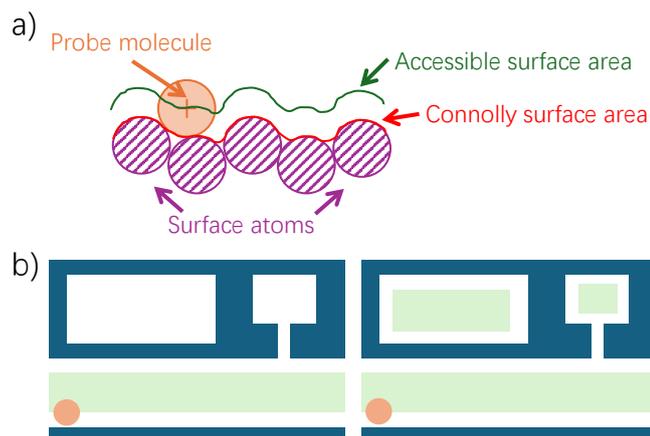


Figure 2.10: Illustration of the difference between the a) Connolly and accessible surface areas, and b) network/solvent-accessible (left) and total (right) surface areas. Key: probe molecule – orange, accessible surface area – green, Connolly surface area – red, surface atoms – purple, volume occupied by the material – blue, unoccupied volume – white, volume occupiable by the centre of the probe – pale green. Figure a) reprinted and adapted with permission from T. Düren, F. Millange, G. Férey, K. S. Walton and R. Q. Snurr, *J. Phys. Chem. C*, 2007, **111**, 15350–15356. Copyright 2007 American Chemical Society. Figure b) reprinted and adapted with permission from L. Sarkisov, R. Bueno–Perez, M. Sutharson and D. Fairen–Jimenez, *Chem. Mater.*, 2020, **32**, 9849–9867. Copyright 2020 American Chemical Society.

2.3.2 Solubility

There are a number of theories that can be used to describe the solubility of molecules, defined as the relative ability of a molecule to dissolve and interact within a solvent, with a general principle of ‘like dissolves like’, indicating that molecules of similar chemistry are likely to dissolve within one another.¹⁷¹ This section describes some of the common solubility theories.

The first solubility theory described is the solvent polarity index, which assigns, as the name suggests, a relative solubility value to each of a wide variety of solvents, with a larger polarity index corresponding to a more polar solvent. Following the ‘like dissolves like’ principle, a molecule of benzene, a non-polar molecule in which all of the dipoles within the system cancel each other out, would be fully miscible within non-polar solvents such as toluene, and relatively insoluble in polar solvents such as N,N-dimethylformamide (solvent polarity indices of 2.4 and 6.4, respectively).^{172,173}

A second solubility theory is the Flory–Huggins solution theory, which describes the ability of a polymer composed of x monomers to interact with the solvent (Equation 2.43) on a lattice with N sites, each occupied by a solvent molecule, N_1 , or one of the x monomers in the polymer, xN_2 .¹⁷⁴

$$N = N_1 + xN_2 \quad (2.43)$$

The volume fraction, a measure of how much of the total volume is composed of the solvent and polymer, ϕ_1 and ϕ_2 , respectively, is given in Equations 2.44 and 2.45, respectively.¹⁷⁴⁻¹⁷⁶

$$\phi_1 = \frac{N_1}{N} \quad (2.44)$$

$$\phi_2 = \frac{xN_2}{N} \quad (2.45)$$

The Gibbs free energy change of mixing the polymer and solvent can then be calculated as ΔG_{mix} (Equation 2.46).^{175,176}

$$\Delta G_{mix} = RT[n_1 \cdot \ln \phi_1 + n_2 \cdot \ln \phi_2 + n_1 \cdot \phi_2 \cdot \chi_{12}] \quad (2.46)$$

Here, the first two terms of the equation are entropic terms, evaluating the overall volume taken up by each of the two components. The third term is an enthalpic contribution, taking into account the total number of solvent-solvent, monomer-monomer and monomer-solvent interactions, and the energy of each. R is the gas constant and T is the constant temperature utilised throughout. n is the number of moles, calculated from the number of molecules, and χ is the polymer-solvent interaction parameter, obtained from the Hildebrand solubility parameters, described next.^{175,176}

The Hildebrand solubility parameter is, similarly to the solvent polarity index, a single value assigned to indicate the solubility of one material relative to another. Following the same general principle as the solvent polarity index, 'like dissolves like', two materials will be fully miscible if they share similar Hildebrand solubility parameters.^{177,178} The principle behind the Hildebrand solubility theory is to consider the forces required to remove a molecule from one phase, the 'monomer phase', and place it within a second phase, the 'solvent phase'. This is equivalent to the vaporisation of a system, where a system in the liquid state, a condensed phase, absorbs energy to separate the molecules into the gas state by breaking the attractive van der Waals intermolecular forces between molecules. The Hildebrand solubility parameter, δ , can be evaluated using the enthalpy of vaporisation, ΔH_{vap} , the energy required to move a molecule from the condensed phase into the gas phase, along with the molar volume of the material in the condensed phase, V_m , (Equation 2.47).^{177,178}

$$\delta = \sqrt{\frac{\Delta H_{vap} - RT}{V_m}} \quad (2.47)$$

Whilst both the solvent polarity index and Hildebrand solubility parameters are designed as simple approaches to calculate the relative solubility of a

molecule within a solvent using a single parameter, they can appear hard to tune. Another widely used solubility theory is the Hansen solubility approach, which splits the total Hildebrand solubility parameter, δ , into three components, δ_d , δ_p , and δ_h , representing the dispersion forces, polar interactions, and hydrogen bonding interactions, respectively (Equation 2.48).^{177,179} Whilst the dispersion forces are derived from the forces acting on each atom in the system and therefore cannot be tuned by modifying the reaction conditions, the polar and hydrogen bonding components can be tuned by characteristics of the solvent, for example, the polarity.¹⁸⁰

$$\delta^2 = \delta_d^2 + \delta_p^2 + \delta_h^2 \quad (2.48)$$

The Hansen solubility parameters of a polar molecule are calculated by firstly assigning the Hildebrand solubility parameter of the closest non-polar alternative, composed solely of dispersion forces, as the dispersion component of the polar molecule. Subtracting the square of the dispersion component from the square of the Hildebrand parameter yields the sum of the squares of the polar and hydrogen bonding components, referred to as δ_a^2 (Equation 2.49).^{177,179}

$$\delta_a^2 = \delta^2 - \delta_d^2 = \delta_p^2 + \delta_h^2 \quad (2.49)$$

The polar component can then be calculated using Equation 2.50,

$$\delta_p^2 = 37.4 \cdot \text{dipole moment} \cdot \sqrt{V_m} \quad (2.50)$$

with the hydrogen bonding component calculated as the remainder.¹⁸⁰

The three Hansen solubility parameters of a molecule can be considered as a three-dimensional (3D) coordinate to plot in the ‘Hansen space’.¹⁷⁷ The closer in Hansen space the points of two molecules are, the more miscible and soluble the two will be.¹⁸⁰ The absolute difference between the Hildebrand solubility parameters of a substance and solvent, $|\delta_{\text{diff}}|$, can be used to establish the relative compatibility of the material within that solvent, where a $|\delta_{\text{diff}}|$ value of less than one indicates a good solvent, a value between one and three indicates an intermediate solvent, and a value of greater than three is a poor solvent.¹⁸⁰ To improve the match between two molecules, the Hansen polar and hydrogen bonding components can be tuned as mentioned above.

2.3.3 X-ray Photoelectron Spectroscopy

X-ray photoelectron spectroscopy (XPS) is a surface sensitive characterisation technique used to determine the elemental composition of a material, with the exception of hydrogen and helium atoms, which cannot be detected using this

method.¹⁸¹ The elemental composition is obtained by irradiating a solid sample with a focussed beam of monochromatic x-rays of a known wavelength.¹⁸¹ As a result of the irradiation, electrons are emitted from the surface of the material via the photoelectric effect. From the kinetic energy of the emitted electrons, the binding energy of each can be calculated. Each peak in the spectrum can then be identified by its binding energy, which corresponds to the electronic configuration of the emitted electrons.¹⁸¹

2.3.4 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is perhaps the most widely used characterisation technique in the analysis of organic molecules. This is used to establish the connectivity within a molecule by exploiting the intrinsic angular momentum known as nuclear spin, \mathbf{I} , present in isotopes such as ^1H and ^{13}C , which both have a spin quantum number, I , of $\frac{1}{2}$. Crucially, to be NMR active, an isotope must have $I > 0$. The magnetic quantum number, m_I , dictates the orientation of the nuclear spin, with $2I+1$ possible orientations of m . Each orientation is degenerate in the absence of a magnetic field, however, upon application of an external magnetic field, B_0 , the degeneracy is lifted *via* the Zeeman interaction, with the energy given in Equation 2.47. γ refers to the gyromagnetic ratio of the nucleus.¹⁸²

In the presence of the external magnetic field, the different orientations of the magnetic dipole moment, μ_z , have different energies, with the lower energy configuration, $m_I = +\frac{1}{2}$, occurring when μ_z is aligned with B_0 , known as the 'spin up' configuration, and the higher energy 'spin down' configuration, $m_I = -\frac{1}{2}$, occurring when μ_z is anti-parallel to the external magnetic field. The two states have an energy difference of ΔE , which is dependent on the gyromagnetic ratio, γ and the strength of B_0 (Figure 2.11). Transitions between adjacent spin states are allowed providing that the frequency of the transition, known as the Larmor frequency, corresponds to Equation 2.51,

$$\nu = \frac{\gamma \cdot B_0}{2\pi} \quad (2.51)$$

where ν corresponds to the frequency of the electromagnetic radiation.¹⁸²

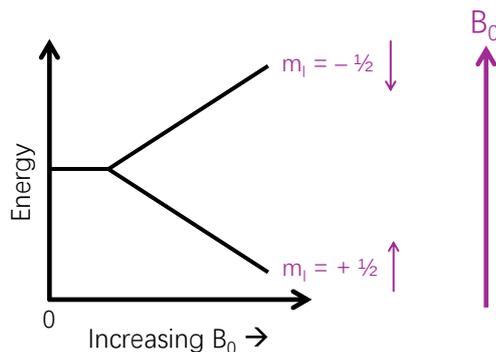


Figure 2.11: Illustration of the nuclear spins in the absence and presence of an external magnetic field. Figure inspired by reference 182.

Whilst the Larmor frequency of an NMR-active nucleus is dependent on the gyromagnetic ratio and external field (Equation 2.51), which will both be constant for each NMR-active nucleus of the same elemental composition within the sample, not every nucleus will experience exactly the same magnetic field.¹⁸² This is due to the local chemical environment of each atom varying slightly depending on the electronic distribution, where some nuclei are more shielded from the external magnetic field than others, creating an effective magnetic field, B_{eff} , around each atom. This creates the concept of chemical shielding, σ . For example, NMR-active nuclei in close proximity to electron-withdrawing groups, which pull electron density towards themselves, are ‘deshielded’, and experience a larger B_{eff} , in contrast to those in close proximity to electron-donating groups.¹⁸²

For ease of analysis, the chemical shieldings are often plotted relative to a reference molecule of known chemical shielding, such as tetramethylsilane (TMS), which is chosen as it has one unique atomic environment for each of carbon, hydrogen, and silicon. This creates a spectrum of the chemical shift, δ (Equation 2.52),¹⁸²

$$\delta = 10^6 \cdot (\sigma_{\text{ref}} - \sigma) \quad (2.52)$$

of each NMR-active atom of the studied isotope (*e.g.* ^{13}C) within the material.¹⁸² σ corresponds to the chemical shielding of the studied atom and σ_{ref} is the chemical shielding of TMS. The chemical shift of each atomic environment can be characterised using data tables to assess the structure of the molecule.¹⁸²

For insoluble materials such as polymers and glasses, which cannot be dissolved within deuterated solvents to obtain their NMR spectra, solid-state NMR (ssNMR) can be utilised.¹⁸³ Whilst this allows for a wider range of materials to be analysed using NMR, it is more challenging to interpret for two

main reasons, both of which cause the resonances in solid samples to broaden compared to solution-state NMR. Firstly, as ssNMR is obtained for a static, solid sample, any structural disorder in the material will be visible due to nuclei in slightly different environments yielding different chemical shifts.¹⁸³ For example, the three hydrogen atoms on a methane molecule will correspond to different chemical shifts due to the inability to tumble and interchange positions throughout the experiment. Secondly, the larger number of ‘crystallites’ within, for example, powdered samples, increases the likelihood that not every crystal will be precisely aligned to the magnetic field, leading to a larger variety of chemical shieldings. This is not ideal, as broadening of the resonances lowers the intensity due to the total integrated NMR signal remaining fixed despite the breadth. Additionally, there is a greater probability that individual resonances will overlap, making resolution of each individual peak more challenging.¹⁸³ Therefore, ssNMR is a useful technique to analyse amorphous materials, yet it is unable to be used to fully resolve the structures without being used as part of a characterisation toolkit composed of a number of techniques.

2.3.5 Fourier-Transform Infrared Spectroscopy

FT-IR spectroscopy, undertaken by irradiating a sample with IR radiation, which has a wavelength in the region of 10^{-5} m, is a frequently used characterisation technique to identify functional groups within molecules. A non-linear molecule with N atoms has $3N - 6$ vibrational degrees of freedom. Chemical bonds within any material vibrate at any temperature above 0 K.¹⁸⁴ However, if the incoming IR radiation corresponds to the same frequency as one of the $3N - 6$ degrees of freedom within the molecule, the bond will absorb energy from the IR radiation, allowing the bond to be excited to higher-energy vibrational states.¹⁸⁵ This corresponds to a peak in the FT-IR spectrum providing that the vibration of a bond causes a change in the molecular dipole moment.^{184,185}

The FT-IR spectrum is plotted as percentage transmittance, corresponding to the percentage of IR radiation emitted, as a function of the wavenumber of each vibration.^{184,185} Characteristic wavenumbers for different chemical environments are widely reported in data tables and can then be used to identify the functional groups within the material based on the spectrum. Typically, this is done by beginning at the higher-wavenumber end of the spectrum, identifying the key peaks, before moving towards the lower-wavenumber end.¹⁸⁵ The fingerprint region, which occurs in the range of $1500\text{--}600\text{ cm}^{-1}$,¹⁸⁵ is unique to each particular molecule. However, this often contains a large number of peaks, which could occur due to a large range of chemical environments, making this challenging to interpret.¹⁸⁵

2.3.6 X-ray Diffraction

X-ray diffraction (XRD) is an incredibly powerful technique used to establish the atomic positions and unit cell parameters of crystalline materials by taking advantage of the similar length scales of x-ray wavelengths and interatomic distances. Similarly to XPS, the technique involves firing a monochromatic beam of x-rays at a sample, which then diffracts the x-rays, giving a 'diffraction pattern' that when solved by computer, yields the atomic positions and unit cell. The sample is rotated during the data collection phase in order to analyse all possible orientations of the crystal.¹⁸⁶

There are two types of sample that can be utilised for XRD: single crystals, and powders. Of these, perhaps the more practical option is powder x-ray diffraction (PXRD), as this does not have the requirement to grow a uniform crystal, which, depending on the material, can take hours to months. Additionally, single-crystal XRD is only able to distinguish the atomic positions and unit cell of the studied crystal, which may not be ideal if the chosen crystal is not representative of the entire sample. In contrast, PXRD is a bulk technique, assessing the crystallinity, bulk purity, and polymorph of the sample, as this incorporates multiple, smaller crystals, ground into a powder.¹⁸⁶

For a reflection to be visible, and therefore observed on the XRD spectrum, Bragg's Law (Equation 2.53)¹⁸⁷

$$n\lambda = 2d \cdot \sin(\theta) \quad (2.53)$$

must be satisfied. This states that the reflected beams must be in-phase and interfere constructively. At any other angle, the reflected beams are out-of-phase, so will cancel out due to interfering destructively.¹⁸⁷

However, whilst crystalline materials have very clearly defined planes within the material due to the periodic nature, each of which correspond to clear diffraction peaks in the powder pattern that can be assigned to specific environments within the sample, this does not apply to all materials. The lack of order within amorphous materials such as those studied within this thesis means that XRD is not the most useful characterisation technique to study amorphous materials. For these materials, there are effectively an infinite number of chemical environments within the structure, leading to one very broad, low intensity hump on the powder pattern, corresponding to the limited short-range ordering, which cannot be identified further.¹⁸⁸ It is possible to calculate PXRD patterns of simulated materials, including the amorphous materials studied here, using software such as Mercury¹⁸⁹ or VESTA,¹⁹⁰ which can be compared to the experimental PXRD patterns. However, it is important to consider that the simulated sample contains far fewer atoms than the

experimental, so the simulated spectrum is likely to be subject to artificial crystallinity, resulting in sharper diffraction peaks compared to the experimental.⁴⁵

2.3.7 Microscopy

Microscopy is a useful tool to visualise the shape and size of particles making up chemical systems. The most common types of microscopy used to visualise chemical systems are scanning electron microscopy (SEM) and transmission electron microscopy (TEM).¹⁹¹ Both techniques utilise a beam of electrons to interact with the sample under high vacuum, and the interaction of the electron beam with the sample generates an image of the particles composing it.¹⁹¹ The main difference between the two techniques is the type of image acquired, which differs due to which electrons are detected. For SEM, the three-dimensional image of the system is generated by collecting the electrons which reverberated off the surface of the material, leading to an image of the material's surface.¹⁹¹ In contrast, TEM collects two-dimensional images of the bulk of the sample, including the morphology and crystal structure, using transmitted electrons which have travelled through the sample before reaching the detector. Additionally, TEM allows for greater resolution of the material (maximum magnification: SEM = \sim 1-2 million times, TEM = $>$ 50 million times).¹⁹¹

2.4 Methodology

A number of software packages were utilised to collect and analyse the data reported within this thesis. Each is summarised below.

2.4.1 Ambuild

Ambuild is an open-source molecular simulation code developed by Trewin to simulate the formation of amorphous microporous polymers such as the materials studied within this thesis.^{19,26,38,45,95,104,192} Ambuild is written in python for easy accessibility and outsourcing to other software codes such as HOOMD-blue and Poreblazer (described later).^{19,162,163,193,194} Ambuild is able to exploit GPU architecture for increased simulation speed, and system size, which is particularly important when considering the simulation of amorphous materials, which do not have a defined unit cell. This allows for a larger number of repeat models to be generated for each material, each a unique 'snapshot' of the experimental sample, giving a more representative model to use for comparison to the real-world polymer framework.^{19,38} Typically, four to five repeat structural models are generated of each material simulated,

allowing for a range of structures to be assessed and the resulting porosity properties averaged.

The driving force behind the development of Ambuild is its ability to model the full synthetic reaction and mimic the catalytic cycle underlying the experimental synthesis of each material considered. This means that by mimicking the experimental conditions as closely as possible, the models generated by Ambuild are not simply designed to match the density or porosity properties of the experimental material. Instead, the polymer itself is formed in a comparable way throughout, allowing for the formation of intermediates and alternative ring features that may occur as a result of the kinetically controlled reaction.^{19,38,192}

Catalytic mechanisms and synthetic reactions that have previously been modelled using Ambuild, many included within this thesis, include: Sonogashira-Hagihara;^{19,38} Friedel-Crafts;²⁶ acid-catalysed cyclotrimerisation;⁴⁵ Eglinton,¹⁹⁵ Yamamoto⁹⁵ and condensation.²³ Materials that have been modelled using Ambuild to date include hypercrosslinked polymers,²⁶ conjugated microporous polymers,^{19,23,38} porous aromatic frameworks,¹⁹⁶ covalent triazine frameworks⁴⁵ and organically synthesised porous carbon.¹⁶

Ambuild is designed to add pre-specified molecular building blocks into a unit cell. Examples of these input files for a methane building block are given in Figure 2.12.¹⁹² Each building block represents one molecular type, input as a .car file containing the atomic positions, atom and forcefield types, and charges.¹⁹² Each .car file has a corresponding .csv file, specifying the reactive sites of that building block. This is done using the end group/cap atom notation (Figure 2.13), where the end group is the atom that will form part of a new bond with another end group that meets the specified bonding rules, and the cap atom is an atom bonded to the end group that is lost on forming the new bond.^{19,38,192} For example, to form a polymer from methane building blocks, the carbon atom is the end group, and the four hydrogen atoms are all classed as cap atoms.

The structure of each building block is drawn, and the geometry optimised using the Forcite package within Materials Studio version 5,¹⁵⁸ as described later. The PCFF forcefield is utilised for the geometry optimisation of building blocks.^{115,192} To run Ambuild, a list of forcefield parameters and all of the building blocks are added to the the working directory along with the python input file. Optimisation of the Ambuild simulation cell occurs using rigid bodies, where each building block is held fixed, and only new interactions formed between blocks are optimised. This occurs by increasing the force constant values (typically by a value of two) from the original values of the PCFF forcefield in order to reflect the rigidity of the system.^{115,192} The rigid body approx-

imation reduces the computational expense of the simulation whilst maintaining a chemically sensible system, due to the pre-optimisation of the building blocks within Materials Studio.^{158,192}

With the input building blocks generated, the user specifies the size of the unit cell, in Ångström, using the *boxDim* command. The unit cell can be specified as cubic or cuboidal. Building blocks can then be added to the unit cell using the *seed* command. This adds building blocks of the specified type and quantity (*e.g.*, to match the experimental stoichiometric ratios) into the cell either at random, or into a given region of the cell, referred to as a *zone*. Bonding rules can be specified to dictate which end groups can bond, providing that they are within the specified bond margin and bond angle margin.¹⁹² There is also an option, as described in Chapter 3 section 3.3, where two building blocks can bond to a catalyst molecule, and upon doing so, can form a bond to each other, before breaking their respective bonds to the catalyst. Building blocks can be deleted and re-seeded at random to allow for increased statistical sampling of the material, and multiple models are typically generated for better comparison to experiment.¹⁹²

a) !BIOSYM archive 3
 PBC=OFF ← Specify periodic boundary conditions?
 Materials Studio Generated CAR File
 !DATE Tue Jan 22 10:35:54 2013

C1	0.000000	0.000000	0.000000	XXXX	1	c	C	0.000
H1	0.000000	0.000000	1.089000	XXXX	1	h	H	0.000
H2	1.026719	0.000000	-0.363000	XXXX	1	h	H	0.000
H3	-0.513360	-0.889165	-0.363000	XXXX	1	h	H	0.000
H4	-0.513360	0.889165	-0.363000	XXXX	1	h	H	0.000

end
end

Atom index Absolute xyz coordinates Forcefield type Partial charges

b)

Type	endGroup	capAtom	Dihedral	delAtom
a	0	1	-1	-1
a	0	2	-1	-1
a	0	3	-1	-1
a	0	4	-1	-1

Type Binding site type
endGroup Index number of atom forming part of new bond
capAtom Index number of atom lost on forming new bond
Dihedral Specify the atom defining the dihedral angle around the bond between two endGroups? If -1 or omitted, this will not be undertaken.
delAtom Delete this atom on forming a new bond? If -1 or omitted, this atom will not be deleted.

Figure 2.12: Ambuild input files for a methane building block. a) .car file, b) .csv file. More information about the terminology used within Ambuild is given in reference 192.

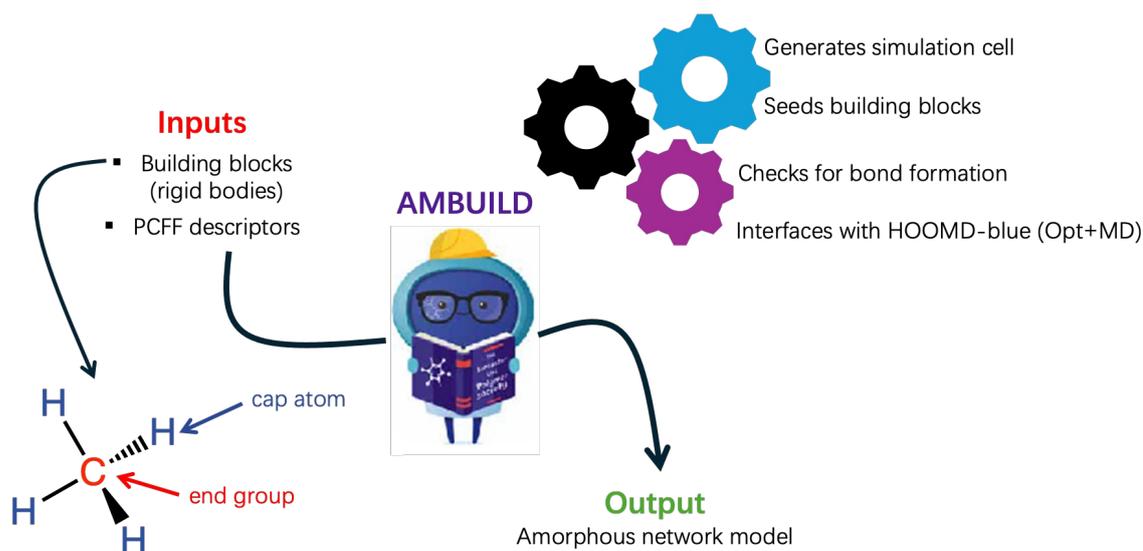


Figure 2.13: The Ambuild polymer formation process. The end group and cap atom notation is illustrated for a methane building block.

Polymers can be generated within Ambuild *via* a number of different mechanisms. The two types used within this thesis are the *growBlocks* and *zipBlocks* steps (Figure 2.14). The *growBlocks* step involves seeding at least one building into the simulation cell, and then directly ‘growing’ subsequent building blocks onto existing blocks within the cell.¹⁹² The atom defining the dihedral angle around the bond between two end groups can be defined within the .csv file for each building block. The acceptance criteria within Ambuild for the *growBlocks* step states that a new bond can form between an additional building block and an existing fragment within the simulation cell providing that i) there is at least one unreacted end group of the specified type, and ii) that there is sufficient free volume within the cell for the additional building block to be added within the maximum number of attempts.¹⁹²

The *zipBlocks* step works by seeding multiple building blocks into the simulation cell, and temporarily loosening the specified bond margin and bond angle margin criteria. These loosened criteria form a cone around available end groups within the cell, with a length of the extended bond margin and a radius at the apex of the extended bond angle margin. Ambuild then assesses whether any new bonds can form with the extended criteria, *i.e.*, whether two end groups are within the same cone, forms any new bonds, and runs a full rigid-body geometry optimisation.¹⁹² Within the *zipBlocks* method, neither of the atoms forming part of the new bond are removed from the cell to form the bond. The bond simply forms as a result of the two end groups being within the same cone during the *zipBlocks* step. Following the bond formation, the rigid-body geometry optimisation allows the atoms within the new bond to find a lower energy conformation, minimising the associated forces.¹⁹²

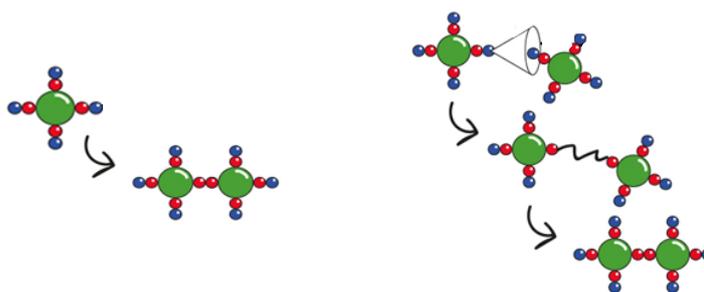


Figure 2.14: Cartoon representation of the Ambuild *growBlocks* (left) and *zipBlocks* (right) steps. End groups are shown in red and cap atoms are shown in blue.

The acceptance criteria within Ambuild for a *zipBlocks* step means that for a bond to occur, two end groups of the predetermined type must be within the specified bond margin and bond angle margin. If this is the case, the bond will form without any extra tests being required.¹⁹²

The cutoff parameters used for the bond margin and bond angle margin criteria must be carefully determined to only allow bonds to form between building blocks that would optimise to form a chemically sensible structure. This means that the cutoff margins must be small enough to prevent a large number of bonds forming simultaneously, as in the case of the hybrid coarse-grained model described within section 3.5 of Chapter 3.¹⁰⁴ This was also the case during the NPT MD steps within the artificial synthesis of the hypercrosslinked polymer described in Chapter 4,²⁶ where as a result of the decreasing cell volume, a larger number of building blocks were within close enough proximity to react. Reducing the bond margin and bond angle margin decreased the number of competing forces from each of the bonds attempting to be minimised during the same geometry optimisation cycle. In contrast, when large numbers of solvent molecules were included within the simulation cell, as in the remainder of Chapter 3, the cutoff margins for each *zipBlocks* step were relatively large, to allow the building blocks, which were often held apart by solvent molecules, to be able to react with each other and form the polymer network.^{19,38} The bond margin and bond angle margin criteria for each of the systems described within this thesis are discussed with justification within the relevant results chapters.

Whilst these two methods to build a polymer within Ambuild are both perfectly reasonable, the approach taken depends on the chemistry of the reaction. For example, the *growBlocks* method was utilised within Chapter 4 section 4.3 to replicate the polystyrene synthesis, as the rapid reaction of building blocks would occur readily within radical building blocks.²⁶ In contrast, the *zipBlocks* method was used within Chapter 3 section 3.3 to mimic the Sonogashira-Hagihara reaction to form CMP-1, where the polymer forms *via* diffusion of building blocks to find one another in the presence of solvent.^{19,38}

The standard HOOMD-blue integration timestep of 0.0001 ps was utilised within the geometry optimisation and molecular dynamics steps of the Ambuild simulation.¹⁹²⁻¹⁹⁴ The simulated polymerisation was classed as 'complete' when no new bonds had formed during the last twenty *zipBlocks* tests. This typically corresponded to an approximate MD runtime of 20 ns. Whilst a longer simulation would allow for a greater quantity of structural sampling to occur, in the absence of new bond formation, the decision was made to prioritise generating additional structures to compare to the experimental material.

Molecular dynamics and geometry optimisation were utilised throughout the Ambuild simulations. These can be specified within the Ambuild input file using the *mycell.runMD* and *mycell.optimiseGeometry* commands, respectively. The MD ensemble was changed between NPT and NVT by editing the *integrator* specified within the Ambuild file. The systems studied within this thesis required frequent geometry optimisation and MD throughout the structure generation in order to allow a chemically sensible structure to form. The frequent MD throughout the structure generation increased the quantity of structural sampling incorporated throughout, leading to a better representation of the experimental materials. It is a standard procedure to use a combination of NVT MD to prevent changes in the unit cell volume during the statistical sampling, and NPT MD to remove unstable pore voids from the material that would not occur experimentally.

The PCFF was utilised throughout the Ambuild simulations as a well-established forcefield to model amorphous polymers.^{19,38,41,104,115} Geometry optimisation and molecular dynamics were outsourced to HOOMD-blue.^{193,194} As a result, the potential forms chosen for the Ambuild simulations needed to be available options within this code. To set up the molecular mechanics forcefield to describe each system, Ambuild uses the harmonic bond and angle potentials and the periodic dihedral potential available within HOOMD-blue,^{193,194} given in Equations 2.54-2.56, respectively.^{106-108,193,194} The harmonic bond potential, $V(r_{ij})$, of the bond between atoms i and j , r_{ij} , is calculated as a function of the force constant k_{bond} and the equilibrium bond length r_0 (Equation 2.54).¹⁰⁷

$$V(r_{ij}) = \frac{k_{\text{bond}}}{2}(r_{ij} - r_0)^2 \quad (2.54)$$

The harmonic angle potential, $V(\theta_{ijk})$, of the angle between atoms i , j and k , θ_{ijk} , is given as a function of the force constant k_{angle} and the equilibrium angle θ_0 (Equation 2.55).¹⁰⁸

$$V(\theta_{ijk}) = \frac{k_{\text{angle}}}{2}(\theta_{ijk} - \theta_0)^2 \quad (2.55)$$

The periodic dihedral potential, $V(\phi_{ijkn})$, of the dihedral angle of the planes traversing atoms i , j and k , and j , k , and n , ϕ_{ijkn} , is calculated using the force constant, k_{dihedral} , sign factor, d , angle multiplicity factor, n , and equilibrium dihedral angle, ϕ_0 (Equation 2.56).¹⁰⁶

$$V(\phi_{ijkn}) = \frac{k_{\text{dihedral}}}{2} (1 + d \cdot \cos(n \cdot \phi_{ijkn} - \phi_0)) \quad (2.56)$$

The intermolecular forces between pairs of atoms located within different molecules are described using the Lennard-Jones potential (Equation 2.57).^{53,197}

$$V_{\text{Lennard-Jones}}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.57)$$

This considers both the repulsive interactions arising due to the Pauli principle as the electron density of the molecules enter close proximity to each other (the first term in Equation 2.57), and the attractive interactions between the non-bonded atoms within the molecules, arising due to van der Waals forces (the second term in Equation 2.57). The Lennard-Jones potential of atoms i and j , r_{ij} , located within two different molecules, is then calculated as a function of the strength of the interaction between the two atoms, indicating the depth of the energy well, ϵ_{ij} , and the i - j distance where the potential energy equals zero, σ_{ij} .⁵³ The van der Waals cutoff radius utilised throughout the Ambuild simulations, unless specified otherwise, was set to 10 Å, as a reasonable distance to consider the significant interactions whilst reducing excessive computational expense.¹⁹

Partial charges within the system are calculated using the Gasteiger¹⁹⁸ approach. This approach was chosen as it is commonly used within the scientific community and is available within a wide range of codes. Additionally, there are open-source software packages available that are able to calculate Gasteiger partial charges, such as OpenEye,^{199,200} allowing for the possibility for these to be incorporated within the Ambuild code in the future to improve automation and allow for charges to be updated throughout the simulation, for example, during the simulated charging and discharging of a battery system. The Gasteiger approach¹⁹⁸ calculates molecular charges based solely upon the orbital electronegativities, which occur due to the molecular connectivity. This emphasises the importance of establishing the correct atom typing within the input structure.¹⁵⁸ Additionally, the Gasteiger approach requires all hydrogen atoms to be explicitly defined within the model to ensure that the calculated charges are representative of the structure.^{198,199} Firstly, the initial partial charges are calculated as the formal charge of each atom, with the exception of conjugated rings, where the partial charges are set as the average of the formal charges of each atom in the ring.¹⁹⁹ Orbital electroneg-

activities, χ_i , are calculated as a function of the orbital occupation (Equation 2.58)

$$\chi_i = a_i + b_i Q + c_i Q^2, \quad (2.58)$$

where Q is the total charge of the atom and a_i , b_i and c_i are dependent on the atom and hybridisation type.^{158,198} This algorithm is iterative, where the electron density is shifted to equalise the relative orbital electronegativities of each sigma bond.¹⁹⁹ Once the initial charges have been calculated, the partial charges are modified to equalise the orbital electronegativities of each atom in a sigma bond.^{158,198} An example of this is the sigma bond between two atoms, i and j , with electronegativities of χ_i and χ_j , respectively. j is the more electronegative atom, so the electron density would be more localised on atom j . For the n^{th} iteration of the Gasteiger algorithm, the charge distribution would be expressed as Equation 2.59:

$$\Delta q_i = \frac{\chi_j - \chi_i}{\chi_i^+} \cdot \left(\frac{1}{2}\right)^n, \quad (2.59)$$

Where χ_i^+ is the electronegativity of the empty orbital.^{158,198} The factor of $(\frac{1}{2})^n$ is a damping factor to limit the transfer of charge between individual atoms in a given iteration. The charges are calculated for each sigma-bonded atom pair in a single iteration, and the new charges are then fed back into the second equation to recalculate the orbital electronegativities. This is continued until all changes in atomic charges are within a specified convergence criteria.^{158,198}

In contrast, the QEq method,^{201,202} which is an alternative method used to calculate partial charges, considers both the atomic electronegativities and molecular geometry within the algorithm. As this approach relies on a more complex algorithm due to the additional dependence on the geometry, this proves challenging on systems of more than ~ 200 atoms.¹⁵⁸ This makes the QEq approach unsuitable to be used to update partial charges throughout an Ambuild simulation, which typically comprises of ~ 200 monomer building blocks, along with thousands of solvent molecules.¹⁵⁸

The QEq approach, as for the Gasteiger approach, is based on equilibrating atomic orbital electronegativities with respect to a charge distribution.^{158,201,202} The energy of a system with respect to charge can be written as a Taylor expansion (Equation 2.60),

$$E_i(Q) = E_{i0} + Q_i \left(\frac{\delta E}{\delta Q}\right)_{i0} + \frac{1}{2} Q_i^2 \left(\frac{\delta^2 E}{\delta Q^2}\right)_{i0} + \dots, \quad (2.60)$$

where E_{i0} relates to the neutral atom.^{158,201,202} By considering the energy

of a system when charged ± 1 , corresponding to the ionisation potential (IP) and electron affinity (EA), the orbital electronegativity of neutral atom i , χ_i^0 , and Coulombic repulsion between two electrons sharing an orbital, J_{ii}^0 , can be given as Equations 2.61 and 2.62, respectively^{158,201,202}

$$\left(\frac{\delta E}{\delta Q}\right)_{i0} = \frac{1}{2}(IP - EA) = \chi_i^0 \quad (2.61)$$

$$\left(\frac{\delta^2 E}{\delta Q^2}\right)_{i0} = IP - EA = J_{ii}^0. \quad (2.62)$$

The energy of the isolated atom can therefore be written as a function of charge using Equation 2.63^{158,201,202}

$$E_i(Q) = E_{i0} + \chi_i^0 \cdot Q_i + \frac{1}{2} J_{ii}^0 \cdot Q_i^2. \quad (2.63)$$

However, isolated atoms are not representative of many molecular systems, and so Coulombic terms are introduced to better describe the interactions between neighbouring atoms (Equations 2.64 and 2.65).^{158,201,202}

$$E(Q_1 \dots Q_N) = \sum \left(E_{i0} + \chi_i^0 \cdot Q_i + \frac{1}{2} J_{ii}^0 \cdot Q_i^2 \right) + \sum_{i < j} Q_i \cdot Q_j \cdot J_{ij}, \quad (2.64)$$

$$E(Q_1 \dots Q_N) = \sum (E_{i0} + \chi_i^0 \cdot Q_i) + \frac{1}{2} \sum_{i,j} Q_i \cdot Q_j \cdot J_{ij}. \quad (2.65)$$

N is the number of atoms in the model and J_{ij} describes the Coulombic interaction between atoms i and j . Numerically differentiating Equation 2.63 gives an atomic scale chemical potential, given in Equation 2.66^{158,201,202}

$$\chi_i(Q_1 \dots Q_N) = \frac{\delta E}{\delta Q_i} = \chi_i^0 + \sum_j J_{ij} \cdot Q_j. \quad (2.66)$$

The optimal charges will occur when this chemical potential is identical throughout, given as Equation 2.67^{158,201,202}

$$\chi_1 = \chi_2 = \dots = \chi_N. \quad (2.67)$$

The partial charges were then used to calculate the Coulombic interactions between non-bonded atoms i and j as a function of the dielectric permittivity of a vacuum, ϵ_0 , the partial charges, q_i and q_j , the charge of an electron, e , and

the distance between atoms i and j , r_{ij} (Equation 2.68).⁵³

$$V_{\text{Coulombic}}(r_{ij}) = \frac{1}{4\pi\epsilon_0} \times \frac{q_i \cdot q_j \cdot e^2}{r_{ij}} \quad (2.68)$$

2.4.2 HOOMD-blue

HOOMD-blue, developed by Glotzer,^{193,194} was utilised as the geometry optimisation and molecular dynamics engine throughout the Ambuild simulations. The default number of optimisation cycles, MD cycles, and the integration timesteps of one million, one hundred thousand and 0.0001 ps were used throughout, and the HOOMD-blue temperature factor set to 55.0, unless otherwise specified.¹⁹² The `mode_minimize_rigid_fire` integration algorithm was used within HOOMD-blue as a way to minimise the energy of rigid body systems.^{203,204} This works by minimising the energy of a group of particles whilst keeping the remaining cell contents in fixed positions.^{203,204} Throughout the MD simulations, HOOMD-blue integrates the degrees of freedom in the model as a function of time.²⁰⁵ To do this, an integrator (*e.g.* the FIRE integration algorithm), the MD ensemble, the forces of each atom and any constraints in the system are specified.²⁰⁵ NVT and NPT molecular dynamics were undertaken using the Nosé–Hoover thermostat and the Martyna–Tobias–Klein (MTK) barostat-thermostat, respectively.^{145,193,194,206}

In some instances, as in the NPT MD simulation of the hypercrosslinked polymer discussed in Chapter 4 section 4.3, it was necessary to reduce the integration timestep used within the geometry optimisation and MD. The purpose of this was to allow the forces acting on each atom to readjust and become easier to compute by HOOMD-blue. The large forces occurring within the simulation are hypothesised as being due to the non-bonding van der Waals interactions between pairs of atoms located within different molecules in the system, modelled using the Lennard–Jones equation (Equation 2.57).^{53,197} The generic shape of this potential, $V_{\text{Lennard-Jones}}$, as a function of distance, r_{ij} , is given in Figure 2.15. It can be seen that to the left of the plot, at small r_{ij} values, the corresponding potential energies, $V_{\text{Lennard-Jones}}$, are approximately vertical.^{53,197} As the interatomic forces are proportional to both the potential energies and the distance between atoms, to minimise the forces in this high energy region and ensure a small movement of the atoms as a function of time, a small timestep is required. Then, moving towards larger r_{ij} values, the associated potential energy and forces decrease, meaning that a larger timestep can be utilised. This is comparable to umbrella sampling, where a decreased step size is used when approaching the minimum energy structure to allow the minimum energy structure to be identified.²⁰⁷

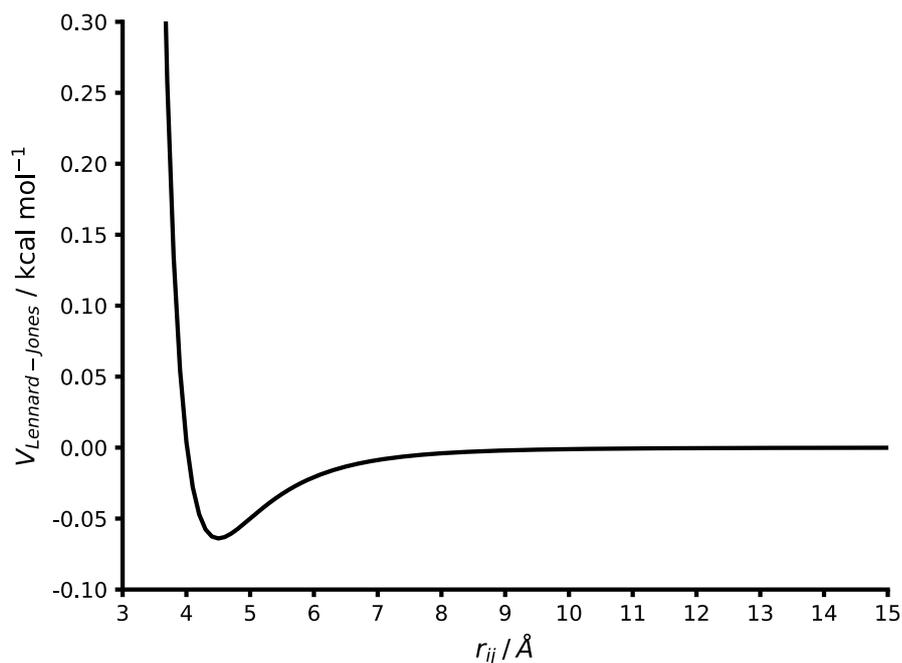


Figure 2.15: The Lennard–Jones potential, showing the potential energy between pairs of atoms located within different molecules in the system, $V_{\text{Lennard-Jones}}$, as a function of the distance between atoms, r_{ij} . The epsilon and sigma values of $0.064 \text{ kcal mol}^{-1}$ and 4.01 \AA , respectively, were taken from the PCFF for a ct-ct (carbon triple bond) interaction. To show the minimum energy well, the y-axis was cut off at an energy of $0.30 \text{ kcal mol}^{-1}$, however the real energy is much larger (approximately $4 \times 10^{18} \text{ kcal mol}^{-1}$).

HOOMD-blue in particular was chosen as, like Ambuild, it is written in Python, allowing for easy integration between the two codes, making it more facile for the structure to be taken from Ambuild, optimised using HOOMD-blue, and then fed back into Ambuild. Additionally, HOOMD-blue is GPU-based, enabling for larger simulation cell sizes to be modelled within the same timescale as would be possible using a CPU-based code, or for more simulations to be obtained within the same timescale.^{193,194} This is in contrast to other MD engines such as DL_POLY, which is written in Fortran and CPU-based.²⁰⁸ However, it is worth noting that there are downsides to using HOOMD-blue as the optimisation and MD engine for the materials studied within this thesis, as it is written for physical systems, which makes interpretation of some of the input parameters challenging. An example of this is the temperature, which is specified as a factor concerning the thermal energy within the system, with the actual temperature being dependent on the kinetic energy of the simulation cell at a particular time and the Boltzmann constant.²⁰⁹ In contrast, DL_POLY specifies each parameter, including the temperature, in the CONTROL input file as an absolute value.²⁰⁸

2.4.3 Poreblazer

Poreblazer was developed by Sarkisov^{162,163} as a tool to simulate the porosity of materials such as metal-organic frameworks, but can also be applied to disordered materials such as those discussed in this thesis. The user specifies as inputs an .xyz file, an input.dat file, a defaults.dat file, and a UFF.dat file. The .xyz file contains the atomic coordinates of the system. The input.dat file specifies the name of the .xyz file and unit cell dimensions. The defaults.dat file contains a range of inputs including the forcefield utilised for the calculation (UFF by default), Lennard-Jones sigma values for helium and nitrogen atoms, the largest anticipated pore size, and the number of samples per atom. Finally, the UFF.dat file is composed of a list of atom types and their corresponding Lennard-Jones sigma and epsilon values, and molecular weights.^{162,163}

Poreblazer generates the total and network-accessible surface areas and pore volume, and the simulated pore size distribution of a modelled material.^{162,163} The surface areas are conceptually considered as rolling a probe molecule across the surface of the material and calculating the area that the centre of the probe travels. *In silico*, this is actually based on a Monte Carlo approach where each atom is described by a sphere of diameter σ , where σ is the sum of the atom diameter and probe diameter. Probe points are randomly placed on the surface of each sphere and tested for overlaps with framework atoms, and the area of each sphere (containing only the percentage of the area that does not overlap with framework atoms) is calculated. The overall surface area comes from the sum of the areas of each atom and is then converted from \AA^2 to $\text{m}^2 \text{g}^{-1}$.^{162,163} The pore volume is also calculated using a Monte Carlo approach. The network accessibility of a material to a particular probe size is acquired by splitting the unit cell into cubelets of a pre-specified size.^{162,163} A probe particle is placed within each cubelet, and if there are no overlaps with any of the atoms in the system, the cubelet is classed as 'occupiable'. Once all of the cubelets have been assessed, the network accessibility is classified. This is done by identifying the largest cluster of consecutive occupied cubelets using the Hoshen-Kopelman cluster labelling algorithm,²¹⁰ and analysing whether this cluster forms a continuous path from one side of the simulation cell to the other. If a continuous path is found, the network is deemed accessible.^{162,163}

Poreblazer calculates the geometric pore size distribution.^{162,163} A geometric pore is defined as the largest spherical pore that can form without any overlaps with atoms in the system. Poreblazer calculates the PSD using a Monte Carlo approach, where the pore size range is split into a number of bins using the specified largest anticipated pore size and number of bins.^{162,163} A test point is placed within each bin and tested for overlap with any atoms in the

system. Following this, the largest pore that can form that contains the test point without overlapping with any atoms is identified.^{162,163} Once calculated, the value of the bin corresponding to that pore radius and all smaller radii is increased by one, indicating at least one pore of that size within the system. Once complete, the bin distribution is normalised.^{162,163} This corresponds to the cumulative pore volume function, $V_p(\mathbf{r})$, the free volume that can be occupied by spheres of radius less than or equal to \mathbf{r} . The PSD function, $dV_p(\mathbf{r}) d\mathbf{r}$, is obtained as the derivative of $V_p(\mathbf{r})$.^{162,163} Poreblazer 4.0 was used throughout Chapters 3, 4 and 6 to generate simulated porosity data of the generated Ambuild models.^{162,163}

2.4.4 Materials Studio

Materials Studio 5.0,¹⁵⁸ managed by BIOVIA (formally Accelrys), is a powerful software suite used for drawing, visualising and performing quantum mechanical and/or molecular dynamics calculations on chemical systems. Modules included within Materials Studio 5.0 include Gaussian, CASTEP, Forcite and Discover, along with means to generate polymer structures.¹⁵⁸ Materials Studio 5.0 was used throughout to generate the Ambuild building block input files. Each was generated by drawing the structure in Materials Studio and obtaining the PCFF typing of each atom using the Discover module. The structures then underwent a molecular mechanics-based geometry optimisation using the PCFF forcefield within the Forcite molecular dynamics module. The Forcite module was chosen to optimise the building blocks due to the convenience of drawing the structure, optimising the geometry using the forcefield required for Ambuild, and exporting the structure all within Materials Studio, allowing for a .car file to be generated. This ensures that the building block is given within a filetype that contains all of the required information for Ambuild to occur.¹⁵⁸ The partial charges of each atom were calculated using the Gasteiger approach.^{158,198} In addition to being able to perform calculations on structures, Materials Studio 5.0 also contains a range of analysis tools, including modelling atom volumes and surfaces, generating concentration profiles, and isotherm simulation. Materials Studio 5.0 was used in Chapter 3 section 3.4.3 to generate the densities and surface areas of the sliced Solvent/Monomer Phase Interface Systems.^{38,158}

2.4.5 DL_POLY

Developed by Daresbury Laboratories, DL_POLY 4.0²⁰⁸ was used in Chapter 4 section 4.3 to simulate the isomerisation and diffusion simulations. DL_POLY is a classical molecular dynamics code written in Fortran, and is free to academic institutions who acquire a license to use it for non-commercial research.²⁰⁸ DL_POLY requires three input files: CONTROL, CONFIG and FIELD.²⁰⁸

The CONTROL file contains the commands telling DL_POLY the nature of the calculation to complete and the parameters required to complete it, such as the MD ensemble, temperature, pressure, number of steps, integration timestep, and cutoff parameters. The CONFIG file contains the dimensions and periodicity of the unit cell and the atomic coordinates, velocities, forces and forcefield types.²⁰⁸ The FIELD file contains forcefield information for the system, specifying all of the required equilibrium bond lengths, angles, dihedrals and pair potentials. The CONFIG file in each case was generated directly from the Ambuild output file and the parameters used in the FIELD file for the HCP were obtained from the Ambuild simulation.^{26,208} This was possible due to the in-built functionality within Ambuild to write the structural output as a CONFIG file and create a FIELD file containing all required forcefield parameters. The *traj* command can be added to the CONTROL file to enable the production of a HISTORY file, composed of a list of 'frames'. Each frame is the CONFIG file for a particular time in the simulation, and the sequential addition of frames to the HISTORY file allows for the generation of a trajectory.²⁰⁸

2.4.6 Visual Molecular Dynamics

Visual Molecular Dynamics (VMD)²¹¹ is a visualisation tool designed to analyse structures of biological systems, however, it can also be used to visualise other materials such as those described within this thesis. VMD is able to input a large number of file types, including the HISTORY files generated by DL_POLY.^{208,211} In addition to visualisation, VMD can also be used to monitor bond lengths, angles, and dihedrals (as undertaken in Chapter 4 section 4.3) and to compute radial distribution functions, as in Chapter 4 section 4.4.2. VMD was used within Chapter 4 section 4.4.2 to collect the diffusion trajectories of azobenzene and carbon dioxide within the HCP throughout the DL_POLY simulations.^{26,211}

2.4.7 Gaussian

Gaussian 09²¹² is a very common electronic structure code utilised to generate properties of molecules.²¹² Gaussian is able to compute single-point energies; geometry optimisation using quantum mechanical or molecular mechanics approaches; vibrational frequencies; nuclear magnetic resonance, and absorption, among others.²¹² Within this thesis, a range of DFT calculations were undertaken using Gaussian with a variety of exchange–correlation functionals and basis sets. Chapter 4 section 4.4.2 used Gaussian to calculate the binding energy of carbon dioxide interacting with each isomer of azobenzene.²⁶ Chapter 5 section 5.4.1 used Gaussian 09 to optimise and calculate the NMR spectra of small fragments that may be expected within the OSPC-1

material.⁹⁵ Chapter 6 section 6.4.2 used Gaussian 09 to calculate the IR spectra of fragments optimised using HOOMD-blue within the Ambuild model of CTF-1.⁴⁵

2.4.8 CASTEP

CASTEP is a materials modelling code available for free to academic institutions with a license.²¹³ CASTEP is designed to model periodic systems using quantum mechanical approaches such as DFT and HF. Plane-wave basis sets and pseudopotentials are utilised to model the properties of the studied systems.²¹³ Calculations that can be undertaken using CASTEP include geometry optimisation; single-point energies; nuclear magnetic resonance; molecular dynamics; band structure, and optical spectroscopy.²¹³ A variety of XC functionals can be used within CASTEP including LDA,⁷⁷ GGAs such as PBE,⁸⁶ meta-GGAs such as rSCAN,²¹⁴ and hybrid functionals such as B3LYP.^{79,85,90,91,213}

The first step to a CASTEP calculation involves inputting a 'guess' structure which approximates the structure of a material and the unit cell parameters.²¹⁵ It may be advisable, as was undertaken in this thesis, to optimise the geometry of the guess structure before computing the NMR data, in order to minimise the forces on the system and fully relax the structure and unit cell parameters.²¹⁵ Firstly, the energy of the system is minimised with respect to the basis set at the current geometry and the forces on each atom computed.²¹⁵ This may take up the bulk of the calculation expense, depending on how close a fit to the ground state the guess structure is. The NMR calculation then begins. A Gauge Including Projector Augmented Wave (GI-PAW)²¹⁶ calculation is undertaken, which utilises the plane-wave basis set, pseudopotentials, and the Projector Augmented Wave (PAW) approach²¹⁷ to calculate the NMR tensors of the system.²¹⁵ Ultrasoft pseudopotentials were used in the models described within this thesis.²¹⁸ CASTEP version 8.0²¹³ was used within Chapter 5 sections 5.4.2 and 5.4.3 to run geometry optimisation^{66,68,86,97,213,219-221} and NMR^{66,68,86,97,213,215,216,218,220,222} calculations of fragments of the OSPC-1 material, and OSPC-1 in different crystalline topologies.^{95,213} PBE⁸⁶ was utilised as the XC functional throughout the CASTEP simulations within this thesis due to the trade-off in computational accuracy and expense.²¹⁵

Chapter 3

Artificial synthesis of a conjugated microporous polymer as a case study for rationalising the porosity properties of CMP materials

Never teach a pig to sing. It wastes your time and annoys the pig.

Seen on a greetings card

3.1 Summary

Conjugated microporous polymers (CMPs) are fully amorphous, conjugated polymers with a range of applications due to the inherent flexibility in the synthetic strategies and monomers used to make them. Due to the amorphous nature of these materials, their properties are highly dependent on the synthetic conditions, such as the choice of reaction solvent. Whilst this has been known for a number of years, little research has explored why the choice of reaction solvent can influence the porosity of the resulting material. This led to the work described in this chapter and reported by Mollart and Trewin,³⁸ where an artificial synthesis approach is used to rationalise the porosity of CMP materials synthesised in solvents of varying polarities. Following the full synthetic methodology and catalytic mechanism, the first CMP material, CMP-1, was modelled using four solvents with a range of polar-

ity values. The models incorporated varying quantities of solvent molecules and different starting configurations and were assessed both before and after network formation. This led to the rationalisation of the differing porosity properties of CMP materials as occurring due to varying phase separation dependent on the polarity of the solvent relative to the growing polymer framework. However, the maximum system size that can be modelled using atomistic molecular mechanics is a limitation. To fully understand the larger-scale effects occurring within these frameworks such as phase separation and gelation, alternative approaches, such as coarse grain methodologies, must be employed to model systems with a greater quantity of atoms.

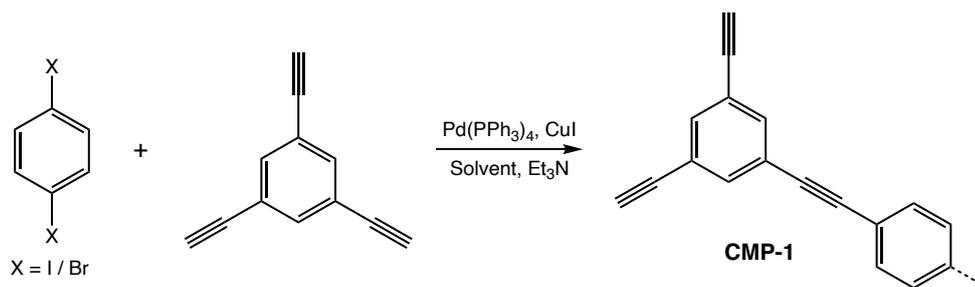
3.2 Introduction

CMPs are made from lightweight elements such as carbon, hydrogen, nitrogen, and oxygen, linked together by alternating single and double bonds, giving a fully π -conjugated polymer skeleton.^{5,20,223,224} Inefficient packing of the polymer in the solid state gives rise to permanent, three-dimensional micropores throughout. CMP materials are also fully amorphous by definition due to being formed under kinetic control, meaning that there is no clearly defined crystal structure, leading to a very broad powder x-ray diffraction pattern with no clear peaks to identify.^{5,20,223,224}

CMPs have applications in a variety of areas, including, but not limited to, energy storage,^{16,23,225} photocatalysis,²²⁶ molecular separations,²²⁷ adsorption,²²⁸ CO₂ capture,²²⁹ catalysis,²³⁰ sensing,²³¹ and cancer treatments,^{223,232} which means many papers have been reported since the initial study in 2007 by Jiang and co-workers (Scheme 3.1, Figure 3.1).^{13,223}

The first CMP, CMP-1, is also now commercially available,²³³ showing the increased interest in these materials. Example CMP materials with applications in energy storage include pyrene-based CMPs, which can show a tailored band gap depending on the monomers employed within the synthesis,²²⁶ aza-CMP, which shows promise as a supercapacitor,^{22,23} and organically synthesised porous carbon (OSPC)-1, which has an application as a battery anode for lithium-ion batteries (Table 3.1).¹⁶

CMPs are made up of nodes connected by linkers (Figure 3.2a). Each node and linker building block is classified with the notation of C_n , where n is the number of available bonding sites, with a minimum n of 2 (Figure 3.2b).²⁰ For self-condensation reactions, both building blocks have the same number of bonding sites, *e.g.*, $C_2 + C_2$, whereas for cross-coupling reactions, one of the building blocks has a larger number of bonding sites, *e.g.*, $C_2 + C_3$, which was used in the first reports of CMP materials.^{13,20,224,234-237}



Scheme 3.1: Reaction scheme to form CMP-1, first reported by Jiang and co-workers.¹³ Scheme reproduced from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

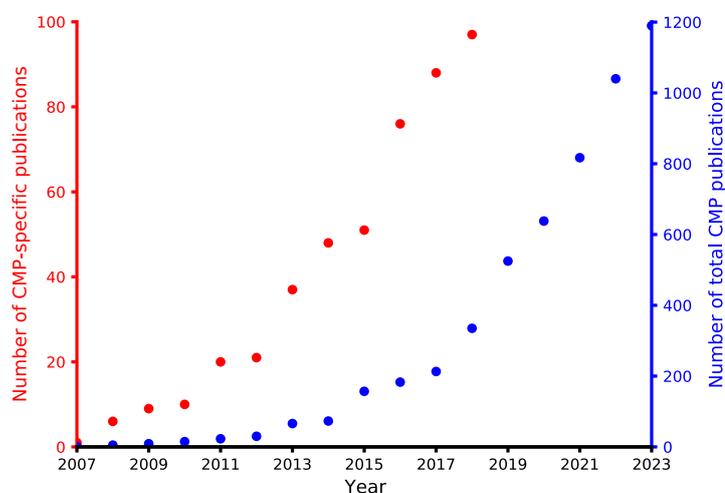
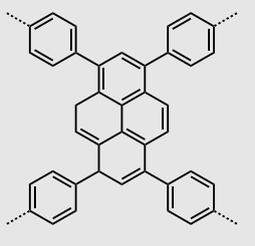
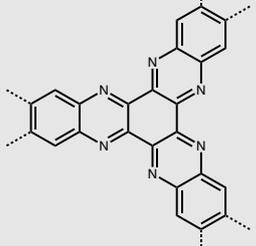
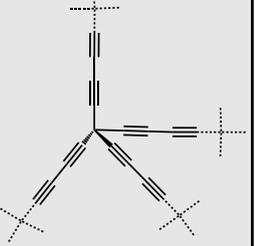


Figure 3.1: Plot showing the number of 'CMP papers' over time. Key: red – number of papers published between 2007–2018, not considering papers that only briefly mention CMPs as part of the introduction and focussing on those that refer back to references 13, 224, 238, with the data obtained from reference 223 using Web of Science; blue – total number of papers published between 2007–2023, with the data accessed on 29th January 2024 using the search term of 'conjugated microporous polymer' within Google Scholar.

The properties of the resulting CMP material can be tuned by the properties of the node and linker used in the synthesis, including the length of the linker and the nodal dimensionality.²⁰ In metal-organic and covalent organic frameworks, a longer linker is able to hold the nodes further apart in three-dimensional space, giving an increased porosity.²⁰ However, for CMPs, the opposite trend is observed (Figure 3.3), where the increased flexibility within the linkers as they increase in length causes a higher quantity of network interpenetration within the structure.^{5,20,223,224} This involves nearby polymer chains intertwining, which leads to an increase in van der Waals intermolecular forces and so lowers the system energy but also decreases porosity. In comparison, shorter, more rigid linkers and nodes with more connecting groups help to hold the structure open by reducing the degree of network interpenetration available within the structure.^{5,20,223,224}

Table 3.1: Examples of CMPs with applications in energy storage.^{16,22,226} Table reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

CMP name	Pyrene-based CMPs	Aza-CMP	OSPC-1
Structure			
Specific property	Tuneable band gap	Formation of electrical double layer	Electron conduction
Application example	Photocatalysis	Supercapacitor	Battery anode

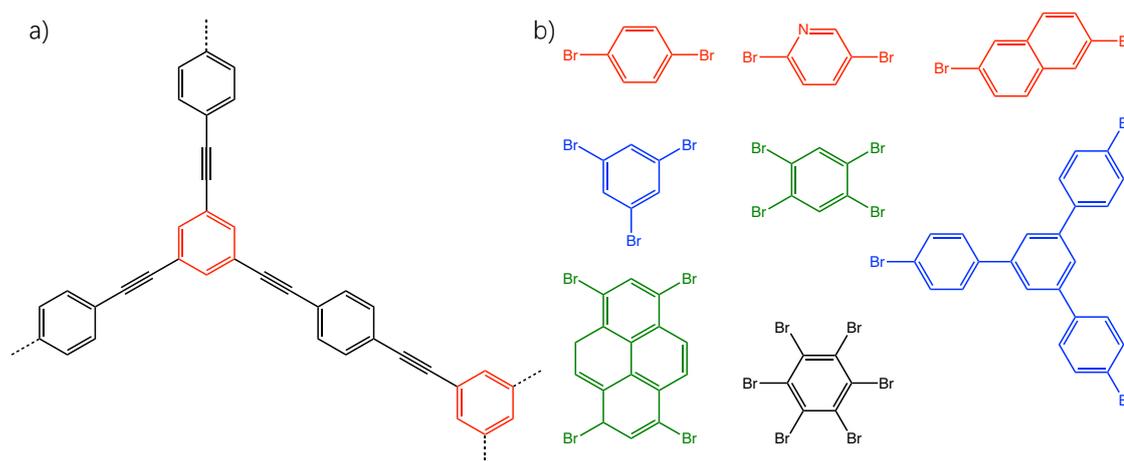


Figure 3.2: CMP building blocks. a) Illustration of the node and linker in CMP-1, with the nodes coloured red and the linkers black. b) Example brominated aryl CMP building blocks showing the C_n nomenclature. Key: C_2 building blocks – red, C_3 – blue, C_4 – green, C_6 – black.^{20,224}

Nodal dimensionality (Figure 3.4) is the deviation of the reactive groups within the framework from one plane, for example, aza-CMP forms as a flat sheet, whilst other CMPs such as porous organic polymer (POP)-3, have some reactive groups that distort from the plane, giving a more three-dimensional structure.²⁰ In the case of POP-3, which is formed from a trimerisation reaction of 1,3,5-triethynylbenzene, the high connectivity and strain of the monomer leads to a rigid, three-dimensional polymer framework.²³⁹ The nodal dimensionality can be assessed computationally by calculating the in-plane and out-of-plane node-linker angles of fragments consisting of two nodes connected by a linker. The angles will give an idea of how two- or three-dimensional

the network is, bearing in mind that this is calculated for small fragments and these angles will vary in the overall amorphous material.²⁰ Having a more three-dimensional network will prevent the polymer from packing as efficiently in the solid state, which may lead to an increase in porosity providing that the increased dimensionality does not cause an increase in the network interpenetration observed.^{5,20,223,224} The connectivity of the node and linker is also important as this can influence the porosity of the polymer. For example, a polymer made from the reaction of homocoupled 2,2',7,7'-tetrabromo-9,9'-spirobifluorene gives a Brunauer-Emmett-Teller (BET) surface area of 1275 m² g⁻¹. Reacting this node with *para*-, *meta*-, and *ortho*-dibromobenzene linkers gives resulting CMPs with BET surface areas of 887 m² g⁻¹, 361 m² g⁻¹ and 5 m² g⁻¹, respectively. The reduction in the surface area was attributed to the increased steric hindrance of the linker.^{224,240}

One of the key features of CMPs is the synthetic diversity in the coupling reaction employed, as shown in Schemes 3.2 and 3.3, which gives rise to a large library of different monomers to use in the synthesis.^{13,223,226,241,242} This increases the number of possible applications due to the flexibility in the synthetic conditions and differing properties of the various monomers that can be used.^{16,23,226,243} Each reaction requires different synthetic conditions, however, many require group 10 catalysts such as palladium or nickel, which can increase the cost of the materials when considering scale-up applications.^{223,224} CMPs can also be modified post-synthesis, for example, the reaction of a radical thiol-yne with a CMP-1-type structure, given in Scheme 3.4.^{224,244}

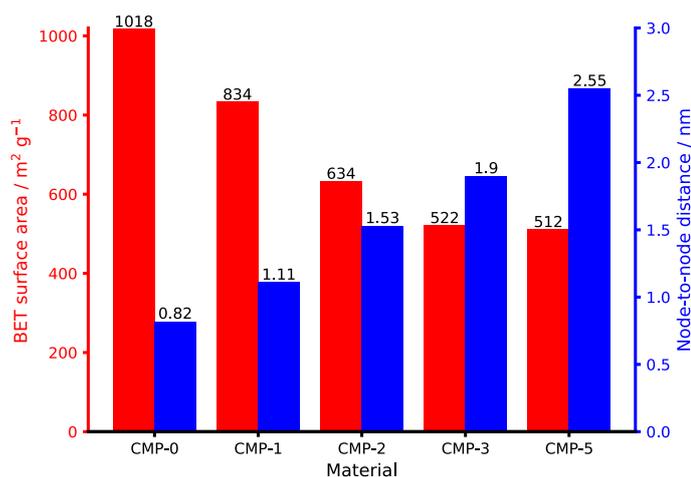


Figure 3.3: Plot showing the influence of node-to-node distance (linker length, blue) on the experimental BET surface area (red) of the resulting CMP material for the original family of poly(aryleneethynylene) CMPs.^{13,236} Data reprinted (adapted) with permission from J.-X. Jiang, F. Su, A. Trewin, C. D. Wood, H. Niu, J. T. A. Jones, Y. Z. Khimyak and A. I. Cooper, *J. Am. Chem. Soc.*, 2008, **130**, 7710–7720. Copyright 2008 American Chemical Society.

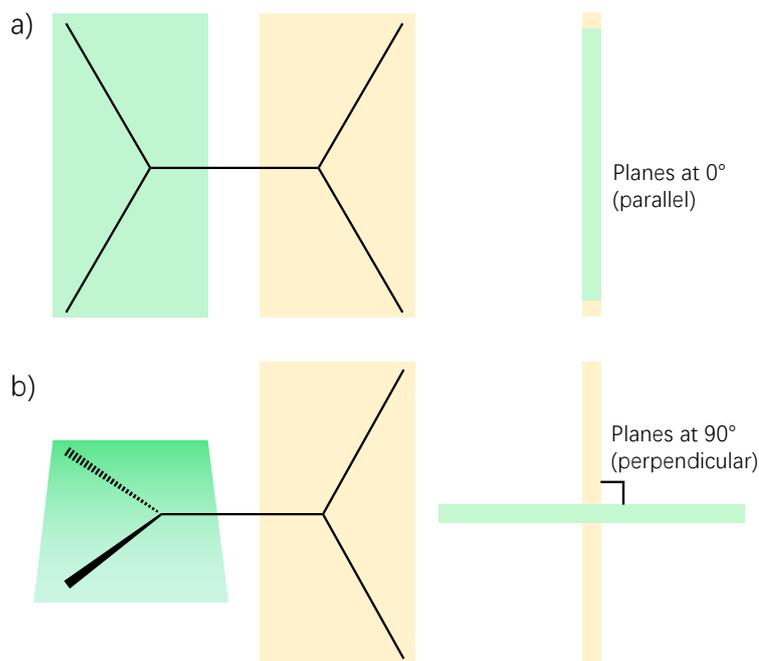
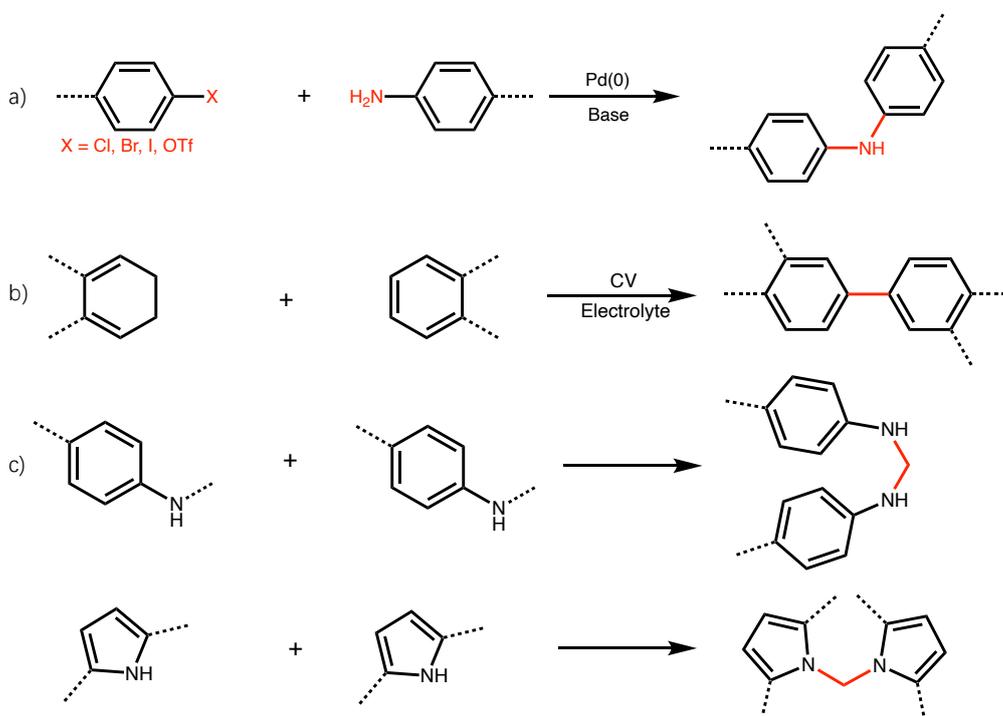
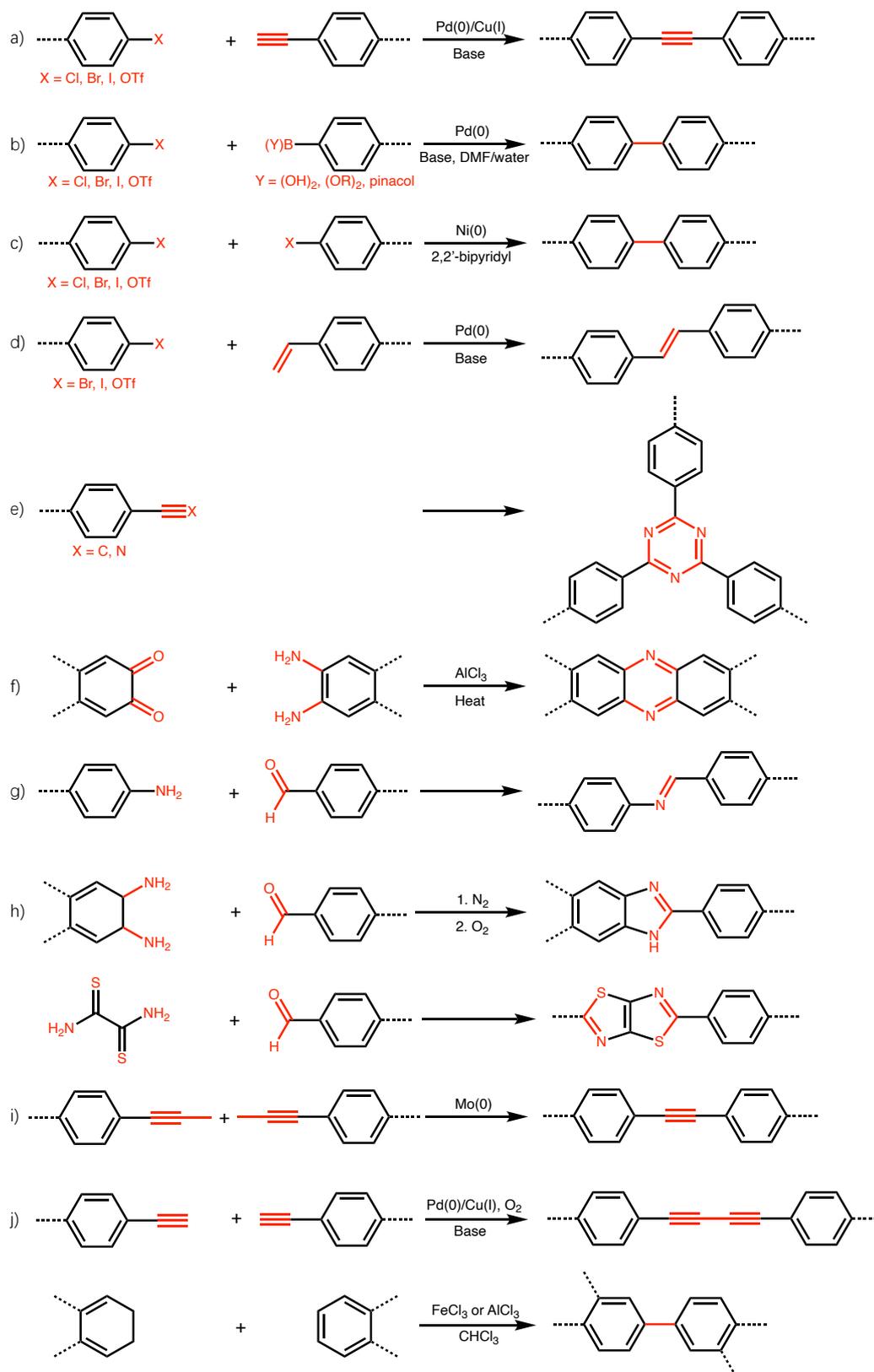


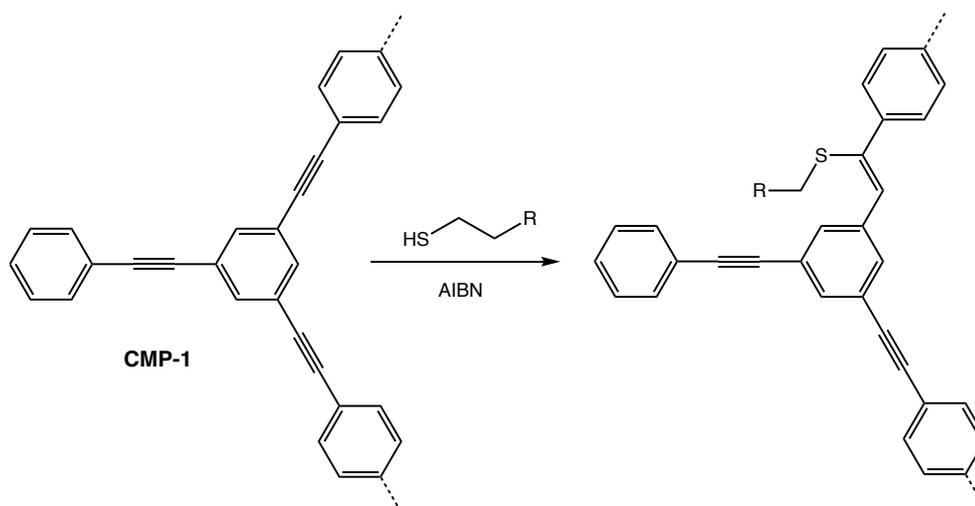
Figure 3.4: Illustration of a nodal dimensionality of a) 0° (as in aza-CMP) and b) 90° (as in POP-3). The two planes comprising the dihedral are shown in green and yellow, with a two-dimensional representation on the right. Figure adapted from reference 20 with permission from the Royal Society of Chemistry.



Scheme 3.2: Reaction mechanisms employed to synthesise CMP materials. a) Buchwald–Hartwig, b) electropolymerisation, c) hypercrosslinking linear polymers. Scheme adapted with permission from reference 223, licensed under CC-BY. Copyright 2020 American Chemical Society.



Scheme 3.3: Reaction mechanisms employed to synthesise CMP materials. a) Sonogashira–Hagihara, b) Suzuki–Miyaura, c) Yamamoto, d) Heck, e) cyclotrimerisation, f) phenazine ring fusion, g) Schiff-base, h) heterocycle linkages, i) alkyne metathesis, j) oxidative coupling. Scheme adapted with permission from reference 223, licensed under CC-BY. Copyright 2020 American Chemical Society.



Scheme 3.4: Example reaction scheme showing the post-synthetic modification of CMP-1 by reaction with a radical thiol-yne. Figure reprinted (adapted) with permission from B. Kiskan and J. Weber, *ACS Macro Lett.*, 2012, **1**, 37–40. Copyright 2012 American Chemical Society.

Many CMPs are formed as insoluble (often brown) powders due to their hypercrosslinked polymer backbones, which can make them challenging to analyse and limit applications to those that keep the materials in the solid state.^{13,234–237,241} However, some examples are known of soluble CMPs that can be used to form thin films,^{245,246} membranes,²⁴⁷ sponges,^{248,249} and hydrogels.^{223,250} The hypercrosslinked structure also means that these materials are chemically and thermally stable, even under acidic conditions.^{13,234–237,241}

The first CMPs, as seen in Scheme 3.1, were synthesised using a Sonogashira-Hagihara (Figure 3.5) reaction of a C₃ node with C₂ linkers: 1,3,5-triethynyl benzene with 1,4-diiodobenzene in the case of CMP-1, with a respectable BET surface area of 834 m² g⁻¹. A 1.5 molar excess of alkyne was utilised as this gave the maximum surface areas obtained.¹³ The Sonogashira-Hagihara scheme involves the oxidative addition of the aryl halide to the Pd(II) catalyst. This is followed by transmetalation of a copper-substituted alkyne, formed by the copper(I) side cycle involving the reaction of the alkyne with copper halide in the presence of base, onto the palladium catalyst. At this point, both monomers are bonded to the Pd catalyst molecule. Then, an isomerisation step places the two monomer ligands *cis* to one another, before reductive elimination yields the cross-coupled product and regenerates the Pd(II) catalyst.^{19,251} The presence of copper within the reaction gives an increased reactivity,²⁵² although copper-free Sonogashira-Hagihara reactions are known, in some cases giving rise to higher surface areas than those with copper iodide present.^{223,253} It was also found that the 1,3,5-triethynylbenzene node was able to react *via* alkyne-alkyne homocoupling to produce poly(phenylene butadiynylene) CMPs under the same synthetic conditions (Scheme 3.5).²⁵⁴ Ad-

ditionally, other catalysts such as iron and nickel centres have been explored for Sonogashira–Hagihara reactions,^{255–258} reducing the associated cost by replacing the expensive palladium catalyst with a cheaper alternative.

It was later found that these CMPs could also be synthesised using dibromo linkers instead of diiodo, which was beneficial as despite the lower reactivity of the bromine monomers,^{234,259} they are more commonly available which reduces the cost when considering scale-up applications and the ease of synthesis of new materials. The lower reactivity of the dibromo monomer was supported by solid-state nuclear magnetic resonance (ssNMR) spectroscopy and elemental analysis, which indicated a larger proportion of unreacted halide end groups (and a lower degree of polycondensation) in the CMP-1 synthesised from the dibromo monomer compared to the diiodo.^{13,234,236}

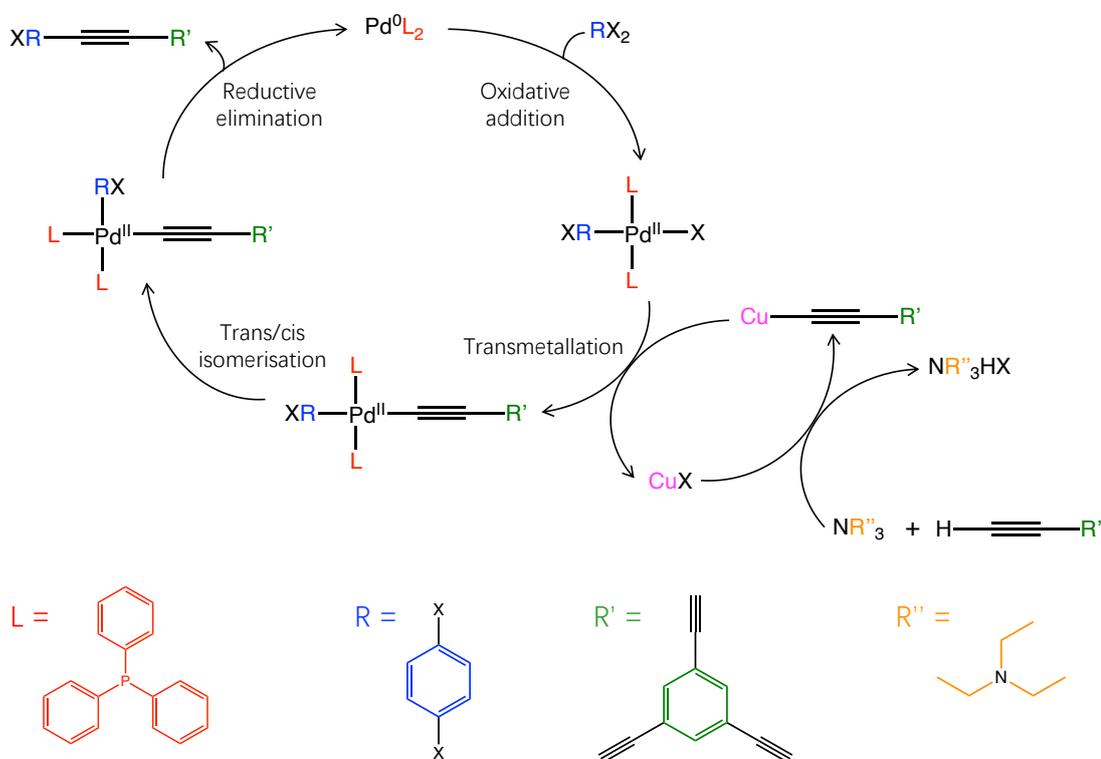
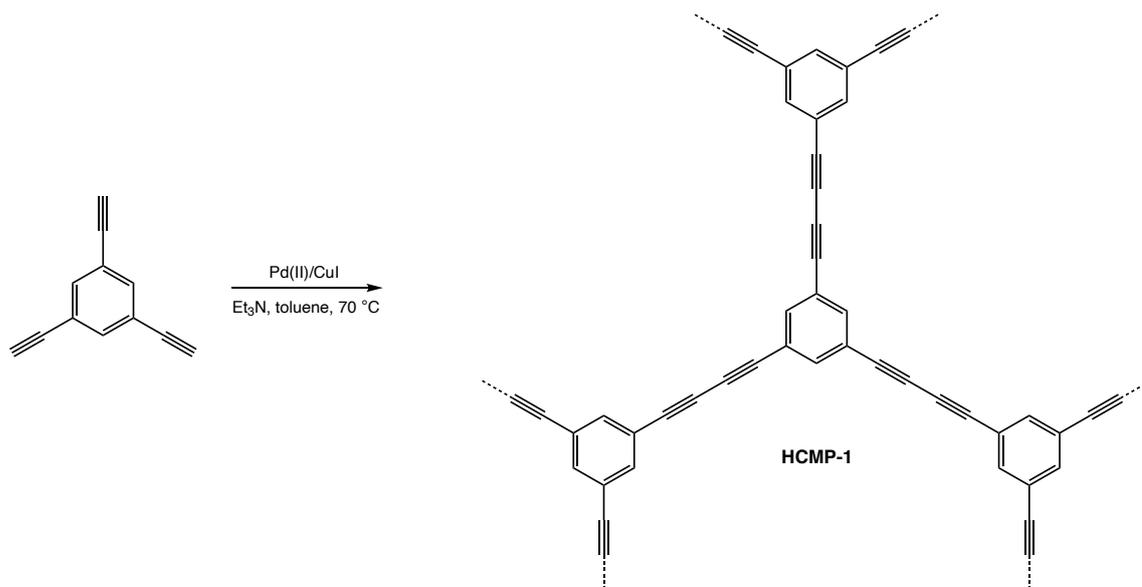


Figure 3.5: The Sonogashira–Hagihara reaction scheme and the ligands used in the synthesis of CMP-1. In the case of CMP-1, R is a phenyl ring with a *para*-substituted halogen atom (Br/I, to match the X group), R' is a phenyl ring with ethynyl groups in the 3- and 5- positions, R'' is an ethyl group, and L is a triphenylphosphine ligand. Figure reproduced and adapted from reference 104, which contributes to the work in this chapter, with permission from Institution of Chemical Engineers (IChemE) and the Royal Society of Chemistry.



Scheme 3.5: Reaction scheme to synthesise homocoupled CMP HCMP-1.²³⁶ Scheme replicated with permission from reference 223, licensed under CC-BY. Copyright 2020 American Chemical Society.

For example, in the ssNMR spectra, the theoretical ratio of alkyne to aromatic peaks assuming complete polymerisation of the alkyne groups would be 0.40, the ratio in the CMP-1 network formed from the diiodo monomer was 0.27, and the CMP-1 formed from the dibromo monomer had a ratio of 0.16. Elemental analysis of the CMP-1 network formed from the diiodo monomer gave a halide weight percentage of 2.27 wt.%, whilst the network formed from the dibromo monomer yields 3.64 wt.% Br.^{13,234,236} This makes the presence of copper iodide even more important, as this increases the reactivity of the system.²⁵² The choice of halide also influences the porosity, with the CMP-1 network synthesised using a diiodo monomer exhibiting more type I behaviour in the nitrogen sorption isotherm compared to that from the dibromo monomer, which resembled type IV more closely (Figure 3.6).^{13,234,236} This was proposed as being due to the lower reactivity of the dibromo monomer leading to a difference in the phase separation behaviour and interparticulate porosity within the structures, where the degree of solvation of the material may differ depending on the structure.^{13,234,236}

The experimental network formation mechanism of CMP-1 synthesised in *N,N'*-dimethylformamide (DMF) was discussed in detail by Laybourn and co-workers,²³⁷ where the framework was analysed after varying reaction times. It was found that at early reaction times (< 40 minutes), no insoluble polymer was observed.

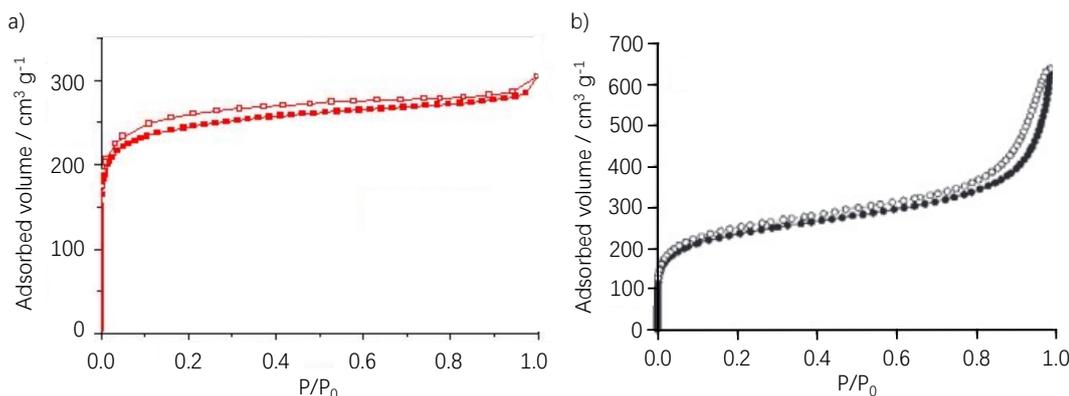


Figure 3.6: Isotherms of CMP-1 synthesised in toluene solvent using 1,3,5-triethynylbenzene and a) 1,4-diiodobenzene, or b) 1,4-dibromobenzene. Figure a) reproduced (adapted) from reference 13 with permission from John Wiley and Sons. Copyright 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. Figure b) reprinted (adapted) with permission from R. Dawson, A. Laybourn, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Macromolecules*, 2010, **43**, 8524–8530. Copyright 2010 American Chemical Society.

The peak intensity of oligomers visible in the solution state NMR decreased over time (Figure 3.7a).²³⁷ This implied that the oligomers had grown to a point where they became insoluble and not observable using solution state NMR as the reaction proceeded.²³⁷ On further increasing the reaction time, a gel was formed, which then led to a brown precipitate. This was accompanied by a decrease in the percentage of bromine observed by energy-dispersive x-ray spectroscopy (EDX) (60 minutes = 4.8 wt.%, 1080 minutes = 2.9 wt.%), which suggested an increasing degree of polycondensation over time.²³⁷ It was noted that the relatively low percentage of bromine after 60 minutes suggested that a large quantity of alkyne-bromine cross-coupling had already occurred by this point.²³⁷ The ratio of polymerised alkyne ($\sim 2200\text{ cm}^{-1}$, 91.5 ppm) to unreacted alkyne ($\sim 2100\text{ cm}^{-1}$, 82.4 ppm) in the IR and NMR spectra, respectively (Figure 3.7b and c), increased as the reaction proceeded, suggesting a decrease in the number of unreacted alkyne end groups.²³⁷

As the alkyne monomer was in 1.5 molar excess compared to the dibromo monomer, the small quantity of remaining unreacted alkyne end groups indicated that additional reactions were occurring as well as the expected Sonogashira-Hagihara cross-coupling.^{237,251} The presence of a shoulder in the NMR spectrum at $\sim 137\text{ ppm}$, which was ascribed to alkyne-alkyne homocoupling in homocoupled CMP materials synthesised using very similar conditions to Sonogashira-Hagihara,²⁵⁴ suggested that this may be occurring at later stages in the reaction once all of the dibromo monomer had fully reacted and the percentage of end groups had decreased.²³⁷ This alkyne-alkyne homocoupling was proposed as enhancing the microporosity in the material by holding the structure open *via* additional hypercrosslinking.²³⁷

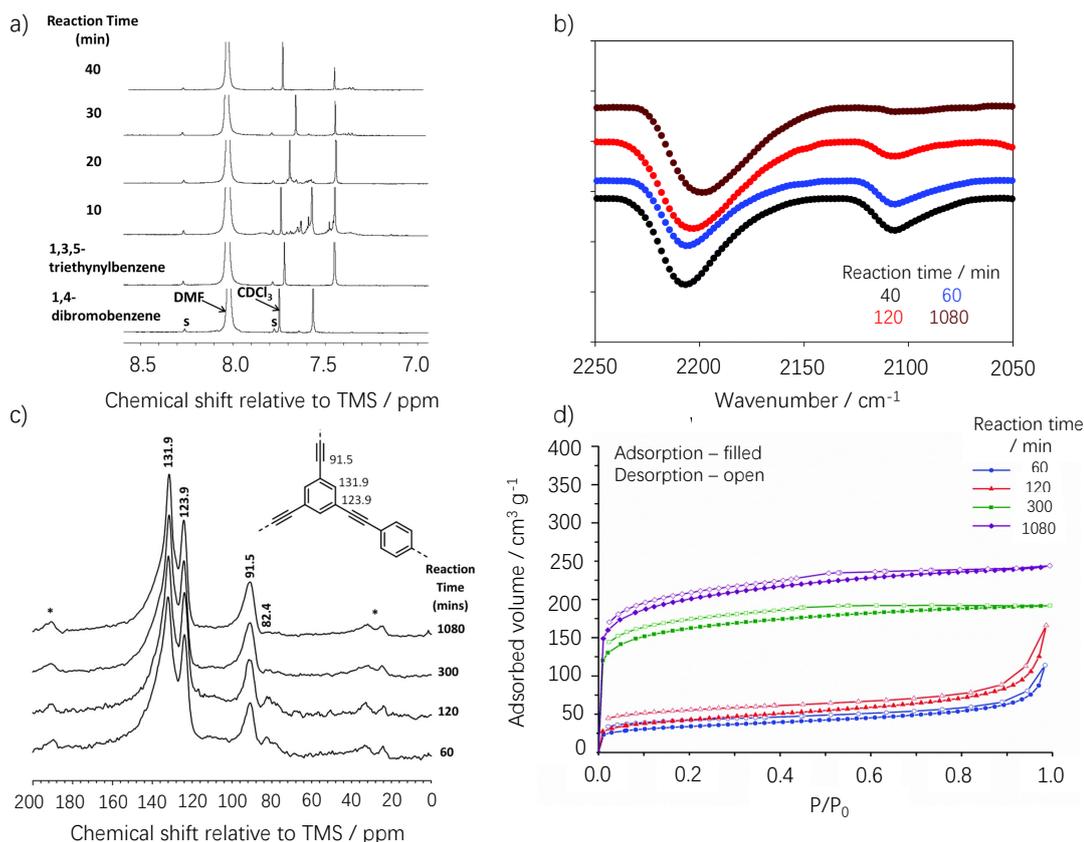


Figure 3.7: Experimental CMP-1 data reported by Laybourn and co-workers in reference 237 as a function of reaction time. a) ^1H solution-state NMR of soluble monomers and intermediates. b) Fourier-transform infrared spectra of CMP-1. Key: 40 minutes – black, 60 minutes – blue, 120 minutes – red, 1080 minutes – dark red. c) $^{13}\text{C}\{^1\text{H}\}$ HPDEC MAS ssNMR spectra of CMP-1, recorded at an MAS rate of 10 kHz, with the structure and identification of each peak given in the inset. d) Nitrogen adsorption-desorption isotherms recorded for CMP-1 at 77 K. Key: adsorption – filled symbols, desorption – open symbols, 60 minutes – blue circles, 120 minutes – red triangles, 300 minutes – green squares, 1080 minutes – purple offset squares. Reproduced and adapted from reference 237 with permission from the Royal Society of Chemistry.

The porosity differences in the materials analysed after different times can be observed in the nitrogen adsorption-desorption isotherms, given in Figure 3.7d. Whilst all appear to have some type I character, indicating the formation of a monolayer, with some hysteresis on desorption, the isotherms taken at earlier reaction times have a smaller uptake at low pressures ($\frac{P}{P_0} < 0.1$), indicating less microporosity within the material.²³⁷ The porosity in the materials at these early stages was rationalised as arising from adsorption into the interparticulate regions due to the increased uptake at high pressures ($\frac{P}{P_0} > 0.9$), with a $V_{0.1}/V_{\text{Tot}}$ ratio (indicating the quantity of microporosity within the material) of 0.22 for the polymer collected after 60 minutes.²³⁷ The polymers analysed after reaction times of less than 300 minutes gave observed pore sizes collected using non-local density functional theory (NL-DFT) of 30–45

nm, which fall into the mesopore range of 2–50 nm defined by the International Union of Pure and Applied Chemistry (IUPAC).^{2,237}

After longer reaction times, a much higher quantity of nitrogen was adsorbed at low pressures, along with a plateau in the uptake at high pressures, which indicated that these polymers had an increased degree of microporosity, with a $V_{0.1}/V_{\text{Tot}}$ ratio of 0.77 for the polymer analysed after 300 minutes and observed pore sizes in the region of 2 nm, which are within the IUPAC micropore range of 0–2 nm.^{2,237} These findings were confirmed by scanning electron microscopy (SEM), with larger, fused polymer particles with rough edges observed after longer reaction times, which correlated with the increased porosity as the reaction proceeded.²³⁷

On analysing all of the observed findings, Laybourn and co-workers established a network formation mechanism for CMP-1, where at early stages of the reaction, both monomers were soluble within the solvent and able to interact to form small, soluble oligomers.²³⁷ Further reaction of these oligomers resulted in their increasing size to such an extent that they were no longer soluble in the solvent. At this point, the oligomers precipitated out of solution as insoluble clusters, forming the observed gel-like phase.²³⁷ Once the 1,4-dibromobenzene was consumed, the remaining unreacted alkyne end groups reacted further *via* homocoupling between the alkyne units within the clusters to enhance the microporosity of the structure, resulting in the clusters combining to form the insoluble CMP-1 product (Figure 3.8).²³⁷

The CMP-1 network formation mechanism report²³⁷ was followed by a computational study of CMP-1 synthesised with DMF solvent by Thomas and co-workers.¹⁹ This work was termed as an ‘artificial synthesis’, as it followed each step in the synthesis of CMP-1,²³⁷ taking into account the experimental quantities of each reagent, and mimicked each step in the Sonogashira-Hagihara catalytic cycle,²⁵¹ using the Ambuild code.¹⁹ Each monomer and catalyst building block had specified end group and cap atoms to define the bonding, as shown in Figure 3.9, and in the case of the 1,3,5-triethynylbenzene, it was assumed that the copper side cycle within the Sonogashira-Hagihara mechanism had already occurred, leaving all of the alkyne units fully terminated with copper atoms.¹⁹ Bonding rules were then defined to allow unreacted 1,4-dibromobenzene (DBB) and copper-terminated 1,3,5-triethynylbenzene (TEB) monomers to bond to the catalyst, and to allow monomers that had already bonded to the catalyst to form bonds to one another and break their respective bonds to the catalyst.¹⁹

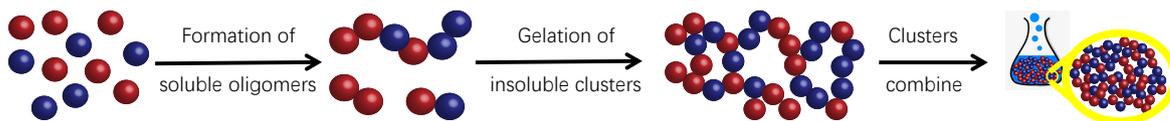


Figure 3.8: Experimental CMP-1 network formation mechanism reported by Laybourn and co-workers in reference 237. The two monomers used in the synthesis are given in red and blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

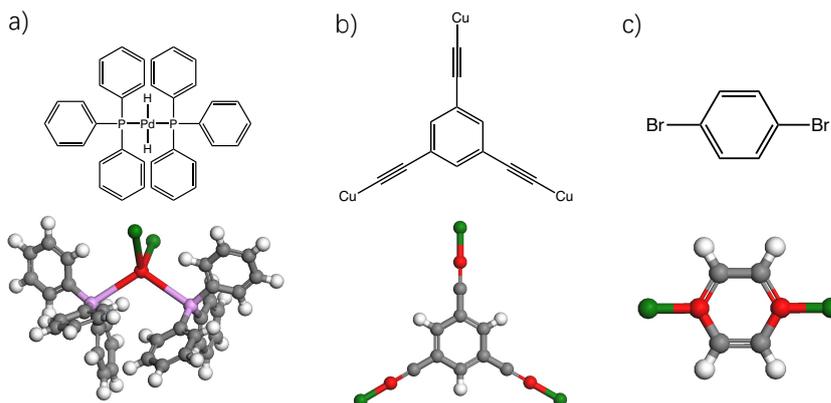


Figure 3.9: Building blocks used in the artificial synthesis of CMP-1. a) Catalyst, b) copper-terminated 1,3,5-triethynylbenzene (TEB), c) 1,4-dibromobenzene (DBB). End group atoms are shown in red and cap atoms are shown in green. The hydrogen cap atoms added to the palladium end group in the catalyst are simply there as placeholders to define the bonding and do not have any associated intermolecular forces. The idealised structures are also included on the top row for clarity. Figure reprinted (adapted) with permission from J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon and A. Trewin, *J. Phys. Chem. B*, 2020, **124**, 7318–7326, which contributes to the work in this chapter. Copyright 2020 American Chemical Society.

In each case, 100 DBB, 100 TEB and 4 catalyst building blocks were seeded into the minimum cell size that could incorporate all of the monomers, catalyst, DMF solvent and triethylamine (TEA).¹⁹ The quantities of DMF and TEA were varied to mimic the different regions of solvation observed within the heterogeneous, phase separated CMP-1 network, where four levels of solvation were modelled, containing 3, 13, 130 and 1300 DMF, and 1, 7, 70, and 700 TEA building blocks, referred to as solvation schemes 3, 13, 130 and 1300 from now onwards.¹⁹ Following the seeding process, a network generation stage occurred, allowing the system to form new bonds to build the polymer network *via zipBlocks* steps, geometry optimisation, and NVT (constant number of molecules, cell volume and temperature) molecular dynamics (MD), using HOOMD-blue as the geometry optimisation and MD engine and the polymer consistent forcefield (PCFF) throughout.^{19,115,193,194} Please refer to Chapter 2 section 2.4.1 for a full explanation of the parameters and terms used within the Ambuild code.

As part of the network generation, unreacted DBB and TEB building blocks were removed from the cell, and randomly re-seeded following an MD step, in order to increase the statistical sampling within the material.¹⁹ The network generation proceeded until the frameworks had completely reacted, judged by no new bonds forming during the last twenty *zipBlocks* steps. Four repeat models were generated for each level of solvation to generate more samples of the amorphous framework for better comparison to the experimental results reported by Laybourn and co-workers.^{19,237}

Following the network generation, a variety of desolvation strategies were modelled to understand the extent of desolvation that occurs in the real-world material, where some desolvation strategies involved the removal of solvent only (strategy 5), others involved the removal of solvent and unreacted monomers and catalyst blocks (strategies 4 and 6), and the remaining strategies 1–3 removed solvent, unreacted monomers and catalyst, and oligomers of up to three building blocks.¹⁹ Geometry optimisation and NVT MD occurred throughout to allow the cell contents to rearrange during the desolvation process. Strategy 6 also incorporated alkyne-alkyne homocoupling after the removal of building blocks, which was found experimentally to be a viable mechanism occurring alongside the Sonogashira–Hagihara and increased the ratio of microporosity within the material.^{19,237} After desolvation, a workup step occurred to simulate the reaction of the CMP-1 material with methanol, which replaced any accessible, unreacted copper end groups with hydrogen atoms.¹⁹ This was simulated using *Ambuild growBlocks* steps, which bond the end group of a hydrogen molecule to the unreacted carbon end group in TEB, losing the copper and hydrogen cap atoms of the TEB and hydrogen building blocks, respectively.¹⁹ Finally, a cell equilibration step occurred to allow the cell dimensions to reduce after the desolvation and workup processes, which removed any unstable pore voids within the simulation models that would not exist in the real-world system.¹⁹

It was observed that when using the experimental (stoichiometric) quantity of solvent molecules²³⁷ in the artificial synthesis (solvation scheme 1300), there was no evidence of polymer formation, even on running the network generation for an extended number of steps.¹⁹ However, all of the other degrees of solvation studied (solvation schemes 3, 13 and 130) followed the same general trend, as shown in Figure 3.10, of an initial plateau in the mass of the largest block in the system once the network generation began, followed by a steady increase at approximately step 20 as the system reacted to form small oligomers.¹⁹ This increase continued until approximately step 40, at which point the oligomers had grown into larger clusters which began to react with one another, forming a gel.¹⁹ From step 120, there was a sharp increase in mass corresponding to the clusters reacting to form a fully reacted polymer

system that crossed the periodic boundary.¹⁹ This general trend matched the experimental network formation mechanism proposed by Laybourn and co-workers.^{19,237}

It was also established that the average length of the gel phase (70 steps for solvation scheme 3, 128 steps for solvation scheme 130) increased with the amount of solvent. This was proposed as being due to the increased quantity of solvent hindering the diffusion of reactants throughout the simulation cell and slowing the rate of reaction, which may explain why no polymer was formed at the experimental degree of solvation.¹⁹

The desolvated models were analysed for each level of solvation to establish which was the most appropriate desolvation strategy to use in each case by comparing the maximum pore diameter to the diameter of the fragments that would be removed.¹⁹ For the CMP-1 system made using solvation scheme 3, the system was very highly condensed and dense, leading to almost no reduction in the simulation cell size and small pores.¹⁹ This system also exhibited no microporosity, meaning that the only building blocks that would be able to diffuse through the pore structure and be removed from the cell upon desolvation would be the DMF and TEA solvent molecules (desolvation strategy 5). This was also the case for solvation scheme 13, and so for these models, desolvation strategies 1–4, and 6, could be ignored.¹⁹ Solvation schemes 3 and 13 are therefore a good match to the CMP-1 samples analysed by Laybourn and co-workers after short reaction times, which have very small amounts of microporosity.^{19,237}

Solvation scheme 130, with the larger quantity of solvent present, formed a more open framework structure that was less dense and more microporous, with larger pores allowing both the solvent and unreacted catalyst and monomer building blocks to diffuse through the pore structure of the material and be removed upon desolvation.¹⁹ This meant that desolvation strategies 4 and 6, which allow removal of solvent and unreacted building blocks, with a maximum reduction in the cell volume of 75% and alkyne-alkyne homocoupling incorporated into desolvation strategy 6, were most appropriate for solvation scheme 130.¹⁹ These models are more comparable to the CMP-1 networks analysed after longer reaction times by Laybourn and co-workers, which contained a much higher quantity of microporosity arising from the larger, fused particles seen after longer reaction times.^{19,237}

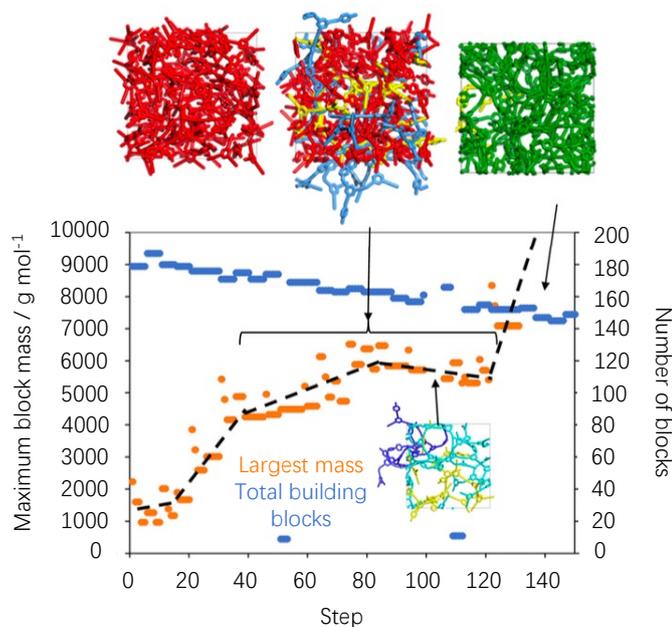


Figure 3.10: Plot showing the change in mass of the largest block in the system (orange, with a black dashed trendline to guide the eye) and the total number of building blocks (blue) for Model 1 of solvation scheme 3 as a function of the reaction coordinate. The insets show the structure of the polymer at various steps in the reaction, with monomers shown in red, small oligomers in yellow, larger oligomers in blue and the polymerised framework in green. Solvent and catalyst building blocks are omitted for clarity and are not included in the total number of building blocks. Figure reprinted (adapted) with permission from J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon and A. Trewin, *J. Phys. Chem. B*, 2020, **124**, 7318–7326, which contributes to the work in this chapter. Copyright 2020 American Chemical Society.

However, there was no one clear model that was a perfect match to the experimental, even when considering the fact that CMP-1 is an amorphous system and therefore challenging to model.¹⁹ This was rationalised by considering that these networks are heterogeneous and therefore are likely not to be equally solvated throughout.¹⁹ Instead, it is possible that the monomers and catalyst would phase separate out of the solution phase and react to form a polymer particle with a dense central core region leading to very small pores and limited porosity, as in solvation scheme 3. This core would then be surrounded by less dense, more porous regions moving away from the central core towards the solution phase, for example, regions composed from solvation schemes 13, 130 and 1300.¹⁹

A spherical model was therefore developed to test this, made from a central core of CMP-1 artificially synthesised using solvation scheme 3 and desolvated using strategy 5 (removal of solvent only), surrounded by a middle layer of CMP-1 made using solvation scheme 13 and desolvated using strategy 5.¹⁹ The outer shell of this sphere was composed of CMP-1 artificially synthesised using solvation scheme 130 and desolvated using strategies 4

and 6 (removal of solvent and unreacted building blocks, with alkyne-alkyne homocoupling also incorporated into strategy 6).¹⁹ The percentages that each layer of the spherical particle contributed to the overall were determined by fitting to the experimental density, micropore volume and BET surface area, and elemental analysis.¹⁹ The contributions of each layer to the overall sphere differed depending on the experimental technique that the overall sphere was compared to, and it was judged that fitting to the experimental BET surface area gave the best overall fit to the other characterisation techniques (the micropore volume, density, and pore size distribution (PSD)).¹⁹ In this model (Figure 3.11), the percentages with respect to particle volume of the central core, made from solvation scheme 3 and desolvation strategy 5, middle layer, made from solvation scheme 13 and desolvation strategy 5, and outer shell, made from solvation scheme 130 and desolvation strategies 4 and 6, were 40%, 36% and 24% (14% desolvation strategy 4 and 10% desolvation strategy 6), respectively.¹⁹ Upon desolvation, it would be expected that these spherical particles would be able to fuse together *via* either alkyne-alkyne homocoupling or alkyne-bromine cross-coupling.¹⁹ The simulated PSD is given in Figure 3.12 and the comparison of the properties of the simulated spherical model to experiment are given in Table 3.2.

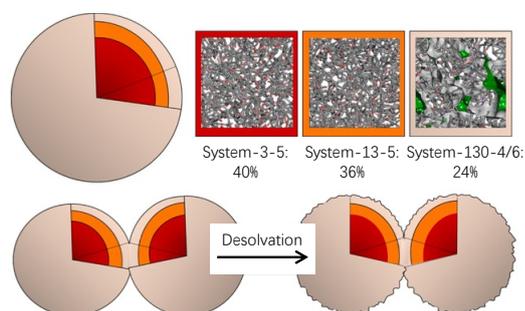


Figure 3.11: Sphere model generated to rationalise the experimental porosity of CMP-1 reported by Laybourn and co-workers.²³⁷ Examples of each structure are shown with the respective surface areas mapped on in green. Key: central core – red, middle layer – orange, outer shell – beige, system-3-5 – CMP-1 synthesised using solvation scheme 3 and desolvated using strategy 5, system-13-5 – CMP-1 synthesised using solvation scheme 13 and desolvated using strategy 5, system-130-4/6 – CMP-1 synthesised using solvation scheme 130 and desolvated using strategies 4 and 6. Figure reprinted (adapted) with permission from J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon and A. Trewin, *J. Phys. Chem. B*, 2020, **124**, 7318–7326, which contributes to the work in this chapter. Copyright 2020 American Chemical Society.

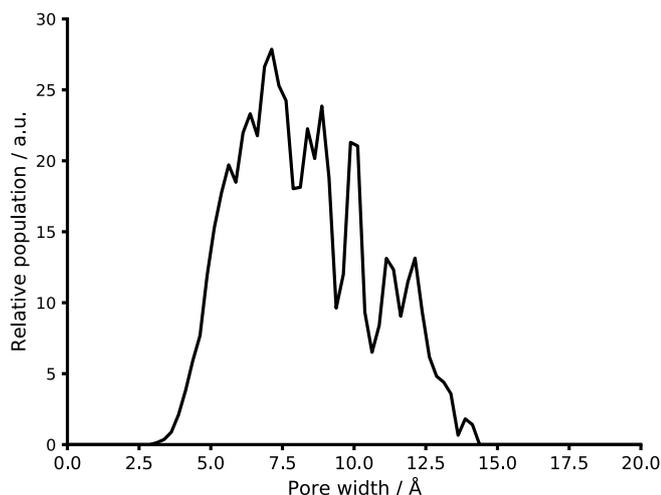


Figure 3.12: Simulated pore size distribution obtained on fitting to the experimental BET surface area, which is taken from reference 237. Figure reprinted (adapted) with permission from J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon and A. Trewin, *J. Phys. Chem. B*, 2020, **124**, 7318–7326, which contributes to the work in this chapter. Copyright 2020 American Chemical Society.

Table 3.2: Comparison of the experimental properties of CMP-1 and the properties of the simulated spherical model given in Figure 3.11. The experimental data is taken from references 234, 235, 237. Table reprinted (adapted) with permission from J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon and A. Trewin, *J. Phys. Chem. B*, 2020, **124**, 7318–7326, which contributes to the work in this chapter. Copyright 2020 American Chemical Society.

Property	Experiment	Simulation
Accessible surface area / m ² g ⁻¹	564	557
Micropore volume / cm ³ g ⁻¹	0.23	0.31
Bulk density / g cm ⁻³	1.13	0.99
wt.% C	80.90	67.18
wt.% H	3.48	2.67
wt.% Remaining	15.58	30.18

It was noted that whilst four levels of solvation were modelled, the real system is likely composed of an almost infinite range of levels of solvation due to the heterogeneous nature of the amorphous CMP-1 material.¹⁹ The simulated PSD on fitting to the experimental BET surface area is given in Figure 3.12 and a comparison of the experimental material and sphere model is given in Table 3.2.^{19,237}

The choice of reaction solvent has also been found to affect the morphology of CMP materials.²⁶⁰ Tan and co-workers reported the influence of solvent choice on the morphology of two-dimensional CMPs polymerised from 1,3,5-triethynylbenzene and 1,3,5-tribromobenzene, with the polymers syn-

thesised using toluene, *p*-xylene and mesitylene classified as CMP-T, CMP-X and CMP-M, respectively.²⁶⁰ The polymers were synthesised using similar conditions to CMP-1, still relying on a Sonogashira–Hagihara mechanism, however, in this instance the 1,4-dibromobenzene linker was replaced with 1,3,5-tribromobenzene, giving the two-dimensional sheet-like structure with some short-range order. Another difference between the two synthetic protocols was that Tan and co-workers added the catalysts to the reaction before heating.²⁶⁰ The structures were analysed using SEM and transmission electron microscopy (TEM) (Figure 3.13), and it was seen that the CMP-T network was made up of sub-micron spheres of uniform size, with films around the outer edges.²⁶⁰ In contrast, CMP-X is made from hollow rod-like structures, and CMP-M from rods with planar thin films around the edges.²⁶⁰ It should be emphasised that in each of the polymers studied by Tan and co-workers, all reaction conditions were identical apart from the solvent choice, indicating that this is the sole factor influencing the morphology in this work.²⁶⁰

A similar study by Dawson and co-workers found that the reaction solvent also influences the porosity of CMPs.²³⁵ In this work, a series of CMPs were synthesised, all using a 1,3,5-triethynylbenzene node and an aromatic dibromo linker (Scheme 3.6).²³⁵ As in the previous reports of CMPs synthesised by the Cooper group, a 1.5 molar excess of alkyne was used to make the polymers *via* a Sonogashira–Hagihara mechanism.^{13,234–236} Each CMP was synthesised in four solvents of varying polarities: DMF, 1,4-dioxane, tetrahydrofuran (THF) and toluene²³⁵ (respective polarity indices of 6.4, 4.8, 4.0 and 2.4, where the polarity index is a standard measure of how polar each solvent is. A higher polarity index indicates a more polar solvent, *e.g.*, water has a polarity index of 10.2).¹⁷²

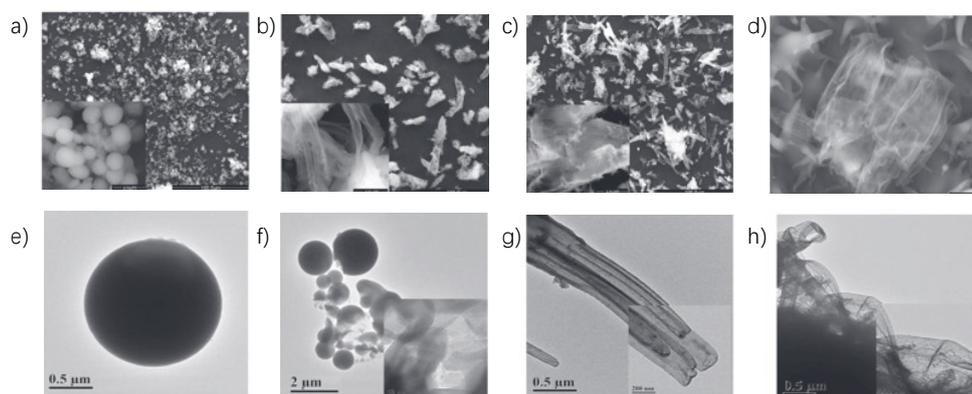
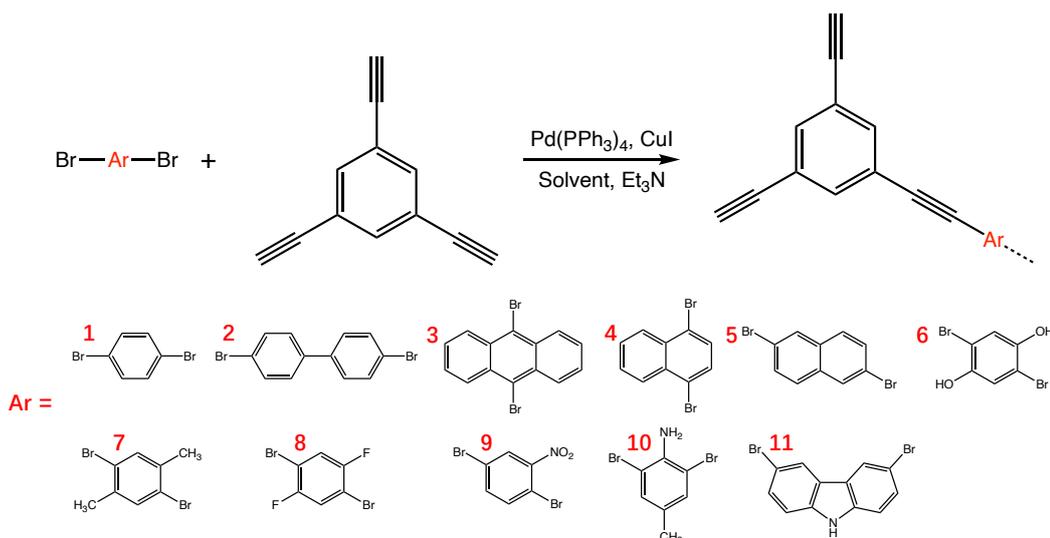


Figure 3.13: Morphology of the CMPs synthesised by Tan and co-workers.²⁶⁰ a)–d) SEM, with a scale bar of 100 μm . a) CMP-T, b) CMP-X, c) CMP-M, d) thin film observed in CMP-X. e)–h) TEM, with the insets showing an expanded view. e) CMP-T, f) CMP-T, g) CMP-X, h) CMP-M. Figure reproduced (adapted) from reference 260 with permission from John Wiley and Sons. Copyright 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.



Scheme 3.6: Generic reaction scheme for the CMPs synthesised by Dawson and co-workers.²³⁵ The various aromatic co-monomers are shown underneath. Figure reprinted (adapted) with permission from R. Dawson, A. Laybourn, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Macromolecules*, 2010, **43**, 8524–8530. Copyright 2010 American Chemical Society.

The BET surface area of each CMP was then calculated (Table 3.3), and it was found that generally, the CMPs showed a maximum surface area and micropore volume ($V_{0.1}$, calculated at $\frac{P}{P_0} = 0.1$) when synthesised using DMF, and a minimum surface area and micropore volume when synthesised using toluene.²³⁵ CMPs synthesised using THF and 1,4-dioxane showed comparable properties in the majority of cases, with the surface areas generally being slightly higher using THF and the micropore volume being slightly higher using 1,4-dioxane.²³⁵ This suggested that the original solvent choice of toluene^{13,234,236} was not ideal for CMP synthesis, as this gave CMPs with lower surface areas and micropore volumes. It was proposed that the increase of surface area when using DMF was due to the increase in microporosity, which was supported by the nitrogen adsorption-desorption isotherms (Figure 3.14) showing a type I isotherm when the CMP was synthesised in DMF or 1,4-dioxane, which typically offered the higher micropore volumes.²³⁵ In comparison, a type IV isotherm was observed when using THF or toluene, which supported the higher prevalence of mesopores within the CMP framework when using these solvents.²³⁵

Table 3.3: Brunauer–Emmett–Teller surface areas (S_{BET}), micropore volumes calculated at $\frac{P}{P_0} = 0.1$ ($V_{0.1}$), total pore volumes calculated at $\frac{P}{P_0} = 0.99$ (V_{Tot}), and ratio of microporosity ($V_{0.1}/V_{\text{Tot}}$). The structure of each monomer is given in Scheme 3.6. Table reprinted (adapted) with permission from R. Dawson, A. Laybourn, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Macromolecules*, 2010, **43**, 8524–8530. Copyright 2010 American Chemical Society.

CMP network	Solvent	$S_{\text{BET}} / \text{m}^2 \text{g}^{-1}$	$V_{0.1} / \text{cm}^3 \text{g}^{-1}$	$V_{\text{Tot}} / \text{cm}^3 \text{g}^{-1}$	$V_{0.1}/V_{\text{Tot}}$
1	Toluene	867	0.33	0.99	0.33
	THF	941	0.36	1.24	0.29
	1,4-Dioxane	609	0.23	0.36	0.64
	DMF	837	0.32	0.45	0.71
2	Toluene	204	0.09	0.43	0.21
	THF	734	0.28	0.68	0.41
	1,4-Dioxane	744	0.29	0.55	0.53
	DMF	744	0.29	0.55	0.53
3	Toluene	136	Not reported	Not reported	Not reported
	THF	716	0.28	1.05	0.27
	1,4-Dioxane	573	0.22	0.60	0.37
	DMF	550	0.21	0.37	0.57
4	Toluene	344	0.09	0.17	0.53
	THF	522	0.20	0.52	0.38
	1,4-Dioxane	580	0.22	0.67	0.33
	DMF	742	0.28	0.70	0.40
5	Toluene	436	0.08	0.13	0.62
	THF	532	0.21	0.49	0.43
	1,4-Dioxane	638	0.25	0.83	0.30
	DMF	599	0.23	0.44	0.52
6	Toluene	761	0.27	1.73	0.16
	THF	847	0.33	0.69	0.48
	1,4-Dioxane	778	0.32	0.65	0.49
	DMF	1043	0.40	0.71	0.56
7	Toluene	682	0.25	1.06	0.24
	THF	985	0.38	1.09	0.35
	1,4-Dioxane	1022	0.40	0.88	0.45
	DMF	899	0.34	0.75	0.45
8	Toluene	690	0.27	0.52	0.52
	THF	639	0.25	0.41	0.61
	1,4-Dioxane	869	0.39	0.61	0.64
	DMF	1260	0.48	0.88	0.55
9	Toluene	247	0.07	0.18	0.39
	THF	335	0.13	0.17	0.76
	1,4-Dioxane	823	0.40	0.60	0.67
	DMF	967	0.37	0.94	0.39
10	Toluene	542	0.20	0.97	0.21
	THF	994	0.38	0.97	0.39
	1,4-Dioxane	611	0.23	0.47	0.49
	DMF	653	0.25	0.36	0.69
11	Toluene	779	0.29	1.40	0.21
	THF	1014	0.39	0.99	0.39
	1,4-Dioxane	727	0.28	0.74	0.38
	DMF	1056	0.40	0.77	0.52

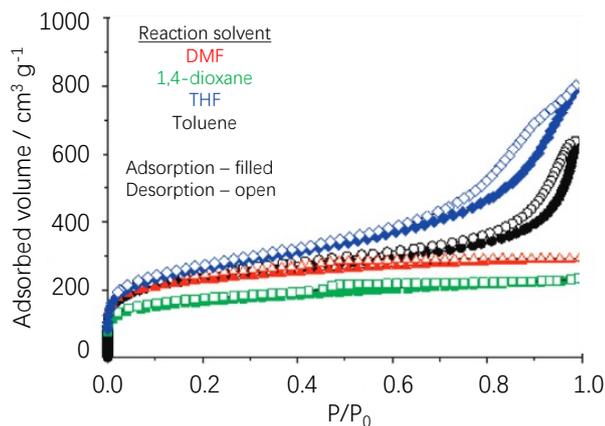


Figure 3.14: Experimental nitrogen adsorption-desorption isotherms for CMP-1 synthesised in varying solvents by Dawson and co-workers. Key: adsorption – filled symbols, desorption – open symbols, DMF – red triangles, 1,4-dioxane – green squares, THF – blue offset squares, toluene – black circles. Figure reprinted (adapted) with permission from R. Dawson, A. Laybourn, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Macromolecules*, 2010, **43**, 8524–8530. Copyright 2010 American Chemical Society.

The reaction solvent choice also influenced the degree of polycondensation within each material. The studied CMPs had a lower terminal to quaternary alkyne ratio when synthesised in DMF compared to toluene (CMP-1 terminal to quaternary alkyne ratio: DMF = 0.16, toluene = 0.40).²³⁵ This indicated a lower degree of polycondensation in toluene compared to DMF, which could explain why these CMPs were more microporous when synthesised using DMF, as the more densely polymerised network would be less able to collapse into a dense framework upon desolvation.²³⁵

Following this, Mollart and Trewin rationalised the differences in CMP porosity with respect to reaction solvent choice by further analysing the experimental porosity data reported for the CMPs originally synthesised by Dawson and co-workers.^{235,261} Firstly, the microporous surface areas were calculated for each material (Table 3.4) by multiplying the total BET surface area by the micropore volume, as the degree of microporosity was believed by Dawson and co-workers to strongly influence the differences in total surface area.^{235,261}

The polarity of each monomer was established by obtaining the molecular dipole calculated from Mulliken charges, as given in Table 3.5, using the B3LYP/6-31G model chemistry in Gaussian 09.^{79,85,90,91,94,152,212,262,263} From analysing the structures and polarities of each dibromo monomer relative to 1,3,5-triethynylbenzene, four groups were established. Each group was considered independently of the others to rationalise the observed trend in each case. The microporous surface areas were plotted against the solvent polarity indices for DMF, 1,4-dioxane, THF and toluene (6.4, 4.8, 4.0 and 2.4, respectively, Figure 3.15).^{172,261}

Table 3.4: Experimental microporous surface areas obtained from the total surface areas reported by Dawson and co-workers.²³⁵ The structure of each monomer is given in Scheme 3.6. Table reproduced and adapted from reference 261, which contributes to the work in this chapter, with permission from the PCCP Owner Societies.

CMP network	Microporous surface area of the CMP network / m ² g ⁻¹			
	Toluene	THF	1,4-Dioxane	DMF
1	286	273	390	594
2	43	301	394	394
3	Not reported	193	212	314
4	182	198	191	297
5	270	229	191	311
6	122	407	381	584
7	164	345	460	405
8	359	390	556	693
9	96	255	551	377
10	114	388	299	451
11	164	395	276	594

Solvent polarity →

Table 3.5: Molecular dipoles calculated from Mulliken charges and absolute differences with respect to 1,3,5-triethynylbenzene of dibromo monomers 1–11. The structure of each monomer is given in Scheme 3.6. Table reproduced and adapted from reference 261, which contributes to the work in this chapter, with permission from the PCCP Owner Societies.

Dibromo monomer	Molecular Dipole Moment / Debye	Absolute difference in dipole moment relative to 1,3,5-triethynylbenzene / Debye
1	0.0001	0.0001
2	0.0002	0.0000
3	0.0003	0.0001
4	0.7045	0.7043
5	0.0000	0.0002
6	0.0017	0.0015
7	0.0000	0.0002
8	0.0001	0.0001
9	4.1278	4.1276
10	0.3503	0.3501
11	3.1943	3.1941
1,3,5-Triethynylbenzene	0.0002	0.0000

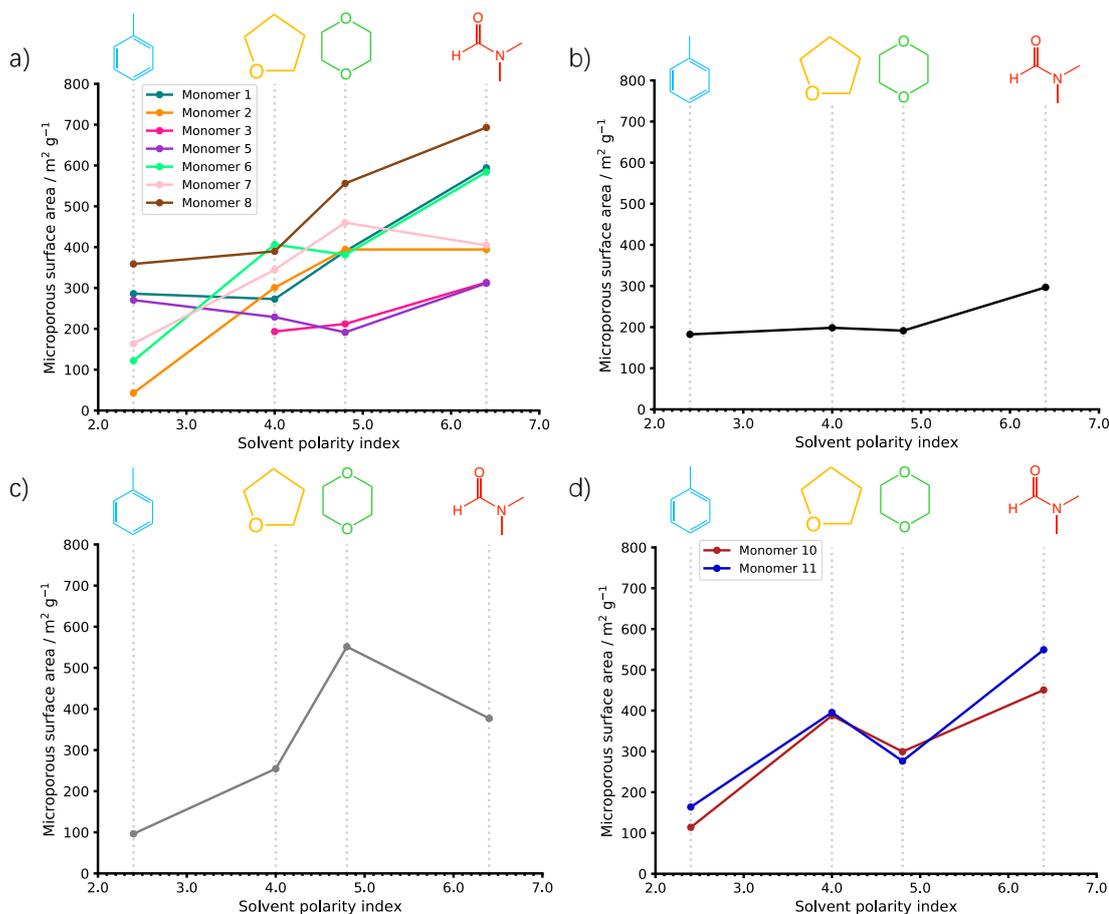


Figure 3.15: Plots of the microporous surface area of CMPs 1–11 as a function of solvent polarity index. The structures of each solvent are shown for clarity at the polarity indices indicated by the pale grey lines. a) Group 1, consisting of dibromo monomers 1, 2, 3, 5, 6, 7, 8, b) group 2, consisting of dibromo monomer 4, c) group 3, consisting of dibromo monomer 9, d) group 4, consisting of dibromo monomers 10 and 11. Key: CMP-1 – teal, CMP-2 – orange, CMP-3 – dark pink, CMP-4 – black, CMP-5 – purple, CMP-6 – green, CMP-7 – pale pink, CMP-8 – brown, CMP-9 – dark grey, CMP-10 – burgundy, CMP-11 – blue. The structure of each monomer is given in Scheme 3.6. Figure reproduced and adapted from reference 261, which contributes to the work in this chapter, with permission from the PCCP Owner Societies.

Group 1, given in Figure 3.15a) was composed of dibromo monomers (DBMs) with no nitrogen groups present (DBMs 1, 2, 3, 5, 6, 7, and 8), leading to a very similar polarity compared to 1,3,5-triethynylbenzene (absolute dipole moment difference < 0.1 Debye). Group 2 (Figure 3.15b) contained only one DBM, DBM-4, which was aromatic as for the Group 1 monomers and follows the same general trend but was classed as its own group due to being of medium polarity (absolute dipole moment difference = 0.7043 Debye). Groups 3 and 4, shown in Figure 3.15c and d, respectively) both contained high polarity monomers due to containing nitrogen groups (group 3 = nitro-functionalised DBM 9, group 4 = amine-functionalised DBMs 10 and 11).²⁶¹

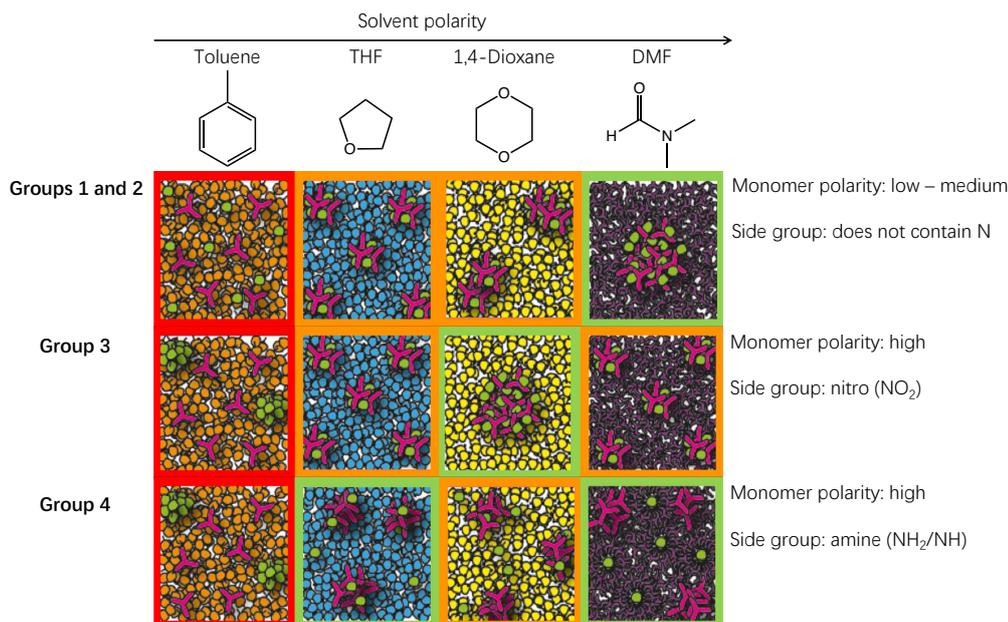


Figure 3.16: Summary of the phase separation behaviour of each group as a function of solvent polarity. The maximum surface area of each group is indicated by the green box, the minimum surface area of each group is indicated by the red box, with the intermediate surface areas given by the orange boxes. Key: DBM – green circles, TEB – pink, toluene – orange circles, THF – blue circles, 1,4-dioxane – yellow circles, DMF – purple circles. In the case of group 4, a maximum surface area in THF is also indicated as this would arise if the monomer was not able to interact *via* hydrogen bonding with the DMF solvent. Figure reproduced and adapted from reference 261, which contributes to the work in this chapter, with permission from the PCCP Owner Societies.

From analysing the plots of microporous surface area against solvent polarity index (Figure 3.16), it was established that the maximum surface areas for each polymer occurred due to full co-phase separation of the reactants from the solvent, unless, in the case of group 4, hydrogen bonding is able to occur between the amine group of the monomer and the DMF solvent.²⁶¹ In contrast, the minimum surface areas arose due to the monomers being separated, either by both monomers being fully miscible within the solvent, as for groups 1 and 2, or by the DBM being phase separated, whilst the TEB is miscible within the solvent, as for groups 3 and 4.²⁶¹

It was also found by Chen and co-workers that the BET surface area and porosity of the polymer can be modified by addition of inorganic salts to the reaction mixture during the synthesis, in a method termed ‘Bristol Xi’an Jiao-tong’ (BXJ).²⁴¹ This was first applied to improve the porosity of CMPs synthesised using the Buchwald–Hartwig catalytic mechanism, with a hypothesis that the lack of inorganic salt incorporated within this mechanism gives the limited surface area and degree of microporosity compared to alternative mechanisms such as Sonogashira–Hagihara, Friedel–Crafts alkylation and Yamamoto.^{241,251,264–266} Different salts were utilised to assess which was optimal

for use in the synthesis, and it was established that salts with smaller ionic radii were better able to tune the porosity due to the enhanced electronegativity.²⁴¹ However, it was later found that this approach can also be used to improve the BET surface area and degree of microporosity of CMPs synthesised using alternative mechanisms, such as CMP-1.¹⁸⁰ It was proposed that the addition of salt to the reaction mixture increases the porosity of the material by tuning the Hansen solubility parameters (HSPs) of the solvent to give a better match to those of the polymer, as shown in Figure 3.17a.^{180,241} This results in phase separation of the polymer from solution at a later stage in the reaction, giving a more microporous polymer with an enhanced BET surface area (CMP-1 surface area and pore volume before salt addition = $886 \text{ m}^2 \text{ g}^{-1}$ and $0.58 \text{ cm}^3 \text{ g}^{-1}$, after salt addition = $1148 \text{ m}^2 \text{ g}^{-1}$ and $0.80 \text{ cm}^3 \text{ g}^{-1}$, respectively).¹⁸⁰ Additionally, a higher uptake of nitrogen is observed in the nitrogen adsorption-desorption isotherm and the PSD of the resulting CMP-1 material is localised within the micropore region, given in Figure 3.17b and c, respectively.¹⁸⁰

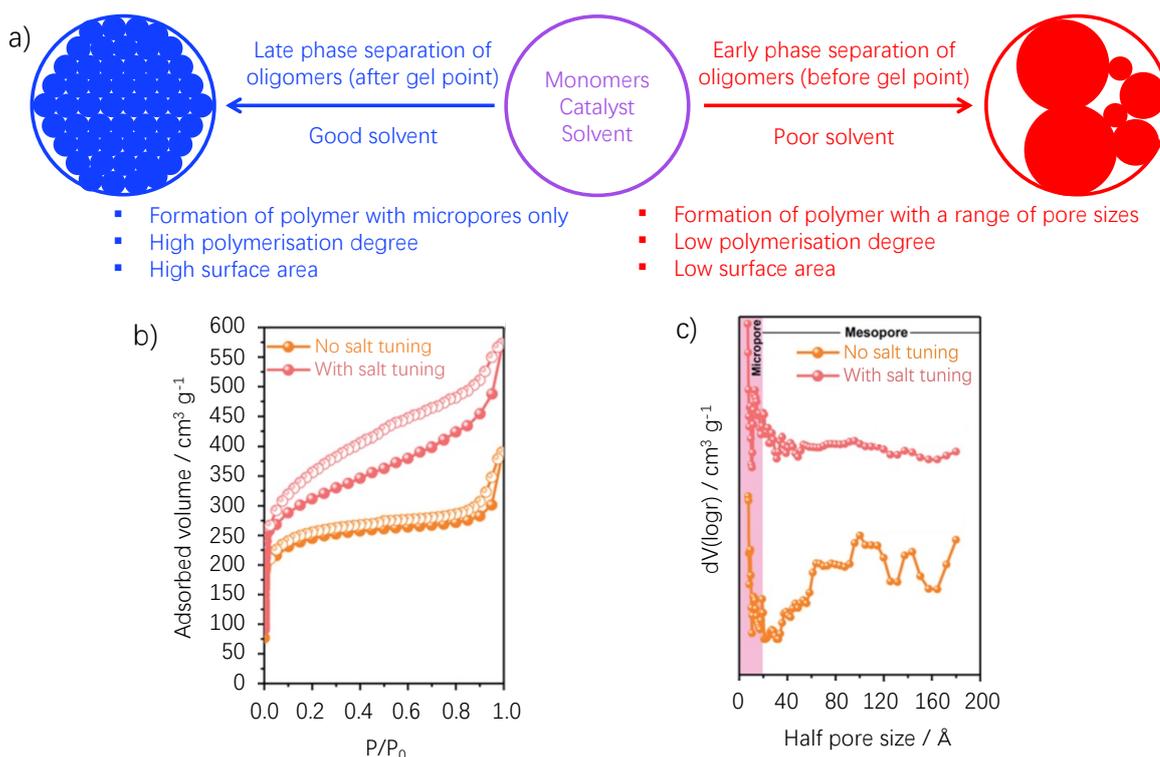


Figure 3.17: Influence of utilising the BXJ approach to tune CMP porosity. a) Cartoon illustration indicating the influence of salt tuning on the phase separation of the resulting CMPs, b) nitrogen adsorption-desorption isotherm and c) NL-DFT pore size distribution for CMP-1 before and after salt tuning with 0.5 mmol NaF. Key: adsorption – filled symbols, desorption – open symbols, before salt tuning – orange, after salt tuning – pink. The pink rectangle in c) indicates the micropore region of 0–20 \AA . Figure a) adapted with permission from reference 241, licensed under CC-BY. Copyright 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. Figure b) reproduced and adapted from reference 180 with permission from the Royal Society of Chemistry.

However, despite the reports suggesting that the porosity of CMP materials differs when synthesised in alternative reaction solvents due to a change in the phase separation of the material as a result of the solvent choice,^{180,235,241} it was still not fully understood why this is the case. This led to the work reported in this chapter, where the artificial synthesis strategy first reported by Thomas and co-workers¹⁹ was adapted to model CMP-1 in the same four solvents studied experimentally by Dawson and co-workers.²³⁵ There is a small number of existing experimental studies reporting the porosity of a CMP material in multiple solvents,^{180,235,241} making it necessary to choose from this library which material would be best to simulate. CMP-1 was chosen again as an ideal test case system to model due to the existing ability of Ambuild to model the Sonogashira-Hagihara catalytic mechanism commonly applied in CMP synthesis, allowing comparison with the previous artificial synthesis study and numerous experimental reports; and the relatively large number of solvents assessed experimentally.^{13,19,38,234,235,237} Enhancing the porosity of a CMP framework by judicious choice of reaction conditions is a key criterion for maximising the potential applications of the material, as larger surface areas correspond to increased power densities within battery electrodes^{267,268} and a maximised pore volume,²²⁴ leading to more facile diffusion of charge carriers throughout a battery and/or an increased gas uptake.³⁸

The aim of this chapter was to understand the influence of reaction solvent choice on the porosity of CMP materials. This was done by using the Ambuild code¹⁹ to generate representative structural models of CMP-1, prepared using the same solvents used experimentally. The full Sonogashira-Hagihara catalytic mechanism was modelled, and the quantities of solvent and initial starting configurations of building blocks were varied to assess the phase separation within the models. The simulated porosity of each model was then calculated using Poreblazer,^{162,163} allowing for the comparison of the properties of the simulated materials to experiment.

3.3 Methodology

The previous successful modelling of CMP-1 using the Ambuild code,¹⁹ which mimicked the experimental synthesis by Laybourn and co-workers,²³⁷ used DMF as the solvent as it was established empirically by Dawson and co-workers²³⁵ that this often gave rise to the resulting CMP with the highest surface area. However, when looking at the total surface area of CMP-1, it is relatively insensitive to the solvent choice used in the synthesis (DMF - $837 \text{ m}^2 \text{ g}^{-1}$, 1,4-dioxane - $609 \text{ m}^2 \text{ g}^{-1}$, THF - $941 \text{ m}^2 \text{ g}^{-1}$, toluene - $867 \text{ m}^2 \text{ g}^{-1}$).²³⁵ In contrast, the percentages of micro- and mesoporosity that contribute to the total surface area do vary significantly depending on the reaction solvent choice, as

seen in Figure 3.18, and solvent polarity, given in Table 3.6. The microporous and mesoporous surface areas were calculated by multiplying the total surface area by the degree of microporosity within the system, $V_{0.1/Tot}$, to acquire the microporous surface area, and then subtracting this from the total to obtain the mesoporous surface area. The differences in the ratio of micro- to mesoporosity when synthesised in each solvent meant that it was important to understand why this arises through computational modelling in order to help with the rational design of new CMP materials in the future.³⁸

The network generation procedure used to simulate CMP-1 in each solvent was the same as reported by Thomas and co-workers,¹⁹ where once seeded, the cell contents were able to diffuse throughout the cell *via* HOOMD-blue geometry optimisation (integration timestep = 0.0001 ps, number of optimisation cycles per output step = 1 million, van der Waals cut-off = 10 Å) and NVT MD (integration timestep and van der Waals cut-off as above, number of MD cycles per output step = 1 million, temperature factor = 55.0).^{193,194}

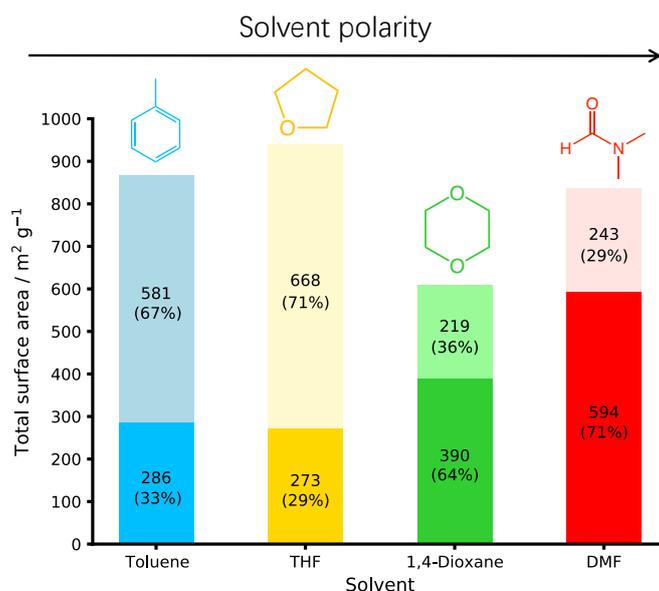


Figure 3.18: Plot of the experimental surface area of CMP-1 synthesised in DMF - red, 1,4-dioxane - green, THF - yellow and toluene - blue. The microporous contribution (percentage of the total surface area given in brackets) to the total surface area is indicated by the darker shades and the mesoporous contribution by the paler shades. Data obtained from reference 235. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Table 3.6: Comparison of the polarity indices,¹⁷² Hildebrand solubility parameter, and its Hansen solubility components¹⁷⁹ for each solvent. Table reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Solvent	Toluene	THF	1,4-Dioxane	DMF
Polarity index	2.4	4.0	4.8	6.4
Hildebrand solubility parameter / MPa ^{0.5}	18.2	19.4	20.5	24.8
Dispersion component / MPa ^{0.5}	18.0	16.8	19.0	17.4
Polar component / MPa ^{0.5}	1.4	5.7	1.8	13.7
Hydrogen bonding component / MPa ^{0.5}	2.0	8.0	7.4	11.3

The input file for the network generation procedure is given in Appendix A Figure A.1, and the PCFF forcefield was used throughout.¹¹⁵ Unreacted monomer end groups could bond to the catalyst *via* a *zipBlocks* step if within the predetermined bond length and bond angle margins of 10 Å and 90°, respectively.¹⁹ These margins are so large due to the presence of solvent throughout the simulation cell making it more challenging for the monomer and catalyst building blocks to diffuse and react. After the network generation was complete, judged by no new bonds forming during the last twenty *zipBlocks* steps, the systems were desolvated, and the porosity properties were calculated using Poreblazer 4.0.^{38,162} Please refer to Chapter 2 section 2.4.1 for a full explanation of the parameters and terms used within the Am-build code.

Three categories of models were studied within this work. Firstly, the approach of Thomas and co-workers,¹⁹ where the number of solvent and TEA molecules was varied (modelling 1%, 33%, 67% and 100% of the experimental stoichiometry, Table 3.7). The quantities of catalyst and monomers were kept constant throughout, with the simulation cell contents seeded at random into the smallest possible cell size, referred to as the ‘Degree of Solvation Systems’.³⁸

Secondly, the quantities of all building blocks were kept constant throughout, using 100% of the experimental stoichiometry of solvent and TEA. The starting configuration of the monomers and catalyst relative to the solvent and TEA were changed to assess the influence of phase separation on the system, referred to as the ‘Phase Separated Systems’.³⁸ The Phase Separated Systems were composed of: the fully mixed large cell configuration; the four clusters configuration; the two clusters configuration, and the one small cluster configuration. In the fully mixed large cell configuration, all of the monomer and

catalyst building blocks were randomly seeded. In the four clusters configuration, the monomers and catalyst were split into four clusters, each containing 25 DBB, 25 TEB and 1 catalyst building blocks, equally spaced throughout the unit cell. The two clusters configuration was made up of two clusters, each containing 50 DBB, 50 TEB and 2 catalyst building blocks. The one small cluster seeded all of the monomers and catalyst building blocks into the centre of the simulation cell. In each case, the solvent and TEA molecules were then seeded into the remaining cell volume.³⁸

Finally, the ‘Solvent/Monomer Phase Interface Systems’ were studied, with the monomers and catalyst molecules seeded into the centre of the simulation cell and surrounded by solvent and TEA molecules, referred to as ‘one large clusters’.³⁸ The quantities of monomers and catalyst were scaled up by a factor of 15 times the experimental stoichiometry in order to better assess the interface between the monomers and solvent. It was found that the catalyst molecules were favourable in both phases and so could diffuse freely throughout the system (Figure 3.19).³⁸

Table 3.7: Number of building blocks seeded into the Ambuild models. The Phase Separated Systems have a respective quantity of solvent of 100%. Table reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Building block	Quantity of solvent with respect to experimental stoichiometry				
	1%	33%	67%	100%	100% (15x scaled monomers and catalyst)
Catalyst	4	4	4	4	60
DBB	100	100	100	100	1500
TEB	100	100	100	100	1500
TEA	11	359	717	1076	1076
Toluene	23	782	1565	2347	2347
THF	31	1027	2055	3082	3082
Dioxane	29	978	1956	2934	2934
DMF	32	1076	2153	3229	3229

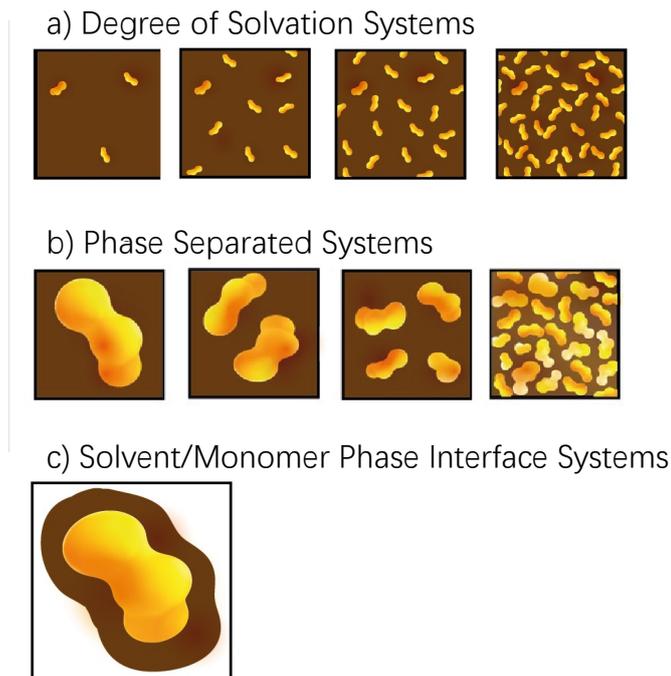


Figure 3.19: Cartoon representation of a) the Degree of Solvation Systems (left to right: 100% solvent and TEA quantity, 67%, 33%, 1%), b) Phase Separated Systems (left to right: one small cluster, two clusters, four clusters and fully mixed large cell configurations) and c) Solvent/Monomer Phase Interface Systems. Key: solvent and TEA – brown, monomers and catalyst – yellow. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

The Phase Separated Systems and Solvent/Monomer Phase Interface Systems were all seeded into larger unit cells of length 200 Å to better study the interface between phases. Before desolvation, a short NPT (constant number of molecules, cell volume and pressure) cycle (integration timestep = 0.00000005 ps, number of MD cycles = 50000 per output step, van der Waals cut-off = 10 Å, HOOMD-blue temperature factor = 1)^{193,194} was undertaken to optimise the cell dimensions whilst giving the system a small amount of energy. These systems were analysed both after the network generation and desolvation in order to compare the final polymer networks, and also after diffusion only (Appendix A Figure A.2). This was composed of a loop of geometry optimisation (integration timestep = 0.0001 ps, number of optimisation cycles per output step = 1 million, van der Waals cut-off = 10 Å) and NPT MD (integration timestep = 0.0001 ps, number of MD cycles per output step = 3 million, van der Waals cut-off = 10 Å, HOOMD-blue temperature factor = 55.0, pressure = 1 bar),^{193,194} in order to assess the potential energy of each system.³⁸ As in the previous artificial synthesis study,¹⁹ four repeat models of CMP-1 in each solvent for each category were simulated in order to increase the structural diversity within the computational models, giving a better match to the real-world system.³⁸

After network generation, the Degree of Solvation Systems were desolvated using strategies 1–6 as described by Thomas and co-workers¹⁹ in order to better compare to the previous report. However, the Phase Separated Systems and Solvent/Monomer Phase Interface Systems were desolvated using an adapted method which incorporated the possibility for alkyne-alkyne homocoupling at an earlier stage in the reaction, before the removal of solvent and unreacted monomers in size order (reflecting the increased difficulty of fragments to diffuse throughout the pore structure as they increase in size). This was believed to be more comparable to the experimental network formation mechanism reported by Laybourn and co-workers,²³⁷ where alkyne-alkyne homocoupling occurs once the aryl bromide has been completely consumed and before the solvent is removed, in order to allow the homocoupling to influence the pore structure of the CMP network before collapse upon desolvation.³⁸ Additionally, the NVT MD within the desolvation and cell workup processes was replaced with NPT MD in order to allow the cell dimensions to equilibrate throughout, using the same parameters as described above for the NPT cycle of these systems after network generation.³⁸

3.4 Results and Discussion

3.4.1 Degree of Solvation Systems

The first thing to analyse with the Degree of Solvation Systems is the influence of the number of solvent molecules, or degree of solvation, on the mass of the largest block in the system (Figure 3.20). Here, it can be seen that in all cases, as the number of solvent molecules increases from 0 to approximately 75 (indicated by the purple rectangle), the mass of the largest block increases. This was rationalised by the need to have some solvent within the system in order to allow the monomers and catalyst to diffuse throughout and increase the likelihood of reacting.³⁸ However, on adding additional solvent into the reaction mixture, the solvent hinders the ability of the building blocks to diffuse, which inhibits bond formation. This means longer reaction times were required in order to see the same degree of polycondensation, as reported by both Laybourn and co-workers experimentally, and Thomas and co-workers computationally.^{19,237} As this trend occurs in each of the solvent systems studied, this indicates that it is not only the type of solvent that can influence the polymer formation, but also the presence and quantity of solvent itself. Whilst it is expected that there will be an ideal degree of solvation to maximise polymer formation in each solvent, this is very challenging to predict and was beyond the scope of the current work.³⁸

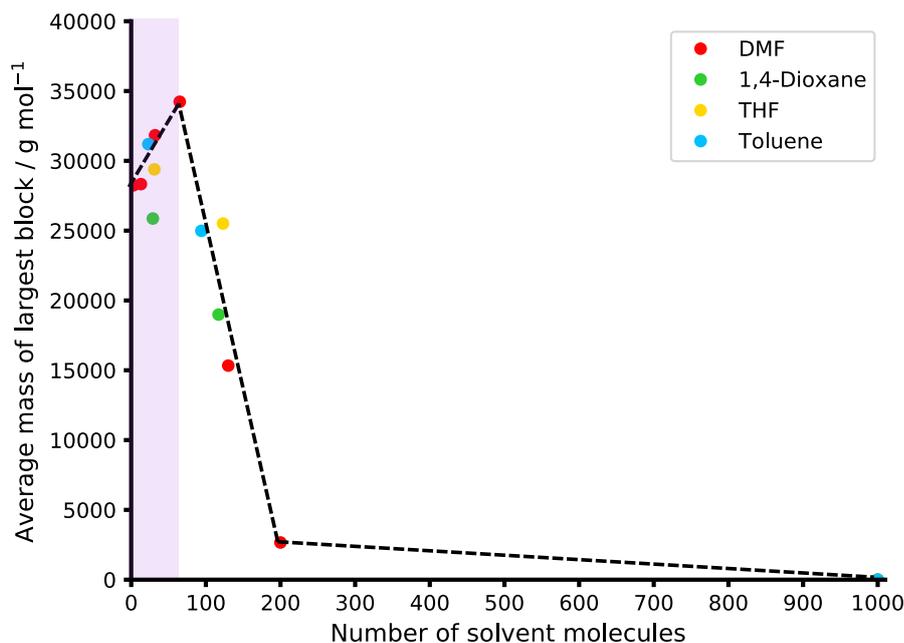


Figure 3.20: Plot of the mass of the largest fragment block in the system as a function of degree of solvation. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. The black dashed line is added as a trendline to guide the eye and is based on the DMF points, of which there are the greatest number. The purple rectangle indicates the initial increase in mass with increasing solvation. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

However, whilst predicting the ideal degree of solvation in each system is beyond the scope of this work, it was possible to establish trends based on the solvent polarity (Figure 3.18). As previously noted by Mollart and Trewin,²⁶¹ CMP-1 is a group 1 material composed from two non-polar monomers with no nitrogen functionality, meaning that as the solvent polarity increases, it would be expected that the monomers would be less soluble and miscible in the solvent. Therefore, both monomers should be fully miscible in toluene, partially co-phase separated in THF and 1,4-dioxane, to a larger extent in 1,4-dioxane due to the higher solvent polarity, and fully co-phase separated in DMF.^{38,261}

The surface area of the Degree of Solvation Systems is plotted against the degree of solvation in Figure 3.21, with the full porosity data of each of these systems reported in Appendix A Tables A.1-A.4. It can be seen that there is an approximately logarithmic relationship between the surface area and solvent concentration, regardless of the choice of solvent.³⁸ Other than this, no clear trend can be established between the solvent polarity and surface area as the degree of solvation changes. From the previous report by Mollart and Trewin,²⁶¹ it would be expected that as the solvent polarity increases, for example, in the case of DMF, there would be a worse match between the polarity of the solvent and growing polymer, as judged by larger differences in

the polarity indices and HSPs of the polymer and solvent.³⁸ This would lead to greater phase separation of the building blocks out of solution and therefore a lower concentration of solvent within the reaction mixture and a denser, less porous CMP framework with a lower surface area.³⁸

Contrastingly, a lower polarity solvent such as toluene has a better match in polarity with the growing polymer, and so it would be expected that the greater miscibility of the growing polymer within the solvent would lead to the polymer phase separating from solution at a later stage in the reaction.³⁸ This would lead to a higher concentration of solvent within the reaction mixture, giving a less dense, more open, and homogeneous polymer network with a higher surface area and a larger degree of microporosity. This hypothesis was supported by Chen and co-workers, who found that on synthesising CMP-1 in toluene with BXJ salt tuning, the BET surface area of the CMP increases from $886 \text{ m}^2 \text{ g}^{-1}$ to $1148 \text{ m}^2 \text{ g}^{-1}$, along with a localisation of the pores into the micropore region (Figure 3.17).^{38,241}

In the previous artificial synthesis report by Thomas and co-workers,¹⁹ the experimental surface areas for CMP-1 were rationalised by considering a spherical particle, with a very dense, low porosity central core, modelled by a system with a small number of solvent molecules present relative to the number of monomer and catalyst building blocks. This was surrounded by regions with increasing porosity as a result of increasing solvation and decreasing phase separation of the monomers from the solvent moving away from the polymer core towards the solvent phase.¹⁹

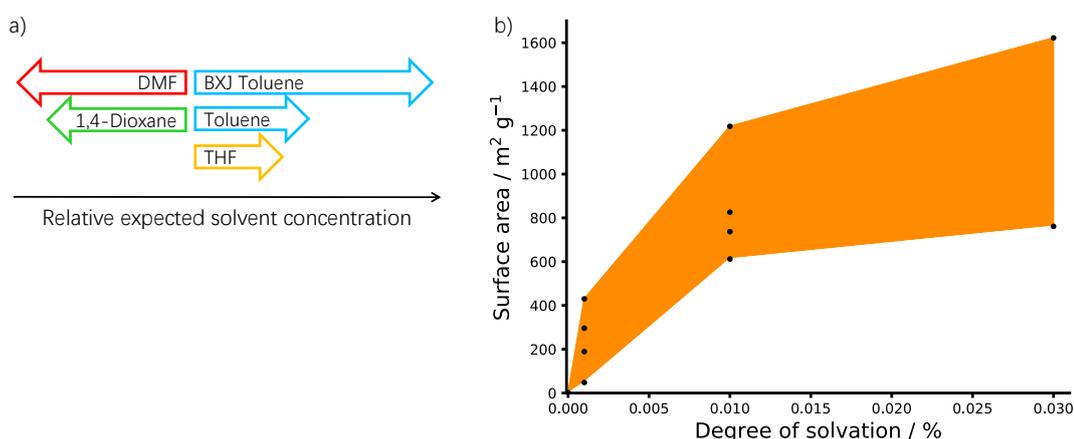


Figure 3.21: a) The relative expected quantity of solvent within each framework based on assessing the HSPs and polarity index of each solvent compared to the growing polymer are indicated by the coloured arrows. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. b) Plot of the degree of solvation against the surface area of the resulting CMP-1 material. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Similarly, in the Degree of Solvation Systems modelled here, it was seen that the surface area of the polymer increases moving from the centre of polymer particle (Figure 3.22), where there is little to no solvent, outwards towards the solvent phase, regardless of which reaction solvent is used.³⁸ The models generated using low degrees of solvation were desolvated using strategy 5, which allows the removal of solvent only due to the small size of the maximum pore diameters.

Contrastingly, the models simulated at higher degrees of solvation, with larger available pore sizes, were desolvated using strategy 6. This incorporates the removal of solvent, unreacted monomers and catalyst, and allows alkyne-alkyne homocoupling to occur.³⁸ Despite all of the polymers showing increasing surface areas with increasing solvation, there were differences between the frameworks synthesised in the different solvents. Those synthesised in toluene and 1,4-dioxane had surface areas at high levels of solvation that had decreased compared to the previous points at intermediate solvation. This was proposed as being due to the increased level of solvation increasing the amount of mesoporosity in the framework, which collapses upon desolvation to give a denser, less porous polymer than those obtained at intermediate solvation.³⁸ The CMP-1 frameworks synthesised in DMF and THF did not form a polymer at high levels of solvation, meaning that the oligomers formed were removed upon desolvation, leading to no resulting polymer surface area at high solvation. Despite this, it was not possible to determine the ideal ratio of each degree of solvation making up the overall CMP-1 polymer particle when synthesised in each solvent, and so it was necessary to explore the phase separation behaviour occurring within these frameworks.³⁸

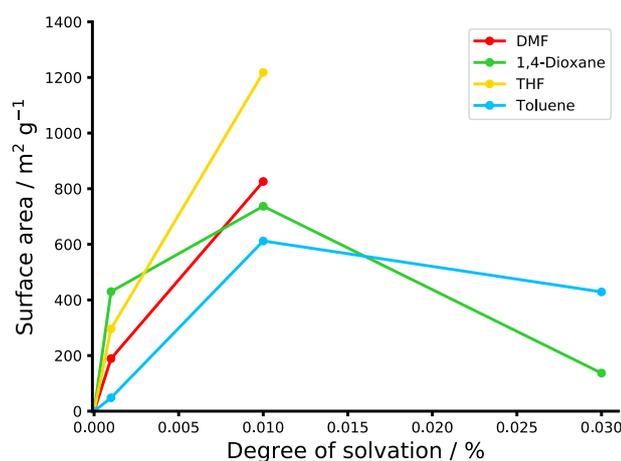


Figure 3.22: Plot of the pathway from the centre of a CMP-1 particle, with a low degree of solvation, towards the solvent phase at a higher degree of solvation against the surface area of the resulting CMP-1 material. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

3.4.2 Phase Separated Systems

The phase separation occurring within the CMP-1 framework was assessed using the Phase Separated Systems, where the starting configuration of the cell was varied to model a range of different miscibilities of the monomer and catalyst building blocks within the solvent.³⁸ Initially, the potential energy of each configuration before any bond formation occurred was measured as a function of MD cycle number throughout a loop consisting only of geometry optimisation and NPT molecular dynamics. This allowed the assessment of how the potential energy was affected by diffusion of the building blocks from each starting configuration. Also included for comparison were the one large clusters making up the Solvent/Monomer Phase Interface Systems.³⁸

The potential energies were plotted following the process given in Figure 3.23, which shows the DMF two clusters and DMF one large cluster configurations as examples. The potential energies were obtained from the Ambuild simulation as a result of the HOOMD-blue^{193,194} NPT MD. The potential energy per solvent molecule of each model was then plotted as a function of MD cycle number, as shown in Figures 3.23a and 3.23c. Then, the average potential energy per solvent molecule was calculated across repeat models 1–4 for each MD cycle number and plotted. Finally, a linear trendline of the average potential energy per solvent molecule was plotted in order to analyse the change over time. Figures 3.23b and 3.23d show the average potential energy per solvent molecule and linear trendline for the DMF two clusters and DMF one large cluster configurations.

It can be seen from Figure 3.23a that the DMF one large cluster configuration shows very similar potential energies from one repeat model to the next. This was the general trend across all of the starting configurations for each solvent apart from the DMF two clusters configuration. The DMF two clusters configuration, given in Figure 3.23c, in contrast, shows one repeat model (model 4), which has an energy of approximately 10 kJ mol⁻¹ greater than models 1–3. This shows the importance of obtaining multiple repeat structures per system for these amorphous frameworks, in order to consider structural diversity within the material.

Following the above process, linear trendlines of the potential energies per solvent molecule of each starting configuration for each solvent were compared in Figure 3.24.³⁸ By plotting the potential energy per solvent molecule, rather than the total potential energy, the systems simulated in each solvent could be considered to assess the influence of starting configuration and solvent type on the potential energy of the system.³⁸ The final potential energies of each system at the end of the simulation were plotted against the ‘degree of phase separation’ in Figure 3.25. This was defined as the surface area of

each of the clusters relative to the number of 1,4-dibromobenzene molecules within the system (Equation 3.1),³⁸

$$\text{Phase separation} = \frac{\text{Cluster length}^2 \times 6 \text{ faces} \times \text{number of clusters}}{\text{Number of 1,4-dibromobenzene molecules}} \quad (3.1)$$

which allowed the incorporation of the one large clusters from the Solvent/Monomer Phase Interface Systems within the plot.³⁸ For the fully mixed large cell, there are effectively an infinite number of clusters, and so the degree of separation would be infinite, but it is plotted as a value of 250 in Figure 3.25 so each configuration could clearly be shown on the same plot. In all cases, the most favourable system with respect to the potential energy was the one large cluster configuration, and the two clusters configuration was the least favourable.³⁸ However, the respective changes going from one configuration to the next differ with the solvent choice. The energy of the two clusters configuration compared to the neighbouring one small cluster and four clusters was far larger when DMF was the solvent employed compared to the less polar 1,4-dioxane, THF and toluene.³⁸

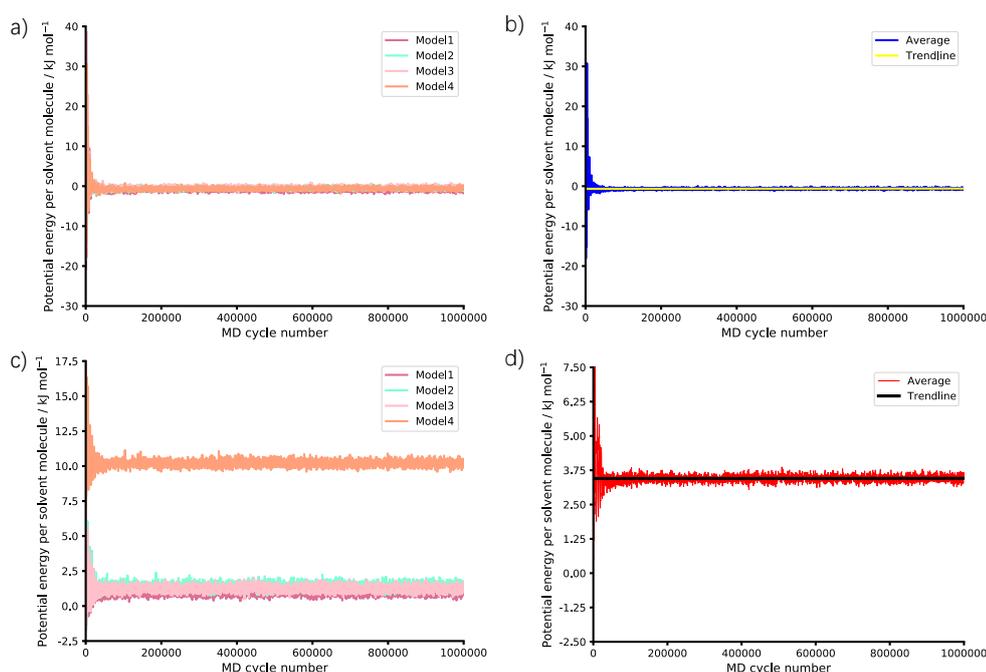


Figure 3.23: Example plots of the potential energy per solvent molecule of the Phase Separated Systems as a function of the MD cycle number after geometry optimisation and NPT molecular dynamics only. a) and c) show repeat models 1–4 individually, b) and d) show the average potential energy across models 1–4 with a linear trendline added. a) and b) DMF one large cluster, c) and d) DMF two clusters. Key: model 1 – pale red, model 2 – green, model 3 – pink, model 4 – orange, two clusters average – red, two clusters trendline – black, one large cluster average – blue, one large cluster trendline – yellow.

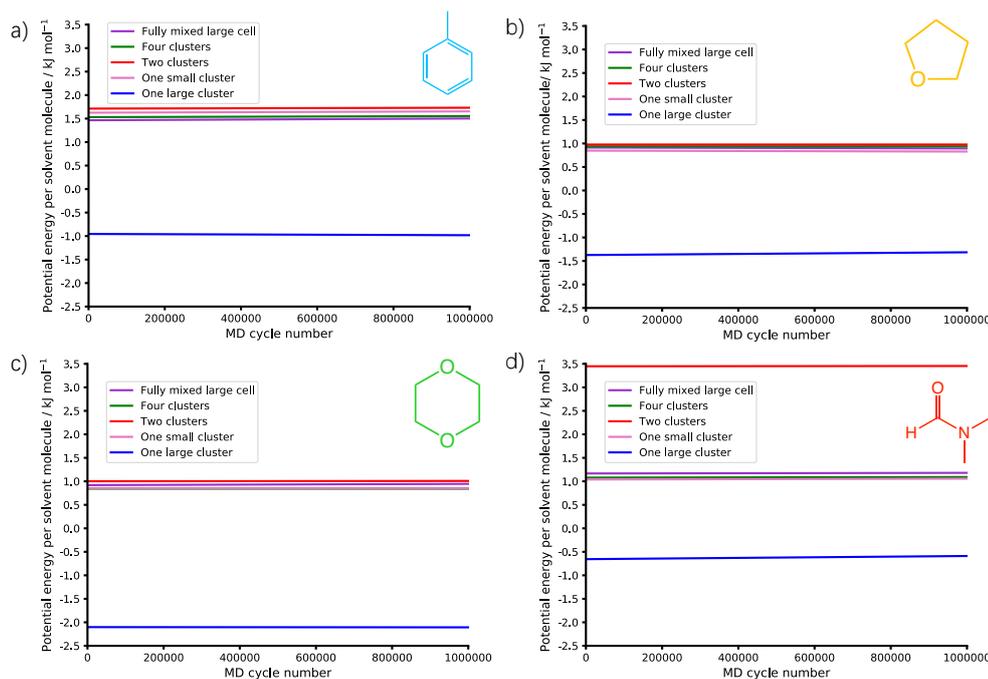


Figure 3.24: Linear trendlines of the potential energy per solvent molecule of the Phase Separated Systems as a function of the reaction timestep after geometry optimisation and NPT molecular dynamics only, averaged across repeat models 1–4. The structures of the solvent molecules are added for clarity. a) Toluene, b) THF, c) 1,4-dioxane, d) DMF. Key: fully mixed large cell – purple, four clusters – green, two clusters – red, one small cluster – pink, one large cluster – blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Following this initial assessment of the potential energy of the Phase Separated Systems after diffusion only, each system underwent the network generation, and desolvation, homocoupling, cell workup and equilibration protocol to generate the polymer networks and compare the resulting porosity of each. Table 3.8 presents a summary, with the individual data for each model given in Appendix A Tables A.5–A.8.³⁸ It should be noted that as was found in the previous artificial synthesis report by Thomas and co-workers,¹⁹ the high level of solvation incorporated within the Phase Separated Systems (100% of the experimental stoichiometry) gave rise to an increase in the available pore sizes.¹⁹ The obtained surface areas were also very large compared to the original report and experimental data for CMP-1,^{13,19,234–237} which was due to an artefact in the system setup, meaning that whilst the relative surface areas in each configuration could be considered, the absolute values were not representative of the real-world system.³⁸

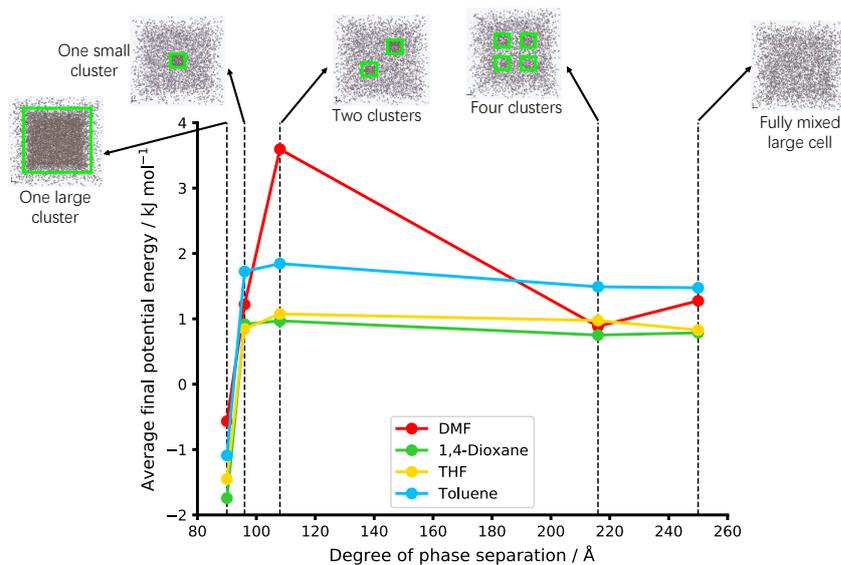


Figure 3.25: Plot of the final potential energy per solvent molecule of the Phase Separated Systems and Solvent/Monomer Phase Interface Systems as a function of the degree of phase separation (cluster surface area relative to the number of 1,4-dibromobenzene molecules, Equation 3.1), averaged across repeat models 1–4. Examples of each configuration are given in the insets, with the green boxes given around each cluster for clarity. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Table 3.8: Summary of the porosity of each Phase Separated System artificially synthesised in toluene, THF, 1,4-dioxane and DMF, averaged across repeat models 1–4 for each configuration. Cluster SA relative to number of DBB blocks – degree of phase separation as defined in Equation 3.1, with the value for the fully mixed large cell configuration being effectively infinite, PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Solvent	Configuration	Cluster SA relative to number of DBB blocks	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
Toluene	Fully mixed large cell	1E+15	17.25	24.93	7.72	10971	0.13	31.57
	Four clusters	216	8.47	13.68	1.82	2900	0.51	20.99
	Two clusters	108	11.37	19.12	2.82	4679	0.38	22.54
	One small cluster	96	11.65	18.70	2.37	4625	0.35	22.17
THF	Fully mixed large cell	1E+15	17.98	27.67	8.93	11439	0.12	32.77
	Four clusters	216	8.31	14.64	1.70	3770	0.45	21.31
	Two clusters	108	12.36	19.68	2.49	4569	0.36	22.51
	One small cluster	96	13.72	20.54	3.01	5311	0.31	23.94
Dioxane	Fully mixed large cell	1E+15	19.33	26.82	9.37	11636	0.11	33.24
	Four clusters	216	13.60	22.64	3.34	5865	0.31	24.87
	Two clusters	108	9.41	15.36	1.97	3714	0.46	21.12
	One small cluster	96	16.85	26.65	3.96	5776	0.25	25.31
DMF	Fully mixed large cell	1E+15	19.78	29.90	10.21	11928	0.09	34.44
	Four clusters	216	9.30	18.49	2.20	4733	0.39	22.47
	Two clusters	108	13.90	21.20	3.59	4829	0.37	23.98
	One small cluster	96	17.66	26.89	4.20	5282	0.27	25.59

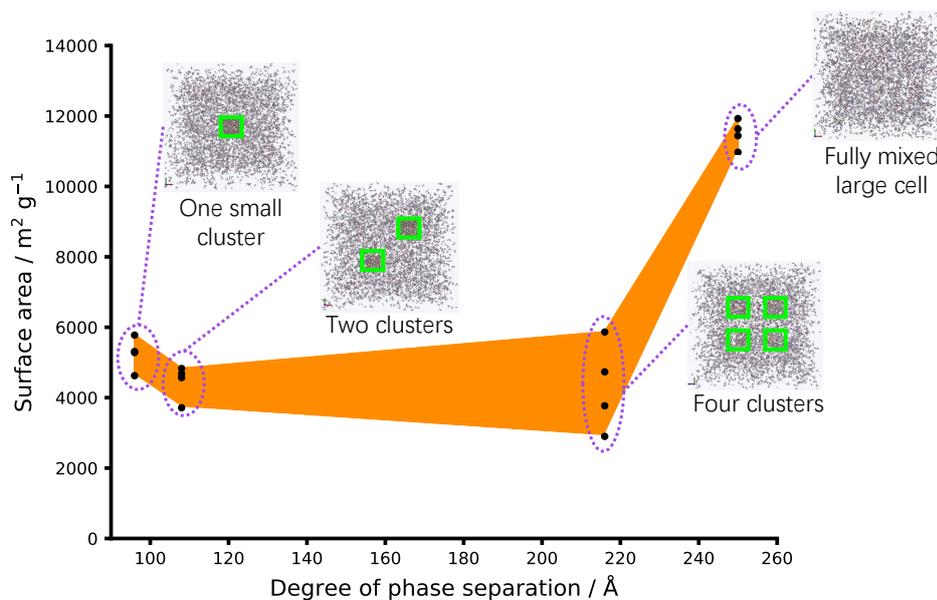


Figure 3.26: Plot of the surface areas of the Phase Separated Systems as a function of degree of phase separation, averaged across repeat models 1–4. A larger degree of phase separation indicates a larger degree of solvent mixing. Examples of each configuration are given in the insets, with the green boxes given around each cluster for clarity. The absolute values are not representative and are just used for comparison purposes between configurations. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

In all solvents, the configuration giving rise to the models with the highest surface areas and micropore volumes was the fully mixed large cell (Figure 3.26), whilst all of the remaining configurations displayed similar values that were approximately half of the network-accessible surface areas of the fully mixed large cell configuration. This trend matches well with the previous hypothesis of Chen and co-workers, where an increased monomer solubility in the solvent leads to an increased quantity of solvent within the reaction mixture and a polymer with a higher resulting surface area.^{38,241}

3.4.3 Solvent/Monomer Phase Interface Systems

The Solvent/Monomer Phase Interface Systems, with a scaled-up quantity of monomers and catalyst building blocks, are the closest in size to the experimental CMP-1 polymer particles. The experimental SEM data suggests that these amorphous, spherical particles have diameters in the range of 150–450 nm,^{13,234–237} however, the current, fully atomistic approach described here is limited to modelling the microscale region only, with pore sizes in the range of 0–2 nm, meaning that particles of the same scale as the experimental material cannot be modelled using the current approach.^{2,38} Even at this size of model, analysis proved challenging, and it was not possible to obtain porosity information of these systems using Poreblazer 4.0.¹⁶² Instead, porosity analysis of

these materials was undertaken using Materials Studio 5.0,¹⁵⁸ *via* a process described in more detail later. Bearing this in mind, it is possible to consider these systems as the interface between the central polymer core, where it is unlikely, due to the very small pore sizes and high density, that solvent will be able to interact with the CMP material, and the solvent phase.³⁸

Firstly, as in the case of the Phase Separated Systems, the structures of the Solvent/Monomer Phase Interface Systems modelled in each solvent were analysed before the network generation could occur, after diffusion of the building blocks only.³⁸ Reflecting on the polarity of the solvent in comparison to the non-polar CMP-1 monomers (Figure 3.18), both monomers would be expected to be fully miscible in toluene, as the least polar solvent (polarity index = 2.4),¹⁷² meaning that there should be the least amount of phase separation within this solvent. Increasing the solvent polarity to THF, 1,4-dioxane and DMF (polarity index = 4.0, 4.8 and 6.4, respectively)¹⁷² would give a decreasing miscibility of the monomers within the solvent and an increasing phase separation.²⁶¹ This would suggest that as the solvent polarity increases, the quantity of solvent spread throughout the simulation cell would decrease and the solvent would reside in more localised, phase separated regions.³⁸

To test this, concentration profiles of each Solvent/Monomer Phase Interface System were generated using Python after diffusion of the building blocks pre-network generation.³⁸ These were designed to analyse the location of solvent throughout the simulation cell by taking slices in each of the three Cartesian dimensions and counting the number of solvent molecules in each slice (Figure 3.27). The first step involved taking a slice through each plane of the Solvent/Monomer Phase Interface Systems with fractional lattice ranges of 0.45–0.55 in two Cartesian dimensions, by firstly collecting all solvent molecules that fit within the lattice range in one direction, then keeping those that also fit within the lattice range for the second axis making up the plane (Figure 3.27a–c).³⁸ Following this, fragments of fractional lattice size 0.1 were taken along the third axis (with a lattice range of 0–1, Figure 3.27d). The number of solvent molecules within each fragment were then counted.

The lattice ranges used were: $x = 0.45\text{--}0.55$, $y = 0\text{--}1$, $z = 0.45\text{--}0.55$ (slice A), $x = 0\text{--}1$, $y = 0.45\text{--}0.55$, $z = 0.45\text{--}0.55$ (slice B) and $x = 0.45\text{--}0.55$, $y = 0.45\text{--}0.55$, $z = 0\text{--}1$ (slice C).³⁸ The solvent molecules were collected by identifying a ‘unique’ typing per solvent molecule, given as the oxygen atom in the DMF and THF molecules and the sp^3 carbon atom in toluene. In the case of 1,4-dioxane, each atom type appeared more than once per molecule. To prevent bias by selecting one oxygen atom over the other, the profile was generated for the first oxygen atom per molecule, then for the second, and the final profile contained the average number of solvent molecules per frag-

ment across both oxygen atoms. To allow comparison between the different solvents, the number of solvent molecules per fragment was normalised according to Equation 3.2.³⁸ The Python script utilised to generate the concentration profiles is given in Appendix A Figure A.3-A.4.

$$\text{Normalised count} = 100 \times \frac{\text{Solvent molecules in the fragment}}{\text{Total number of solvent molecules}} \quad (3.2)$$

The concentration profiles for the Solvent/Monomer Phase Interface Systems after diffusion of building blocks only are given in Figures 3.28–3.31.³⁸ To better compare the concentration profiles, they were also plotted in four dimensions (Figure 3.32), where the xyz coordinates correspond to the points on the plot, and the fourth dimension (colour) corresponds to the normalised count of solvent molecules within each fragment, averaged across repeat models 1–4 for each fragment in each solvent.³⁸ These gave an indication of where the solvent resided within each Solvent/Monomer Phase Interface System and show that there is solvent throughout each model. This adds weight to the hypothesis that larger models are required to visualise the region at the very centre of the CMP-1 polymer particles where it is expected that no solvent will be able to interact due to the limited porosity.³⁸

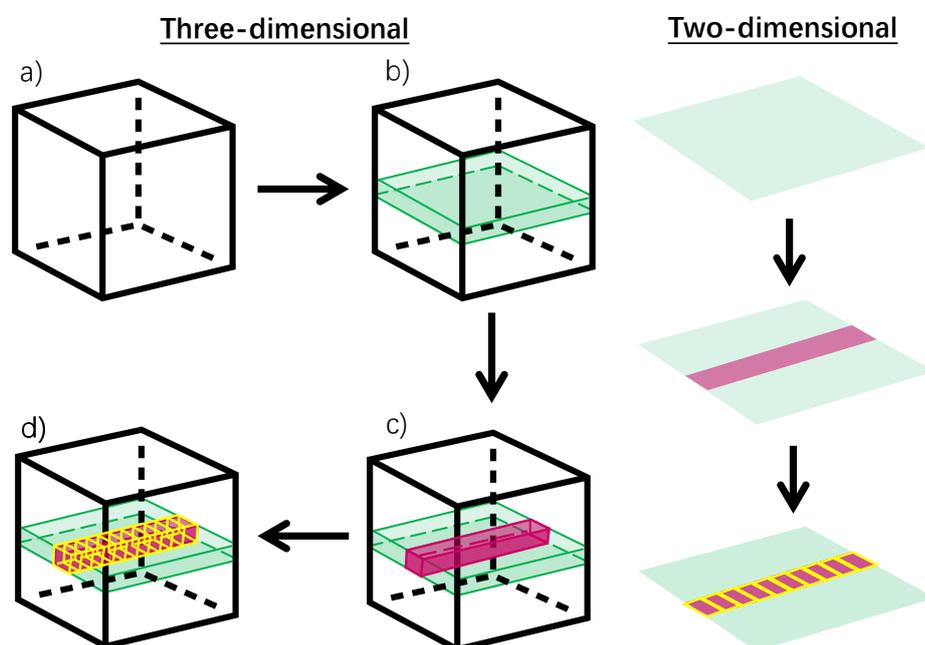


Figure 3.27: Cartoon representation of the methodology used to acquire the concentration profiles for the Solvent/Monomer Phase Interface Systems after diffusion of building blocks only. The process composed of: a) take the whole system, b) collect the coordinates of the solvent molecules that are within the first axis range (green), c) keep those that are also within the second axis range (pink), and d) sample fragments of fractional lattice size 0.1 (yellow), counting the number of solvent molecules within each. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

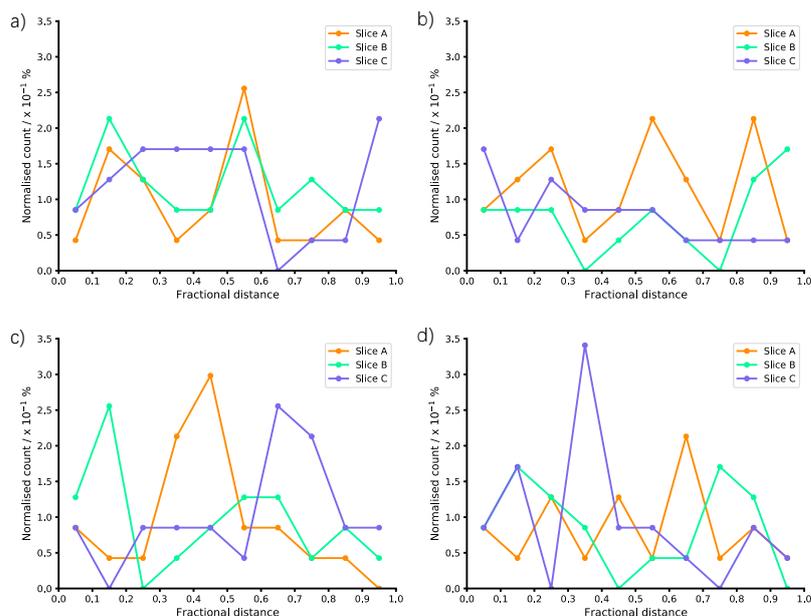


Figure 3.28: Concentration profiles showing the normalised number of solvent molecules as a function of fractional distance throughout the unit cell for the CMP-1 Solvent/Monomer Phase Interface Systems generated after diffusion of building blocks only within toluene solvent. a) Model 1, b) Model 2, c) Model 3, d) Model 4. Key: slice A – orange, slice B – green, slice C – purple. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

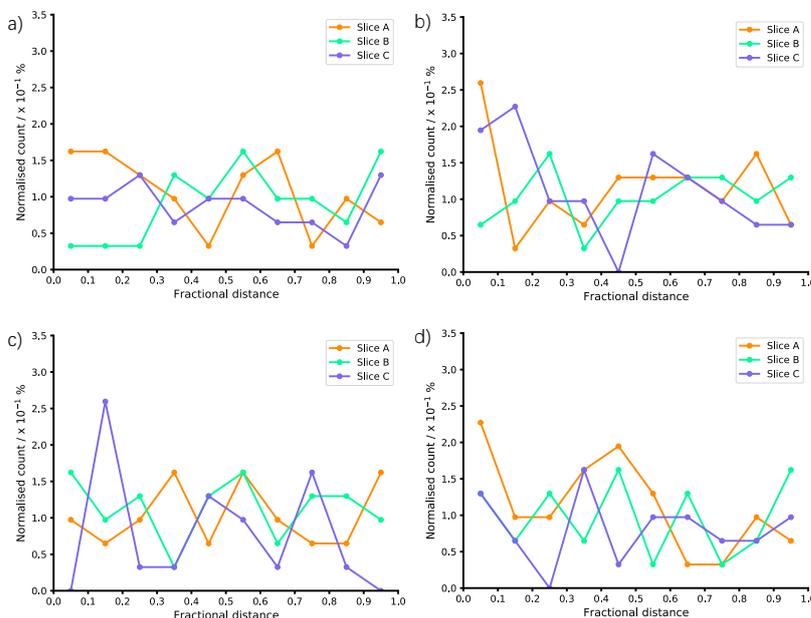


Figure 3.29: Concentration profiles showing the normalised number of solvent molecules as a function of fractional distance throughout the unit cell for the CMP-1 Solvent/Monomer Phase Interface Systems generated after diffusion of building blocks only within THF solvent. a) Model 1, b) Model 2, c) Model 3, d) Model 4. Key: slice A – orange, slice B – green, slice C – purple. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

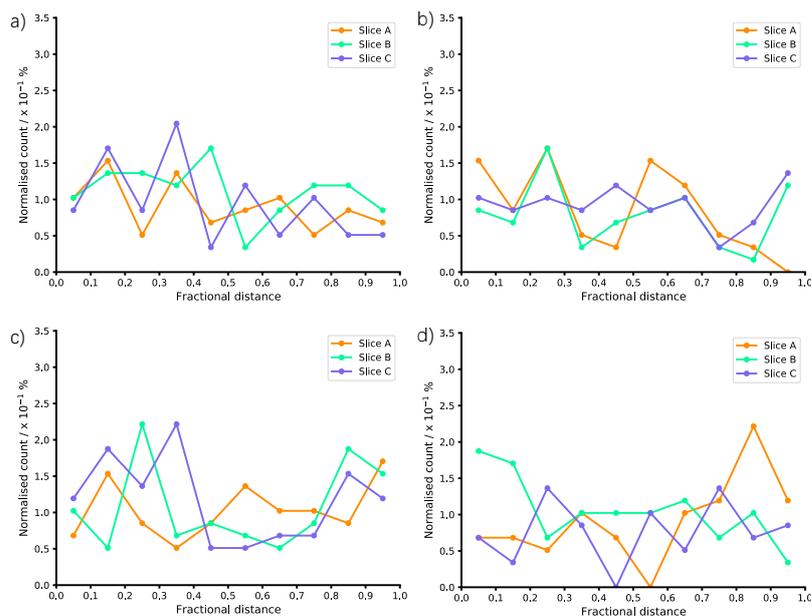


Figure 3.30: Concentration profiles showing the normalised number of solvent molecules as a function of fractional distance throughout the unit cell for the CMP-1 Solvent/Monomer Phase Interface Systems generated after diffusion of building blocks only within 1,4-dioxane solvent. a) Model 1, b) Model 2, c) Model 3, d) Model 4. Key: slice A – orange, slice B – green, slice C – purple. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

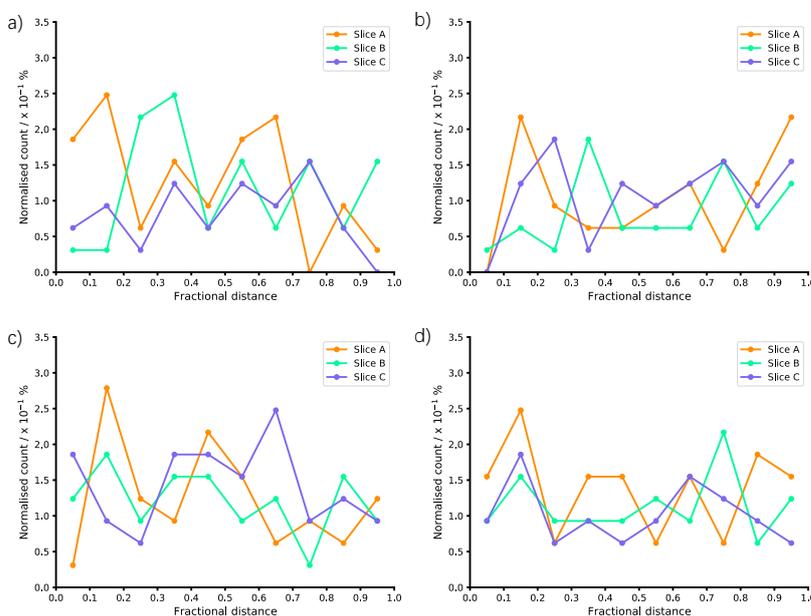


Figure 3.31: Concentration profiles showing the normalised number of solvent molecules as a function of fractional distance throughout the unit cell for the CMP-1 Solvent/Monomer Phase Interface Systems generated after diffusion of building blocks only within DMF solvent. a) Model 1, b) Model 2, c) Model 3, d) Model 4. Key: slice A – orange, slice B – green, slice C – purple. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

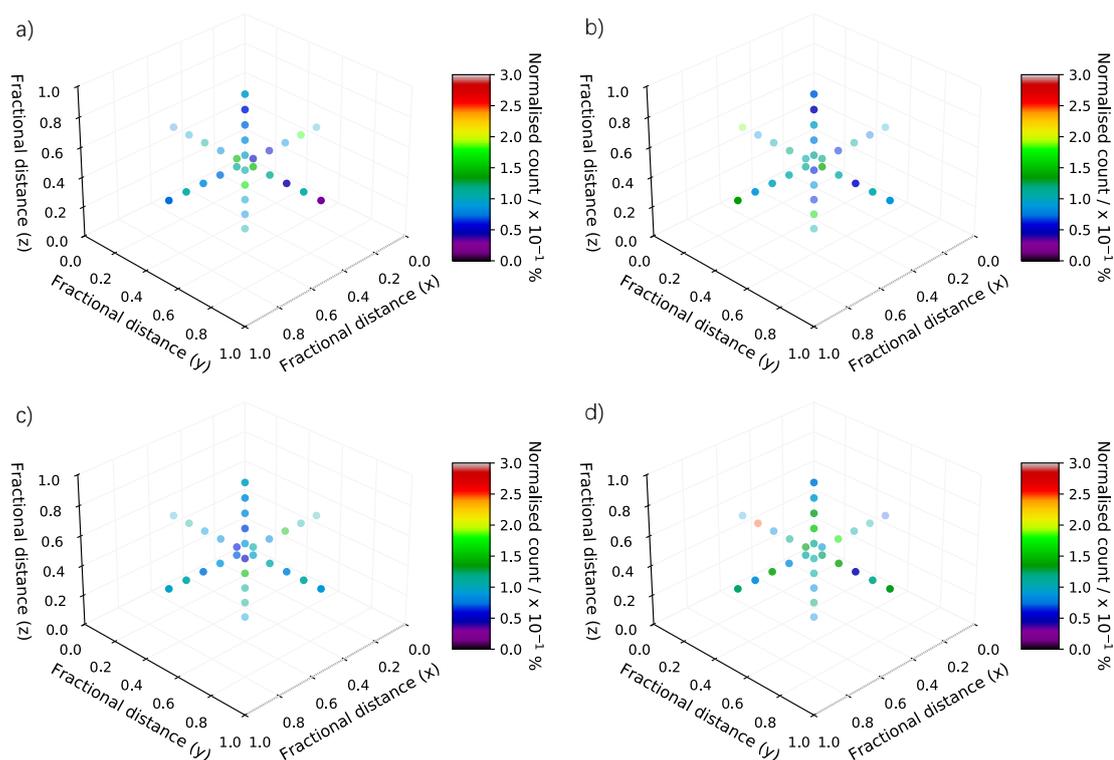


Figure 3.32: Four-dimensional concentration profiles of the CMP-1 Solvent/Monomer Phase Interface Systems generated after diffusion of building blocks only. The points correspond to the xyz coordinates of the centre of each sampled bin and the fourth dimension (colour) corresponds to the normalised count of solvent molecules within each bin. The count is averaged across repeat models 1–4. a) Toluene solvent, b) THF solvent, c) 1,4-dioxane solvent, d) DMF solvent. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

The concentration profiles were also plotted by multiplying the raw count of solvent molecules (before normalisation) within each bin by the molecular volume of each solvent to give a solvent volume (Figure 3.33) in order to assess how much of the total cell volume is made up of solvent relative to the quantity of CMP.³⁸ The molecular volume of each solvent was acquired by drawing each solvent molecule in Materials Studio 5.0,¹⁵⁸ and running a Forcite optimisation at the ‘fine’ level of quality, using the PCFF forcefield with the charges being optimised using the Gasteiger approach.¹⁹⁸ The Connolly volume occupied by the van der Waals surface was then acquired, using a probe radius of 0 Å and an ultra-fine grid interval of 0.15 Å. The molecular volumes of each solvent were given as: toluene - 0.657 cm³ g⁻¹, THF - 0.646 cm³ g⁻¹, 1,4-dioxane - 0.585 cm³ g⁻¹, DMF - 0.634 cm³ g⁻¹.³⁸ Figure 3.33 was then plotted, with the points corresponding to the solvent volume of each of slices A, B, and C across all four models per structure, per fractional distance. This corresponded to a horizontal line of points at, for example, a fractional distance of 0.05, each representing the solvent volume of one of the slices of a model within the fractional distance range of 0–0.1. In some cases, where

the solvent volumes were identical for multiple datapoints in the same lattice range, the points would overlap.³⁸

Whilst at first glance the profiles given in Figure 3.33 appear similar, given the relatively small number of models sampled and limited simulation cell size compared to the amorphous real-world material, trends can be established depending on the solvent polarity.³⁸ For example, toluene, as the least polar solvent, where the smallest degree of phase separation of the monomers out of solution would be expected, shows a clear increase in the solvent volume in the central region of the cell, whilst DMF, as the most polar solvent, shows a clear decrease (Figure 3.33a and d, respectively).³⁸

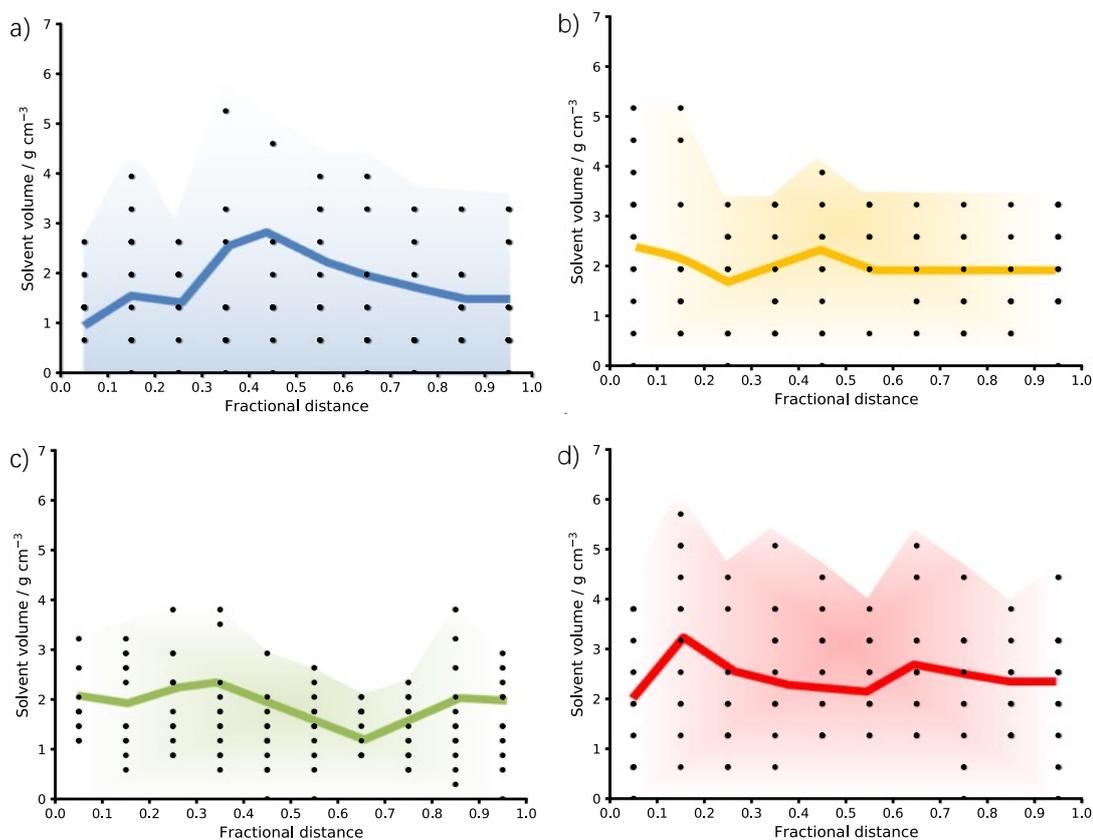


Figure 3.33: Concentration profiles showing the occupied solvent volume as a function of fractional distance throughout the unit cell for the CMP-1 Solvent/Monomer Phase Interface Systems generated after diffusion of building blocks only, including the solvent volumes for slices A–C of all of repeat models 1–4. The shaded regions indicate the general shape of each plot, and the line, which is the average solvent volume across all points (where some overlap) for each fractional distance, is included as a trend to guide the eye. a) Toluene solvent, b) THF solvent, c) 1,4-dioxane solvent, d) DMF solvent. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

The fractional distance corresponding to the maximum number of solvent molecules for each slice of each of the Solvent/Monomer Phase Interface Systems was plotted against the solvent polarity (Figure 3.34).³⁸ This shows a very clear indication of the degree of phase separation occurring with each reaction solvent, which matches the predicted hypothesis. In toluene, the least polar solvent (polarity index = 2.4),¹⁷² there is an almost uniform distribution of solvent throughout the simulation cell. Increasing the solvent polarity to THF (polarity index = 4.0)¹⁷² shows a unimodal distribution of solvent in a fractional distance range of 0.05–0.65, with one additional point at 0.95.³⁸ Considering that the system is periodic, this point would add to the unimodal distribution across the periodic boundary. With 1,4-dioxane as the solvent (polarity index = 4.8),¹⁷² there is a bimodal distribution of solvent throughout the edges of the simulation cell, spanning towards the centre, with no solvent present in the fractional distance range of 0.5–0.7. When the most polar solvent studied, DMF (polarity index = 6.4)¹⁷² is employed, there are very discrete regions of solvent throughout the simulation cell, localised in fractional distance ranges of 0.1–0.35 and 0.6–0.8, with one additional point at 0.95.³⁸ This clearly demonstrates that as the solvent becomes more polar, giving a worse polarity match to the non-polar monomers, the amount of phase separation increases within the CMP-1 framework. The observed differences in phase separation occurring within the polymers due to differing miscibilities of the monomers within each solvent change the homogeneity within the structure. This gives rise to differences in the percentages of micro- and mesoporosity that contribute to the total surface area of each material.³⁸

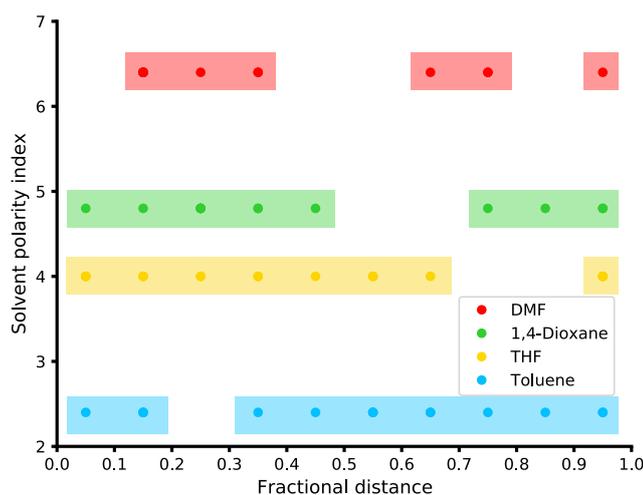


Figure 3.34: Plot of the fractional distance corresponding to the maximum number of solvent molecules within each slice of each Solvent/Monomer Phase Interface System. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

The Solvent/Monomer Phase Interface Systems were also assessed post-artificial synthesis, after the network generation, and desolvation, homocoupling, workup and cell equilibration processes. The hypothesis of Thomas and co-workers¹⁹ was that the CMP-1 framework was composed of spherical particles that fuse upon desolvation, with increasing pore sizes and surface area moving from the centre of each polymer particle outwards towards the solution phase.¹⁹ It would be expected that the smaller pores within the polymer core would correspond to a denser structure in comparison to the more porous outer regions.³⁸ To assess this, slices were taken through each model after network generation and post-artificial synthesis. The slices were obtained using Materials Studio 5.0,¹⁵⁸ moving from the edge of each model into the centre. The fractional lattice ranges used were $a = 0.45-0.55$, $b = 0-1$, $c = 0.45-0.55$ (slice A) and $a = 0-1$, $b = 0.45-0.55$, $c = 0.45-0.55$ (slice B).³⁸ After network generation, the simulation cell length was 200 Å, corresponding to a slice of 20 Å × 200 Å × 200 Å for slice A and 200 Å × 20 Å × 200 Å for slice B, with the centre of the model being at 100 Å. Post-artificial synthesis, the cell volume had decreased due to the removal of building blocks freeing up space within the cell. This was removed during the NPT MD to dispose of any unstable pore voids that would not occur in the real-world material.^{19,38} The centre of the simulation cell post-artificial synthesis occurred at approximately 45 Å, with the ratio of the slice compared to the whole model staying consistent with the slice after network generation. The slice was sampled by taking fragments of size 0.1 using Materials Studio 5.0.¹⁵⁸ The slicing protocol is shown in Figure 3.35, and the Connolly, smoothed solvent surface, and smoothed solvent-accessible surface areas are given in Appendix A Tables A.9-A.10.³⁸

The plot of density against distance from the edge of the cell after network generation (Figure 3.35) shows little change depending on the solvent, which is expected as the space within each model will be filled efficiently.³⁸ The relaxation of the cell volume during the desolvation protocol caused large decreases in the volume within regions of the cell previously composed of mostly solvent with less dense regions of polymer in comparison to those that contained fully polymerised CMP-1 material.³⁸ Figure 3.36 shows the density of each of the Solvent/Monomer Phase Interface Systems post-artificial synthesis as a function of distance from the edge of the cell. It can be seen that the average density of each model is in the region of 0.8 g cm⁻³ that has been previously reported for CMP-1 materials,^{13,234-237} with some values above and below due to the amorphous nature of the framework and position within the slice.³⁸ Additionally, Figure 3.36 shows a reduction in void space after the cell equilibration, which removed empty regions of space within the cell, such as those around the cell edge in Figure 3.35.

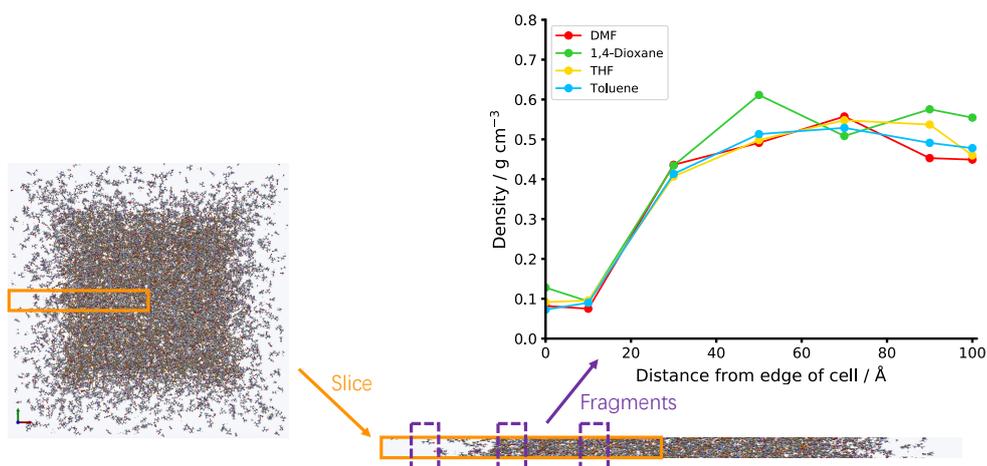


Figure 3.35: The slicing protocol used to analyse the Solvent/Monomer Phase Interface Systems after the network generation procedure. The plot of average density is shown relative to the distance from the edge of the simulation cell. 0 Å corresponds to the edge of the simulation cell and 100 Å corresponds to the centre. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Comparing the density of each fragment after network generation to that after network generation, desolvation, homocoupling, workup and cell equilibration (Figure 3.37) indicates that the density increased throughout after desolvation due to the removal of solvent and unreacted building blocks.³⁸ The plot after desolvation, homocoupling, workup and cell equilibration shows a similar trend to that after network generation but is compressed into a smaller cell volume. This trend was rationalised by a ‘crumple zone’ analogy.³⁸ This analogy indicated that the centre of each polymer cluster, which was already of high density after network generation due to a higher degree of polymerisation, was not able to increase much in density after desolvation, homocoupling, workup and cell equilibration. This was due to the space already being efficiently filled by the rigid CMP-1 polymer framework.³⁸ In comparison, the regions nearer the edge of the simulation cell, which were of lower density after the network generation due to being less polymerised, were able to condense more efficiently, leading to a greater volume of space that could be filled upon desolvation to give a denser network.³⁸

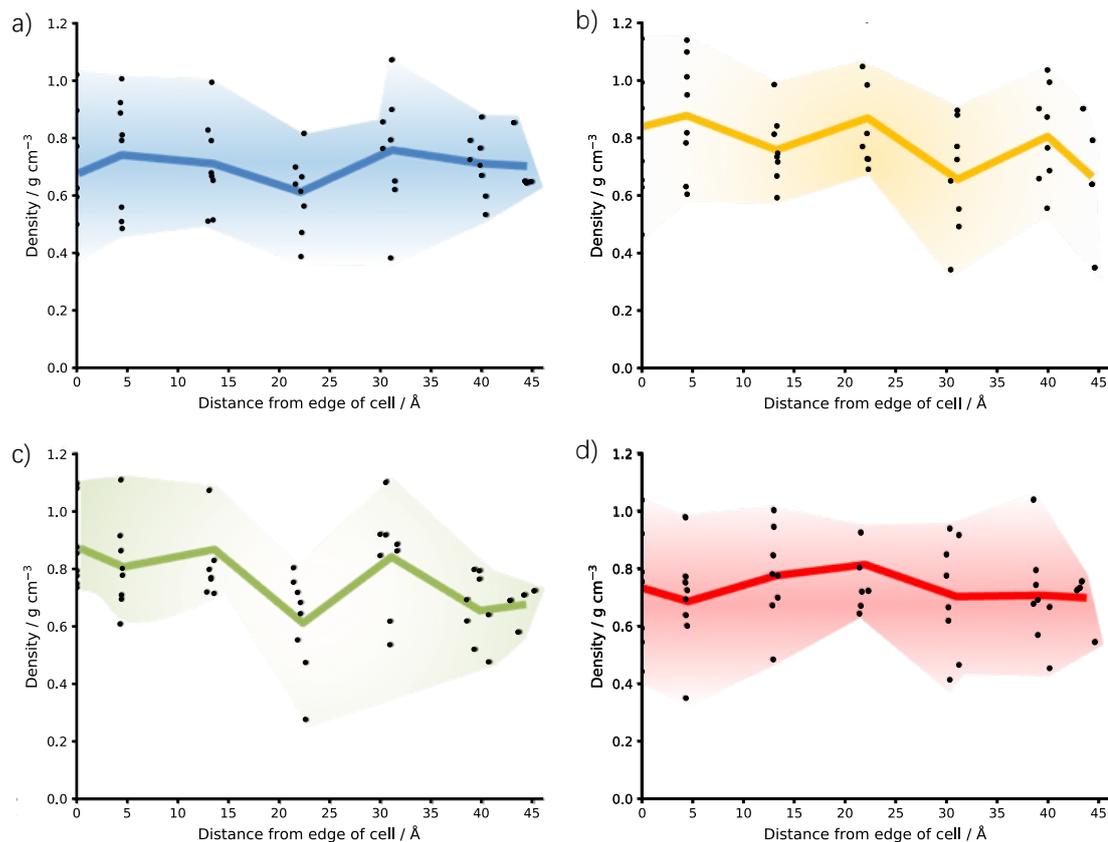


Figure 3.36: Plot of the density of the Solvent/Monomer Phase Interface Systems as a function of distance from the edge of the simulation cell post artificial-synthesis. 0 Å corresponds to the edge of the simulation cell and 45 Å corresponds to the centre. The densities for slices A and B of all of repeat models 1–4 are included to show the range of values obtained. The shaded regions indicate the general shape of each plot, and the line, which is the average density across all points (where some overlap) for each distance, is included as a trend to guide the eye. a) Toluene, b) THF, c) 1,4-dioxane, d) DMF. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

The hypothesis at this point was that the degree of phase separation of the monomers and growing polymer framework relative to the solvent, influenced by the varying solubility within each solvent as the polarity changes, causes the differences in the percentages of micro- and mesoporosity making up the total porosity of the CMP-1 material (Figure 3.18 shows the experimental data).³⁸ This appeared to be correct from the solvent volume concentration profile given in Figure 3.33 and the resulting maximum solvent count as a function of fractional distance given in Figure 3.34, which indicated that the CMP-1 frameworks simulated in toluene and THF contained a more uniform spread of solvent throughout the simulation cell in comparison to those simulated with 1,4-dioxane and DMF. To further test this hypothesis, PSDs were obtained for the Solvent/Monomer Phase Interface System fragments analysed after network generation, desolvation, homocoupling, cell workup and

equilibration, shown in Figure 3.38.³⁸ Due to the nature of the slicing protocol to obtain each fragment, which was necessarily small in comparison to the overall model in order to allow the pore sizes to be calculated using Poreblazer 4.0,¹⁶² the overall pore width range obtained is very small, and is not indicative of the range of pore sizes that would be expected within these materials. It is also not necessarily the case that the centre of the cell after network generation remains in the same place after the desolvation protocol, where an increased quantity of solvent in one region of the simulation cell could shift the centre of the network away from the centre of the cell.³⁸ Instead, this approach offers a comparison of the trends seen when synthesised using the various solvents studied, and the respective ability of each to pack efficiently upon solvent removal.³⁸

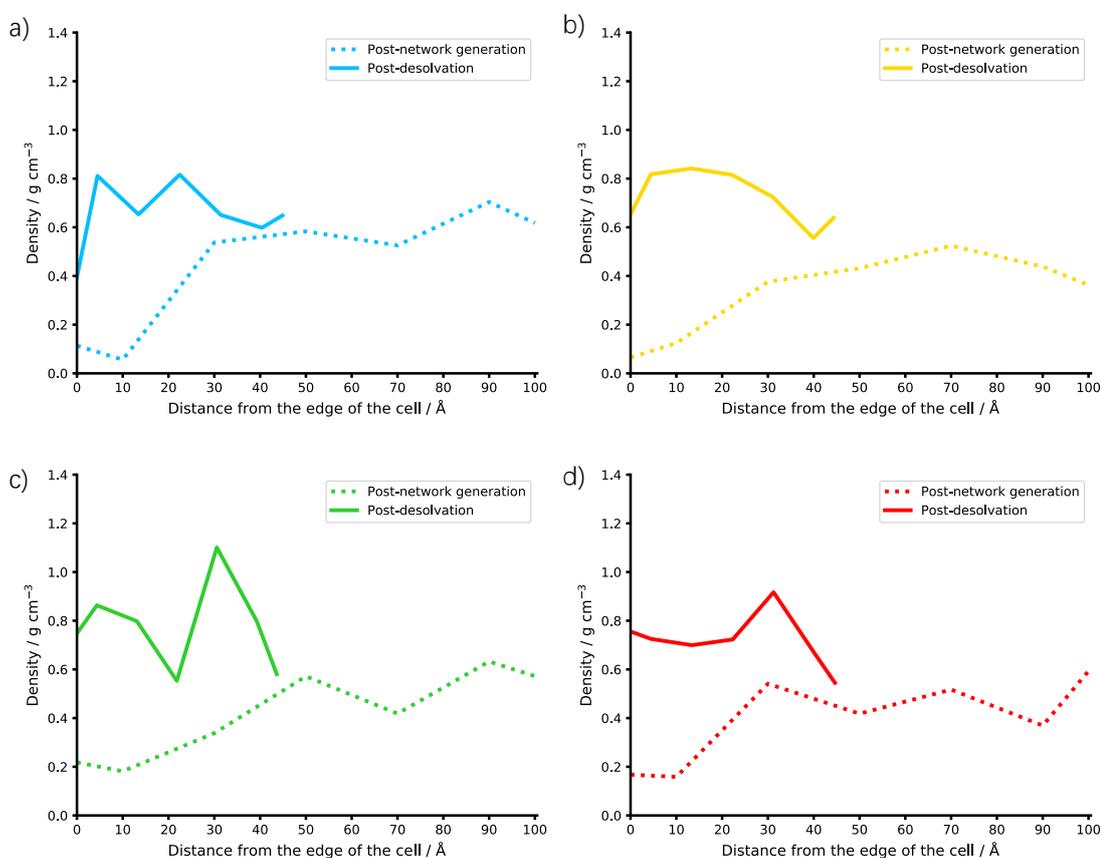


Figure 3.37: Examples of the crumple zone effect observed with respect to the density of the Solvent/Monomer Phase Interface Systems as a function of distance from the edge of the simulation cell after the network generation (dotted line) and network generation, desolvation, homocoupling, workup and cell equilibration protocol (solid line), where 0 Å corresponds to the edge of the simulation cell, 100 Å corresponds to the centre of the cell after network generation and 45 Å corresponds to the centre of the cell post-artificial synthesis. a) Model 3 slice A, synthesised in toluene, b) Model 1 slice A, synthesised in THF, c) Model 3 slice A, synthesised in 1,4-dioxane, d) Model 4 slice A, synthesised in DMF. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

Each PSD was obtained by running fragments 0–6 obtained for sampling the density and porosity of the Solvent/Monomer Phase Interface Systems after network generation, desolvation, homocoupling, workup and cell equilibration, for slices A and B of repeat models 1–4, through Poreblazer 4.0.¹⁶² This gave eight pore size distributions per fragment, per solvent.³⁸ As each cell had reduced to a slightly different cell volume during the desolvation, homocoupling, workup and cell equilibration protocol, this resulted in very small differences in the pore widths sampled, so these were averaged across the eight datasets. The intensities of each were summed to represent the total contribution of pores within each fragment.³⁸

At first glance, the PSDs of the CMP-1 models simulated in toluene, 1,4-dioxane and DMF appear similar, with pores within the range of 2.0–6.5 Å, whilst the pores of the models simulated using THF span over a slightly wider range of 2.0–7.0 Å.³⁸ The individual pore sizes for each fragment do change slightly within each range due to the amorphous nature of the material and the location of each fragment within the cluster. The relative differences in pore sizes between models can be analysed further. For DMF, which is the most polar solvent studied (polarity index = 6.4),¹⁷² it would be expected from the hypothesis of Chen and co-workers²⁴¹ that this would correspond to a lower concentration of solvent throughout the reaction mixture due to a worse match between the polarity of solvent and polymer, leading to a high degree of microporosity throughout the material.³⁸ After desolvation, the low quantity of solvent throughout the reaction mixture means that a relatively small number of solvent molecules can be removed from the framework. This corresponds to only a small amount of space that can be packed upon desolvation, leading to the microporosity remaining high throughout, regardless of which fragment is studied.³⁸

As the solvent polarity decreases to 1,4-dioxane and THF (polarity index = 4.8 and 4.0, respectively),¹⁷² increasing quantities of solvent will be expected throughout the material during the network generation, leading to increased void spaces that can be filled upon desolvation. These correspond to a small shift to larger pore sizes as the solvent polarity decreases from DMF to 1,4-dioxane to THF, which is observed in the spectra.³⁸ Following this logic, it would be expected that the largest pore sizes would be observed for the CMP-1 models synthesised using toluene, the least polar solvent studied (polarity index = 2.4),¹⁷² However, this is not seen in the spectrum, and in fact, the PSDs of the models synthesised using THF, the second least polar solvent, have the highest observed pore sizes.³⁸

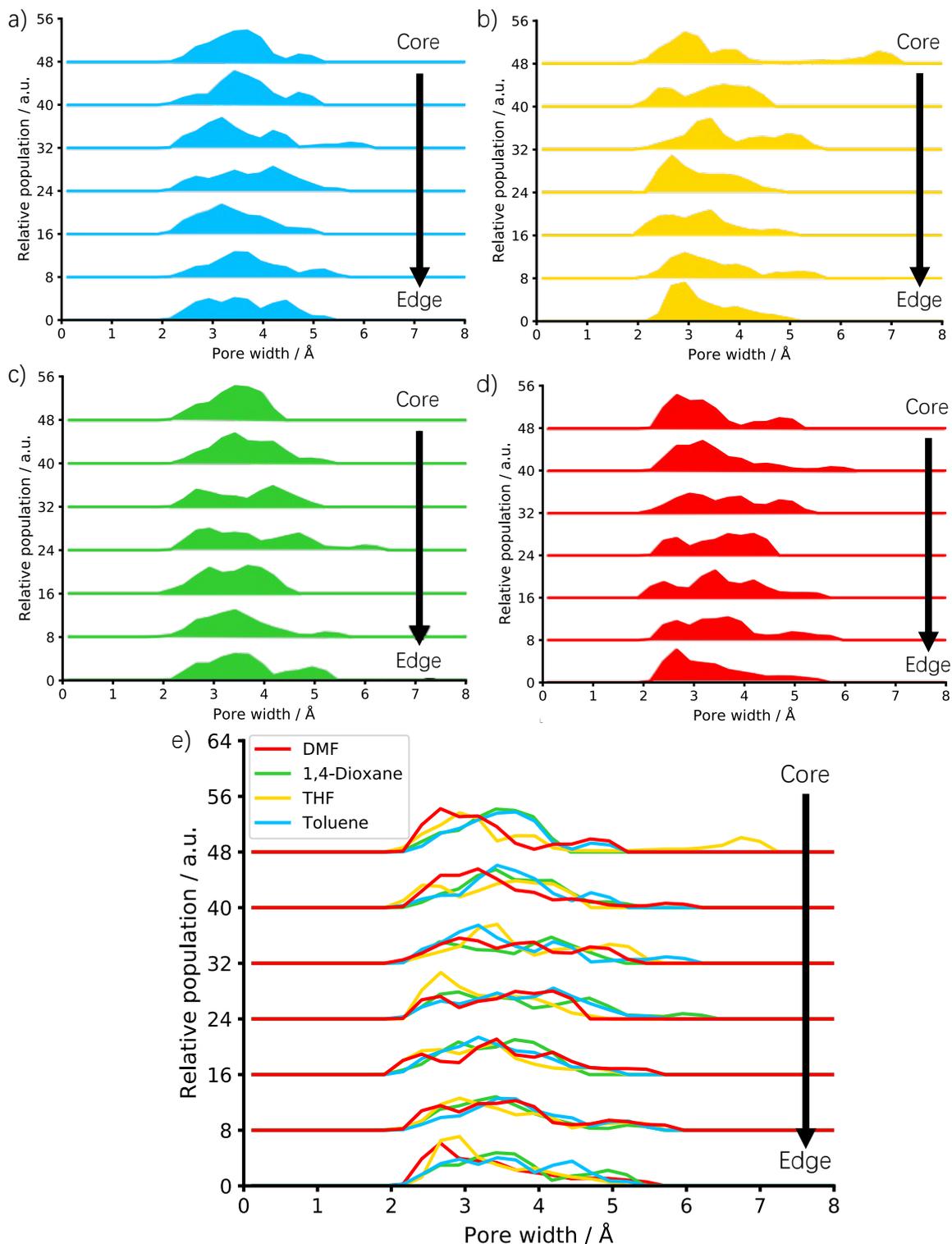


Figure 3.38: Pore size distributions acquired using Poreblazer 4.0¹⁶² for the sampled fragments of the Solvent/Monomer Phase Interface Systems after network generation, desolvation, homocoupling, workup and cell equilibration. The bottom stack represents the edge of the fragment, and the top stack represents the centre. a) Toluene, b) THF, c) 1,4-dioxane, d) DMF, e) overall. Key: DMF – red, 1,4-dioxane – green, THF – yellow, toluene – blue. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

This result was rationalised by the suggestion that after the network generation, it would be expected that a decreasing solvent polarity would lead to a higher concentration of solvent within the reaction mixture and increased pore sizes throughout for all of the solvents studied. However, an increase in pore size would lead to increasing flexibility within the polymer framework as the CMP-1 linkers are held further apart, reducing the van der Waals intermolecular forces available within the structure.³⁸ Upon desolvation, homocoupling, workup and cell equilibration, this would lead to more of a significant collapse in a more mesoporous and flexible framework such as the one synthesised using toluene, in comparison to those synthesised in more polar solvents. Therefore, THF appears to be the optimal solvent to maintain a network structure with slightly smaller pores than those synthesised in toluene during the network generation, which have a slightly higher degree of rigidity preventing as much collapse of the structure upon desolvation. This trend matches the experimental trend in mesoporosity within CMP-1 synthesised in different solvents (Figure 3.18), where THF has the highest quantity (71% mesoporosity), followed by toluene, 1,4-dioxane and DMF (67%, 36% and 29% mesoporosity, respectively).³⁸

To test this theory, slices through each of the xy , xz and yz planes were obtained for Model 1 of the Solvent/Monomer Phase Interface Systems after network generation using Materials Studio 5.0 (Figure 3.39).¹⁵⁸ These plotted the density of the solvent relative to that of the CMP-1 framework. In order to definitely confirm this, however, a new analysis technique which allows the porosity analysis of larger models than the very small sliced fragments studied here must be possible.³⁸

It would be expected that a higher degree of microporosity throughout the material, arising due to a higher solvent polarity giving a smaller quantity of solvent within the reaction mixture during the network generation, would lead to small pockets of solvent spread throughout the CMP-1 volume, whilst a lower degree of microporosity would correspond to larger, more distinct regions of pure solvent volume, and pure CMP-1 volume.³⁸ This is observed within the plots in Figure 3.39, where, as the solvent polarity increases going from toluene to THF to 1,4-dioxane to DMF, the regions of solvent relative to CMP transition from being large and reasonably well-separated to smaller and spread more consistently throughout the material. The larger regions of solvent observed in the less polar solvents such as toluene will correspond to greater regions of space that can be contracted upon desolvation to remove unstable pore voids within the material that would not occur experimentally, leading to a greater collapse of the structure upon desolvation.³⁸

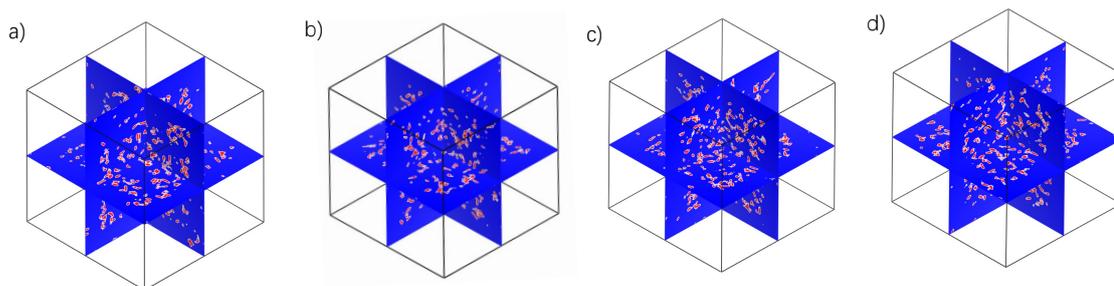


Figure 3.39: Plots of the density in the xy , xz and yz planes of the Solvent/Monomer Phase Interface Systems after network generation for Model 1, where the blue regions indicate the volume of solvent, and the red and white regions indicate the volume occupied by the CMP-1 material. a) Toluene, b) THF, c) 1,4-dioxane, d) DMF. Figure reproduced and adapted from reference 38, upon which this chapter is based, with permission from the Royal Society of Chemistry.

3.5 Conclusions

Conjugated microporous polymers are an amorphous class of hypercrosslinked polymer with a fully π -conjugated skeleton and permanent microporosity throughout due to inefficient packing in the solid state.^{5,20,223,224}

The solvent choice used in the reaction to synthesise CMPs is an important factor that is able to affect the porosity and morphology of the resulting polymer.^{180,235,241,260} This was first reported experimentally in 2010,²³⁵ where a series of CMP materials including CMP-1 were synthesised in four solvents of varying polarities. It was established that generally, CMPs exhibit a higher porosity in more polar solvents such as N,N-dimethylformamide compared to the original, less polar solvent choice of toluene. This was rationalised by the increase in the microporosity within the material with a more polar solvent. However, this was not explored much further until a recent report by Mollart and Trewin, which formed the basis of this chapter.³⁸

This chapter describes the modelling of CMP-1 in the same four solvents used in the original experimental synthesis report and mimicking the full synthetic conditions and catalytic mechanism. It was rationalised that the solvent has little influence on the microporosity of the central core of the CMP-1 particles due to the dense, non-porous nature of this region.³⁸ Instead, it was suggested that the differences in the percentages of meso- and microporosity that make up the total surface area differ depending on the miscibility of the growing polymer framework within each solvent, affecting the homogeneity and resulting pore structure of the material.³⁸

The Solvent/Monomer Phase Interface Systems studied, where a large cluster of CMP-1 monomers and catalyst was surrounded by solvent molecules, allowed a relatively simple way to analyse the material pre-synthesis. This

method established which was the reaction solvent choice that would correspond to the highest degree of microporosity within the framework pre-synthesis. Using computational approaches to assess the optimal reaction conditions pre-synthesis reduces the need for costly experimental solvent screening processes. This is also an attractive prospect from a sustainability angle, and allows the possibility for the design of new CMPs with intriguing properties in the future.³⁸

However, one of the main conclusions from this work was that models with a greater number of atoms are needed to fully assess the structure of the material at a larger scale. Current analysis techniques are time intensive for the number of atoms studied within the Solvent/Monomer Phase Interface Systems, and it would be difficult to use this approach for larger systems.³⁸

As a follow-up, the use of a hybrid coarse-grain methodology, where some atoms are replaced with 'grains' that reflect the symmetry and shape of the molecule, leading to a reduction in simulation time due to the smaller number of atoms modelled, was explored.¹⁰⁴ This enabled the successful modelling of CMP-1 with an increase in the cell volume of up to 64× the original models studied, and an overall speed-up factor of up to 44% when compared to the all-atom equivalent.¹⁰⁴ The future of CMP modelling will require this level of detail in order to fully rationalise structures on the mesoscale.¹⁰⁴

Chapter 4

Artificial synthesis of a hypercrosslinked polymer as a case study for rationalising the gas uptake properties of HCP materials

Everything starts as somebody's daydream.

Larry Niven

4.1 Summary

Hypercrosslinked polymers (HCPs), a type of microporous organic polymer often synthesised using Friedel-Crafts (FC) alkylation reactions, have a wide variety of potential applications. A HCP material synthesised experimentally from waste Styrofoam using an external crosslinker and loaded with 8.9 wt.% azobenzene showed varying pore size distributions (PSDs) and carbon dioxide uptakes depending on the reaction conditions employed and the resulting azobenzene isomer. The work presented in this chapter, in collaboration with colleagues from Edinburgh University and based on the work of Liu and co-workers,²⁶ demonstrates the use of simulations to rationalise these observed experimental properties.

This involved firstly creating representative structural models of the HCP using the Ambuild code,¹⁹ following the full catalytic mechanism. Following

this, the polymers were loaded with *trans*-azobenzene, and the switching between azobenzene isomers simulated using DL_POLY 4.0.²⁰⁸ This allowed for the comparison of the simulated PSDs to experiment. It was established that the observed differences in PSDs were due to the differing dimensions of *cis*- and *trans*-azobenzene. Simulated diffusion of CO₂ throughout the HCP loaded with *cis*- and *trans*-azobenzene demonstrated that the differences in CO₂ uptake were due to the increased quantity of micropores in the presence of *cis*-azobenzene, rather than the ability of CO₂ to diffuse throughout the material.²⁶

4.2 Introduction

HCPs are light-weight, low density materials that swell when immersed in solvent.²⁶⁹ Like the other classes of microporous organic polymers, they are made from abundant elements such as carbon and hydrogen and show good chemical and thermal stability. HCPs have applications in fields such as gas adsorption and storage, molecular separations, sequestration of dyes and toxic elements such as arsenic from water, and catalysis.^{5,24,265,266,270}

Unlike the other classes of microporous organic polymer, HCPs are prepared using a relatively novel coupling chemistry, generally requiring Friedel-Crafts (FC) alkylation in the presence of an anhydrous Lewis acid such as FeCl₃, AlCl₃, or SnCl₄, or a Scholl reaction utilising both a Lewis acid and protic acid.^{5,265,266} They can be synthesised from a wide variety of inexpensive starting materials, without the requirement for any expensive noble metal catalysts, which makes them appealing when considering reaction scale-up.^{5,265,266} There are generally three distinct methods to prepare HCPs: 1) post-crosslinking of a 'ready-made' polymer material, first reported by Davankov and co-workers;^{271,272} 2) direct one-step polycondensation, where bis(chloromethyl)-functionalised monomers are reacted directly to form the HCP,²⁷³ and 3) knitting of small aromatic molecules or, for example, linear styrene-based polymers, with an external crosslinker.²⁷⁴ In all cases, the crosslinking process helps to hold the polymer framework open by forming rigid covalent bonds between the polymer chains, preventing collapse upon desolvation and maintaining a low density material with permanent microporosity throughout.²⁶⁵

The first HCPs, based on polystyrene and styrene-divinylbenzene co-polymer precursors, were developed by Davankov and co-workers, giving them the name 'Davankov-type resins'.^{271,272} This method firstly involves synthesising the polymer precursor *via*, in the case of polystyrene, free radical polymerisation. The resulting polymer is then immersed within a chlorinated solvent

such as 1,2-dichloroethane (DCE), allowing the framework to swell and the polymer chains to untangle, with the newly created space being filled with solvent. At this point, rigid crosslinks form between the polymer chains to hold the framework in an open configuration, which is retained upon desolvation, with the space previously occupied by solvent becoming part of the interconnected pore structure within the material (Figure 4.1).^{265,275}

A variety of different crosslinking agents can be used to synthesise HCPs *via* the post-crosslinking method, such as chlorodimethyl ether (which is carcinogenic and therefore not an ideal material to use),²⁷⁶⁻²⁷⁸ tetrachloromethane,^{279,280} dichloroethylene (DCX),²⁸¹ 4,4'-bis(chloromethyl)biphenyl (BCMBP),²⁸¹ 1,3,5-tris(chloromethyl)-2,4,6-trimethylbenzene,²⁸¹ and di- and triiodoalkanes (Figure 4.2).²⁷⁵ Changing the length, rigidity and functionality of the crosslinking agent (*e.g.*, bifunctional or trifunctional) gives variety within the structure of the HCP itself. For example, a crosslinking agent with long, flexible chains and reduced functionality will lead to a smaller number of long crosslinks, which are likely to rearrange upon desolvation to give an efficiently packed, non-porous and dense structure.²⁶⁶

Many of the first HCPs, synthesised using the post-crosslinking method, were based on polystyrene or styrene-divinylbenzene co-polymers. Hradil and Králová reported post-crosslinking of styrene-divinylbenzene co-polymers such as poly-(divinylbenzene-covinylbenzyl chloride) (DVB-VBC) to give HCP-DVB-VBC, formed from a FC alkylation reaction in the presence of a tetrachloromethane crosslinking agent and a Lewis acid catalyst, with surface areas of up to $1000 \text{ m}^2 \text{ g}^{-1}$. The chloromethyl groups within the polymer structure react as internal electrophiles to form methylene bridges, which give additional crosslinking within the structure and enhance the microporosity (Scheme 4.1).^{265,279} Li and co-workers established that the pore size distribution of the HCP-DVB-VBC materials could be tuned by varying the quantity of divinylbenzene within the structures, with the polymer becoming more microporous with increasing divinylbenzene content, until a totally microporous polymer was reached when the divinylbenzene content exceeded 7%.²⁸²

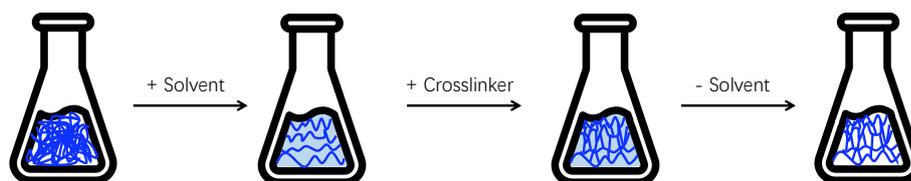


Figure 4.1: Cartoon illustration demonstrating the post-crosslinking of a polymer into a hypercrosslinked polymer material. Figure inspired by reference 275.

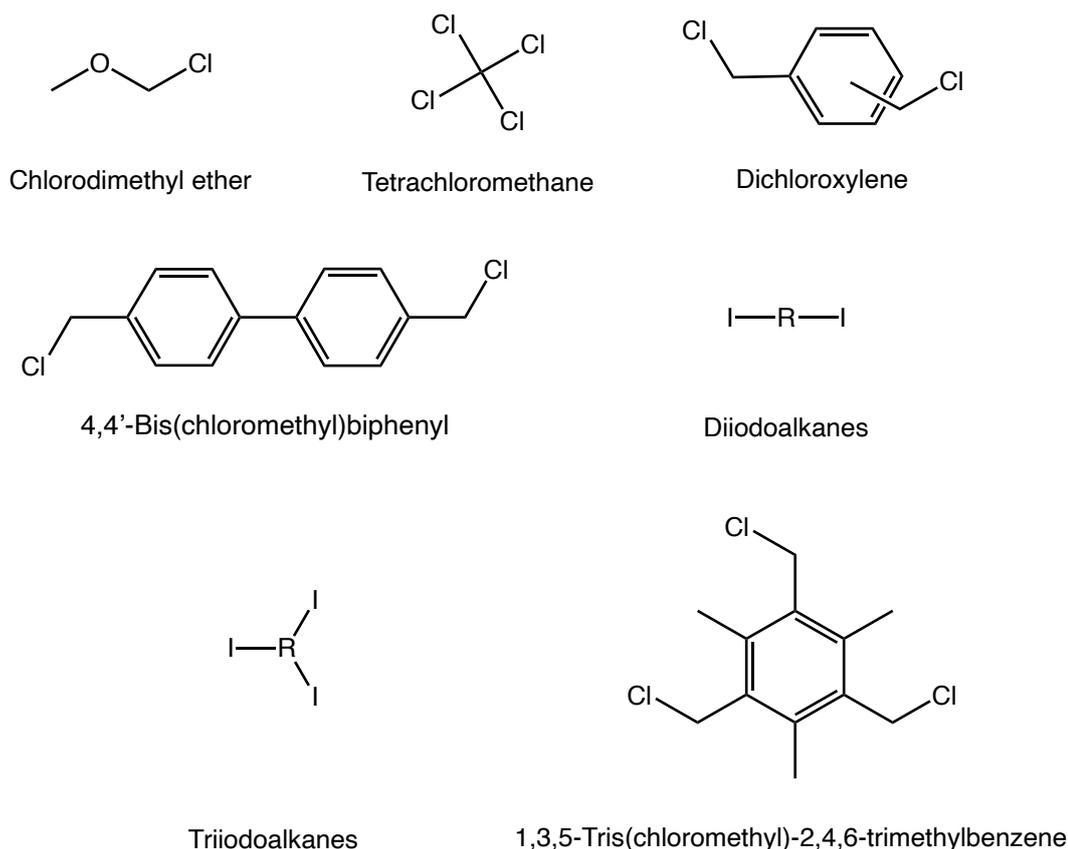
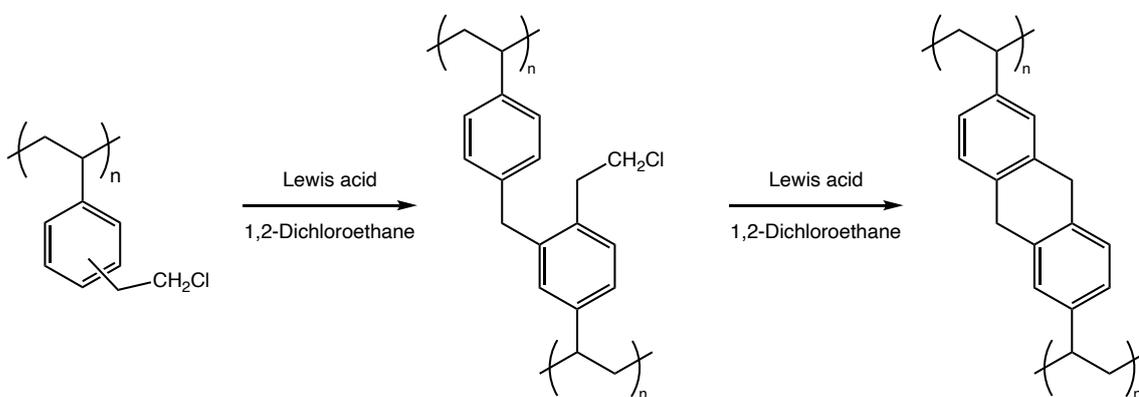


Figure 4.2: Chemical structures of various crosslinking agents that may be used to synthesise HCPs via the post-crosslinking method.^{275–281}



Scheme 4.1: Reaction scheme showing the post-crosslinking of DVB-VBC to form HCP-DVB-VBC. Scheme reprinted (adapted) with permission from J.-H. Ahn, J.-E. Jang, C.-G. Oh, S.-K. Ihm, J. Cortez and D. C. Sherrington, *Macromolecules*, 2006, **39**, 627–632. Copyright 2006 American Chemical Society.

As well as there being a number of choices in the crosslinking agent used in the reaction, the Lewis acid catalyst required for the FC alkylation, solvent, monomer ratios, and crosslinking reaction time were optimised by Ahn and co-workers, leading to a large difference in the obtained surface areas, ranging from 300–2090 m² g⁻¹.²⁸³ They established that the most efficient com-

bination was to use an FeCl_3 catalyst and DCE solvent, with the crosslinking itself occurring rapidly. A 2 mol% DVB-VBC polymer precursor showed an extensive microporous structure (surface area = $1200 \text{ m}^2 \text{ g}^{-1}$) after 15 minutes, with the surface area increasing to $>1800 \text{ m}^2 \text{ g}^{-1}$ after two hours, and a relatively small increase to the maximum surface area of $2090 \text{ m}^2 \text{ g}^{-1}$ after 18 h.^{266,283}

Whilst the post-crosslinking method could be successfully used to synthesise HCPs with surface areas in the region of $2000 \text{ m}^2 \text{ g}^{-1}$,²⁸³ the number of distinct frameworks was limited by the requirements for the polymer precursor to be synthesised before the crosslinking reaction could occur, taking extra time in the synthesis. The limited functionalities available within the monomers that would be suitable for both the free radical polymerisation reaction to synthesise the polymer precursor, and the FC alkylation crosslinking reaction also reduced the number of distinct frameworks. This has led to the vast majority of the HCPs synthesised using the post-crosslinking method being polystyrene- or divinylbenzene-based.^{266,270} HCPs can also be synthesised directly from the monomers using a direct one-step polycondensation reaction, which was revolutionised by Wood and co-workers after an initial exploration yielded polymers with low porosity.^{273,284}

The direct one-step polycondensation reaction involves reacting monomers with bis(chloromethyl) functional groups in a one-pot condensation polymerisation in the presence of the Lewis acid catalyst. The chloromethyl groups act as internal electrophiles to form methylene bridges between adjacent phenyl rings, with HCl lost as a by-product of the elimination reaction.^{265,266,270}

Wood and co-workers first demonstrated this technique with a number of different monomers: DCX (all isomers), BCMBP and 9,10-bis(chloromethyl)anthracene (BCMA).²⁷³ The resulting HCPs had surface areas of up to $1904 \text{ m}^2 \text{ g}^{-1}$ and were largely microporous, with the surface area and pore size of the materials being tailored, as in the case of the post-crosslinked HCPs, by varying the monomer ratios and reaction conditions, with the maximum surface area arising from a 1:3 co-polymer of *p*-DCX and BCMBP. Of the studied monomers, the highest surface area for the homo-polymerised HCP made from one monomer type only was observed for HCP-BCMBP (surface areas: HCP-*p*-DCX = $1391 \text{ m}^2 \text{ g}^{-1}$, HCP-BCMBP = $1874 \text{ m}^2 \text{ g}^{-1}$, HCP-BCMA = $921 \text{ m}^2 \text{ g}^{-1}$). This was postulated as being due to BCMBP acting as a template to assemble neighbouring aryl rings into a *para* conformation, which increases the accessible volume within the material.²⁷³

In addition to acting as the internal electrophile for a self-condensation reaction such as those described by Wood and co-workers,²⁷³ bis(chloromethyl) groups can be used as external crosslinking agents for direct one-step poly-

condensation reactions of co-monomers with different functionality (Figure 4.3).^{266,285-287}

Hydroxymethyl compounds can also be used in place of bis(chloromethyl) to make HCPs by this method, with Luo and co-workers reporting the synthesis of two HCPs from 1,4-benzenedimethanol (BDM) and benzyl alcohol (BA) monomers.²⁸⁸ The FC-catalysed self-condensation reactions gave microporous polymers in both cases, with only a small decrease in the surface area for HCP-BA, with its monohydroxymethyl monomer, compared to HCP-BDM, with a bishydroxymethyl monomer (surface areas: HCP-BDM = 847 m² g⁻¹, HCP-BA = 742 m² g⁻¹). This demonstrated for the first time that mono-functionalised monomers could be used to synthesise porous HCP materials using the direct one-step polycondensation technique (Scheme 4.2).²⁸⁸ The use of hydroxymethyl functional groups in place of the chloromethyl is certainly advantageous from an environmental perspective due to the release of stoichiometric water as a by-product in place of hydrochloric acid.²⁷⁰

Whilst the FC alkylation reaction is the most well-known synthetic route used to make HCPs, it is not the only method that can be used. Li and co-workers reported the use of the Scholl reaction to synthesise a variety of HCP materials from monomers containing an extensive number of different functional groups. Scheme 4.3 gives the generic reaction scheme for a Scholl reaction, and Figure 4.4 gives some example monomers that have been used to prepare HCPs using the Scholl method.²⁸⁹ The Scholl reaction, which involves the formation of a new aryl-aryl bond in place of two aryl-hydrogen bonds in the presence of a Lewis acid and protic acid, has a high activity which makes it suitable for a large range of monomers, including those with high or low electron density, a range of pH environments, fused rings, and heterocycles.^{265,289} As a result of the large number of monomers that can be used for this reaction, surface areas for the resulting HCPs ranged from 636-1421 m² g⁻¹, and the properties of the polymers, such as the CO₂ adsorption capacity, also differed depending on the monomer functionality, with HCPs containing alkaline functional groups having higher CO₂ adsorption capacities.^{265,289}

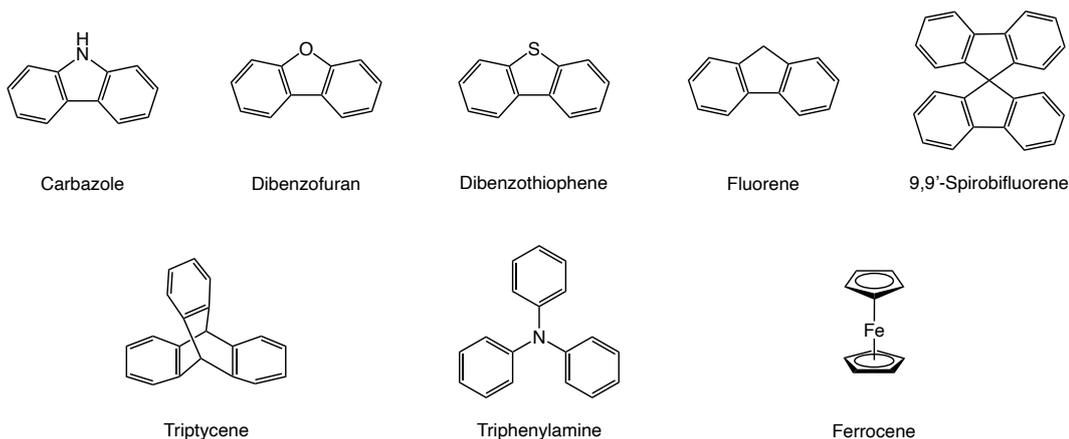
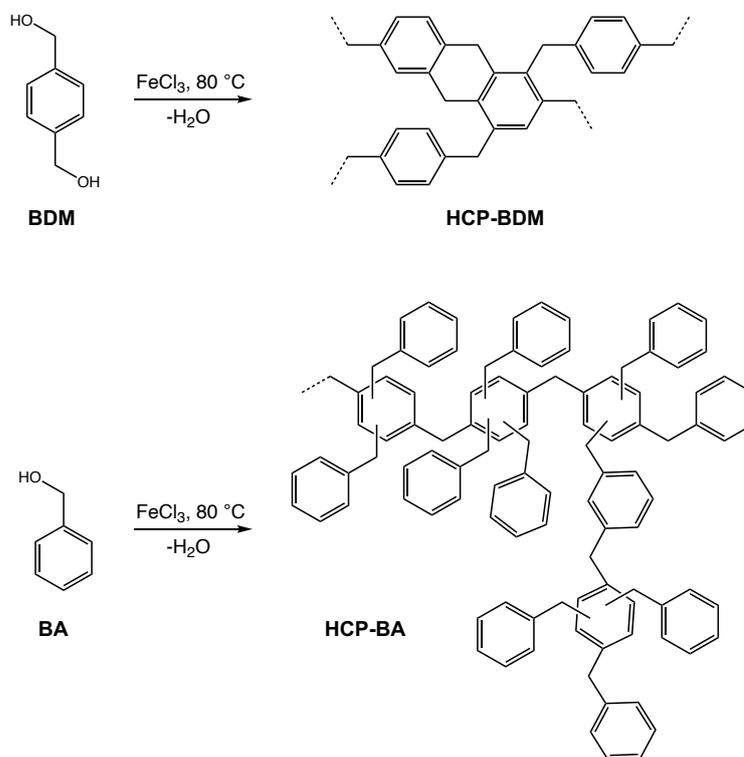
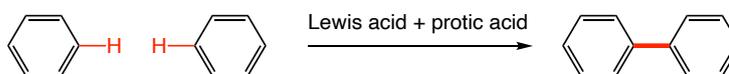


Figure 4.3: Chemical structures of various co-monomers that have been used to synthesise HCPs *via* a direct one-step polycondensation reaction with an external bis(chloromethyl) crosslinking agent. Structures reported within references 285–287. Figure reproduced and adapted from reference 266 with permission from the Royal Society of Chemistry.



Scheme 4.2: Reaction scheme showing the direct one-step polycondensation of HCP-BDM (top) and HCP-BA (bottom). Figure reproduced and adapted from reference 288 with permission from the Royal Society of Chemistry.



Scheme 4.3: Generic reaction scheme for the Scholl reaction used to prepare HCPs. Figure inspired by references 265, 266, and 289.

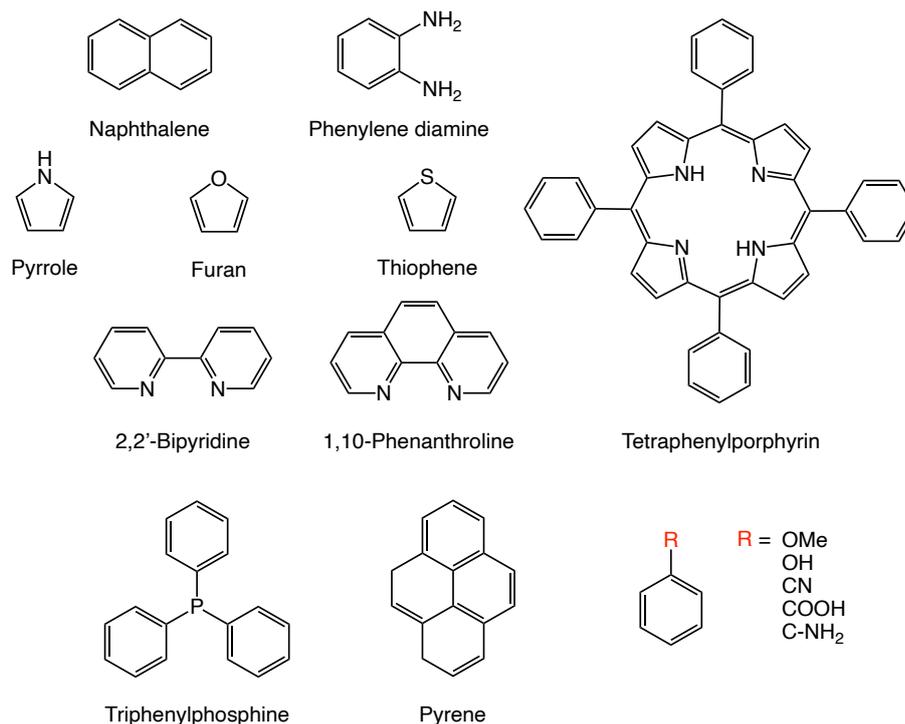
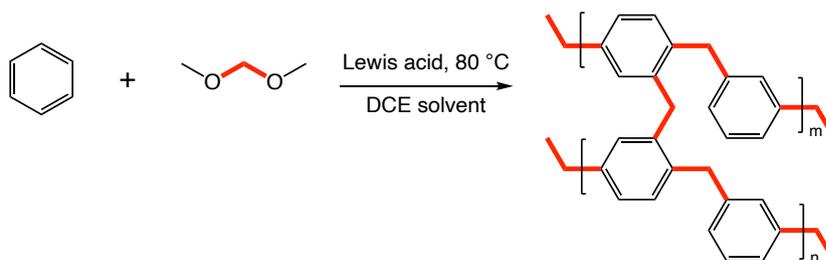


Figure 4.4: Example monomers used to synthesise HCPs using the Scholl method. Figure inspired by references 265 and 289.

The third synthetic method used to make HCPs is the knitting of aromatic monomers in the presence of solvent, an external crosslinker and a Lewis acid catalyst, *via* a simple one-step FC reaction, first reported by Li and co-workers.^{265,266,270,274} The reaction scheme is given in Scheme 4.4 and example monomers used to synthesise HCPs *via* the knitting method are given in Figure 4.5. Due to the presence of the external crosslinker, typically formaldehyde dimethyl ether (FDA), the monomers themselves do not require any specific functionality to form the polymer and therefore a large number of monomers can be used to prepare HCPs *via* this method, which is advantageous compared to the direct one-step FC polycondensation reaction, whilst there is no requirement for a polymer precursor, making this method superior to the post-crosslinking reaction also.²⁶⁵ This is an important feature when considering materials design, as in order to discover new materials with enhanced properties, it is essential to allow flexibility within the number of available starting materials.



Scheme 4.4: Example reaction scheme used to synthesise HCPs using the knitting method. Figure reproduced from reference 266 with permission from the Royal Society of Chemistry.

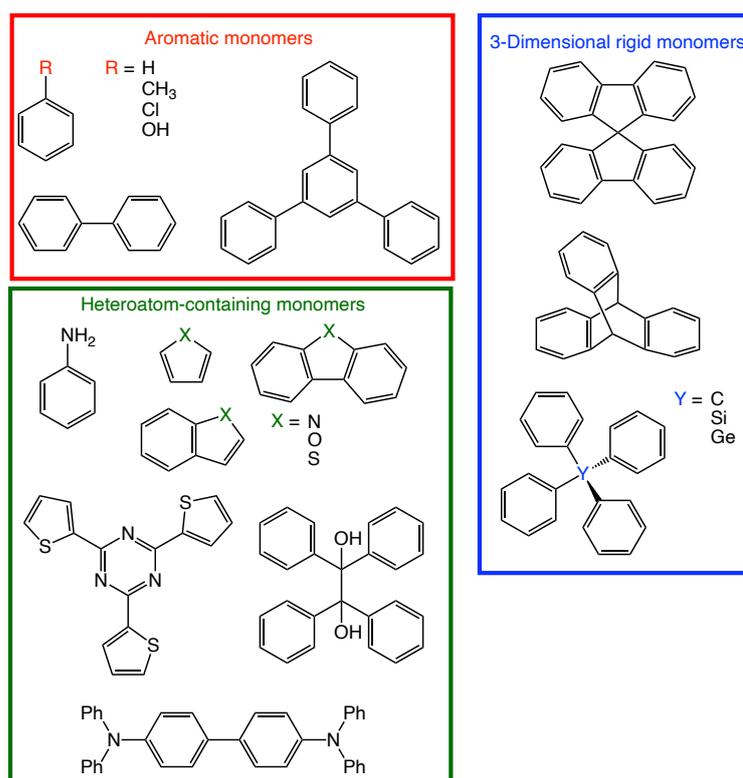


Figure 4.5: Example monomers used to prepare HCPs *via* the knitting reaction.^{274,290–298} Figure reproduced from reference 266 with permission from the Royal Society of Chemistry.

It has been hypothesised that the mechanism to synthesise HCPs using the knitting method occurs by a catalyst molecule reacting with an FDA crosslinker to lower the strength of the binding interaction between the methoxy group and central carbon atom within the FDA. These intermediate FDA carbocations which reside within the DCE solvent then form bonds to the aryl rings in the monomers with the loss of methanol as a by-product, and finally, the added methoxymethyl groups that have bonded to the aryl rings then further react with neighbouring aryl rings to form rigid methylene linkages, giving the HCP.²⁶⁶

The knitting reaction requires milder synthetic conditions in comparison to the post-crosslinking and direct one-step FC polycondensation methods, mak-

ing it appealing for scale-up.²⁶⁵ Additionally, without the hydrochloric acid by-products generated by the chloromethyl group used to make HCPs *via* FC alkylation within the direct one-step polycondensation method, this technique is more environmentally friendly.²⁷⁴

The porosity properties of the resulting HCPs can be tuned by varying the monomer to crosslinker ratio, where in the case of a HCP knitted from a benzene monomer, the maximum surface area obtained of 1391 m² g⁻¹ occurred with a monomer : FeCl₃ catalyst : FDA crosslinker ratio of 1 : 3 : 3.²⁷⁴ Also, whilst it is not necessary to have monomers with specific functionality to make HCPs using the knitting method, additional functionality can be added to the polymer by careful choice of the utilised monomers.²⁶⁵ For example, when phenol was used as the monomer, an increase in CO₂ uptake was observed in comparison to the HCP synthesised from biphenyl monomers (CO₂ uptakes: HCP-phenol = 9.4 wt.%, HCP-biphenyl = 6.9 wt.%).²⁷⁴

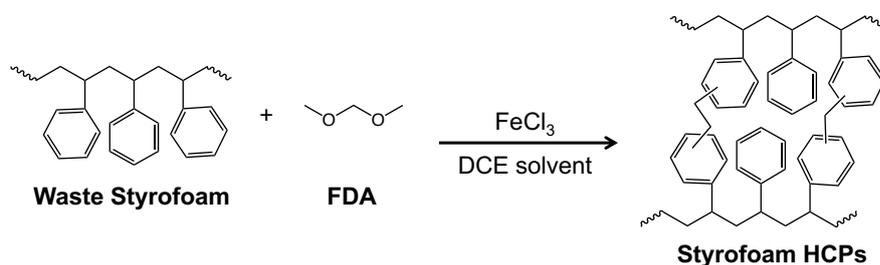
The knitting approach can also be used to efficiently prepare HCPs from linear aromatic polymer chains (such as polystyrene). Ratvijitvech and co-workers reported the synthesis of HCPs prepared by the knitting of linear polystyrene,²⁹⁹ with the surface areas of the resulting HCPs increasing with an increasing degree of polymerisation within the polystyrene precursor. Yang and co-workers then further enhanced the field by knitting polystyrene precursors within a high concentration of solvent to form a solution-processable polystyrene network that could be used as a thin film.³⁰⁰

It has been widely reported that perhaps the most successful of the three methods used to synthesise HCPs is the knitting with an external crosslinker due to its mild and facile synthetic conditions and wide flexibility in the availability of different monomer building blocks.^{265,266,270} Dong and co-workers reported the synthesis of recycled HCPs made from waste Styrofoam (the commercial name for polystyrene) using the knitting with external FDA crosslinker method (Scheme 4.5).²⁴ The reported HCPs showed an initial increase in surface area and pore volume on increasing the ratios of monomer : FeCl₃ catalyst : FDA crosslinker from 1 : 1 : 1 to 1 : 5 : 5, resulting in an increased quantity of crosslinking leading to a larger accessible pore volume within the material, as shown in Figure 4.6. However, on increasing the ratio further to 1 : 8 : 8, the surface area decreased, which was proposed as being due to the extra crosslinking giving rise to a reduction in the accessible pore volume.²⁴ The HCPs, which, due to their recycled precursor materials are able to reduce excess plastic production, showed strong adsorption of a selection of cationic and anionic dye materials due to the π -conjugation throughout the framework and its ability to interact with the various dyes. Additionally, they showed high As(V) adsorption, particularly within an acidic environment

due to the additional chemical stability within the HCP, suggesting that these materials could be suitable for water purification. These materials offer an exciting future prospective for low-cost, chemically stable polymer materials that are applicable to a number of different areas.²⁴

HCPs are also able to adsorb guest molecules within the pore structure of the material, which can be exploited to give the HCPs additional properties. One example guest molecule that can be incorporated within the HCP is azobenzene, a classic photoswitch material that is able to switch between its *trans* and *cis* isomers upon application of light and can be incorporated into the HCP *via* physical steeping.^{26,301} Photoswitch materials like azobenzene are advantageous due to offering a reduction in the energy required for adsorption and desorption processes.³⁰² They have been investigated for use within phase-change materials due to their ability to store thermal energy within the chemical bonds of the less stable isomer. Then, upon converting back to the more stable isomer, the stored energy is released as heat.³⁰³

Figure 4.7 shows a summary of the properties of each of the two isomers of azobenzene. The *trans* isomer is more stable and planar with a lack of dipole moment. The less stable *cis* isomer, accessed by application of UV irradiation (365 nm wavelength) to the *trans* isomer, is more strained due to its angular nature (molecular strain energy = 52.4 kJ mol⁻¹).³⁰³⁻³⁰⁹ *cis*-Azobenzene has a dipole moment of 3.0 D, which is potentially able to interact with the quadrupole on adsorbed CO₂ molecules, leading to an increase in uptake.^{25,26,310,311} Reconversion of the *cis* isomer to the *trans* can be accessed by exposure to visible light (> 400 nm wavelength).³⁰³ Converting between the two isomers also involves a change in the physical dimensions of the molecule itself, with the *trans*-azobenzene molecule being larger in dimension than the *cis* isomer (dimensions: *trans*-azobenzene = 9 Å, *cis*-azobenzene = 5.5 Å).³¹²



Scheme 4.5: Reaction scheme used by Dong and co-workers to synthesise Styrofoam-based HCPs using the knitting method. Figure reprinted (adapted) with permission from X. Dong, A. Akram, B. Comesaña-Gándara, X. Dong, Q. Ge, K. Wang, S.-P. Sun, B. Jin and C. H. Lau, *ACS Appl. Polym. Mater.*, 2020, **2**, 2586–2593. Copyright 2020 American Chemical Society.

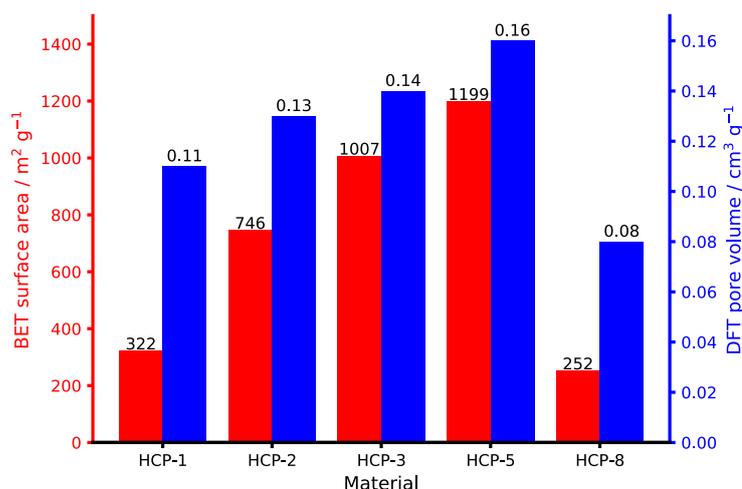


Figure 4.6: Experimental BET surface areas (red) and density functional theory (DFT) pore volumes (blue) of the Styrofoam-based HCP materials reported by Dong and co-workers. Data reprinted (adapted) with permission from X. Dong, A. Akram, B. Comesaña-Gándara, X. Dong, Q. Ge, K. Wang, S.-P. Sun, B. Jin and C. H. Lau, *ACS Appl. Polym. Mater.*, 2020, **2**, 2586–2593. Copyright 2020 American Chemical Society.

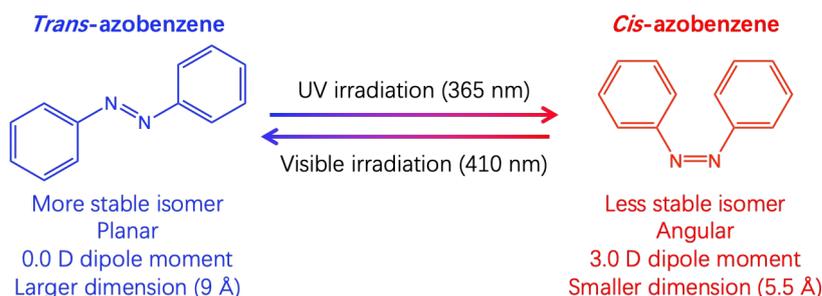
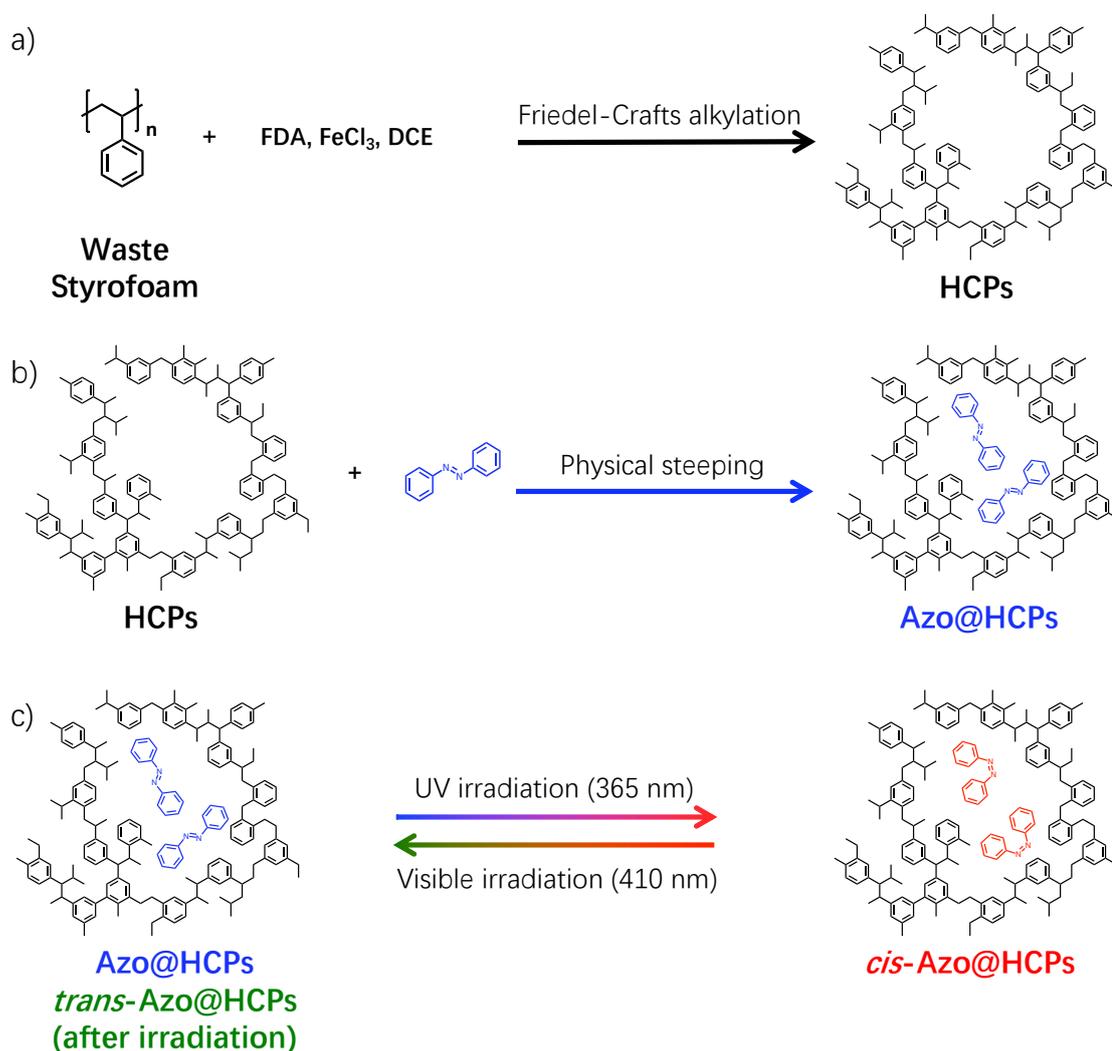


Figure 4.7: Summary of the properties of azobenzene.^{26,301,310–312}

This chapter describes the artificial synthesis of a HCP synthesised experimentally from waste Styrofoam using the knitting with external crosslinker method. Azobenzene was later added into the pores of the material *via* physical steeping. The artificial synthesis of the HCP using Ambuild¹⁹ followed the experimental Friedel-Crafts alkylation reaction.²⁶ Once generated, the simulated HCP models were then loaded with *trans*-azobenzene. The switching between the *trans*- and *cis*-azobenzene isomers within the HCP was then simulated using DL_POLY 4.0²⁰⁸ to understand the effect of the azobenzene isomer on the pore size distribution and CO₂ uptake. This work was undertaken in collaboration with Dr Aotian Liu, Professor Xianfeng Fan and Dr Chen Hon Lau at Edinburgh University, who obtained the experimental data. The author thanks these collaborators for the use of their data within this thesis to support the simulation results.²⁶

4.3 Methodology

In both the artificial and real-world systems, the synthesis involves three distinct steps. These are summarised in Scheme 4.6. Firstly, the HCP (referred to throughout as 'HCPs') is synthesised from waste Styrofoam (the commercial name for polystyrene). Secondly, azobenzene was added into the pores of the HCP as a guest to give 'Azo@HCPs'. Finally, irradiation experiments were undertaken to switch between the two isomers of azobenzene (Azo@HCPs to *cis*-Azo@HCPs by application of UV light, followed by the conversion of *cis*-Azo@HCPs to *trans*-Azo@HCPs using visible light).²⁶



Scheme 4.6: Reaction scheme showing the different materials discussed in this chapter. a) Synthesis of HCPs, b) HCPs to Azo@HCPs, c) Azo@HCPs to *cis*-Azo@HCPs, and d) *cis*-Azo@HCPs to *trans*-Azo@HCPs. Scheme adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

Experimentally, the HCPs were synthesised as reported by Dong and co-workers,²⁴ where the waste Styrofoam material was dissolved within the DCE solvent. The FeCl_3 Lewis acid catalyst and FDA crosslinker were later added to the solution. Following this, the reaction mixture was refluxed for 24 h at 80 °C, stirring throughout. The resulting HCP was then filtered and washed with chloroform, methanol, acetone and water, then dried under vacuum for 12 h at 120 °C.²⁶

The Styrofoam polymer was simulated by firstly seeding one styrene monomer into the centre of an Ambuild simulation cell of size $100 \text{ \AA} \times 100 \text{ \AA} \times 100 \text{ \AA}$. Then, a loop of *growBlocks*, geometry optimisation (integration timestep = 0.0001 ps, number of optimisation cycles = one million, van der Waals cut-off distance = 10 Å), and NVT molecular dynamics (integration timestep and van der Waals cutoff distance as above, total simulated MD time = 4.5 ns, HOOMD-blue temperature factor = 55.0) steps was undertaken.^{193,194} The use of *growBlocks* steps to bond additional styrene monomers to those existing within the simulation cell is appropriate to model the free radical mechanism used to synthesise Styrofoam as the radical combination step will be favourable and occur rapidly.²⁶ Once the polymer had started to form, five additional styrene monomers were seeded into the central region of the simulation cell to increase the structural sampling within the material, preventing the formation of a linear polymer chain, and increase the number of free end groups.²⁶ The Ambuild input file is given in Appendix B Figure B.1. Please refer to Chapter 2 section 2.4.1 for a full explanation of the parameters and terms used within the Ambuild code.

Following the formation of the Styrofoam, the next step was to crosslink the structure using FDA, mimicking the real-world hypercrosslinking conditions. Whilst experimentally the FeCl_3 Lewis catalyst reacts with the FDA to form a carbocation which makes it more electrophilic and hence more reactive, artificially, an approximation was made so that this step had already occurred. This is reasonable due to the favourability of this step to occur, meaning that the FDA carbocation could be directly bonded to the available end groups within the Styrofoam polymer using another loop of *growBlocks*, followed by geometry optimisation (same parameters as above) and NVT MD (same parameters as above apart from the temperature factor, which was reduced to 1.0, and a total simulated MD time of 20 ns).^{193,194} Two distinct bonding rules were required for the hypercrosslinking simulation: the first allowed the FDA carbocations to bond to the Styrofoam polymer, and the second allowed the FDA to link two polymer chains together using a *zipBlocks* step with predetermined bond length and bond angle margins of 10 Å and 90°, respectively (Appendix B Figure B.2). Finally, the loop of *grow FDA*, geometry optimisation, MD and *zipBlocks* was repeated, this time incorporating 0.925

ps NPT MD steps to allow the cell volume to decrease and incorporate additional crosslinking within the structure, with a very small integration timestep of 0.00000005 ps to allow the forces within the system to equilibrate as the dimensions decreased.^{193,194} In addition, the margins of the *zipBlocks* steps were reduced to 4.5 Å and 60°, respectively, to take into account the reduction in cell volume and closer packing of the framework structure (Appendix B Figure B.3). Five repeat polymer structures, referred to as models A–E, were obtained to increase the sampling of the polymer.²⁶ The structures of models A–E are given in Figure 4.8.

The second step in the experimental synthesis was the incorporation of 8.9 wt.% azobenzene into the pores of the HCP material *via* physical steeping. This occurred by dissolving the azobenzene into ethanol, then immersing the HCP into the azobenzene solution within a sealed vessel for 12 h. The resulting Azo@HCPs product was filtered and washed with ethanol, then air-dried for one day.²⁶

To mimic the experimental loading of azobenzene into the HCP, *trans*-azobenzene building blocks were randomly seeded into each HCP model. It was established that twenty-four *trans*-azobenzene molecules equated to the experimental loading of 8.9 wt.%, and so twenty-four building blocks were seeded into the cell. This was followed by geometry optimisation and NVT MD steps (same parameters as for the Styrofoam formation, Appendix B Figure B.4).^{193,194}

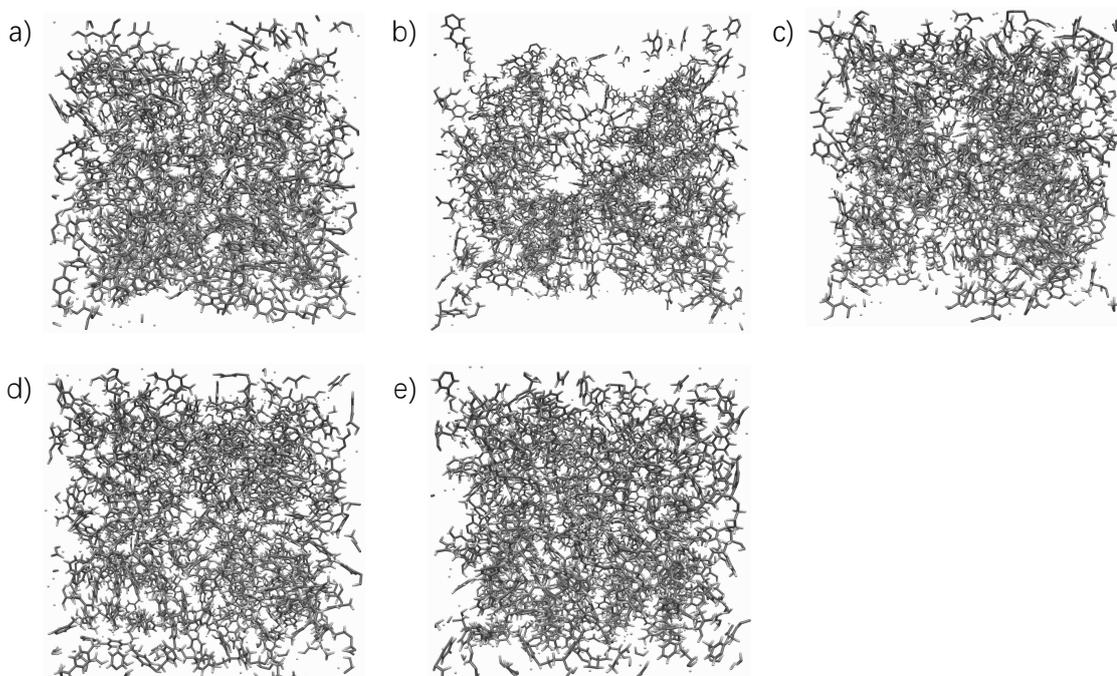


Figure 4.8: Structures of models A–E utilised within this chapter. a) model A, b) model B, c) model C, d) model D, e) model E. Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

The loading process was repeated three times for each of the five pristine HCP models to again increase structural sampling. The azobenzene loading process is not likely to be homogeneous throughout in the real-world system. It may be the case, for example, that the distribution of azobenzene is lower in the micropore region than in the mesopore region. For simplicity, the assumption was made that the loading could be considered homogeneous for the model systems.²⁶

The final stage of the experiment is the irradiation. Azo@HCPs (where the azobenzene was in its more stable *trans* isomer) was irradiated with 365 nm UV light to give *cis*-Azo@HCPs, resulting in a *trans*- to *cis*-isomerisation, and the reverse reaction involved irradiating *cis*-Azo@HCPs with 410 nm visible light to give *trans*-Azo@HCPs. For clarity, the system after loading and before irradiation was named differently to the system post-visible light irradiation, despite both having azobenzene present in its *trans* isomer. Each irradiation cycle occurred for 30 minutes, with a distance between the irradiation source and material of 1 cm.²⁶

There are two proposed mechanisms by which an azobenzene molecule can switch between isomers: rotation of the C-N=N-C dihedral, or by inversion of one of the phenyl rings in the plane of the molecule.³¹³ The inversion pathway typically occurs at lower energy for azobenzene compared to the rotation pathway, suggesting that this is the preferred mechanism of isomerisation for azobenzene.³¹³⁻³²⁰ The barrier to isomerisation of a single azobenzene molecule was calculated by running a scan of the C-N=N-C dihedral of the azobenzene molecule in Gaussian 09,²¹² using the CAM-B3LYP/cc-pVTZ model chemistry.^{96,321} A Grimme dispersion was applied to better describe the interactions within the azobenzene molecule and be more comparable to further Gaussian calculations described later in this chapter.³²² The starting point of the scan was the most stable, optimised *trans*-azobenzene isomer, which had a C-N=N-C dihedral angle of 180°. Steps were then taken to decrease the C-N=N-C dihedral angle in 10° increments, optimising the structure at each step. Once the dihedral angle was in the region of 60°, optimising the geometry to ground state structures proved challenging, and so it was necessary to obtain the remaining structures using other methods, as described in the following paragraph.

The structures with C-N=N-C dihedral angles between 60° and 30° were obtained by using smaller increments of 2° between steps. The structures with dihedral angles of 20° and 10° were obtained by starting a scan of the dihedral from the *cis*-azobenzene isomer, with a dihedral angle of 0°, and increasing the dihedral angle in 10° increments, optimising the structure at each step. This approach was determined to be robust by calculating the energy of the

azobenzene molecule with a C-N=N-C dihedral angle of 30° both from sequentially decreasing the C-N=N-C dihedral angle of the *trans*-azobenzene molecule, and by sequentially increasing the C-N=N-C dihedral angle of the *cis*-azobenzene molecule, with both giving the same energy in Hartree to 8 decimal places.

The barrier to isomerisation in azobenzene was also calculated *via* the inversion pathway by Wazzan,³¹³ using the B3LYP/6-31G(d,p) model chemistry.^{79,85,90,91,93,94,313} This was done by running a scan of the C-N=N angle of the azobenzene molecule, starting from the *trans*-azobenzene geometry and scanning the C-N=N angle in increments of 10° from 100-240°. ³¹³ Due to the symmetrical nature of azobenzene, there was no difference between the α - and β -inversion, meaning that either C-N=N angle could be chosen for the inversion scan.

The relative energy of a single azobenzene molecule is plotted against the C-N=N-C dihedral angle in Figure 4.9a. It can be seen that as expected, the *trans* isomer of azobenzene, which has a C-N=N-C dihedral angle of 180°, was approximately 50 kJ mol⁻¹ more energetically favourable than the *cis*,^{313,320} with a C-N=N-C dihedral angle of 0° (Figure 4.9 a). The transition state occurring between these two isomers occurred when the dihedral angle was in the range of 90-100°. When comparing this to the plot of relative energy against the C-N=N angle and C-N=N-C dihedral angle obtained by Wazzan (Figure 4.9b),³¹³ it could be seen that the energies of the *trans*- and *cis*-azobenzene from the inversion pathway remained relatively constant with those obtained from the rotational pathway. The energy of the inversion pathway transition state had decreased compared to that of the rotation pathway, an indication of the lower energy penalty to switching between isomers using the inversion pathway, as reported in the literature.³¹³⁻³²⁰ The small differences in the energies obtained by Wazzan in comparison to those obtained here are a result of the difference in model chemistry utilised (this work - CAM-B3LYP/cc-pVTZ,^{96,321,322} Wazzan - B3LYP/6-31G(d,p)).^{79,85,90,91,93,94,313}

Computationally, the irradiation steps were simulated using DL_POLY 4.0.²⁰⁸ Within these simulations, the energy at a given dihedral angle ϕ was calculated using the cosine equation (Equation 4.1),

$$U(\phi) = A[1 + \cos(m\phi - \delta)] \quad (4.1)$$

where $U(\phi)$ is the potential energy of the angle of interest ϕ , A is the peak height, m is the multiplicity, and δ is the equilibrium angle (*trans*-azobenzene = 0°, *cis*-azobenzene = 180°).²⁰⁸ Here, the *trans*-azobenzene isomer has a minimum potential energy at 180° and a maximum potential energy at 0°, whilst the *cis*-azobenzene isomer has a minimum potential energy at 0° and a

maximum potential energy at 180° , as seen in Figure 4.10.

Whilst these equilibrium angles seem counter-intuitive considering where the minima and maxima lie on the potential energy surface for each isomer, these are the angles required within the cosine equation to give the correct potential energy surface in each case. The angles can be rationalised by considering that in the *trans*-azobenzene isomer, the largest group bonded to each nitrogen atom (the phenyl ring) point away from the largest group bonded to the other nitrogen atom, meaning that the planes forming the dihedral will cancel, giving an equilibrium angle of 0° . In contrast, in *cis*-azobenzene, the largest group bonded to both nitrogen atoms is pointing in the same direction, meaning that the planes comprising the dihedral are unable to cancel, giving the corresponding equilibrium angle of 180° (Figure 4.11).

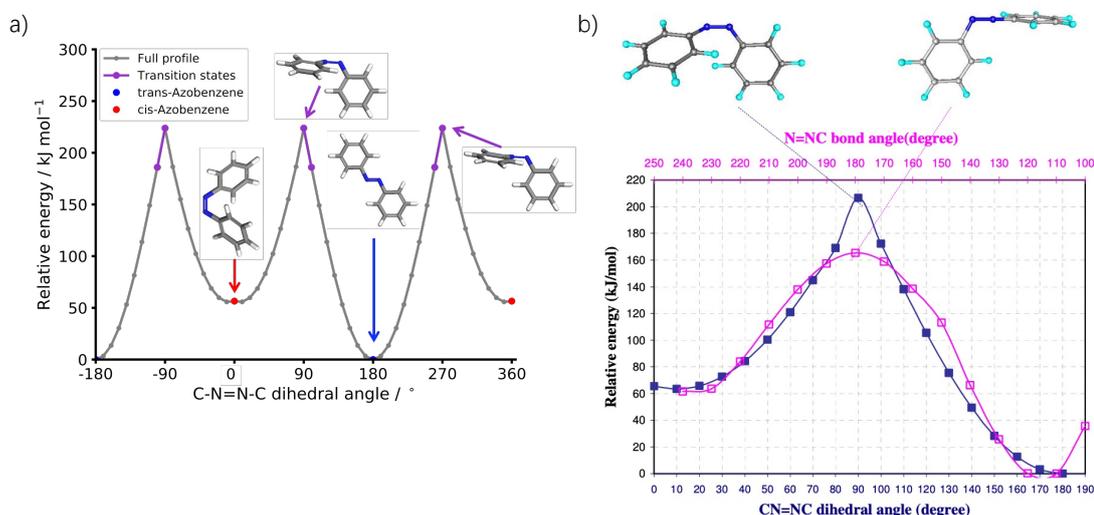


Figure 4.9: a) Plot of the barrier to isomerisation of a single azobenzene molecule *via* the rotation pathway, obtained using the CAM-B3LYP/cc-pVTZ model chemistry and a Grimme dispersion in Gaussian 09.^{96,212,321,322} The structures of the azobenzene with C–N=N–C dihedrals of 0° , 90° , 100° and 180° are given as insets. The energies are taken relative to the *trans*-azobenzene structure. Key: *cis*-azobenzene – red, *trans*-azobenzene – blue, transition state structures – purple, full energy profile – grey. b) Plot of the barrier to isomerisation of a single azobenzene molecule *via* the rotation and inversion pathways, obtained by Wazzan using the B3LYP/6-31G(d,p) model chemistry.^{79,85,90,91,93,94,313} The structures of the transition states for each pathway are given as insets. Key: rotation pathway – blue, inversion pathway – pink. b) was replicated from reference 313.

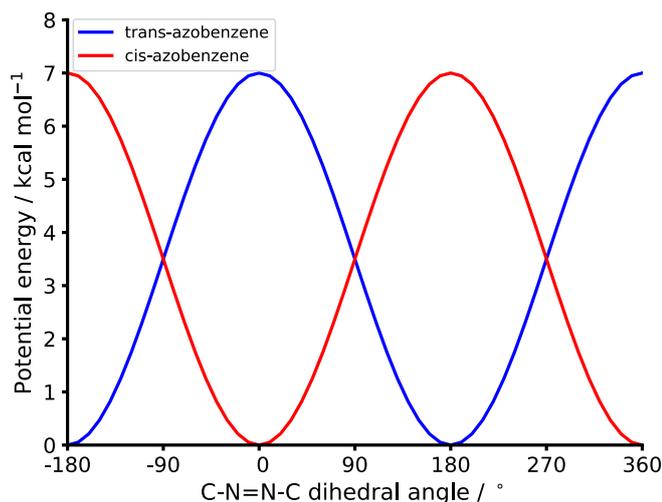


Figure 4.10: Plot of the potential energy of the C-N=N-C dihedral angle in *trans*- and *cis*-azobenzene, obtained using the cosine equation and utilised within the computational DL_POLY 4.0 simulations.²⁰⁸ Key: *trans*-azobenzene – blue, *cis*-azobenzene – red.

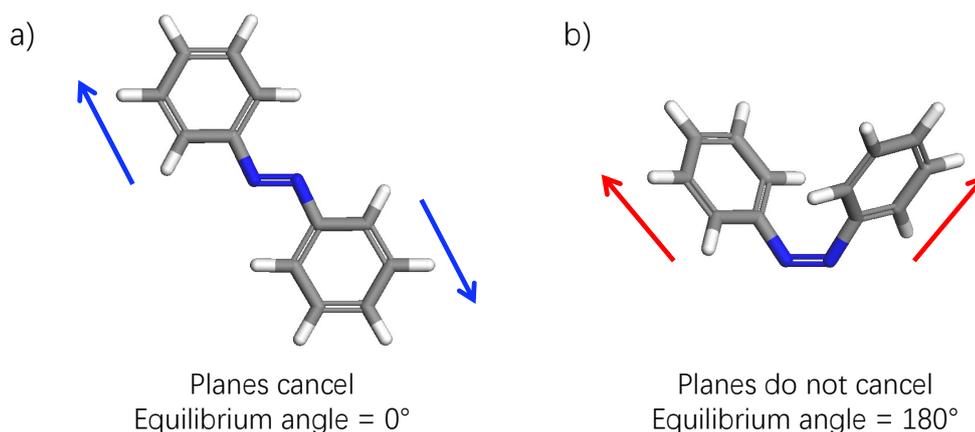


Figure 4.11: The planes comprising the C-N=N-C dihedral angle in a) *trans*-azobenzene and b) *cis*-azobenzene. The two planes are given by the arrows, and the equilibrium angles required to model each isomer using the cosine equation are also given.

The 15 Azo@HCPs structures obtained from the azobenzene loading of the Ambuild HCP models were used as inputs for the DL_POLY 4.0 simulations.²⁰⁸ Firstly, the C-N=N-C dihedral of each azobenzene molecule was switched from *trans* to *cis* by manually editing the input forcefield file so that the equilibrium angle of each C-N=N-C dihedral was set to 180°, with a force constant, corresponding to the *A* value in Equation 4.1, of 3.5 kcal mol⁻¹. This simulated the application of UV light to form *cis*-Azo@HCPs over a timescale of 5.5 ps of NPT MD within DL_POLY 4.0.²⁰⁸ Following this, the dihedrals were switched back to *trans* by manually editing the input forcefield file so that the equilibrium angle of each C-N=N-C dihedral angle was set to 0°, with a force constant of 3.5 kcal mol⁻¹. This mimicked the formation of *trans*-Azo@HCPs by exposure to visible light for a further 5.5 ps of NPT MD within DL_POLY.²⁰⁸

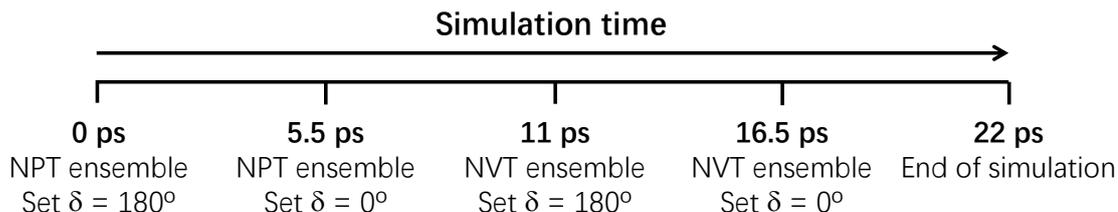


Figure 4.12: Timeline of the simulated irradiation process, specifying the times at which ensembles and C–N=N–C equilibrium angles were changed. δ corresponds to the equilibrium angle utilised in the simulations, set to 0° for *trans*-azobenzene and 180° for *cis*-azobenzene.

This first simulated *trans* to *cis* to *trans* irradiation cycle, which used an NPT MD ensemble, allowed the cell volume to equilibrate for 11 ps during the DL_POLY 4.0 simulation,²⁰⁸ and once this was complete, the process was repeated, using NVT MD.²⁶ It is a standard procedure to use an NPT ensemble to converge the unit cell volume followed by an NVT ensemble to look at the local region of the extended polymer structure. The NVT MD cycle was undertaken using the same equilibrium angle and force constant values within the input forcefield file as in the first irradiation cycle for each isomer.²⁰⁸ The forcefield used to describe all of the remaining atoms in the system remained unchanged throughout the irradiation simulations.²⁶ A summary of the simulated irradiation process, specifying the times at which ensembles and C–N=N–C equilibrium angles were changed, is given in Figure 4.12.

The HCPs and Azo@HCPs models were also submitted to DL_POLY 4.0 for a 11 ps NPT simulation so that the pore size distributions could be obtained under similar conditions.^{26,208} As these systems did not require any switching of the azobenzene isomers, the first NPT MD cycle was sufficient to collect the pore size distributions and surface areas from without requiring a second NVT MD cycle. The input forcefield file of the HCPs models was unchanged as there were no azobenzene molecules present. The input forcefield file of the Azo@HCPs models was modified as above for the *trans*-azobenzene isomer.²⁶

Throughout the first simulated irradiation cycle, used as an equilibration simulation, it would be expected that at 0 ps and 5.5 ps, where the switches in isomer were implemented (0 ps - *trans*- to *cis*- isomerisation, 5.5 ps - *cis*- to *trans*- isomerisation), there would be an immediate sharp shift in the C–N=N–C dihedral angle due to a reassignment of the energy throughout the system on switching the dihedral. Following this reassignment, which would be accompanied by a change in the forces acting on the molecules, the forces would then equilibrate over the remaining timescale until the next change in equilibrium angle corresponding to a shift in isomer. In contrast, the C–N=N–C dihedral angles within the azobenzene molecules in the Azo@HCPs system should remain relatively constant throughout.

The C-N=N-C dihedral angle of the azobenzene molecules was monitored and visualised for model A-1 using VMD.²¹¹ To distinguish between individual azobenzene molecules within this model, each was assigned an alphabetical key in order of the index numbers of each molecule. For example, the azobenzene molecule with the lowest atom index numbers in VMD²¹¹ was assigned as azobenzene molecule A.

The azobenzene molecules were randomly seeded in each case, rather than being placed into the minimum energy binding sites, meaning that some molecules would be less restricted by the rigid HCP structure and find diffusion through the HCP and switching between the *trans*- and *cis*- isomers more facile than others. Additionally, it would be expected that the azobenzene molecules that were originally seeded into higher energy binding sites would experience more of a driving force to diffuse into a lower energy site during the MD simulations, where possible. The location and diffusion of each azobenzene molecule within the HCP was analysed using VMD.²¹¹ Figure 4.13 shows the location and diffusion trajectory throughout the first simulated irradiation cycle of six of the azobenzene molecules within the HCP structure of model A-1.

An example plot of the C-N=N-C dihedral angle as a function of simulation time is given in Figure 4.14 for azobenzene molecule T within HCP model A-1.²⁶ Azobenzene molecule T is used as the example as this is the molecule that most closely follows the expected trend above, with clear changes in the dihedral angle at 0 ps and 5.5 ps. However, it is important to note that not all of the azobenzene molecules could switch between isomers to the same extent, due to the location of azobenzene molecules throughout the HCP in each loading.²⁰⁸ The plots of the remaining twenty-three azobenzene molecules for model A-1 are given in Figures 4.15 and 4.16. The *y*-axes of the plots presented in Figures 4.14-4.16 are limited to a range of 0-180°, as some of the azobenzene molecules were able to cross the periodic boundary during the simulation, resulting in dihedral angles that oscillated between positive and negative values, for example, +180° and -180°, which are seen to both correspond to the *trans*-azobenzene isomer in Figure 4.10. The dihedral angles were therefore plotted as absolute values.²⁶

Figure 4.14 shows that throughout the first simulated irradiation cycle, the azobenzene molecule has a varying C-N=N-C dihedral angle as a function of time. The C-N=N-C dihedral angle of the azobenzene molecule approached 0° after simulated UV irradiation from *trans*- to *cis*-azobenzene, then returned to 180° after simulated visible irradiation from *cis*- to *trans*-azobenzene. In contrast, for the model with no simulated irradiation, which remained as Azo@HCPs throughout, the C-N=N-C dihedral angle of the azobenzene molec-

ule remained relatively constant at 180° throughout, with slight deviations from the idealised angles due to the molecular dynamics undertaken.²⁶

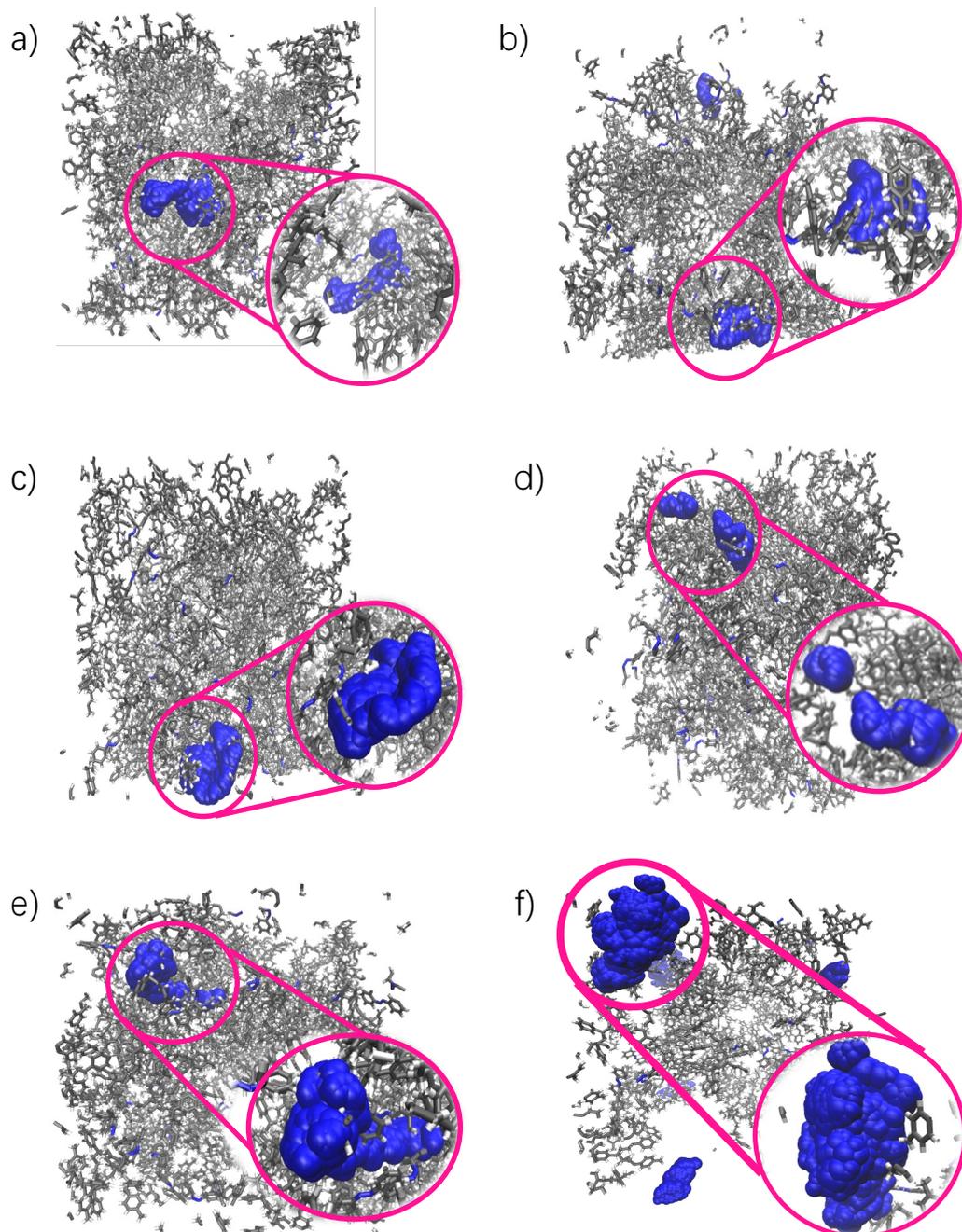


Figure 4.13: Location of six example azobenzene molecules in model A-1 during the first simulated irradiation cycle, with the azobenzene molecules shown as a blue trajectory throughout the 11 ps NPT DL_POLY 4.0²⁰⁸ MD simulation. a) azobenzene molecule T, b) molecule M, c) molecule H, d) azobenzene molecule D, e) molecule K, and f) molecule I.

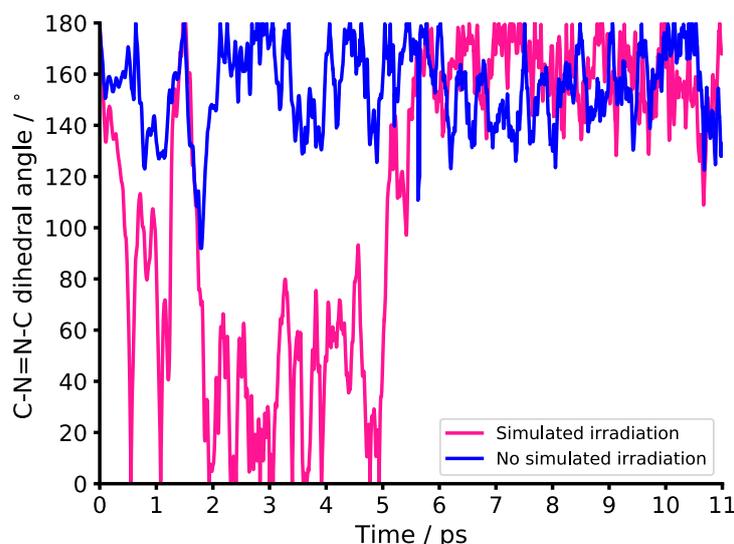


Figure 4.14: Plot of the absolute C–N=N–C dihedral angle for azobenzene molecule T as a function of time for loaded HCP model A-1 after the first simulated irradiation cycle (pink), and for loaded HCP model A-1 with no simulated irradiation, which remains as Azo@HCPs (blue), obtained using VMD.²¹¹ Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

Trends were established between the location of the azobenzene molecule and the resulting plot of the C–N=N–C dihedral angle as a function of simulation time, highlighted for the example structures given in Figure 4.13. Many of the azobenzene molecules, including azobenzene molecule T, were located within the HCP pore structure with additional azobenzene molecules nearby. Here, the C–N=N–C dihedral angle remained at a relatively constant value of 180° when the azobenzene isomer was *trans*, subject to slight deviations on account of the MD undertaken; and the dihedral was able to approach the expected angle of 0° when the isomer was switched to *cis*. This suggested that these azobenzene molecules were able to rotate freely within the HCP pore structure. The relative ability of each molecule to rotate between isomers was dependent on its local environment within the structure, for example, azobenzene molecule D was able to rotate less freely than azobenzene molecule H due to the closer proximity of molecule D to the HCP, leading to a smaller range of observed C–N=N–C dihedral angles for molecule D in comparison to molecule H.

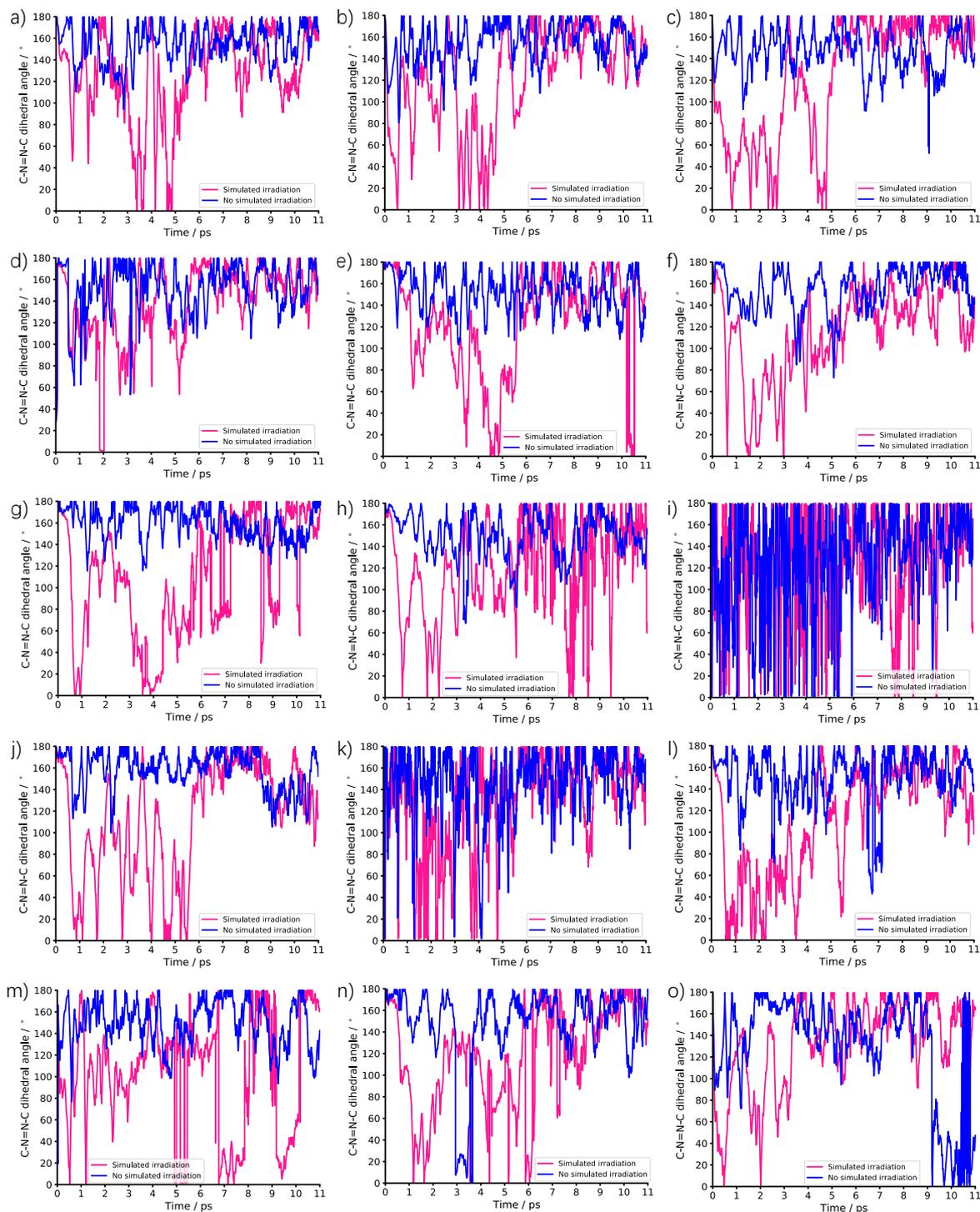


Figure 4.15: Plot of the absolute C–N=N–C dihedral angles for azobenzene molecules A–O as a function of time for loaded HCP model A-1 after the first simulated irradiation cycle (pink), and for loaded HCP model A-1 with no simulated irradiation, which remains as Azo@HCPs (blue), obtained using VMD.²¹¹ a) azobenzene molecule A, b) molecule B, c) molecule C, d) molecule D, e) molecule E, f) molecule F, g) molecule G, h) molecule H, i) molecule I, j) molecule J, k) molecule K, l) molecule L, m) molecule M, n) molecule N, o) molecule O.

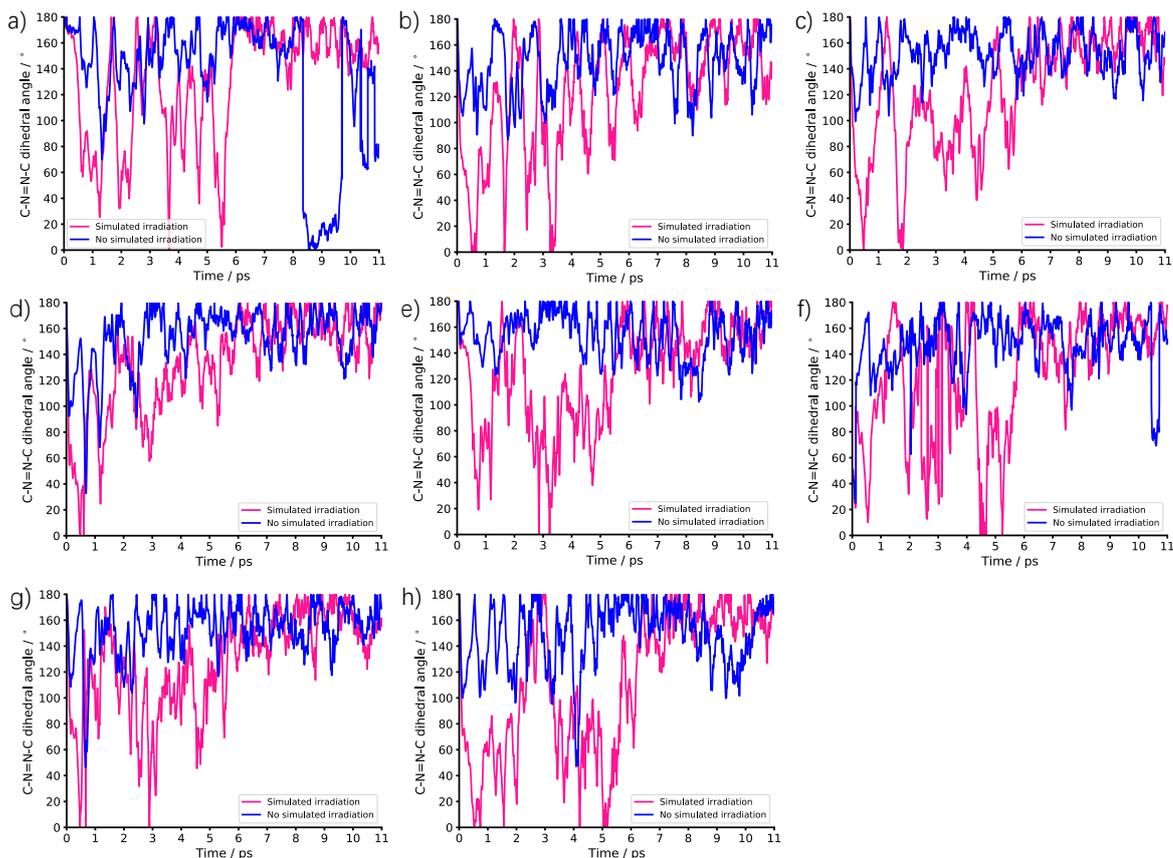


Figure 4.16: Plot of the absolute C–N=N–C dihedral angles for azobenzene molecules P–S, and U–X as a function of time for loaded HCP model A-1 after the first simulated irradiation cycle (pink), and for loaded HCP model A-1 with no simulated irradiation, which remains as Azo@HCPs (blue), obtained using VMD.²¹¹ a) molecule P, b) molecule Q, c) molecule R, d) molecule S, e) molecule U, f) molecule V, g) molecule W, h) molecule X.

In contrast, azobenzene molecule I was located within a large pore further away from the HCP and other azobenzene molecules and appeared to be more isolated. This molecule showed no clear preference for either isomer, and the C–N=N–C dihedral angle varied throughout, suggesting that the azobenzene was rotating freely throughout the simulation, regardless of the equilibrium angle of the molecule. This rotation was presumably an attempt to maximise interactions with the HCP and/or other azobenzene molecules, however, given the relatively large barrier to rotation in azobenzene (Figure 4.9),^{313,320} further work would be required to fully rationalise this.

This suggested that the most ideal switching behaviour within the MD simulations occurred when the azobenzene molecules were located within the HCP pore structure and in close proximity to other azobenzene molecules, for example, azobenzene molecule T. In contrast, when the azobenzene was either too hindered to rotate, for example, azobenzene molecule D, or could rotate too freely, as in the case of azobenzene molecule I, the isomerisation

process did not show the expected behaviour. The DL_POLY 4.0 irradiation simulations were all carried out at 298 K as this was sufficient to allow the azobenzene isomerisation to occur.²⁰⁸ However, using a higher temperature for the simulated switch from *trans*- to *cis*-azobenzene may have induced a more definitive switch in isomers as it would introduce more energy into the system and better mimic the irradiation with high-energy UV light.

4.4 Results and Discussion

4.4.1 Comparison of pore size distribution to experiment

The simulated pore size distributions (PSDs) for models A-E of the HCPs, Azo@HCPs, *cis*-Azo@HCPs and *trans*-Azo@HCPs, collected using Poreblazer after each respective DL_POLY simulation,^{162,208} are given in Figures 4.17–4.21, respectively, with the overall PSDs given in Figure 4.22. Each overall PSD is a combination of those obtained from the repeat models A-E to illustrate the range of pore sizes that would be obtained within an amorphous model.

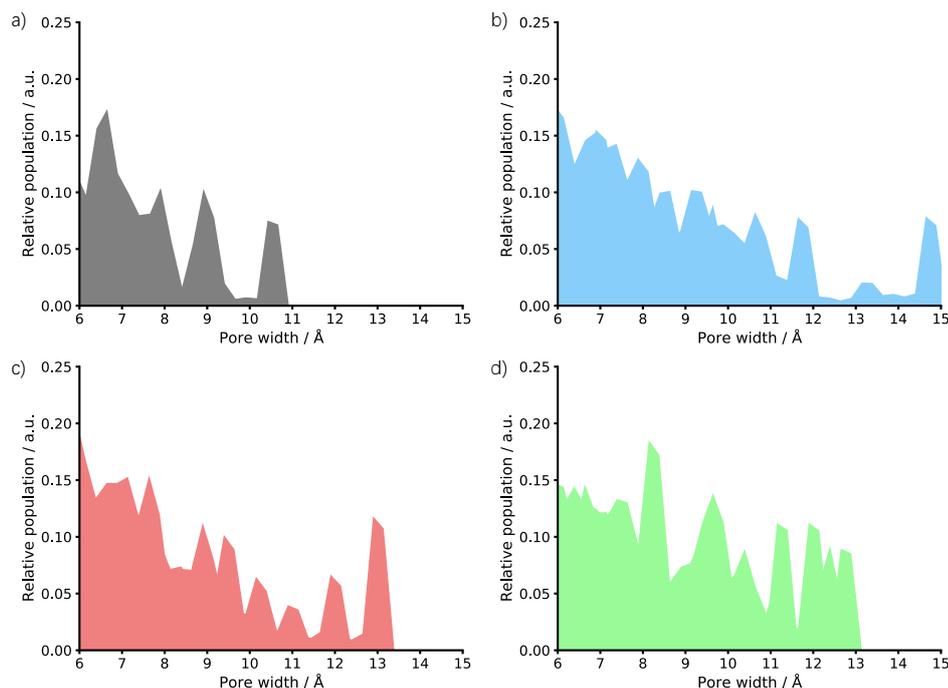


Figure 4.17: PSDs of simulation model A. a) HCPs (gray), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

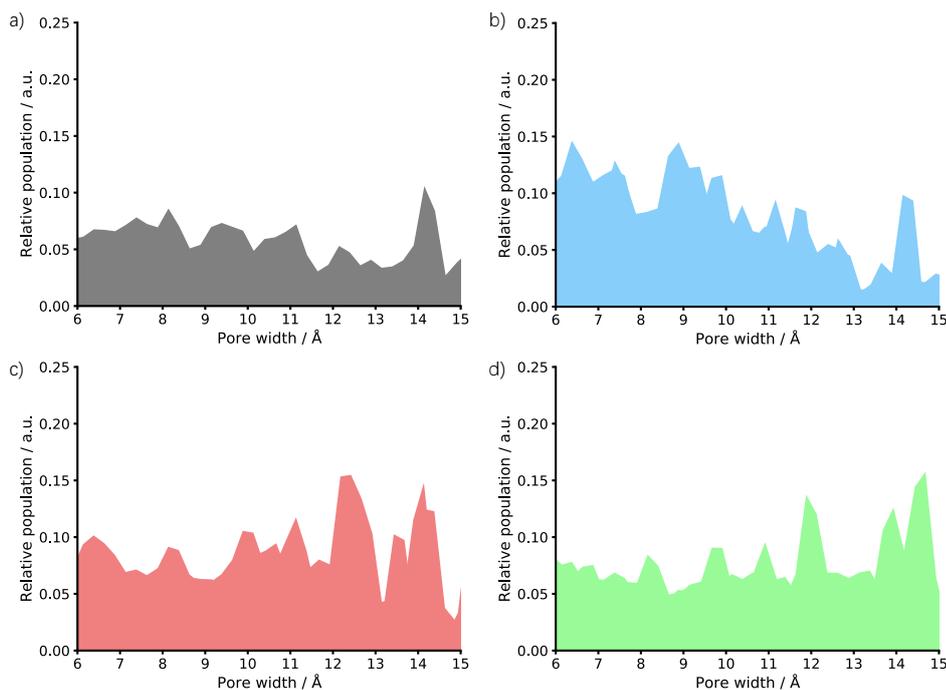


Figure 4.18: PSDs of simulation model B. a) HCPs (gray), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

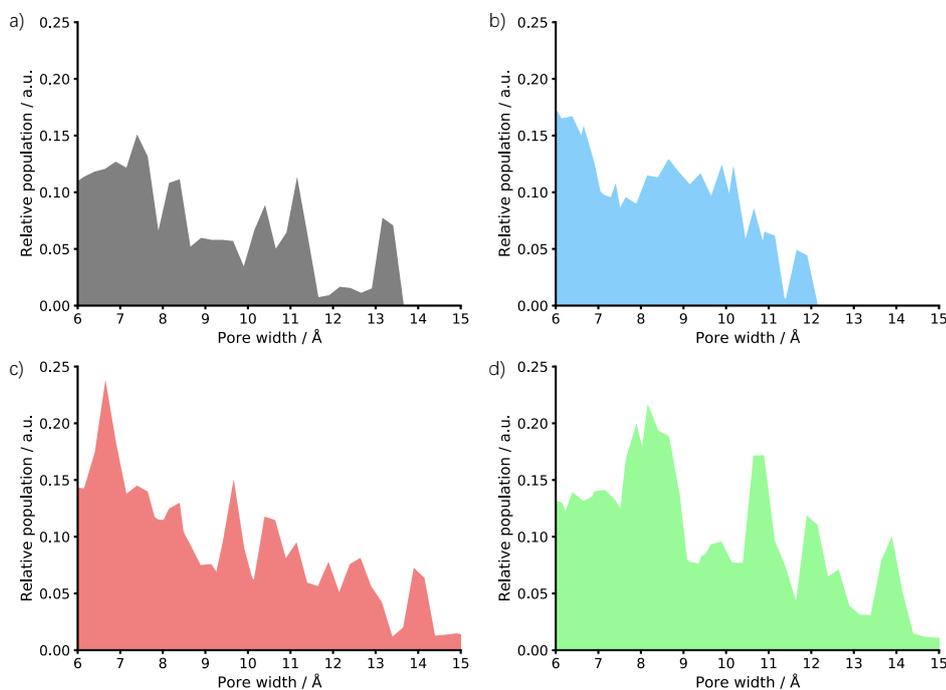


Figure 4.19: PSDs of simulation model C. a) HCPs (gray), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

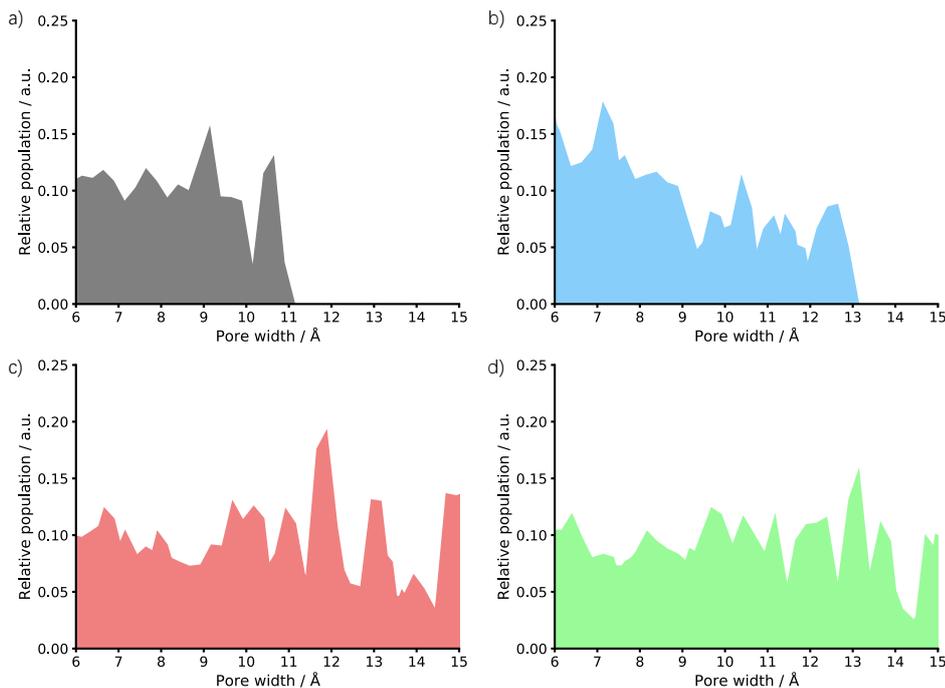


Figure 4.20: PSDs of simulation model D. a) HCPs (gray), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

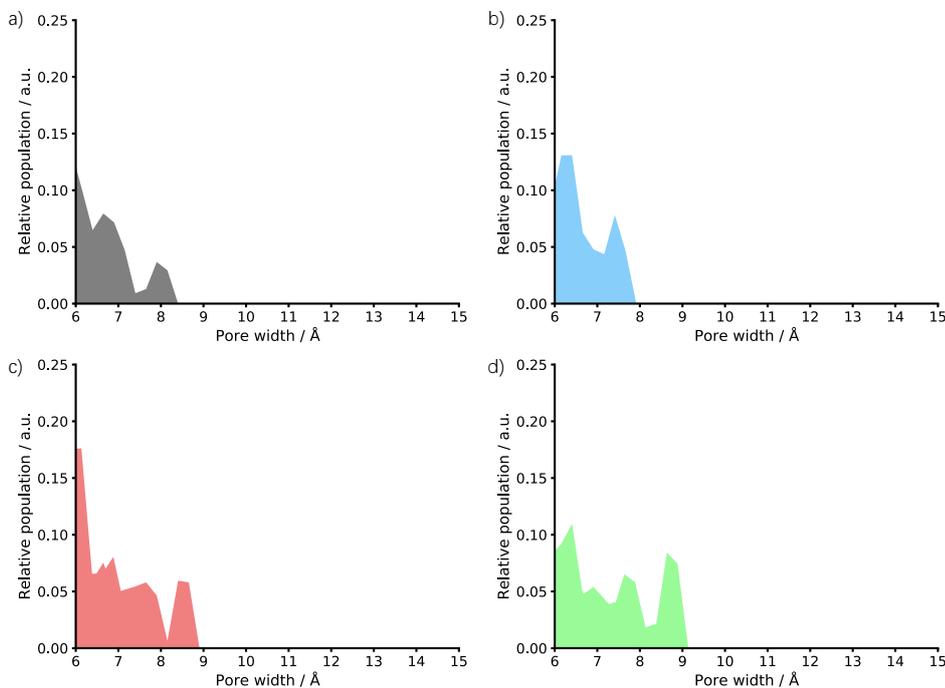


Figure 4.21: PSDs of simulation model E. a) HCPs (gray), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

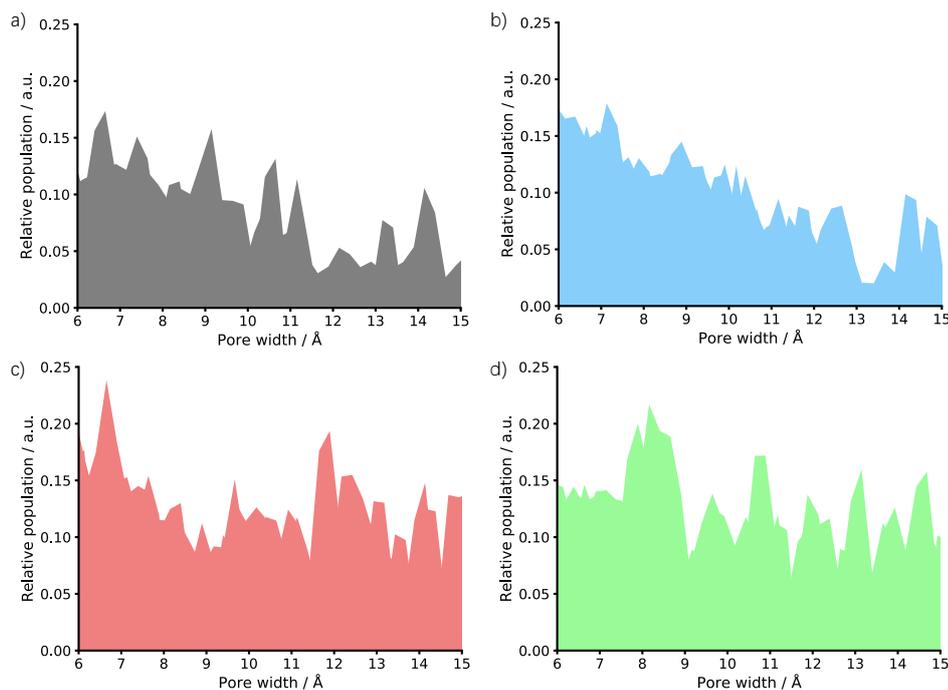


Figure 4.22: Overall PSDs of simulation models A–E. a) HCPs (gray), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

A comparison of the simulated and experimental porosity data are given in Table 4.1. The PSDs of the experimental materials, shown in Figure 4.23, covered a larger pore width range than the simulated PSDs, and showed broader peaks. Both are due to the current modelling approach only being able to model the micropore region. It would be expected that modelling a larger number of samples, in a larger cell volume, would give broader peaks that more closely resemble the experimental. These limitations also gave rise to smaller surface areas and micropore volumes for the simulated models in comparison to experiment.²⁶

It can be seen in Figure 4.23 that pores are present in three regions of the experimental spectrum: 5–10 Å, 13–20 Å and 20–40 Å. At first glance, the experimental spectra appear very similar, however, there are some differences in the experimental pore volumes in the regions of 10–20 Å and 20–40 Å, with reductions in both regions going from HCPs to Azo@HCPs, giving an overall pore volume decline in these regions of 14% and 28%, respectively. This indicates that the azobenzene molecules were residing within these pores, causing a reduction in the respective experimental surface area (HCPs = 1225 m² g⁻¹, Azo@HCPs = 984 m² g⁻¹).²⁶

Table 4.1: Comparison of the averaged simulation and experimental surface areas and micropore volumes for the HCPs, Azo@HCPs, *cis*-Azo@HCPs and *trans*-Azo@HCPs, expressed as actual and relative values. Table adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

System	HCPs	HCPs@Azo	<i>cis</i> -HCPs@Azo	<i>trans</i> -HCPs@Azo
Experimental surface area / m ² g ⁻¹	1225	984	1024	862
Relative experimental surface area / %	100	80	84	70
Simulation surface area / m ² g ⁻¹	1177	830	876	873
Relative simulation surface area / %	100	71	74	74
Experimental pore volume / cm ³ g ⁻¹	1.220	0.957	1.021	0.832
Relative experiment pore volume / %	100	78	84	68
Simulation micropore volume / cm ³ g ⁻¹	0.666	0.481	0.498	0.496
Relative simulation micropore volume / %	100	72	75	74

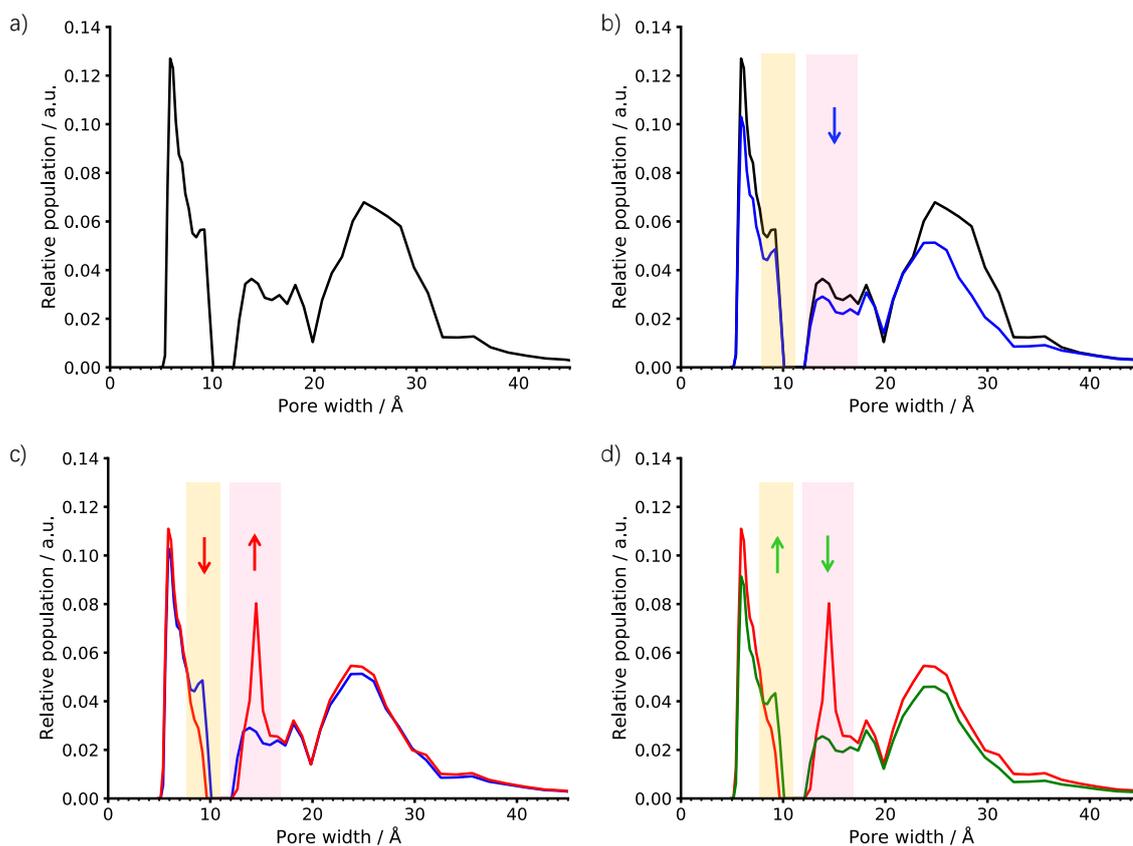


Figure 4.23: Experimental PSDs obtained from nitrogen adsorption at 77 K and a non-local density functional theory (NLDFT) model. a) HCPs (black), b) Azo@HCPs (blue), c) *cis*-Azo@HCPs (red) and d) *trans*-Azo@HCPs (green). The differences in the regions of 10 Å and 14 Å are indicated by arrows corresponding by colour to the relevant material. Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

The experimental pore volume obtained in the 10–20 Å region of the PSD was lower in the Azo@HCPs system compared to the pristine HCPs. However, the pore volume in this region increased upon UV irradiation of the sample to give *cis*-Azo@HCPs, with an overall increase in the experimental pore volume of 26%, giving a higher pore volume than that of the pristine HCPs materials. This corresponded to a respective increase and decrease in the pores in the regions of 14 Å and 10 Å in *cis*-Azo@HCPs. This was rationalised due to the smaller dimensions of the *cis*-azobenzene isomer (5.5 Å) in comparison to *trans*-azobenzene (9 Å). The smaller dimensions of the *cis* isomer require less space within the pore, giving rise to additional free volume within the material for adsorption of gas molecules. This was confirmed by a larger experimental N₂ adsorption for *cis*-Azo@HCPs in comparison to Azo@HCPs (Azo@HCPs = 661 cm³ g⁻¹, *cis*-Azo@HCPs = 697 cm³ g⁻¹).²⁶

Irradiation of the *cis*-Azo@HCPs with visible light to give *trans*-Azo@HCPs causes an overall reduction in the pores in the region of 10–20 Å, with a respective decrease and increase in the pores in the regions of 14 Å and 10 Å due to the larger dimensions of the azobenzene molecules, giving a very similar experimental PSD to that of the Azo@HCPs, indicating that the isomerisation was, indeed, reversible.²⁶

The Connolly surface of model C, using loading 1 for the HCPs@Azo, *cis*-Azo@HCPs and *trans*-Azo@HCPs systems) was calculated using Materials Studio 5.0 (Figure 4.24).¹⁵⁸ This model was chosen as it corresponded to the largest change in pore sizes between each stage in the simulation. A Connolly probe radius of 4 Å was utilised to plot the surfaces as this indicated only the pores with widths greater than or equal to 8 Å, corresponding to the lower bound of the yellow rectangles given in Figure 4.23.²⁶

It can be seen that there were differences between systems depending on the presence and isomer of azobenzene,²⁶ however, it was not trivial to assess the positions of the azobenzene molecules within the pore structure of the HCP. Therefore, the surfaces were also calculated for both the HCP contribution and azobenzene contribution to the total Connolly surface area of each stage in the simulation (Figure 4.25).¹⁵⁸

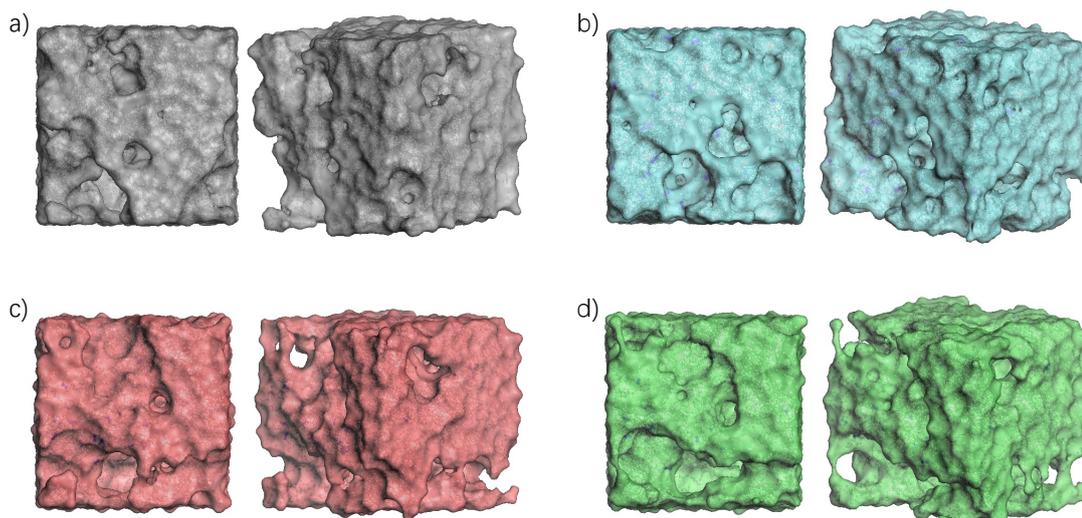


Figure 4.24: Front and side view of the total Connolly surface obtained by Materials Studio 5.0¹⁵⁸ using a probe radius of 4 Å to indicate only pores with widths equal to or larger than 8 Å within the pore size distribution, corresponding to the lower bound of the yellow rectangle in Figure 4.23. a) Model C HCPs (gray), b) model C-1 Azo@HCPs (blue), c) model C-1 *cis*-Azo@HCPs (red) and d) model C-1 *trans*-Azo@HCPs (green).

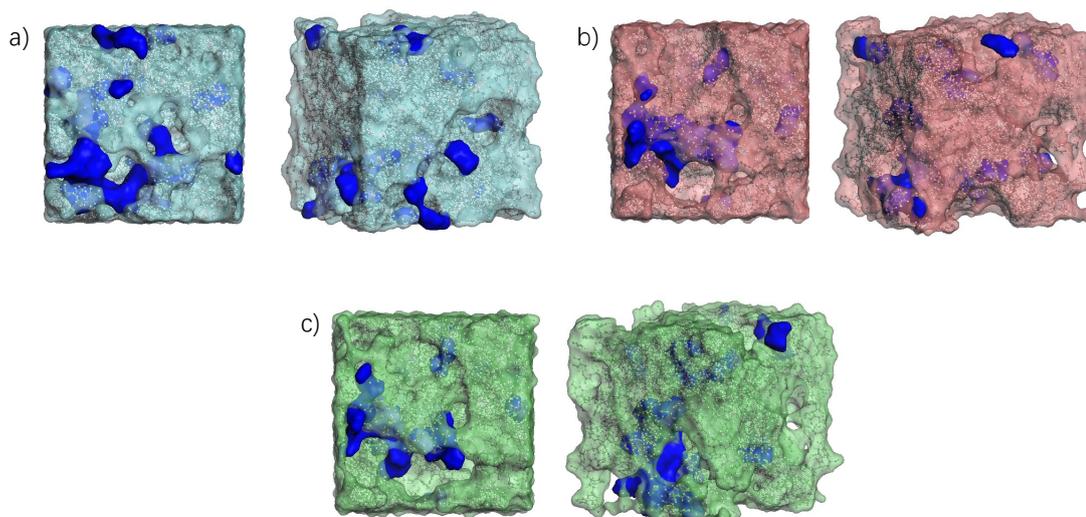


Figure 4.25: Front and side view of the Connolly surface obtained by Materials Studio 5.0¹⁵⁸ using a probe radius of 4 Å to indicate only pores with widths equal to or larger than 8 Å within the pore size distribution, corresponding to the lower bound of the yellow rectangle in Figure 4.23. The contributions of the HCP and azobenzene to the total surface areas given in Figure 4.24 are given as distinct surfaces, with the azobenzene surface shown in dark blue in all cases. a) Model C-1 Azo@HCPs (HCP surface given in pale blue), b) model C-1 *cis*-Azo@HCPs (HCP surface given in red) and c) model C-1 *trans*-Azo@HCPs (HCP surface given in green).

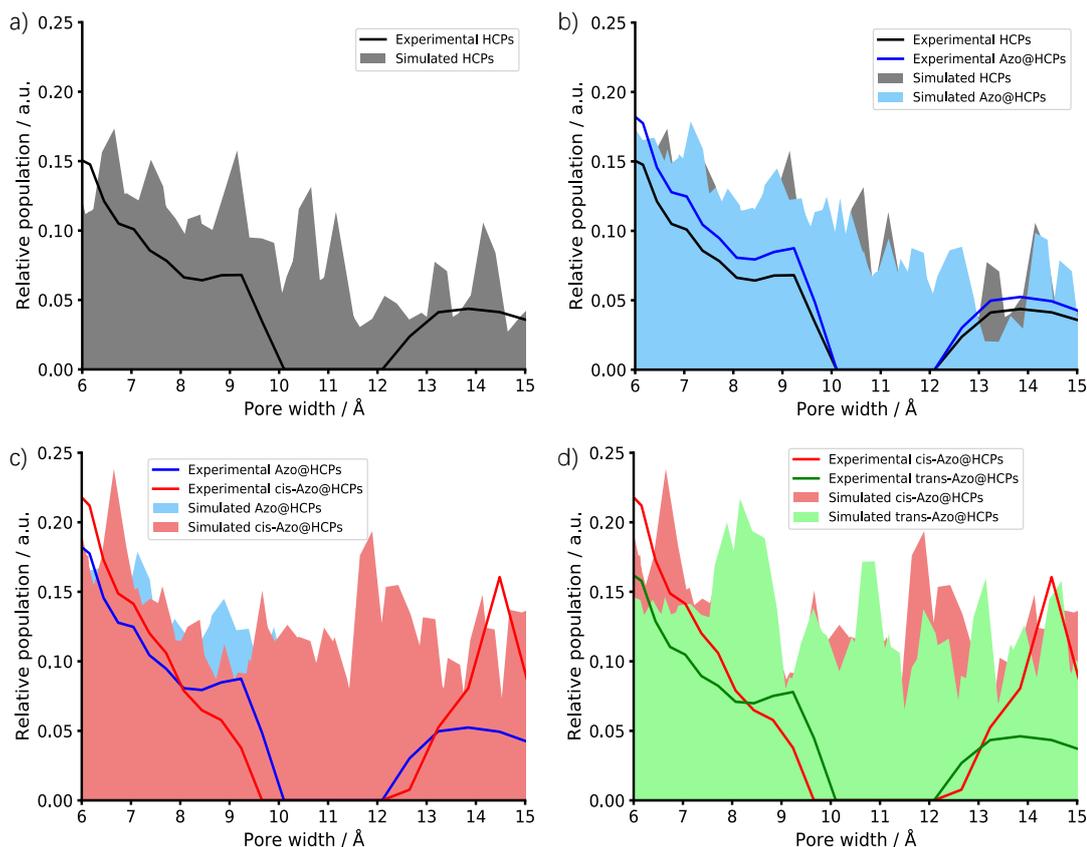


Figure 4.26: Comparison of the overall PSDs of simulation models A–E, with the experimental materials for the equivalent pore width overlaid as lines and the experimental distributions scaled to fit the simulation distributions. Key: simulated HCPs (gray, shaded), simulated Azo@HCPs (blue, shaded), simulated *cis*-Azo@HCPs (red, shaded), simulated *trans*-Azo@HCPs (green, shaded), experimental HCPs (black line), experimental Azo@HCPs (blue line), experimental *cis*-Azo@HCPs (red line), experimental *trans*-Azo@HCPs (green line). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

Figure 4.25 shows that the azobenzene molecules in the Azo@-HCPs system, where the isomer is *trans*, resided close to the larger pores around the edges of the structure. In comparison, the smaller *cis*-azobenzene molecules in the *cis*-Azo@HCPs system (dimensions: *trans*-azobenzene = 9 Å, *cis*-azobenzene = 5.5 Å) were able to diffuse further into the pore structure of the HCP and reside within the smaller pores closer to the centre of the structure, showing a smaller azobenzene surface around the edges of the HCP material. These trends match those obtained from the experimental pore size distributions given in Figure 4.23. Whilst this analysis is useful to establish trends between the structures at each stage of the simulation, the caveat must be made that these surfaces are plotted for one model system only, and to be more robust, this analysis should be repeated across a larger range of model systems.

When the overall simulated PSDs are compared to the experimental in Figure 4.26, similar differences can be observed in the regions of 8–9 Å and 11–13 Å. There was a respective increase and decrease in the pores in the regions of 14 Å and 10 Å on simulating the irradiation of Azo@HCPs with UV light to give *cis*-Azo@HCPs. This is due to the smaller diameter of the *cis*-azobenzene in comparison to the *trans* isomer. The irradiation of *cis*-Azo@HCPs with visible light is simulated to give *trans*-Azo@HCPs, with an increase in the pores of smaller width in the range of 8–9 Å and a decrease in the pores of larger width in the range of 11–13 Å.

4.4.2 CO₂ diffusion and uptake

Whilst the pore size distribution spectra of the materials are sensitive to the presence and isomer of azobenzene, differences are also visible in the experimental CO₂ uptake (Figure 4.27).²⁶ The experimental CO₂ uptake of the pristine HCPs, 46.6 cm³ g⁻¹, decreased upon incorporation of azobenzene into the pores to give Azo@HCPs, with an uptake of 37.7 cm³ g⁻¹. Irradiation of the Azo@HCPs material with UV light at 365 nm to give *cis*-Azo@HCPs increased the experimental uptake to the maximum value of 56.9 cm³ g⁻¹, which is a 51% increase from that of the Azo@HCPs.²⁶ Finally, irradiation of the *cis*-Azo@HCPs material with visible light at 410 nm causes a reduction in the experimental CO₂ uptake to the lowest of the four materials, 33.1 cm³ g⁻¹, a value that is similar to that of the Azo@HCPs, again supporting that the isomerisation reaction is reversible.²⁶ The observed trend of the experimental CO₂ uptake decreasing from HCPs to Azo@HCPs, then increasing from Azo@HCPs to *cis*-Azo@HCPs, and decreasing again going from *cis*-Azo@HCPs to *trans*-Azo@HCPs, matches the observed trend in porosity. However, the percentages of increase and decrease do differ between the two properties, with the HCPs having the highest surface area and micropore volume, and the *cis*-Azo@HCPs having the highest CO₂ uptake.²⁶

Typically, HCPs show an increasing CO₂ uptake and surface area with an increasing number of adsorption sites for the CO₂ to bind to. This is seen for the Azo@HCPs and *trans*-Azo@HCPs, where the presence of azobenzene within the pores of the framework gave a reduction in the pore volume and number of available adsorption sites, and therefore a decreased surface area and CO₂ uptake compared to the HCPs.²⁶ However, the *cis*-Azo@HCPs material showed the highest CO₂ uptake of all of the materials studied, despite the decrease in surface area and pore volume compared to the pristine HCPs. This was proposed as being due to the favourable interaction between the 0.6246 e Å (3 Debye) dipole moment on each *cis*-azobenzene molecule and the -0.8364 e Å² (-13.4±0.4 C m²) quadrupole on the CO₂, along with the increased quantity of micropores within the 10–20 Å region (Figure 4.23).²⁶

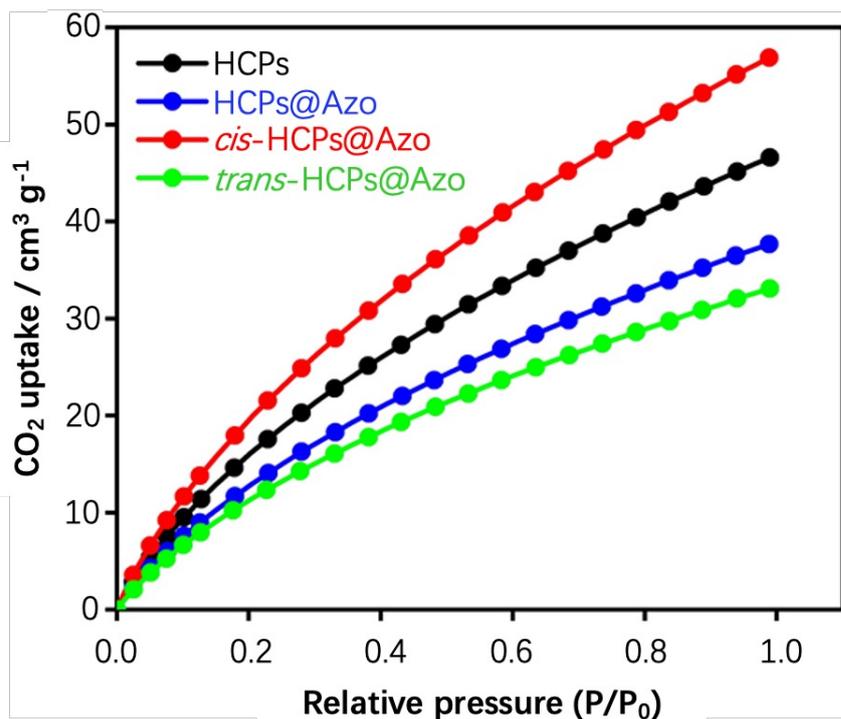


Figure 4.27: Experimental CO₂ uptakes of the HCPs (black), Azo@HCPs (blue), *cis*-Azo@HCPs (red) and *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

To further understand the CO₂ uptake behaviour of these materials, the CO₂ diffusion of the simulated *cis*-Azo@HCPs and *trans*-Azo@HCPs structures was modelled. This was done by randomly seeding one CO₂ building block into each model, with the charges calculated using the QEq approach within Materials Studio 5.0 as: C: 0.8 e, O: 0.4 e, giving a quadrupole across the CO₂ molecule with an overall neutral charge.^{158,201,202} Once seeded, the models were input to DL_POLY for a 1 ns diffusion simulation at 273 K with an NVT ensemble.^{26,208} This gave the diffusion coefficients and root-mean-square deviation (RMSD) values for the CO₂ molecules reported in Table 4.2 (it was assumed that the Azo@HCPs itself would be likely to remain relatively fixed in place in comparison to the CO₂ molecules), and the diffusion trajectories shown in Figure 4.28, visualised using VMD.^{26,211}

Table 4.2 shows that the average diffusion coefficients and RMSD values for the CO₂ molecules were larger for the *trans*-Azo@HCPs models than the *cis*-Azo@HCPs systems, indicating that the CO₂ molecule could travel further through the HCP framework, and at a faster rate, within the *trans*-Azo@HCPs system. This confirms that the increased CO₂ uptake in the *cis*-Azo@HCPs framework was not due to the ability of the CO₂ to diffuse throughout the system, but rather, to the increased quantity of micropores within the 11–13 Å

region for the *cis*-Azo@HCPs system.²⁶ It was hypothesised that this was due to the dipole of *cis*-azobenzene slowing the rate of CO₂ diffusion. This was analysed using radial distribution functions.

Radial distribution functions (RDFs) were then obtained using VMD for the *cis*-Azo@HCPs and *trans*-Azo@HCPs systems (Figure 4.29) after the CO₂ diffusion simulations to analyse the proximity of the nitrogen atoms in the azobenzene to the central carbon atom of the CO₂.²¹¹ It could be seen that the CO₂ molecules resided in closer proximity to the nitrogen atoms of the azobenzene when the isomer was *cis*, with the 3.0 D dipole moment that was able to interact with the quadrupole on the CO₂, in comparison to the *trans* isomer, which does not have a dipole moment. This can also be seen in the trajectories given in Figure 4.30, which show the position of the nitrogen atoms of the azobenzene relative to the CO₂ molecules throughout the diffusion simulations.²⁶

Table 4.2: Simulated CO₂ diffusion coefficients and RMSD values for the *cis*-Azo@HCPs and *trans*-Azo@HCPs models. Table adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

Model	System	C Diffusion coefficient / m ² s ⁻¹	C Sqrt[RMSD] / Å	O Diffusion coefficient / m ² s ⁻¹	O Sqrt[RMSD] / Å
B-1	<i>cis</i> -HCPs@Azo	0.0000	10.4350	2.0862 × 10 ⁻¹²	10.5860
	<i>trans</i> -HCPs@Azo	9.0376 × 10 ⁻¹⁰	23.8790	9.0400 × 10 ⁻¹⁰	23.9400
B-2	<i>cis</i> -HCPs@Azo	7.1444 × 10 ⁻¹⁰	28.0290	7.1994 × 10 ⁻¹⁰	28.1210
	<i>trans</i> -HCPs@Azo	0.0000	23.2240	0.0000	23.3090
C-1	<i>cis</i> -HCPs@Azo	0.0000	27.7920	0.0000	28.0870
	<i>trans</i> -HCPs@Azo	6.0060 × 10 ⁻¹¹	9.7480	5.9406 × 10 ⁻¹¹	9.8760
C-2	<i>cis</i> -HCPs@Azo	0.0000	17.2860	0.0000	17.3590
	<i>trans</i> -HCPs@Azo	1.7886 × 10 ⁻⁹	34.6020	1.7876 × 10 ⁻⁹	34.6420
C-3	<i>cis</i> -HCPs@Azo	3.0470 × 10 ⁻¹⁰	24.6790	3.0954 × 10 ⁻¹⁰	24.7650
	<i>trans</i> -HCPs@Azo	0.0000	16.5050	0.0000	16.5930
E-1	<i>cis</i> -HCPs@Azo	0.0000	2.2783	0.0000	2.3113
	<i>trans</i> -HCPs@Azo	0.0000	9.8309	0.0000	9.8586
E-2	<i>cis</i> -HCPs@Azo	0.0000	10.2050	0.0000	10.2470
	<i>trans</i> -HCPs@Azo	7.4195 × 10 ⁻¹²	9.1716	4.1209 × 10 ⁻¹²	9.1962
E-3	<i>cis</i> -HCPs@Azo	1.1396 × 10 ⁻¹⁰	10.3360	1.2306 × 10 ⁻¹⁰	10.5820
	<i>trans</i> -HCPs@Azo	9.8299 × 10 ⁻¹¹	7.3625	1.0939 × 10 ⁻¹⁰	7.6751
AVERAGE	<i>cis</i> -HCPs@Azo	1.4164 × 10⁻¹⁰	16.3800	1.4433 × 10⁻¹⁰	16.5073
	<i>trans</i> -HCPs@Azo	3.5727 × 10⁻¹⁰	16.7904	3.5806 × 10⁻¹⁰	16.8862

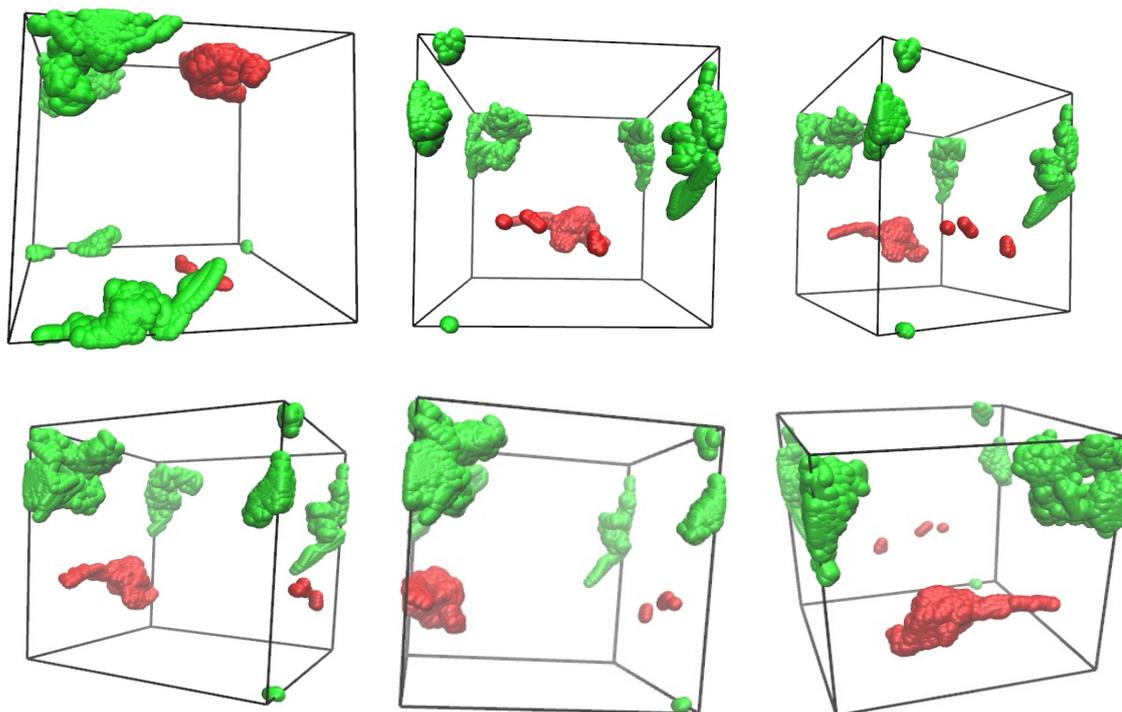


Figure 4.28: Model C-2 simulated CO₂ diffusion trajectories visualised from different angles using VMD with the Azo@HCPs omitted for clarity.²¹¹ Key: trajectory of the CO₂ within *cis*-Azo@HCPs (red) and trajectory of the CO₂ within *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

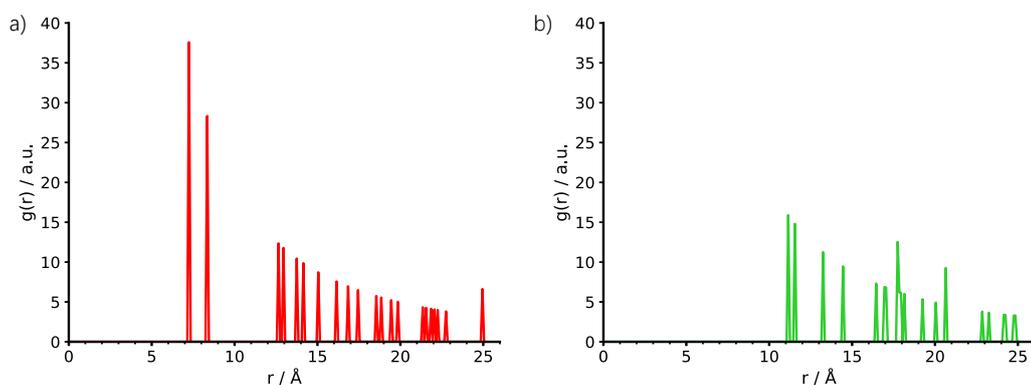


Figure 4.29: Simulated RDFs for model C-2 showing the interaction between the nitrogen atoms of the azobenzene and the carbon atom of the carbon dioxide molecule, obtained using VMD.²¹¹ Key: *cis*-Azo@HCPs (red) and *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

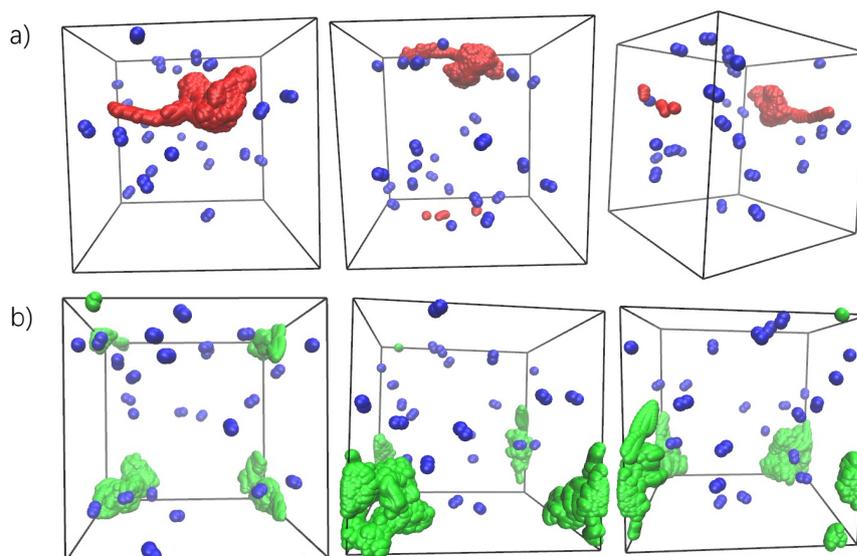


Figure 4.30: Model C-2 simulated CO₂ diffusion trajectories visualised from different angles using VMD with the nitrogen atoms of the azobenzene shown as blue spheres.²¹¹ a) trajectory of the CO₂ within *cis*-Azo@HCPs (red) and b) trajectory of the CO₂ within *trans*-Azo@HCPs (green). Figure adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

To compare the strength of the interactions between the azobenzene and CO₂ molecules, binding energy calculations were undertaken of a single CO₂ molecule in close proximity to a single azobenzene molecule. The CAM-B3LYP/cc-pVTZ model chemistry was used within Gaussian 09, applying a Grimme dispersion to better describe the interactions between the CO₂ and azobenzene molecules.^{96,212,321,322} Each structure was firstly optimised to a minimum-energy ground state geometry, which was confirmed by a harmonic wavenumber analysis, and the binding energies were subsequently calculated as described in Equation 4.2.²⁶

$$E_{\text{binding}} = E_{\text{Azo+CO}_2} - (E_{\text{Azo}} + E_{\text{CO}_2}) \quad (4.2)$$

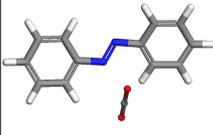
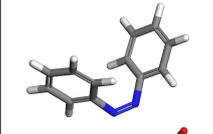
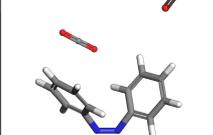
The binding energy of the carbon dioxide molecule interacting with each azobenzene isomer was given as E_{binding} , the energy of the azobenzene and carbon dioxide molecules in close proximity to each other was given as $E_{\text{Azo+CO}_2}$, and the energies of the individual azobenzene and CO₂ molecules were reported as E_{Azo} and E_{CO_2} , respectively.²⁶

The binding energies of CO₂ in close proximity to each of the azobenzene isomers are given in Table 4.3.²⁶ In the case of the less-symmetrical *cis*-azobenzene isomer, the CO₂ molecule was placed in two different locations to assess the effect of the spatial configuration. The binding energy of the CO₂ in relation to the azobenzene molecule was higher, indicating a more favourable interaction, when the carbon dioxide molecule was placed under-

neath the nitrogen atoms of the azobenzene and on the opposite side of the molecule to the phenyl rings.²⁶ This was due to the $E_{\text{Azo}+\text{CO}_2}$ value being lower in this configuration in comparison to that where the CO_2 is placed between the phenyl rings. This may be due to the closer proximity of the CO_2 to both nitrogen atoms of the azobenzene when placed on the opposite side of the molecule to the bulky phenyl rings. Therefore, when comparing the binding energies of the carbon dioxide in proximity to each isomer of azobenzene, the system chosen for the *cis* isomer was the lower energy configuration, where the CO_2 molecule was placed on the opposite side of the azobenzene to the phenyl rings.²⁶

When this comparison was made, it was seen that the binding energy of the *cis*-azobenzene in close proximity to the CO_2 was higher than when the azobenzene isomer was *trans*. This indicated a more favourable interaction between the *cis*-azobenzene and CO_2 , rationalised due to the dipole moment of *cis*-azobenzene being more attractive to the quadrupole of CO_2 in comparison to *trans*-azobenzene, which has no dipole moment.²⁶

Table 4.3: Calculated binding energies of a single CO_2 molecule interacting with a single molecule of azobenzene. Table adapted with permission from reference 26, upon which this chapter is based, licensed under CC-BY. Copyright 2023 The Authors.

System	Azobenzene isomer	Location of CO_2 relative to azobenzene phenyl rings	Binding energy / kJ mol^{-1}	Binding energy / eV
	Trans	Between phenyl rings (symmetrical isomer)	-18.597	-0.193
	Cis	Opposite side to phenyl rings	-20.468	-0.212
	Cis	Between phenyl rings	-18.010	-0.187

4.5 Conclusions

In this work, the artificial synthesis of a hypercrosslinked polymer prepared experimentally by knitting waste polystyrene with an external formaldehyde dimethyl acetal crosslinker in the presence of an FeCl_3 catalyst and dichloroethane solvent is described. Representative structural models of the HCP were simulated using the Ambuild code, following the experimental catalytic mechanism. Azobenzene was then added as a guest molecule, and the

simulated models were used to rationalise the trends in the experimental pore size distributions. The increase in the pores at 14 Å and decrease in the pores at 10 Å when the azobenzene is in its *cis* isomer were rationalised due to the smaller dimensions of *cis*-azobenzene in comparison to the *trans* isomer.²⁶

Furthermore, the experimental 51% increase in CO₂ uptake on going from the Azo@HCPs to *cis*-Azo@HCPs via UV irradiation was rationalised using CO₂ diffusion simulations. These determined that the increase in uptake was not due to the ability of the CO₂ to diffuse throughout the HCP (Figure 4.28) but was due to the increased quantity of micropores in the 11–13 Å region (Figure 4.26). Finally, it was also observed through radial distribution functions and binding energy calculations that the positive dipole moment of *cis*-azobenzene is able to interact more favourably with the quadrupole on the CO₂ molecule in comparison to *trans*-azobenzene, which has no dipole moment.²⁶

This work provides an exciting opportunity to reduce plastic pollution by synthesising HCPs from waste polystyrene materials. The incorporation of azobenzene into the pore structure could be used to enable low-energy carbon capture in the future due to the favourable interactions between the positive dipole moment of *cis*-azobenzene and the quadrupole of carbon dioxide.²⁶

Chapter 5

Artificial synthesis of organically synthesised porous carbon as a case study for rationalising the solid-state NMR spectra of OSPC materials

Curiouser and curiouser.

Lewis Carroll, Alice in Wonderland

5.1 Summary

Organically synthesised porous carbon (OSPC) materials, which are amorphous frameworks assembled from sp^3 carbon nodes connected by sp carbon linkers, show great potential as anode materials for lithium-ion batteries. However, the ^{13}C solid-state nuclear magnetic resonance (ssNMR) spectrum of the first OSPC material, OSPC-1, does not at first glance agree with the structure proposed by x-ray photoelectron spectroscopy and electron energy-loss spectroscopy, instead suggesting a structure composed of sp^2 and sp^3 hybridised carbon.^{16,95} This led to the work described within this chapter and in the report by Mollart and co-workers,⁹⁵ which rationalises the experimental ^{13}C ssNMR spectra of OSPC-1a and OSPC-1b. To differentiate between the two materials, the original OSPC-1 material¹⁶ is referred to as OSPC-1a.

OSPC-1b was synthesised using a one-pot approach, which does not allow for the formation of sp^2 hybridised carbon, rather than the multistep synthetic route of OSPC-1a.^{16,95,323}

This was undertaken by the generation of calculated NMR spectra for a range of molecular fragments of OSPC-1 using Gaussian 09.^{95,212} Following this, a series of periodic models of OSPC-1 in differing crystalline topologies and chains of OSPC-1 were modelled using CASTEP 8.0.²¹³ This allowed the assessment of the influence of crystallinity and chain-chain interactions on the resulting calculated NMR spectra.⁹⁵ Finally, the porosity of OSPC-1 was rationalised by considering a sphere of dense, non-porous OSPC-1, surrounded by a thin layer of open and highly porous OSPC-1. The former is able to mimic the experimental ^{13}C NMR spectra well, whilst the latter mimics the experimental porosity.⁹⁵

5.2 Introduction

As the world enters a digital era, there is a requirement to find renewable energy sources to power devices due to increased demand and dwindling fossil fuel reserves.³²⁴ As a result, there is enhanced interest in renewable batteries, with the most well-researched type being lithium-ion batteries (LIBs).^{16,325} These are currently made commercially using an anode made from graphite, which is cheap with high conductivity ($2\text{--}3 \times 10^3 \text{ S cm}^{-1}$),³²⁶ and a cathode made from lightweight, redox-active transition metal oxides such as LiCoO_2 .³²⁷ The electrodes are separated by a microporous polymer material such as polyethylene,³²⁸ and in the presence of an electrolyte such as LiPF_6 dissolved in a carbonate material such as ethylene carbonate (Figure 5.1a).³²⁹ During battery discharge, lithium ions deintercalate from between the layers of graphite at the anode, diffuse through the separator and intercalate into the transition metal oxide at the cathode. During charging, the reverse process occurs.³²⁷

However, the current commercial set-up has some flaws due to the lack of abundance of lithium³³⁰ and the cost and toxicity of cobalt.³³¹ Additionally, graphite has a limited lifetime for LIBs due to the requirement for the lithium ions to diffuse between the layers of graphite. This proves challenging due to the ionic radius of the lithium ions being relatively large (0.9 \AA)³³² in comparison to the interlayer spacing in graphite (3.336 \AA).³³³ This leads to degradation of the graphite over time due to swelling of the layers upon repeated charging and discharging of the battery.^{16,334} The limited diffusion of lithium through the structure also leads to a build-up of metallic lithium waiting to diffuse between the layers, which can cause short circuit and explosion of the battery if these threads of dendritic lithium reach the cathode (Figure 5.1b).¹⁶

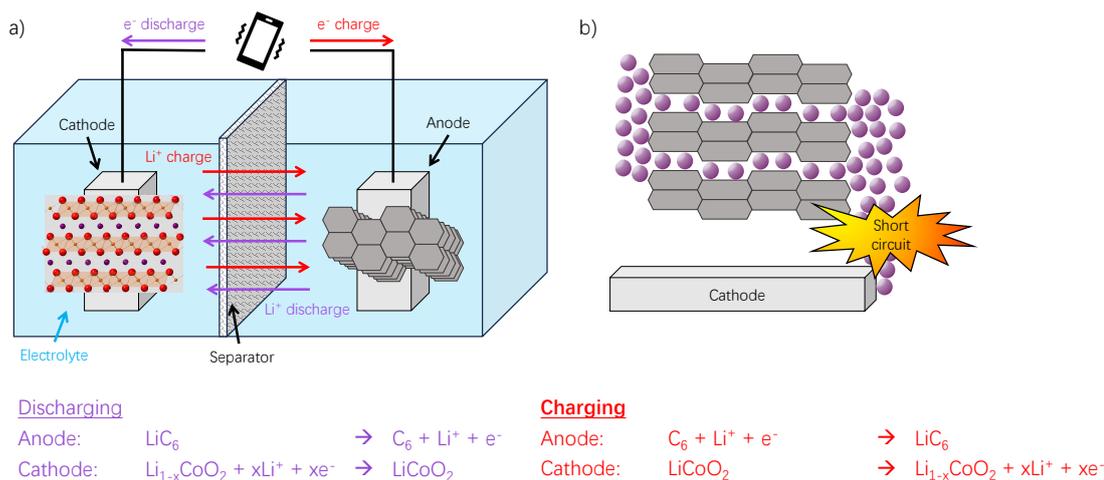
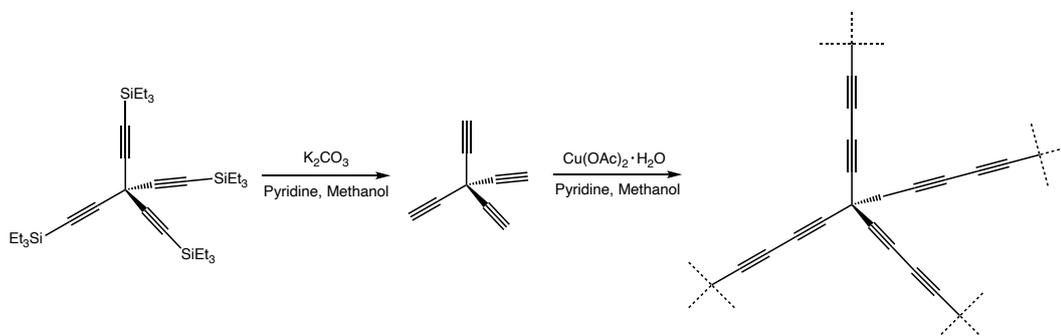


Figure 5.1: a) Lithium-ion battery set-up with a LiCoO_2 cathode and graphite anode, with the charging and discharging half-equations also given.³²⁷ b) Bottleneck effect leading to short-circuit. Key: lithium – purple, CoO_2 polyhedra – orange and red, carbon – grey. Figure a) inspired by reference 327.

To improve the current battery set-up, alternative materials must be sourced. Porous carbon materials are commonly used electrodes in LIBs³³⁵⁻³³⁹ due to their high porosity,³⁴⁰ which often yields high electron conductivity.^{341,342} However, porous carbon was traditionally made either by physical or chemical activation to give activated carbon,^{337,343,344} templating using molecular sieves such as zeolites,^{341,342,345-347} or laser ablation³⁴² to remove the metal from metal carbides, with no clear synthetic route to prepare porous carbon directly.^{16,348,349}

With this in mind, Zhao and co-workers developed a new material, named organically synthesised porous carbon (OSPC)-1 (Scheme 5.1).¹⁶ This material, first reported in 2018, was specifically designed to build a three-dimensional microporous carbon framework directly from tetraethynylmethane building blocks using the radical-based Eglinton homocoupling reaction (Figure 5.2).^{16,350}



Scheme 5.1: Reaction scheme to synthesise OSPC-1. Figure reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

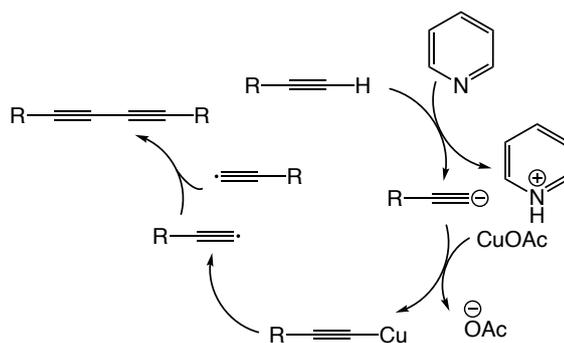


Figure 5.2: Catalytic cycle for the Eglinton homocoupling reaction. Figure adapted from reference 95, upon which this chapter is based.

The Eglinton homocoupling reaction, as shown in Figure 5.2, begins with the alkyne reacting with base to form an alkyne anion, and the protonated base. The anion then reacts with copper(I) acetate, which dissociates to allow the anion to pick up the Cu⁺ cation, leaving the acetate anion as a by-product. The copper-containing alkyne further reacts with copper(II) acetate, yielding an alkyne radical and two equivalents of copper(I) acetate. Finally, termination occurs when two radicals react to give the homocoupled product.^{16,350} Whilst this approach has previously been utilised for the synthesis of metal-organic and covalent organic frameworks, which exhibit permanent porosity due to their clearly defined pore structures, this was the first direct synthesis of porous carbon.¹⁶

OSPC-1 exhibited good chemical and thermal stability.¹⁶ The material was confirmed as amorphous by the powder x-ray diffraction (PXRD) pattern, which showed a broad hump at approximately 20° (Figure 5.3a). The deconvoluted C1s x-ray photoelectron spectroscopy (XPS) peak at 285 eV confirmed that the expected structure had formed. OSPC-1 was composed of *sp*³ carbon nodes (284.5 eV) connected by *sp* carbon linkers (285.8 eV).³⁵¹ The extra peak at 288.4 eV corresponded to carbon bonded to oxygen, rationalised due to the unreacted terminal alkynes at the surface of the material reacting with atmospheric oxygen (Figure 5.3b).¹⁶ The ¹³C solid-state nuclear magnetic resonance (ssNMR) spectrum showed broad peaks at 133 ppm and 53 ppm, along with a peak corresponding to the NMR rotor at 111 ppm (Figure 5.3c). Whilst at first glance the peak at 133 ppm appeared to be in the region expected for *sp*² carbon atoms, there was no peak at 285.05 eV in the XPS spectrum (corresponding to *sp*² carbon).^{16,351,352}

The Electron Energy Loss (EELs) spectrum of OSPC-1 (Figure 5.4a) gave peaks at 285.8 eV and 292.8 eV corresponding to *sp* and *sp*³ carbon atoms respectively.^{16,353–357} The Fourier-transform infrared (FT-IR) spectrum (Figure 5.4b) showed a peak at 2150 cm⁻¹ ascribed to C≡C bonds, indicating complete polymerisation.^{16,358}

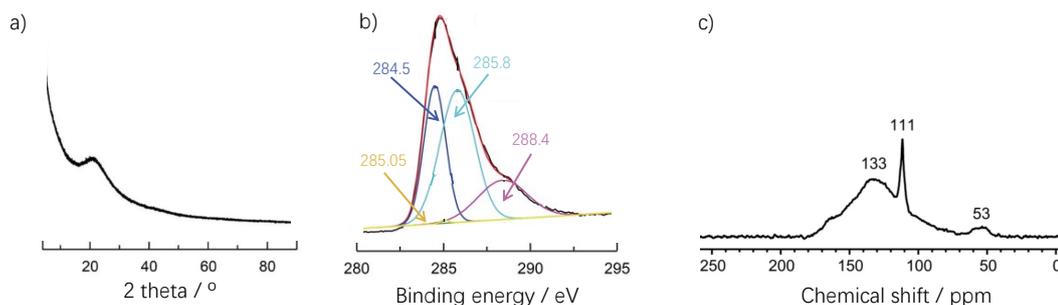


Figure 5.3: Experimental OSPC-1 characterisation. a) XRD spectrum. b) C1s XPS spectrum. c) ¹³C ssNMR spectrum. Figure reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

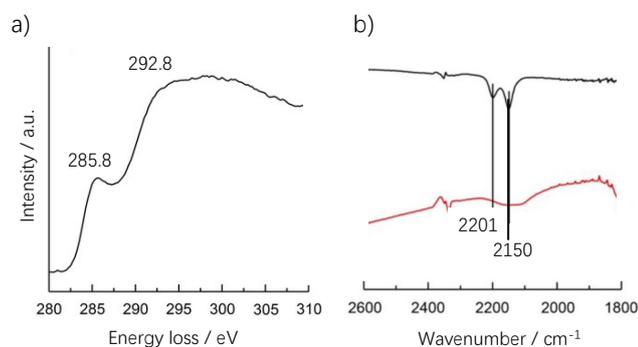


Figure 5.4: Experimental OSPC-1 characterisation. a) EELs spectrum. b) FT-IR spectrum. Key: triethylsilyl (TES)-protected tetraethynylmethane – black, OSPC-1 – red. Figure reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The potential of OSPC-1 to be used as an anode for LIBs was then analysed. The electron conductivity of $1.2 \times 10^{-4} \text{ S cm}^{-1}$ classifies this material as a semiconductor. The conductivity is lower than those of graphene and graphite (up to $1.09 \times 10^5 \text{ S cm}^{-1}$ and $2\text{--}3 \times 10^3 \text{ S cm}^{-1}$, respectively), as the sp^3 carbon nodes hinder the electron conduction.^{16,326,359} It was hypothesised that the high conductivity of OSPC-1 is obtained due to quantum tunnelling of electrons between sp orbitals on neighbouring struts. This explained why this, unlike the insulating diamond allotrope of carbon, can conduct electrons.¹⁶ The reversible lithium-ion capacity of OSPC-1 was calculated as 748 mA h g^{-1} after 100 cycles at a current density of 200 mA g^{-1} . This value exceeded those of fullerene, carbon nanotubes and graphite (2.5 mA h g^{-1} , 224 mA h g^{-1} and 324 mA h g^{-1} , respectively).¹⁶ Following this, overcharge experiments at a current density of 5000 mA g^{-1} for 6 minutes were undertaken to compare an OSPC-1 anode to a graphite anode, after which the samples were analysed using scanning electron microscopy (SEM), a surface-sensitive technique.¹⁸¹

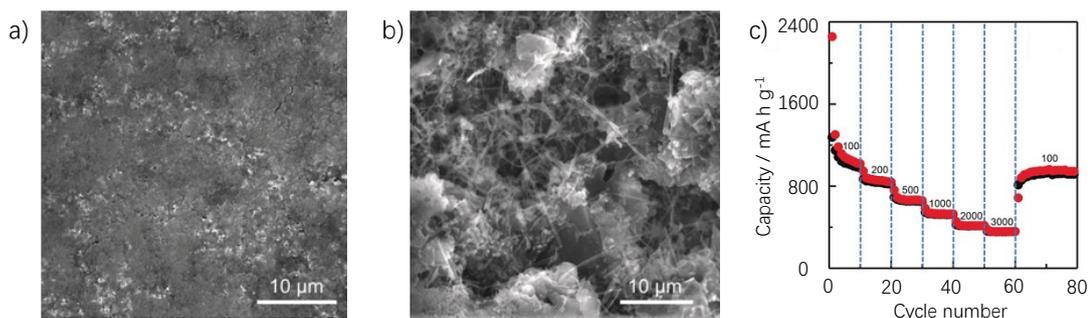


Figure 5.5: SEM after overcharging experiments of the a) OSPC-1 anode and b) graphite anode, with a scale bar of 10 μm. c) Waterfall plot demonstrating the rate capability of the OSPC-1 anode. The current densities at each step in the plot are given as insets, with units of mA g⁻¹. Figure reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

No evidence of dendrites on the surface of OSPC-1 was visible (Figure 5.5a), compared to clear dendrite formation on the surface of the graphite anode (Figure 5.5b). Additionally, the samples were analysed by XRD, a bulk technique, after 100 cycles at a current density of 200 mA g⁻¹, again showing no evidence of dendrites in the OSPC-1 cell and clear dendrite formation in the graphite cell. These findings were rationalised as being due to the larger, three-dimensional pore structure of OSPC-1 allowing more facile diffusion of lithium ions through the material.¹⁶ The rate capability of OSPC-1 was also tested to assess the capacity of the material at current densities ranging from 100–3000 mA g⁻¹. High capacities after ten cycles were observed throughout, even at the highest current density of 3000 mA g⁻¹ (356 mA h g⁻¹). On returning to the lowest current density (100 mA g⁻¹) to begin the second cycle through the waterfall plot, the capacity returned to the initial value of 944 mA h g⁻¹, demonstrating the stability of the OSPC-1 anode to repeated cycling (Figure 5.5c).¹⁶

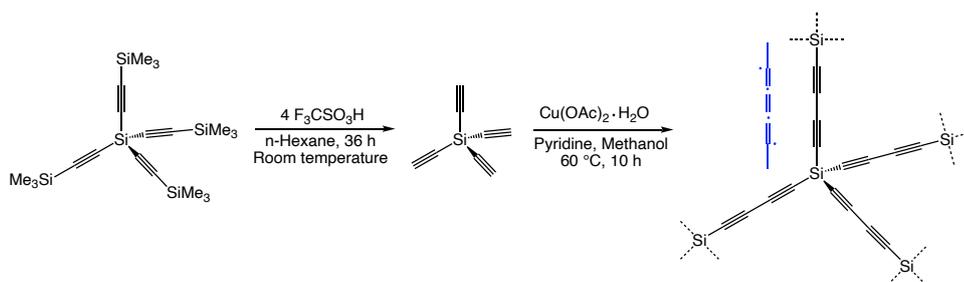
This first report of OSPC-1 established a new synthetic route to a new allotrope of porous carbon composed of *sp* and *sp*³ hybridised carbon atoms that showed good electron conductivity.¹⁶ OSPC-1 appeared to be a promising anode material for LIBs, with a high specific capacity and good rate capability. Additionally, OSPC-1 has the ability to be charged over a large number of cycles at a high rate without any evidence of harmful dendrite formation.¹⁶ Following this, further research assessed the potential of OSPC-1 to be used as an anode for renewable batteries with alternative group 1 and 2 ions, all of which are more abundant (Table 5.1) and less expensive than lithium and are therefore appealing in the future of renewable battery technology.³²⁵ Additionally, since the original 2018 report of OSPC-1,¹⁶ Yang and co-workers have reported similar materials where the *sp*³ carbon nodes were replaced with silicon and germanium atoms to form Si-DY and Ge-CDY, respectively.^{360,361}

Table 5.1: Ionic radii and abundances of lithium, sodium, magnesium, potassium, and calcium ions. Data obtained from references 330 and 332.

Ion	Ionic radius / Å	Abundance / %
Li ⁺	0.90	0.002
Na ⁺	1.16	2.360
K ⁺	1.52	2.090
Mg ²⁺	0.86	2.330
Ca ²⁺	1.14	4.150

Similarly to OSPC-1, these materials showed promising applications as battery anodes for lithium-ion batteries, with high capacities due to the formation of the Si-DY or Ge-CDY film and good capacity retention even after a large number of cycles at a high current density.^{360,361} Si-DY also showed promise as an anode for sodium and potassium-ion batteries, with respectable specific capacities of 253 mA h g⁻¹ after 3500 cycles at 2000 mA g⁻¹ for the sodium-ion battery, and 90 mA h g⁻¹ after 500 cycles at 1000 mA g⁻¹ for the potassium-ion battery. Raman spectroscopy of the Ge-CDY system showed that the structure had not degraded with cycling, suggesting that these materials do not suffer with the same limited stability over time as graphite.^{360,361}

Following on from the reports of OSPC-1, Si-DY and Ge-CDY, a new analogue of Si-DY, termed Si-DA, was synthesised by Dong and co-workers using the Eglinton homocoupling reaction³⁵⁰ (Figure 5.2) used to prepare OSPC-1.³⁶² Si-DA incorporates a persistent radical into the butadiyne linker, as shown in Scheme 5.2, and was the first report of an 'OSPC-type' polymer containing a persistent radical. As for OSPC-1, Si-DY and Ge-CDY, Si-DA was amorphous as judged by PXRD and fully polymerised, with the presence of a C≡C peak at 2150 cm⁻¹ and absence of a C≡H peak at 3284 cm⁻¹ in the FT-IR spectrum. The ¹³C ssNMR spectrum showed a broad resonance centred at 126 ppm which was ascribed to the *sp*² carbon atoms within the open-shell framework.³⁶²



Scheme 5.2: Reaction scheme to synthesise Si-DA. The blue inset shows the proposed diradical open-shell strut within Si-DA. This adapted scheme has been published in *CCS Chemistry* 2022; A Three-Dimensional Silicon-Diacetylene Porous Organic Radical Polymer and is available online at [10.31635/ccschem.022.202202351](https://doi.org/10.31635/ccschem.022.202202351); <https://www.chinesechem-soc.org/doi/10.31635/ccschem.022.202202351>.

Around the same time that the silicon- and germanium-based polymers were reported, Heasman and co-workers reported the theoretical extension of the OSPC-1 family to include new members with varying strut lengths, which is possible within the experimental synthesis.³⁶³ This was the first report of OSPC-0, -2 and -3, with the structures given in Figure 5.6, and the varying strut lengths led to differences in the uptake and diffusion of ions, along with the solid electrolyte interface (SEI) formation.³⁶³

This first exploration suggested that OSPC-0 had the highest overall capacity due to its charging capacity being comparable with that of OSPC-1, whilst its discharging capacity was the smallest of the OSPC family as a result of the limited porosity within OSPC-0.³⁶³ However, also as a result of the limited porosity within this material, it had the slowest diffusion rate under charging conditions. OSPC-3, meanwhile, had the fastest lithium diffusion rate under charging conditions due to the maximised flexibility and porosity, yet had a slightly lower uptake of lithium as given by the lower capacity. This suggested that to maximise uptake, OSPC-0 is the optimal material, whilst to maximise charging rate with a slight reduction in overall capacity, OSPC-3 is the ideal.³⁶³

Whilst the OSPC family of materials shows intriguing electrochemical properties, one of the major hindrances to scaling up these reactions is the complex, multi-step mechanism used to synthesise the polymer frameworks, requiring relatively harsh conditions to synthesise the monomer precursors and generating the product in low yield.^{323,364} To circumvent this challenge, Wang and co-workers established a novel, one-pot synthesis of OSPC-0 (Scheme 5.3, Figure 5.7) based on a nucleophilic condensation reaction. This involves the reaction of commercially available bis(trimethylsilyl)ethyne with tetrabromomethane to sequentially add each linker to the node, in the presence of caesium fluoride, where the fluoride ion acts as the activating agent for the reaction, and diphenyl sulfone solvent.³⁶⁴

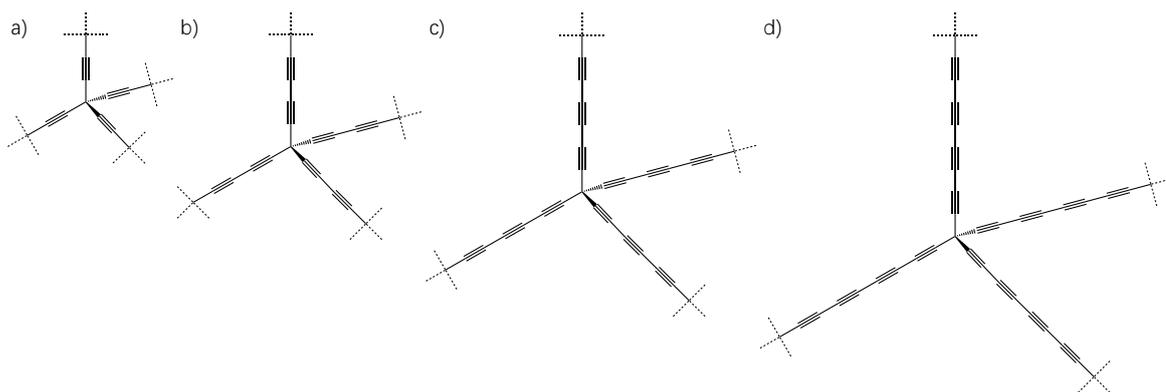
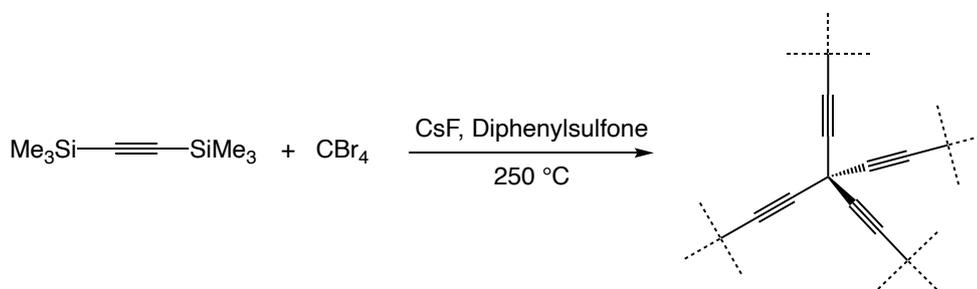


Figure 5.6: The OSPC family. a) OSPC-0. b) OSPC-1. c) OSPC-2. d) OSPC-3. Figure adapted with permission from reference 363, licensed under CC-BY. Copyright 2022 The Authors.



Scheme 5.3: One-pot synthesis of OSPC-0, first reported by Wang and co-workers. Figure reproduced (adapted) from references 95, upon which this chapter is based, from reference 364 with permission from John Wiley and Sons. Copyright 2021 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

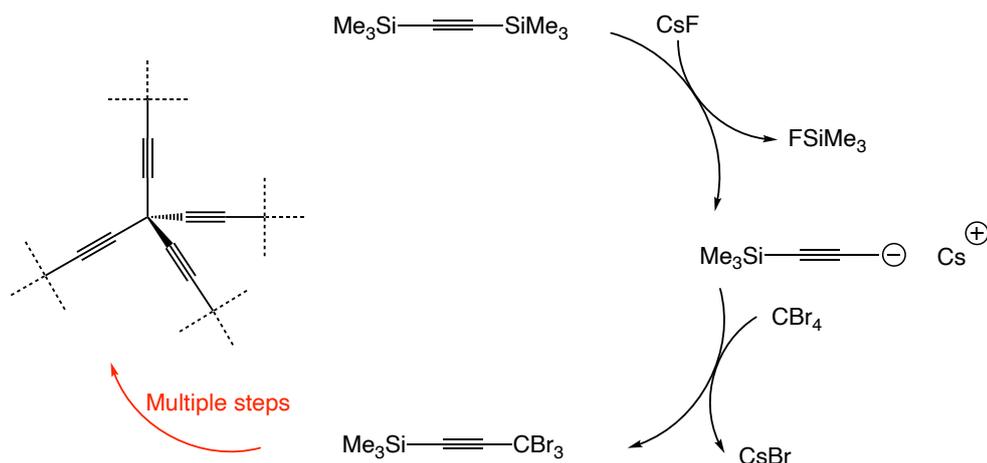


Figure 5.7: Mechanism for the one-pot synthesis. Figure reproduced (adapted) from references 95, upon which this chapter is based, from reference 364 with permission from John Wiley and Sons. Copyright 2021 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

This synthetic method, which occurs by adding all of the reagents for the reaction simultaneously, does not allow for the formation of sp^2 hybridised carbon. The XPS and ^{13}C ssNMR spectra of the one-pot OSPC-0 reported by

Wang and co-workers indicate sp^3 and sp hybridised carbon (peaks centred at 50 ppm and 130 ppm, respectively), consistent with OSPC-1.^{16,364} The one-pot OSPC-0 is also amorphous and microporous, with a type IV isotherm and a BET surface area of $453 \text{ m}^2 \text{ g}^{-1}$.^{16,364}

The one-pot synthetic approach first reported by Wang and co-workers³⁶⁴ was later adapted to produce OSPC-1. This is henceforth referred to as OSPC-1b to differentiate from OSPC-1a, reported in 2018 using the multistep synthetic method.^{16,323} OSPC-0b was also synthesised to analyse the electrochemical properties of OSPC-0b and OSPC-1b for the first time.³²³ Due to the different synthetic routes used to prepare OSPC-1a and OSPC-1b, there are different potential end groups. Terminal alkynes in OSPC-1b could be terminated with bromine atoms and/or trimethylsilyl end groups, which are not possible within the multistep synthesis used to prepare OSPC-1a.³²³

As expected from OSPC-1a and the one-pot OSPC-0 synthesised by Wang and co-workers,^{16,364} OSPC-1b and OSPC-0b both showed only sp^3 and sp hybridised carbon in the deconvoluted C1s XPS peak and ^{13}C ssNMR spectra, with peaks in the ^{13}C ssNMR appearing at 133 ppm (sp C) and 54 ppm (sp^3 C) for OSPC-0b and OSPC-1b (Figure 5.8) and a small peak indicating C-Br interactions in the C1s XPS spectrum. In the OSPC-1b ^{13}C ssNMR spectrum, there was also a peak at 0 ppm ascribed to unreacted trimethylsilyl groups.³²³

The capacities of OSPC-0b and -1b were found to be 440 mA h g^{-1} (OSPC-0b) and 502 mA h g^{-1} (OSPC-1b), after 200 cycles at a current density of 200 mA g^{-1} . The decreased capacity for OSPC-1b compared to OSPC-1a was rationalised by the difference in synthetic route, which could affect the SEI formation of the resulting polymer.³²³

Finally, the lithium dendrite formation behaviour of OSPC-1b in comparison to graphite was assessed by over-discharging experiments, followed by analysis of the half-cell using SEM.³²³ The dendrite formation behaviour of the OSPC-1b and graphite was also assessed using XRD after stress testing the cells at a current density of 5000 mA g^{-1} for 1000 cycles. The SEM images and XRD spectrum showed no evidence of metallic lithium for OSPC-1b, consistent with OSPC-1a, further confirming the lack of dendrite formation for OSPC-1. In comparison, the SEM image of the graphite anode gave clear evidence of lithium in a globular topology, the precursor to lithium dendrite formation.³²³ The XRD spectrum indicated that the graphite lattice had degraded due to the repetitive charging and discharging cycles to allow intercalation of lithium from the structure, and peaks arising from metallic lithium were also observed.³²³

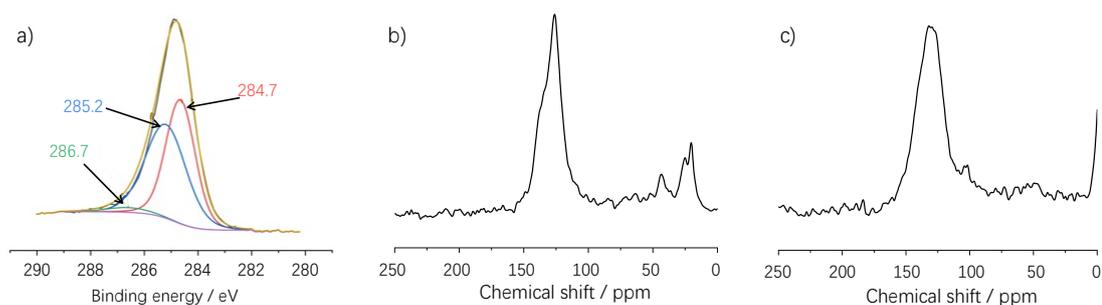


Figure 5.8: a) C1s XPS spectrum of OSPC-1b. b) and c) ^{13}C ssNMR spectra of b) OSPC-0b and c) OSPC-1b, obtained at a magic-angle spinning rate of 25 kHz to enhance the signal to noise ratio. Figure adapted from references 95, upon which this chapter is based, and 323.

Due to a high specific capacity that is twice that of the industry standard for lithium-ion batteries, combined with good chemical and thermal stability and an apparent lack of lithium dendrite formation, OSPC-1 has established itself as a strong contender future anode material for renewable batteries.¹⁶ Similar materials have been reported using OSPC-1 as the inspiration, replacing the sp^3 hybridised carbon node with silicon and germanium atoms, respectively,^{360,361} and the radical properties of a silicon-based OSPC-type framework have also been explored.³⁶² Simulations have reported that alternative ions such as calcium, magnesium, sodium, and potassium, which have larger ionic radii than lithium but are much more abundant, can also intercalate into and diffuse through the OSPC-1 pore structure.³²⁵ The family of OSPC materials can be easily tailored to extend or decrease the length of alkyne struts, tailoring the properties of the resulting material,³⁶³ and a new synthetic method for OSPC-1 and OSPC-0 allows these materials to be prepared using a one-pot approach.^{323,364}

However, despite this, the question still arises regarding the ^{13}C ssNMR spectrum of the OSPC-1 material, which at first glance suggests that the structure is composed of sp^2 and sp^3 hybridised carbon, a direct contradiction to the XPS and EELS spectra of OSPC-1a which characterise the material as sp and sp^3 hybridised carbon; and the one-pot synthetic method of OSPC-1b which does not allow for sp^2 hybridised carbon, yet shows peaks in similar regions of the ssNMR spectrum to OSPC-1a.^{16,323}

The ^{13}C ssNMR peaks for OSPC-1a and OSPC-1b were centred at 54 ppm (peak 1) and 133 ppm (peak 2),^{16,323} with an artefact peak from the NMR rotor at 111 ppm. From analysis of other reported sp^3 ,³⁶⁵⁻³⁷⁰ sp^2 ,^{236,365,367-371} and sp hybridised carbon materials,^{236,365,369,372} peak 1 at 53 ppm and part of peak 2 at ~ 130 ppm were within the range of obtained values for sp^3 carbon.⁹⁵ Peak 2 appeared directly in the range of reported values for sp^2 hybridised carbon. The obtained values for sp carbon appear in multiple distinct regions

of the ^{13}C ssNMR spectrum, yet only cover a small section of peak 1 and the part of peak 2 appearing at low-medium chemical shift values with respect to the tetramethylsilane reference (Figure 5.9, the full list of reported chemical shift values is given in Appendix C Tables C.1 and C.2).^{95,236,365–379}

This chapter demonstrates the use of calculated NMR spectra of representative fragments and periodic models of the OSPC-1 framework to rationalise the experimental ^{13}C ssNMR spectrum and further understand this anode material.^{16,95} Firstly, the NMR spectra of a series of molecular fragments of the proposed OSPC-1 structure were calculated using Gaussian 09.²¹² Following this, the NMR spectra were then obtained for OSPC-1 within a selection of crystalline topologies using CASTEP 8.0.²¹³ As these small fragments and crystalline models were unable to mimic the experimental ^{13}C NMR spectrum for OSPC-1, a series of chains of OSPC-1 were then modelled at varying unit cell dimensions (also using CASTEP 8.0)²¹³ to assess the influence of chain-chain interactions on the resulting NMR spectrum.⁹⁵

It was established that interactions between neighbouring chains were pivotal to rationalising the experimental ^{13}C ssNMR spectrum of OSPC-1.^{16,95} However, the resulting dense and non-porous model was unable to replicate the experimental surface areas of $700\text{--}900\text{ m}^2\text{ g}^{-1}$.^{16,323} This was rationalised by a proposed spherical particle of OSPC-1, containing a majority of dense, non-porous OSPC-1, surrounded by a thin shell of open, highly porous OSPC-1.⁹⁵

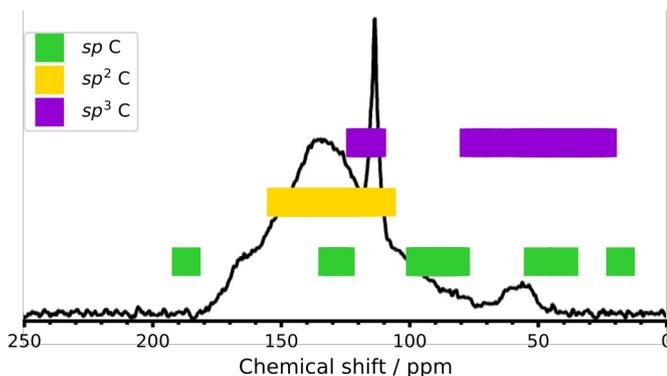


Figure 5.9: Reported chemical shift values for various sp^3 ,^{365–370} sp^2 ,^{236,365,367–371} and sp hybridised carbon materials,^{236,365,369,372} overlaid onto the OSPC-1a ^{13}C ssNMR spectrum reported by Zhao and co-workers.¹⁶ Key: sp^3 carbon – purple, sp^2 carbon – yellow, sp carbon – green. Data obtained from reference 95, upon which this chapter is based. Experimental ssNMR spectrum reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

5.3 Methodology

Within this chapter, three classes of model were generated to rationalise the experimental ^{13}C ssNMR spectrum. Firstly, a series of small fragments of the OSPC-1 structure were modelled. These included the OSPC-1 monomer, dimer and tetramer, macrocyclic rings of OSPC-1, and the OSPC-1 tetramer with the various potential end groups for OSPC-1a and OSPC-1b. These models are referred to as the 'OSPC-1 fragment models'.⁹⁵

Secondly, the OSPC-1 structure was modelled within a range of periodic crystalline topologies. Each topology contained a tetrahedral node, as for the proposed OSPC-1 structure, with varying degrees of network interpenetration. These models are referred to as the 'OSPC-1 periodic models - crystalline topologies'.⁹⁵

Finally, periodic chain models of the OSPC-1 structure were assembled. These were composed of a chain of OSPC-1 lying along the x -axis. The y and z cell dimensions were then systematically varied to assess the influence of chain-chain interactions on the resulting calculated NMR spectra. Following this, another model was generated, composed of two perpendicular chains, with the third unit cell dimension varied systematically. The potential end groups for OSPC-1a and OSPC-1b were also considered within these models, which are referred to as the 'OSPC-1 periodic models - chain structures'.⁹⁵

5.3.1 OSPC-1 fragment models

The Eglinton homocoupling³⁵⁰ of tetraethynylmethane to synthesise OSPC-1a occurred *in situ* from a triethylsilyl (TES)-protected tetraethynylmethane monomer, meaning that the unprotected monomer was not isolated¹⁶ and could not be used for comparison to the calculated models. Instead, the TES-protected monomer (Figure 5.10) was utilised.⁹⁵

The geometry of the TES-protected monomer was calculated by optimising the structure using the B3LYP/6-311G model chemistry within Turbomole 7.5.^{79,85,90,91,380,381} The harmonic frequencies were calculated using the same model chemistry and a ground state was classified based on the result giving no imaginary frequencies.^{79,85,90,91,380} The calculated NMR spectrum was subsequently calculated using the same model chemistry in Gaussian 09,²¹² with the shielding tensors calculated using the single gauge origin, CSGT, and IGAIM (atomic centres as gauge origin) approaches.^{79,85,90,91,212,380,382-384} Spin-spin coupling constants were calculated as described by Deng and co-workers.³⁸⁵ The isotropic chemical shieldings of each carbon atom were then shifted relative to the tetramethylsilane standard using Equation 5.1, where 189.72 ppm corresponded to the isotropic chemical shielding of tetramethyl-

silane calculated within Gaussian 09 using the B3LYP/6-311G model chemistry.^{79,85,90,91,95,212,380}

$$\text{Chemical shift (ppm)} = 189.72 \text{ ppm} - \text{isotropic chemical shielding} \quad (5.1)$$

When comparing the experimental spectrum (Figure 5.10a) to the calculated (Figure 5.10b), both showed environments corresponding to the *sp* carbon atoms in the alkyne struts, the *sp*³ carbon node, and the TES protecting groups (experimental chemical shift values of 101 ppm, 83 ppm, 24 ppm, 7 ppm and 4 ppm, respectively).⁹⁵ The calculated peaks appeared at higher chemical shift values than the experimental, and the calculated chemical shifts were therefore shifted to calibrate the calculated spectrum to the experimental (Equation 5.2, Figure 5.10c). There is no accepted explanation for why the calculated and experimental chemical shifts differ, however, this method of calibration has been reported in references 33 and 386. The values used to calibrate the calculated chemical shifts in Equation 5.2 were obtained by plotting the experimental chemical shift of each carbon atom environment against the calculated, and subsequently calculating the equation of the linear trendline.⁹⁵ It was established that the presence of solvent within the calculated structure had little influence on the NMR spectrum,⁹⁵ therefore, the structures of all remaining models were calculated in the absence of solvent.

$$\text{Calibrated chemical shift (ppm)} = (0.95 \times \text{chemical shift}) - 4.17 \quad (5.2)$$

With the calibration equation in place (Equation 5.2), the NMR spectra of representative fragment models of the OSPC-1 structure were calculated using Gaussian 09,²¹² utilising the same methodology as for the TES-protected monomer, with the geometry optimisation of all remaining fragment models also being undertaken using Gaussian 09.^{95,212}

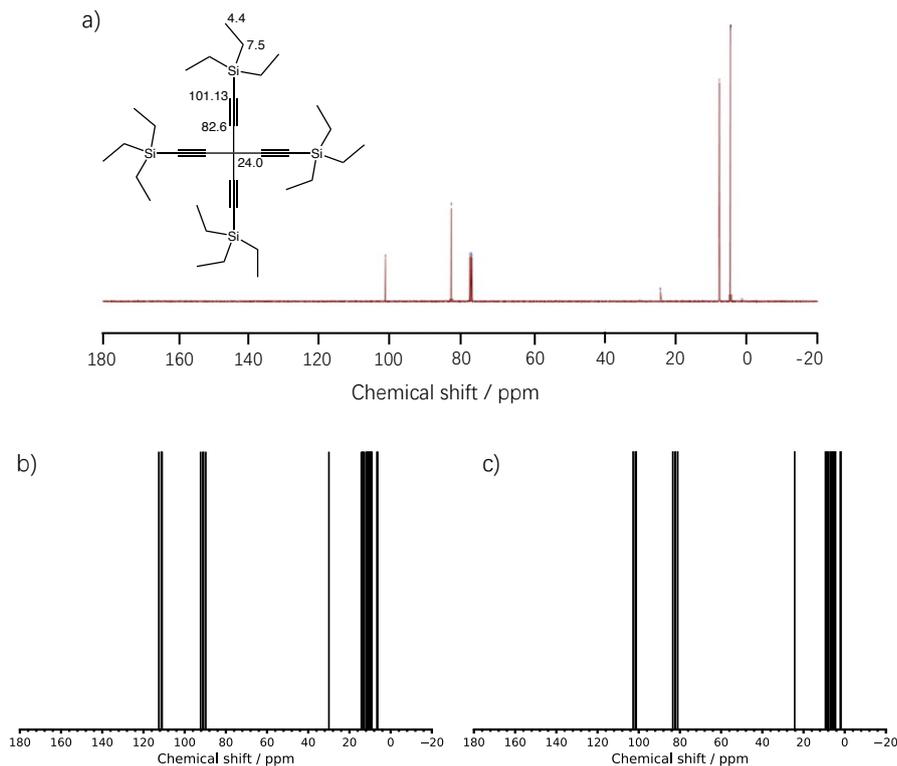


Figure 5.10: ^{13}C NMR spectrum of TES-protected tetraethynylmethane. a) Experimental spectrum, obtained within CDCl_3 solvent.¹⁶ b) NMR spectrum calculated using the B3LYP/6-311G model chemistry within Gaussian 09,^{79,85,90,91,212,380} in the presence of CHCl_3 solvent. c) NMR spectrum calculated using the B3LYP/6-311G model chemistry within Gaussian 09,^{79,85,90,91,212,380} in the presence of CHCl_3 solvent, with each chemical shift value calibrated relative to the experimental TES-protected monomer using Equation 5.2. Figure adapted from reference 95, upon which this chapter is based.

5.3.2 OSPC-1 periodic models

To further understand the OSPC-1 structure, periodic models were generated using CASTEP 8.0,²¹³ a plane-wave density functional theory code. Each structure underwent a geometry optimisation^{66,68,86,97,213,219-221} followed by an NMR calculation^{66,68,86,95,97,213,215,216,218,220,222} using the Perdew-Burke-Ernzerhof (PBE)⁸⁶ exchange-correlation functional. Ultrasoft pseudopotentials were used to describe the core-valence interactions.^{218,219} The all-electron wavefunction for the NMR calculations was constructed using the gauge-included projector augmented wave (GIPAW) algorithm in the presence of a magnetic field.^{215,216} A plane-wave basis set cut-off energy and k-point convergence test was undertaken to ascertain the minimum energy and Monkhorst-Pack grid (the number of k-points required to sample the Brillouin zone)²²⁰ needed to converge the energy and isotropic chemical shieldings. Please refer back to Chapter 2 section 2.2.2 for an explanation of the purpose of the cut-off parameters. For the OSPC-1 models these values were 50 Ry and $1 \times 1 \times 1$, respectively, with a gamma point of $0.5 \times 0.5 \times 0.5$.⁹⁵

As for the fragment models calculated using Gaussian,²¹² chemical shifts were calculated relative to tetramethylsilane, as shown in Figure 5.11, with the reference calculated by placing a single tetramethylsilane molecule into a unit cell of initial length 8 Å. The unit cell length was chosen to be intentionally large relative to the size of the tetramethylsilane molecule to inhibit interactions across the periodic boundaries.⁹⁵ Following the geometry optimisation, which increased the unit cell length to 8.6 Å, the calculated NMR shieldings were compared to the experimental.⁹⁵ Following the comparison, the tetramethylsilane shielding value of 178.92 ppm was utilised. The TES-protected monomer was optimised within unit cell lengths of 20 Å and 30 Å. Following the geometry optimisation, the calculated NMR shieldings were compared to the experimental.⁹⁵ The TES-protected monomer within the 30 Å unit cell was utilised as the calculated NMR shieldings gave a closer match to the experimental. The chemical shifts of the calculated periodic models were calibrated relative to the TES-protected tetraethynylmethane monomer using Equation 5.3. Figure 5.12 shows the calibration of TES-protected tetraethynylmethane.⁹⁵

As for the OSPC-1 fragment models, the values used to calibrate the calculated chemical shifts in Equation 5.3 were obtained by plotting the experimental chemical shift of each carbon atom environment against the calculated, and subsequently calculating the equation of the linear trendline.

$$\text{Calibrated chemical shift (ppm)} = \text{Chemical shift} - 8.67 \quad (5.3)$$

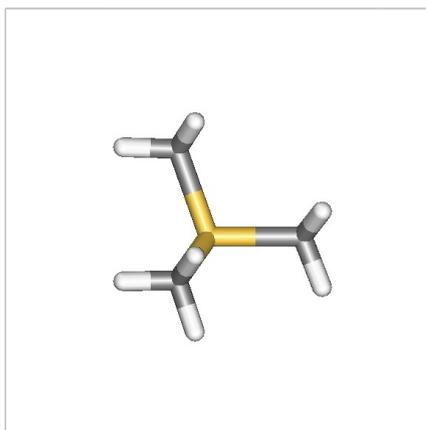


Figure 5.11: Structure of the tetramethylsilane reference. Key: carbon atoms - grey, hydrogen atoms - white, silicon atom - yellow.

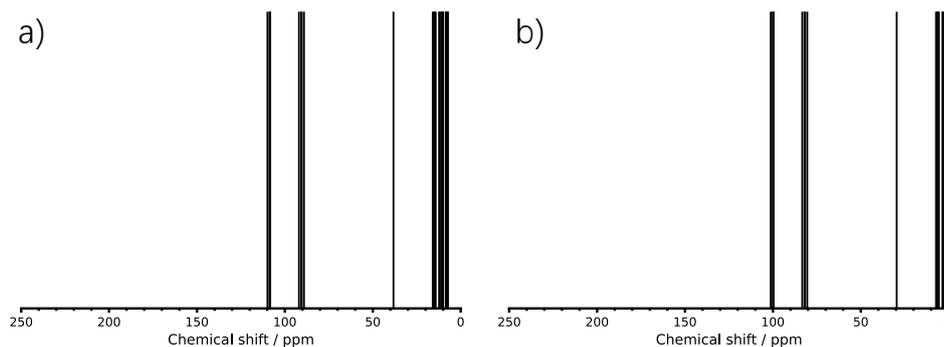


Figure 5.12: a) NMR spectrum of TES-protected tetraethynylmethane calculated using CASTEP 8.0.^{66,68,86,97,213,215,216,218,220,222} b) NMR spectrum of TES-protected tetraethynylmethane calculated using CASTEP 8.0.^{66,68,86,97,213,215,216,218,220,222} Each chemical shift value was calibrated relative to the experimental TES-protected monomer using Equation 5.3. Figure adapted from reference 95, upon which this chapter is based.

5.4 Results and Discussion

5.4.1 OSPC-1 fragment models

The first fragment, the tetraethynylmethane monomer (Figure 5.13a), showed three peaks at 78 ppm (*sp* C adjacent to *sp*³ C), 67 ppm (terminal *sp* C) and 34 ppm (*sp*³ C node), representing the three different carbon environments within the structure. The spectrum of a dimer composed of two bonded OSPC-1 monomers (Figure 5.13b) showed two additional peaks at 72 ppm and 69 ppm, corresponding to the ethynyl struts.⁹⁵

The repeat unit of OSPC-1 is composed of a tetrahedral structure composed of five *sp*³ carbon nodes linked by butadiynyl struts, as shown in Figure 5.13c. The NMR spectrum of the repeat unit showed a small number of peaks in the region of approximately 18–30 ppm, corresponding to the *sp*³ carbon nodes (peak 1 in the experimental spectrum), and a broad set of peaks in the range of 60–90 ppm due to the *sp* carbon atoms (peak 2 in the experimental spectrum).⁹⁵ Of this broad set of peaks, those at lower chemical shifts corresponded to *sp* carbon atoms in the saturated OSPC linkers, whilst the remaining peaks were ascribed to *sp* carbon atoms in the unreacted alkyne end groups. However, whilst the overall shape of this spectrum compares relatively well to the experimental ¹³C ssNMR spectrum of OSPC-1a, the chemical shifts of the repeat unit were lower than the experimental.⁹⁵ This may be due to the model chemistry utilised for the calculated models.

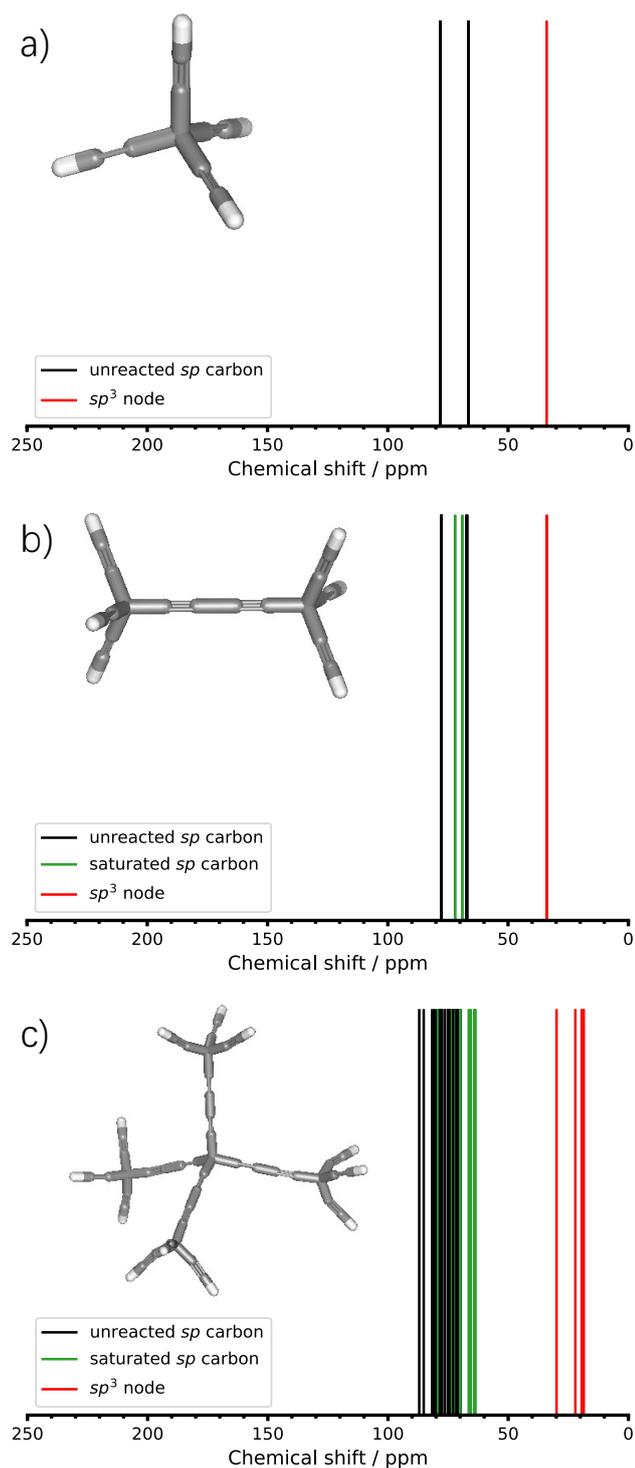


Figure 5.13: Calculated ^{13}C NMR spectra of a) Tetraethynylmethane monomer. b) Dimer of tetraethynylmethane. c) Tetrahedral 5-node repeat unit of OSPC-1. The structures are given as insets. Key: sp^3 carbon node – red, sp carbon in saturated OSPC-1 linker – green, sp carbon in unreacted alkyne – black. Figure adapted from reference 95, upon which this chapter is based.

As has been previously reported by Zhao and co-workers,¹⁶ OSPC-1 is an amorphous framework formed under kinetic control. This means that there is an inherent flexibility in the framework, which will not form the desired ther-

modynamic product throughout, and the node-to-node angles will vary from the idealised 109.5° .⁹⁵ Therefore, to incorporate this flexibility into the modelling approach, a scan of the tetramer was undertaken, varying one node-to-node angle in increments of $\pm 5^\circ$ to generate structures with node-to-node angles in the range of 59.5° – 159.5° . It could be seen that whilst the minimum energy structure corresponded to the repeating unit where all four node-to-node angles were at the expected value of 109.5° , there was a relatively small energy penalty to vary one of the node-to-node angles by $\pm 50^\circ$, along with a slight shift in the peaks within the calculated NMR spectrum.⁹⁵

The small energy cost to varying the node-to-node angle and amorphous nature of OSPC-1 allowed the possibility for a range of structural features to form, as has been previously observed for covalent triazine frameworks and conjugated microporous polymers.^{45,386} These features included macrocyclic rings formed from three to six sp^3 carbon nodes (Figure 5.15). The calculated NMR spectra of these structures showed a shift in the peaks, particularly peak 1 which generally shifted downfield as the size of the macrocyclic ring increased (peak 1: dimer – 28 ppm, 6-node ring – 38 ppm).⁹⁵

In the experimental XPS spectra of OSPC-1a and OSPC-1b,^{16,323} environments other than the idealised sp^3 and sp carbon were observed. Within the mechanisms to synthesise OSPC-1a and OSPC-1b (Figures 5.2 and 5.7 respectively), it is possible that there may be unreacted alkynes or alternative end groups present within the frameworks, and it was therefore important to consider these within the calculated models. The NMR spectra of the OSPC-1 tetrahedral repeat unit, with one node replaced with each of the potential end groups for OSPC-1a and OSPC-1b, are given in Figure 5.16.⁹⁵

For OSPC-1a, alternative end groups included unreacted ethynyl groups, copper atoms that have reacted within the ethynyl group in the first step of the mechanism, carbon radicals that react with the copper-terminated alkyne in the second step of the mechanism, or oxygenated groups such as a ketylenyl radical ($-CCO$) or oxoketene ($-C(O)CO$),³⁸⁷ formed from the reaction of unreacted radicals with atmospheric oxygen.¹⁶ OSPC-1b was synthesised *via* an alternative mechanism and therefore different potential end groups were possible.^{323,364} Within this mechanism, the trimethylsilyl (TMS) end groups are replaced with caesium ions, which then subsequently react with tetrabromomethane, leading to bromine end groups.⁹⁵ A larger basis set, LanL2DZ,^{388–391} was required to model the caesium and copper end groups, with 6-311G³⁸⁰ used for all other atoms for consistency throughout the fragment models. The peaks in each spectrum do have peaks within the expected ranges for OSPC-1a and OSPC-1b, yet none is a perfect match to the experimental, suggesting that larger models would be required to fully rationalise the ssNMR spectra.⁹⁵

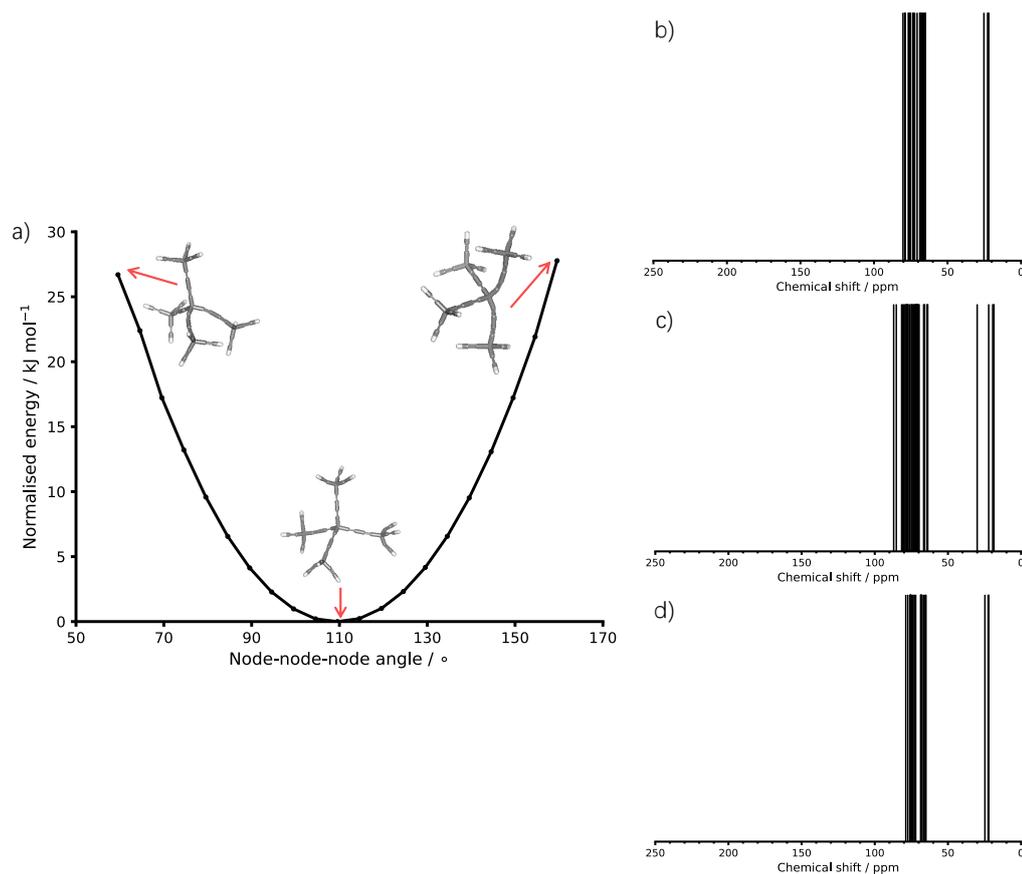


Figure 5.14: a) Plot of the potential energy on scanning one node-to-node angle in increments of $\pm 5^\circ$ to generate structures with node-to-node angles in the range of 59.5° – 159.5° , with the structures at 59.5° , 109.5° and 159.5° given as insets. b)–d) Calculated ¹³C NMR spectra of the tetrahedral 5-node repeat unit of OSPC-1 with a node-to-node angle of b) 59.5° . c) 109.5° . d) 159.5° . Figure adapted from reference 95, upon which this chapter is based.

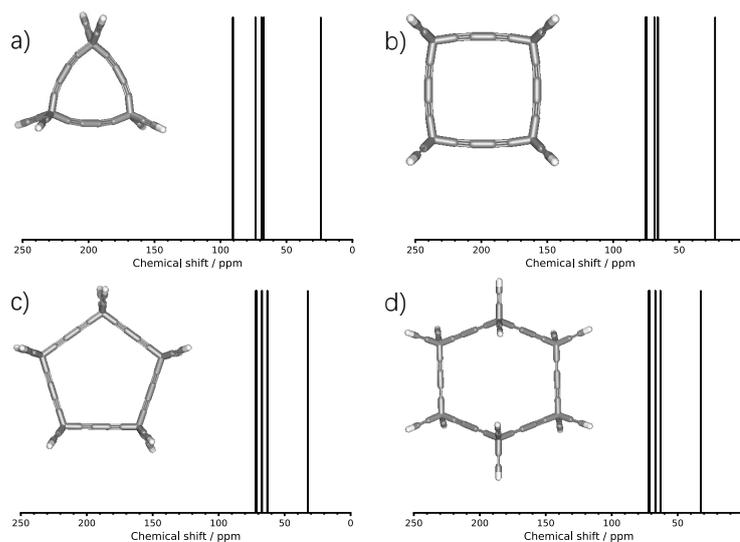


Figure 5.15: Calculated ¹³C NMR spectra of a) 3-node OSPC-1 macrocyclic ring. b) 4-node OSPC-1 macrocyclic ring. c) 5-node OSPC-1 macrocyclic ring. d) 6-node OSPC-1 macrocyclic ring. The structures are given as insets. Figure adapted from reference 95, upon which this chapter is based.

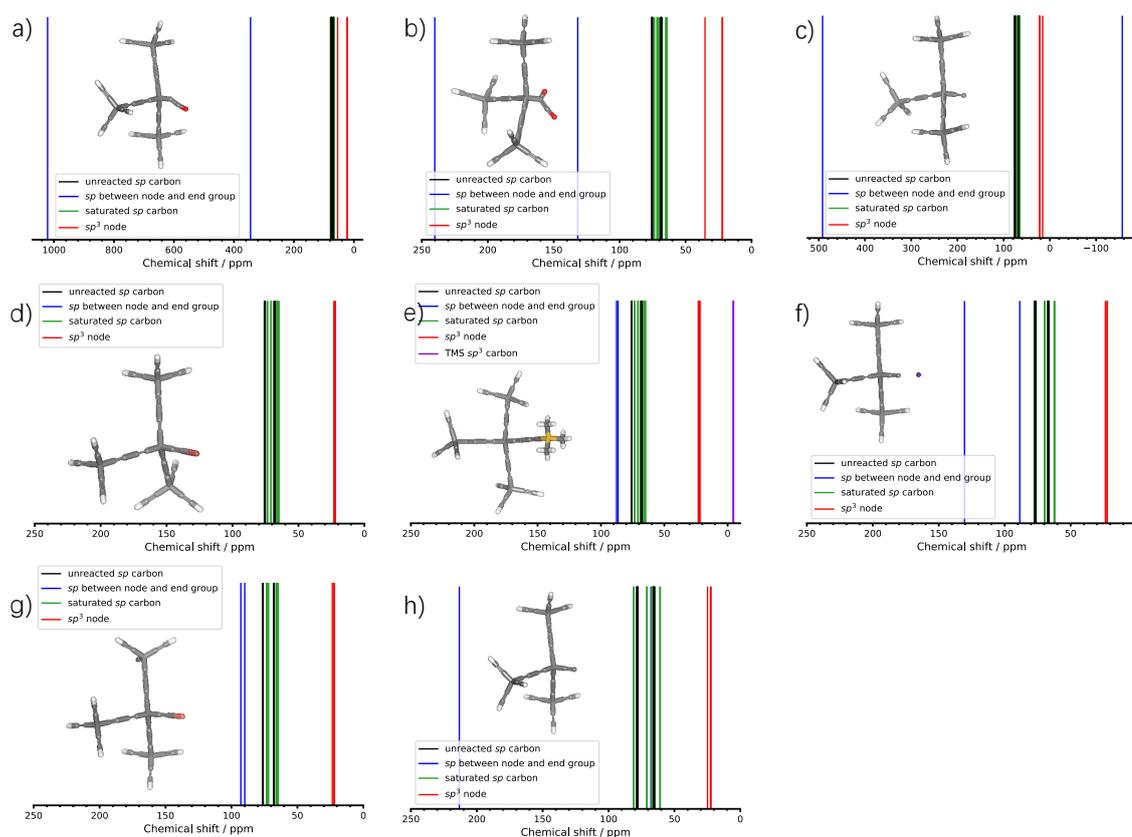


Figure 5.16: Calculated ^{13}C NMR spectra of the repeat unit of OSPC-1 with one node replaced with alternative end groups a) ketenyl radical ($-\text{CCO}$). b) oxoketene ($-\text{C}(\text{O})\text{CO}$). c) radical. d) bromine. e) TMS. f) OSPC-1 anion with a nearby caesium ion. g) OSPC-1 anion with a nearby copper ion. h) OSPC-1 anion. Key: sp^3 carbon – red, sp carbon in saturated OSPC-1 linker – green, sp carbon in unreacted alkyne – black, sp^3 carbon in TMS – purple, sp carbon between OSPC-1 node and alternative end group – blue. The structures are given as insets. Figure adapted from reference 95, upon which this chapter is based.

5.4.2 OSPC-1 periodic models - crystalline topologies

OSPC-1 is an amorphous framework, as shown by a broad hump in the PXRD pattern of OSPC-1a and OSPC-1b indicating a lack of order throughout.^{16,323} However, crystalline models could be generated to assess the relative energies and framework interactions of OSPC-1 within each topology.⁹⁵ OSPC-1 was modelled within the *dia*, *dia-c*, *dia-c3*, *dia-c4*, *dia-c6*, *pts-c* and *unc-c* topologies. A comparison of the structural nets for each of these topologies is given in Figure 5.17.^{392–401} Topologies are often referred to by a three-letter code relating to the full name of the topology, for example, the diamond topology is referred to as *dia*.⁴⁰² Each topology studied contains a tetrahedral node, as in the amorphous framework, with varying degrees of network interpenetration, or catenation. The suffix “-cN” indicates that a network is interpenetrated (or catenated) by N networks, with the “N” value typically omitted when N=2.⁴⁰² This means that the five modelled *dia* topologies, shown in Fig-

ure 5.21a–e, are all composed of OSPC-1 within the diamond topology. The *dia* structure contains one network within the diamond topology, the *dia-c* structure contains two interpenetrated *dia* networks, the *dia-c3* structure contains three interpenetrated networks, and so on. In contrast, the *pts-c* topology is composed of tetrahedral and square vertices. It is named after the shape of the PtS lattice, with two interpenetrated networks.⁴⁰² While there are other reported tetrahedral nets as observed within the Reticular Chemistry Structure Resource (RCSR) database,⁴⁰³ the studied topologies were chosen to give an indication of the relative energies from one topology to the next, rather than to suggest that OSPC-1 is a crystalline framework.

The relative energy per sp^3 carbon tetrahedral node (t-site) of OSPC-1 in each crystalline topology is plotted against the density in Figure 5.18a, where the energies were calculated relative to the lowest energy *dia-c* structure. Generally, the relative energy per t-site decreased with decreasing density, which was an unexpected trait, with energy vs. density plots such as these often showing more favourable structures with increasing densities for crystal structure prediction.⁴⁰⁴ This indicated that for OSPC-1, as has been previously reported,¹⁶ network interpenetration is unfavourable, leading to higher energy structures than the more open frameworks obtained at low densities. Of the seven topologies studied, the *dia*, *dia-c*, *pts-c* and *unc-c* structures retained the OSPC-1 bonding motif, whilst the *dia-c3* structure had some rearrangement, and the *dia-c4* and *dia-c6* structures broke down to form layered materials containing macrocyclic rings.⁹⁵

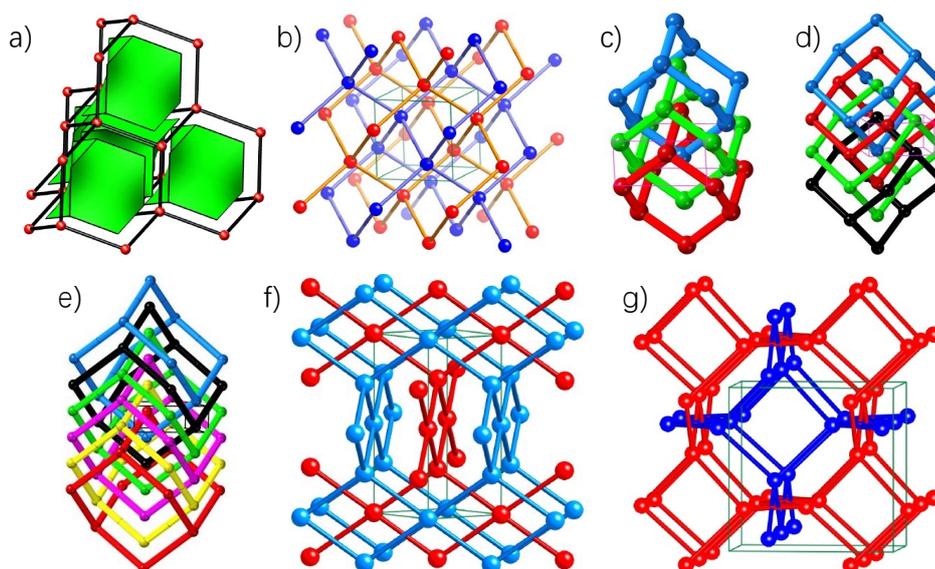


Figure 5.17: Images of the structural nets for the a) *dia*, b) *dia-c*, c) *dia-c3*, d) *dia-c4*, e) *dia-c6*, f) *pts-c* and g) *unc-c* topologies. The intercalated networks b)–g) show each network as a different colour for clarity. a) was taken from reference 392, adapted from references 399 and 400. b)–e) were taken from references 393, 394, 395 and 396, respectively, adapted from reference 401. f) and g) were taken from references 397 and 398, respectively.

The structures and resulting NMR spectra of OSPC-1 in each of the studied topologies are given in Figure 5.18b-h. Of the structures retaining the OSPC-1 bonding motif, the NMR spectrum of OSPC-1 in the *dia* topology most closely resembled that of the fragment models, with a peak at 27 ppm corresponding to the sp^3 hybridised carbon atoms, and a set of peaks ranging from ~ 65 – 73 ppm, ascribed to saturated sp carbon atoms within the OSPC-1 linkers. For the *dia-c*, *pts-c*, and *unc-c* models, the chemical shift ranges of the peaks broaden, showing greater diversity in the chemical environments within the models and giving a better match to the experimental ^{13}C ssNMR spectrum than OSPC-1 in the *dia* topology due to the broadened peak ranges at high chemical shift values.⁹⁵

The solvent-accessible surface areas of OSPC-1 in the *dia*, *dia-c*, *pts-c* and *unc-c* topologies were calculated using Materials Studio 5.0¹⁵⁸ with a probe radius of 1.82 Å.⁹⁵ OSPC-1 in the *dia* and *unc-c* topologies showed high surface areas of $6428\text{ m}^2\text{ g}^{-1}$ and $1295\text{ m}^2\text{ g}^{-1}$, respectively, whilst OSPC-1 in the *dia-c* and *pts-c* topologies had surface areas of zero. The experimental surface areas of OSPC-1a and OSPC-1b of $766\text{ m}^2\text{ g}^{-1}$ (OSPC-1a) and 700 – $900\text{ m}^2\text{ g}^{-1}$ (OSPC-1b) did not match those obtained from the crystalline models.^{16,95,323} However, from calculating the C-C linker distances for each model, it could be seen that a shorter distance corresponded to a better match to the experimental ^{13}C ssNMR spectrum. The C-C linker distances of the topologies were calculated as: *dia* - 11.02 Å , *dia-c* - 5.69 Å , *pts-c* - 4.79 Å , *unc-c* - 2.76 Å . This indicated that linker-linker interactions are a key feature in the OSPC-1 structure, and to replicate the experimental ^{13}C ssNMR spectrum, it was necessary to consider these within the modelling approach.⁹⁵

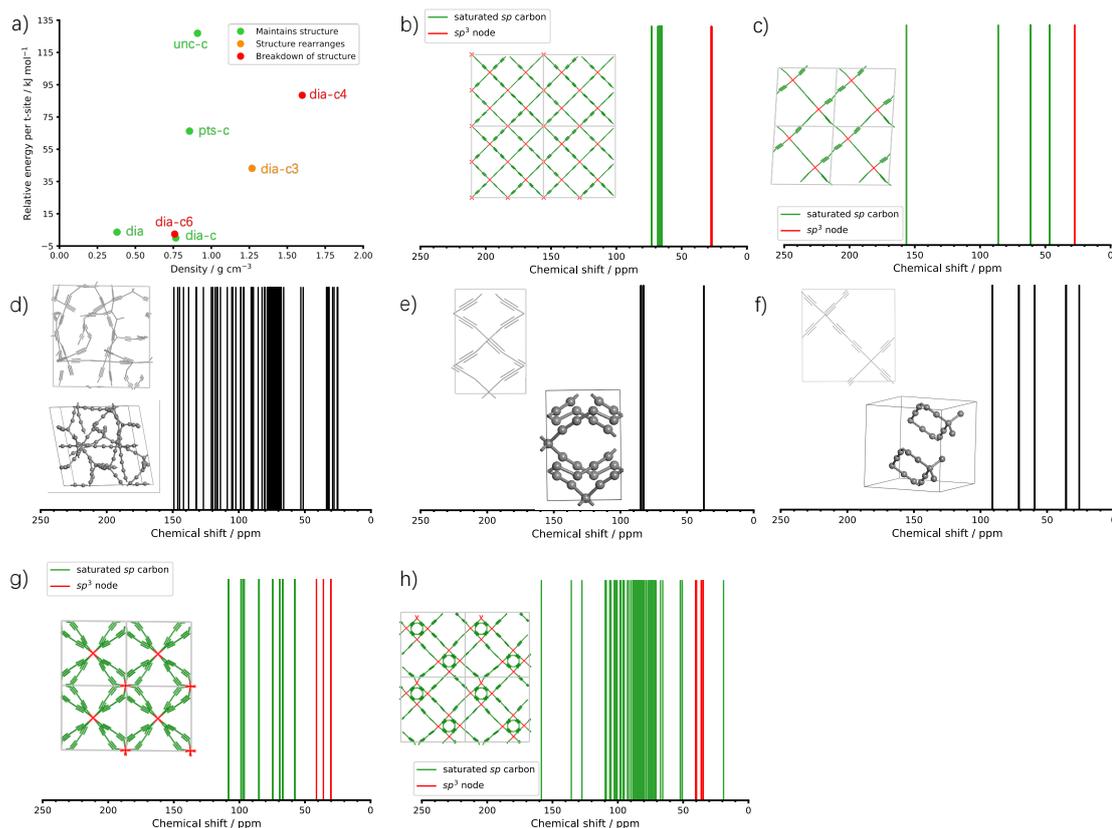


Figure 5.18: a) Plot of the relative energy per t-site of OSPC-1 in each crystalline topology against the density. The energies are calculated relative to the *dia-c* structure, which is given at 0 kJ mol⁻¹. Key: structure retains OSPC-1 bonding motif – green, structure retains the general OSPC-1 bonding motif with some rearrangements – orange, structure breaks down to form a layered material with macrocyclic rings present – red. b)–h) NMR spectra of OSPC-1 in the b) *dia*, c) *dia-c*, d) *dia-c3*, e) *dia-c4*, f) *dia-c6*, g) *pts-c*, h) *unc-c* topologies, calculated using CASTEP 8.0.^{66,68,86,97,213,215,216,218,220,222} Key for b), c), g) and h) (the structures maintaining the OSPC-1 bonding motif): *sp*³ carbon node – red, *sp* carbon in saturated OSPC-1 linker – green. The structures are given in the insets. Figure adapted from reference 95, upon which this chapter is based.

5.4.3 OSPC-1 periodic models - chain structures

Typically, interactions between nuclei in chemical structures would not be expected to alter the NMR chemical shieldings beyond a small number of ‘nearest neighbours’ due to the interactions between more distant atoms being negligible.⁹⁵ Despite this, the crystalline models indicated that interactions between linkers are key to rationalising the experimental ¹³C ssNMR spectrum, which is likely due to the π -conjugation throughout the OSPC-1 structure.⁹⁵

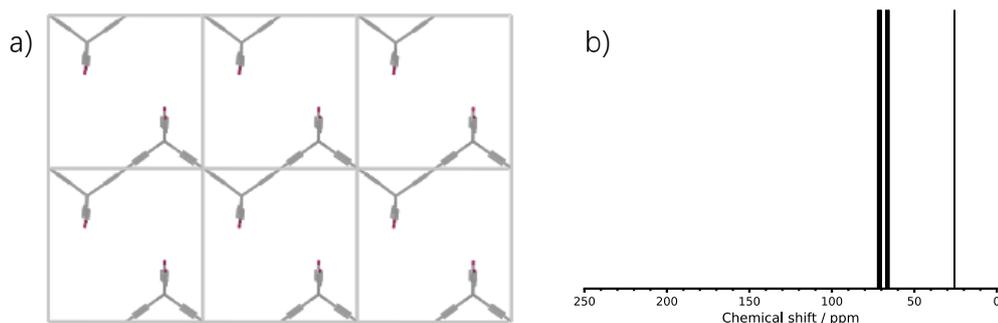


Figure 5.19: a) Structure of the initial periodic chain of OSPC-1 (model A), terminated with hydrogen-capped ethynyl end groups, with cell lengths of $10.9 \text{ \AA} \times 10.9 \text{ \AA} \times 10.9 \text{ \AA}$. The hydrogen atoms are given in plum for better visibility. b) NMR spectrum of model A, calculated using CASTEP 8.0.^{66,68,86,97,213,215,216,218,220,222} Figure adapted from reference 95, upon which this chapter is based.

To assess the influence of these interactions, a one-dimensional infinite chain of OSPC-1 (Figure 5.19a) was simulated by bonding one monomer to each end of the OSPC-1 building block. The remaining unreacted end groups were set as ethynyl groups terminated with hydrogen atoms, rationalising the unreacted alkyne or radical end groups possible within the mechanism to synthesise OSPC-1a, referred to throughout as the ethynyl end group models. The carbon atoms in the OSPC chain were representative of the OSPC-1a and OSPC-1b structures.⁹⁵ The chain models were set up with the chain lying along the x -direction, meaning that this component of the cell length remained fixed throughout.⁹⁵ The unit cell contents and dimensions were optimised using the PCFF forcefield within the Forcite module of Materials Studio 5.0.¹⁵⁸ Following this optimisation, the x -dimension remained unchanged throughout.⁹⁵ Initially, as in the case of model A, interactions between linkers on neighbouring chains were prevented, giving an NMR spectrum with a similar shape to the fragment models, with a peak at 26 ppm (sp^3 C) and a range of peaks between 66–72 ppm (sp C) (Figure 5.19b). Following this, the interactions between OSPC-1 chains as a function of the degree of network interpenetration were assessed by decreasing the cell lengths in the y - and z -directions. The cell lengths were decreased in a systematic matter, optimising the structure at each cell size, then using the optimised structure as the input for the next smallest cell.⁹⁵

The collection of ethynyl end group models and corresponding energy profiles as a function of cell length are given in Figures 5.20–5.21 and Tables 5.2 and C.3–C.13. Each model was organised into a set, where in each set, one of the cell dimensions was decreased systematically. The resulting change in energy was assessed, with the energy in each case calculated relative to the initial energy of the set, corresponding to the largest unit cell size.

For Set 1 (Table 5.2 and Figure 5.20), the z cell length remained at the initial value of 10.9 Å and the y cell length was decreased systematically from an initial value of 10.99 Å to a value of 2.10 Å. It was observed that the energy increased as the y cell length decreased to a maximum energy corresponding to a y cell length of 3.10 Å. This was followed by a subsequent decrease in the energy on further decreasing the y cell length to a minimum energy at a cell length of 2.50 Å. The energy then increased once more at y cell lengths between 2.50 Å and 2.10 Å, corresponding to a breakdown in the OSPC-1 bonding motif as seen for OSPC-1 in the *dia-c3*, *dia-c4* and *dia-c6* topologies, with the formation of different chemical bonds observed.⁹⁵

Table 5.2: Table of the models within Set 1, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å			Structure	Cell length / Å		
	x	y	z		x	y	z
A	10.9832	10.9935	10.9741	FX	10.9196	3.1300	10.9196
B	10.9809	10.0288	10.9557	FY	10.9196	3.1200	10.9196
C	10.9710	9.0382	10.9529	FZ	10.9196	3.1100	10.9196
D	10.9601	8.0352	10.9376	FP	10.9196	3.1000	10.9196
E	10.9728	6.9821	10.9630	FQ	10.9196	3.0000	10.9196
F	10.9461	6.0141	10.9406	FK	10.9196	2.9196	10.9196
FF	10.9196	5.5000	10.9196	PY	10.9196	2.7000	10.9196
FG	10.9196	4.9196	10.9196	QT	10.9196	2.6500	10.9196
FH	10.9196	4.5000	10.9196	QU	10.9196	2.6000	10.9196
FI	10.9196	3.9196	10.9196	QV	10.9196	2.5500	10.9196
FJ	10.9196	3.5000	10.9196	FL	10.9196	2.5000	10.9196
FM	10.9196	3.4000	10.9196	QW	10.9196	2.4500	10.9196
FN	10.9196	3.3000	10.9196	QX	10.9196	2.4000	10.9196
FO	10.9196	3.2000	10.9196	QY	10.9196	2.3500	10.9196
FR	10.9196	3.1900	10.9196	PZ	10.9196	2.3000	10.9196
FS	10.9196	3.1800	10.9196	QZ	10.9196	2.2500	10.9196
FT	10.9196	3.1700	10.9196	RA	10.9196	2.2000	10.9196
FU	10.9196	3.1600	10.9196	RB	10.9196	2.1500	10.9196
FV	10.9196	3.1500	10.9196	QA	10.9196	2.1000	10.9196
FW	10.9196	3.1400	10.9196				

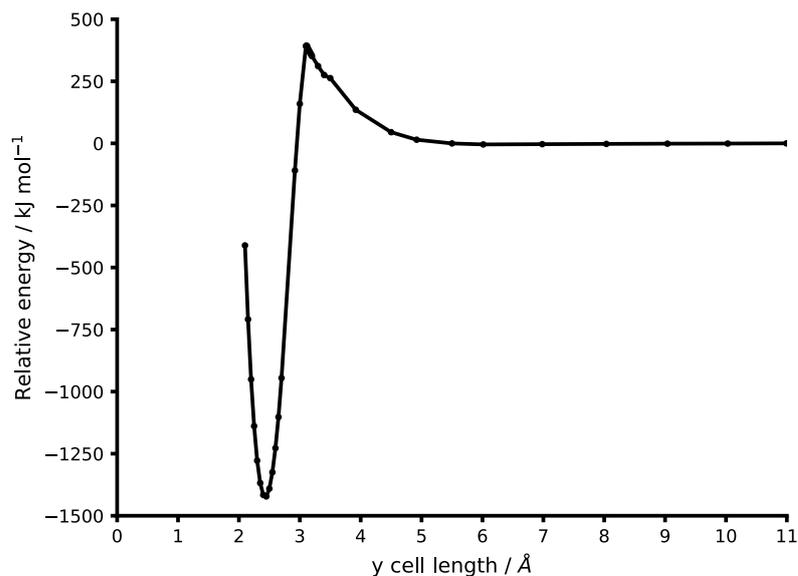


Figure 5.20: Plot of the energy (calculated relative to model A) of the optimised structures within Set 1 as a function of the y cell length. Figure adapted from reference 95, upon which this chapter is based.

Tables of the models included within the remaining Sets 2–12 (and the cell dimensions of each) are given in Appendix C Tables C.3–C.13, with a summary of the unit cell dimensions for Sets 2–12 given in Table 5.3. Sets 2–10 decreased the z cell length sequentially, using a different y cell length in each set.⁹⁵ When moving towards the smaller unit cell dimensions in Sets 6–11, it was necessary to use smaller increments in the y cell length to allow the model to optimise successfully to give a structure that retained the OSPC-1 bonding motif.⁹⁵ For example, for Sets 2–5, the y cell length decreased by an increment of 0.10 Å between Sets, whilst for Sets 6–11, the increment decreased to 0.02 Å between Sets.

The general trend throughout sets 2–10 showed a decrease in the energy on decreasing the z cell lengths from the initial value of 10.92 Å to 6.92 Å, as shown in Figures 5.21a–i. Beyond this point, the energy increased on further decreasing the z cell length. Set 11 then fixed the z cell length at 6.00 Å whilst decreasing the y cell length from 3.12–2.00 Å, giving the most compressed, and highest density one-dimensional periodic chains modelled.⁹⁵ A broadly symmetrical energy profile, shown in Figure 5.21j, was observed, with the minimum energy obtained with a y cell length of 2.40 Å.

To increase the degree of network interpenetration possible within the OSPC-1 structure, set 12 (Figure 5.21k) was composed of two perpendicular chains running along the x and z cell directions. As the y cell length was decreased from the initial value of 21.84 Å, the energy decreased to a minimum energy at a y cell length of ~ 14 Å, with y cell lengths of less than 14 Å corresponding

to a subsequent increase in the energy.⁹⁵ Interestingly, unlike for sets 2–11, set 12 did not show a breakdown of the OSPC-1 bonding motif at small cell sizes. Additionally, the NMR spectra of set 12 showed a broadening of the chemical shift values at decreasing cell sizes, whilst sets 2–11 corresponded to very distinct chemical shifts from one cell length to the next.⁹⁵

Table 5.3: Summary of the unit cell dimensions for Sets 2–12.

Set	Cell length / Å		
	x	y	z
2	10.92	3.50	10.92–5.50
3	10.92	3.40	10.92–5.50
4	10.92	3.30	10.92–6.00
5	10.92	3.20	10.92–6.00
6	10.92	3.18	10.92–6.00
7	10.92	3.16	10.92–6.00
8	10.92	3.14	10.92–6.00
9	10.92	3.12	10.92–6.00
10	10.92	3.10	10.92–6.00
11	10.92	3.12–2.00	6.00
12	21.84	21.84–4.64	21.84

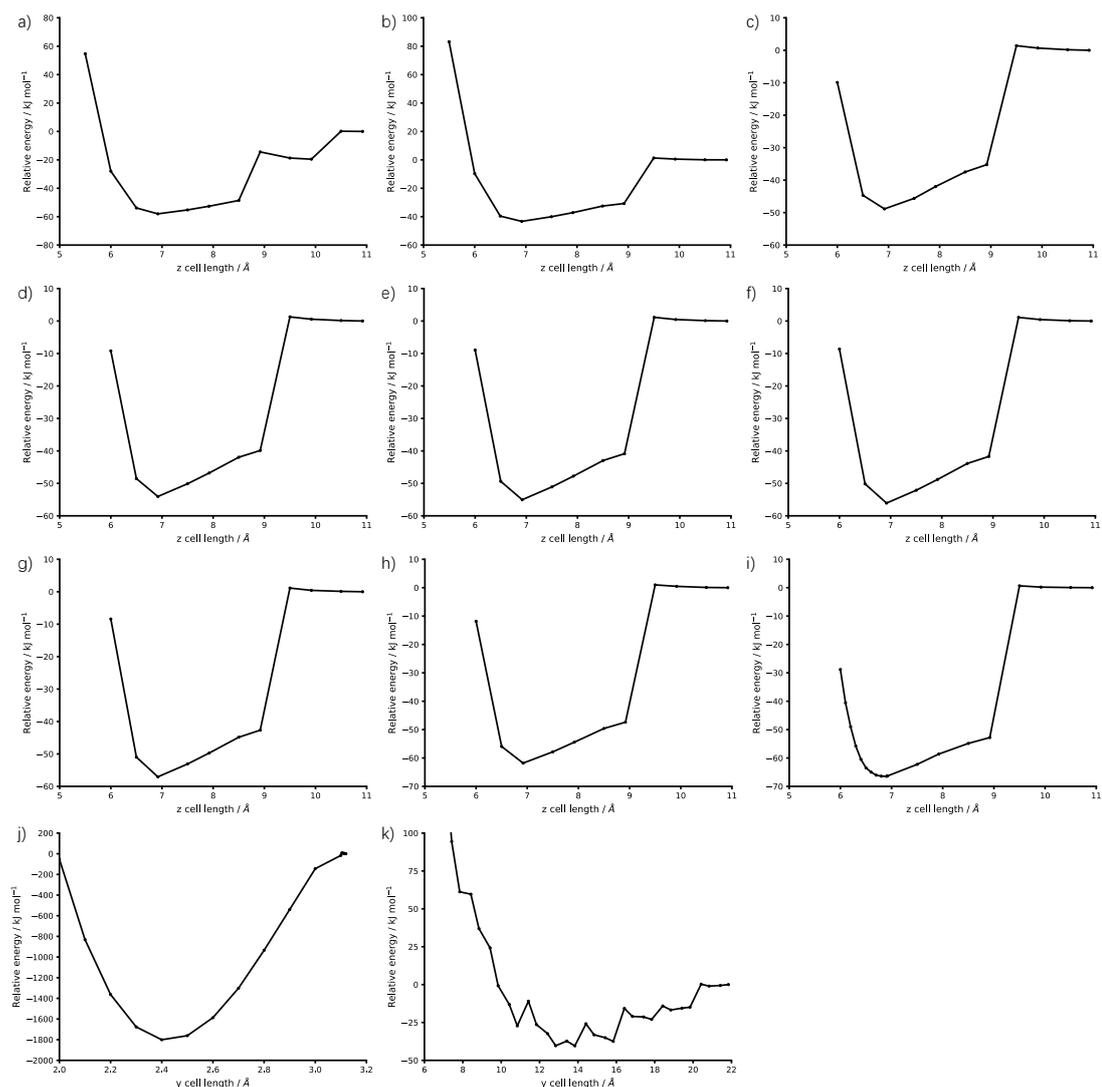


Figure 5.21: Plots of the relative energy of the optimised structures within Sets 2–12 as a function of the cell length. Sets 2–10 decreased the z cell length sequentially. Set 11 then fixed the z cell length at 6.00 Å whilst decreasing the y cell length. Set 12 was composed of two perpendicular chains running along the x and z cell directions, and the y cell length was decreased. a) Set 2, with energies calculated relative to model FJ. b) Set 3, with energies calculated relative to model FM. c) Set 4, with energies calculated relative to model FN. d) Set 5, with energies calculated relative to model FO. e) Set 6, with energies calculated relative to model FS. f) Set 7, with energies calculated relative to model FU. g) Set 8, with energies calculated relative to model FW. h) Set 9, with energies calculated relative to model FY. i) Set 10, with energies calculated relative to model FP. j) Set 11, with energies calculated relative to model MH. k) Set 12, with energies calculated relative to model NM. Figure adapted from reference 95, upon which this chapter is based.

5.4.4 Energy profiles

The models included within the energy profile of the ethynyl end groups one-dimensional periodic chains (Figure 5.22) going from the least dense (model A, dimensions: 10.92 Å × 10.92 Å × 10.92 Å) to the densest model (model QC, dimensions: 10.92 Å × 2.90 Å × 6.00 Å) are given in Table 5.4.⁹⁵

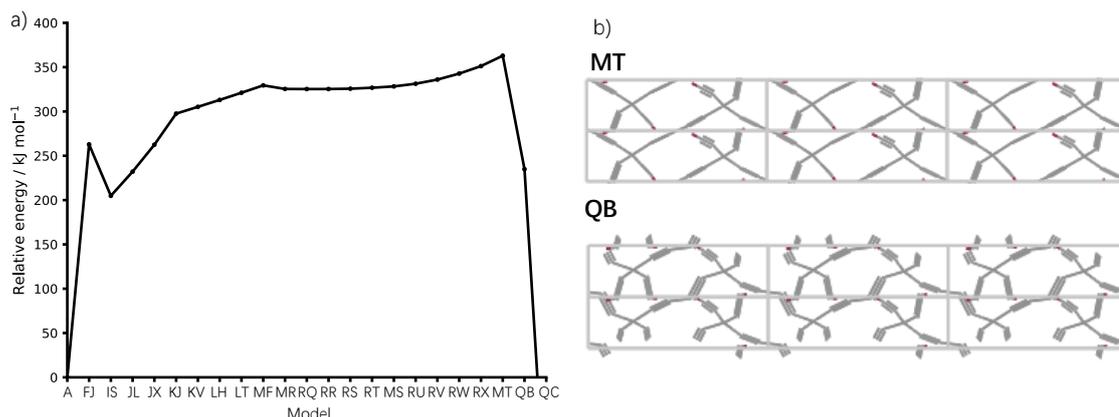


Figure 5.22: Energy profile of the ethynyl end groups one-dimensional periodic chains going from the least dense model A to the densest model QC. a) Plot of the energy (calculated relative to model A) of the optimised structures as a function of the cell density. b) Comparison of the structures of model MT, where the OSPC-1 structure remained intact, and QB, where the bonding motif had broken down. The hydrogen atoms are given in plum for better visibility. Figure adapted from reference 95, upon which this chapter is based.

Table 5.4: Table of the models included within the energy profile of the ethynyl end groups one-dimensional periodic chains going from the least dense model A to the densest model QC. Table adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å			Structure	Cell length / Å		
	x	y	z		x	y	z
A	10.9196	10.9196	10.9196	RR	10.9196	3.1000	6.8000
FJ	10.9196	3.5000	10.9196	RS	10.9196	3.1000	6.7000
IS	10.9196	3.5000	6.9196	RT	10.9196	3.1000	6.6000
JL	10.9196	3.4000	6.9196	MS	10.9196	3.1000	6.5000
JX	10.9196	3.3000	6.9196	RU	10.9196	3.1000	6.4000
KJ	10.9196	3.2000	6.9196	RV	10.9196	3.1000	6.3000
KV	10.9196	3.1800	6.9196	RW	10.9196	3.1000	6.2000
LH	10.9196	3.1600	6.9196	RX	10.9196	3.1000	6.1000
LT	10.9196	3.1400	6.9196	MT	10.9196	3.1000	6.0000
MF	10.9196	3.1200	6.9196	QB	10.9196	3.0000	6.0000
MR	10.9196	3.1000	6.9196	QC	10.9196	2.9000	6.0000
RQ	10.9196	3.1000	6.9000				

Moving from each structure to the next, there was a corresponding decrease in either the y or z cell length, with any increase in the density from model A giving an increase in the energy of the system, as seen in the energy-density plot of OSPC-1 in each of the crystalline topologies (Figure 5.18a). There is a broad minimum between models MF ($y = 3.12 \text{ \AA}$, $z = 6.92 \text{ \AA}$) and MT ($y = 3.10 \text{ \AA}$, $z = 6.00 \text{ \AA}$). Further decreasing the cell volume below that of MT caused a breakdown of the OSPC-1 bonding motif (Figure 5.22b).⁹⁵

To further understand the influence of OSPC-1 chain proximity on the energy of the system, calculations of the interactions within stacked OSPC-1 dimers were undertaken using the B3LYP/6-311G model chemistry in Gaussian 09.^{79,85,90,91,95,212,380} Firstly, a stack of two dimers was assembled (Figure 5.23), with the interaction energy assessed as the dimers were brought into closer proximity. This was done by systematically decreasing the distance between the two dimers in steps of 0.50 Å.⁹⁵ For the two dimers setup, the energies were assessed both on changing the distance between one central *sp* carbon atom in each dimer, and on changing the distance between a central *sp* carbon atom of one dimer and an *sp*³ carbon node of the other. It could be seen that in both profiles, there was a broad minimum energy region at approximately 7.00 Å, with a steep increase in the system energy on bringing the dimers into closer proximity, which does not match the energy pathways of the one-dimensional periodic chain models well.⁹⁵

Following this, a larger three-dimer stack system was modelled (Figure 5.24), calculating the change in energy as the distances between the central *sp* carbon atoms of the outer dimers were systematically decreased in steps of 0.20 Å, with smaller steps of 0.01 Å taken around the energy well at 4 Å.⁹⁵ This system showed increasing energies with decreasing dimer-dimer-dimer distances up to a maximum energy at just above 4.00 Å, followed by a rapid reduction in the energy in the region of 3.75–4.00 Å. Further decreasing the dimer-dimer-dimer distance caused another increase in the energy, corresponding to a breakdown in the OSPC-1 bonding motif as observed for the smallest periodic models and the *dia-c3*, *dia-c4* and *dia-c6* systems.⁹⁵

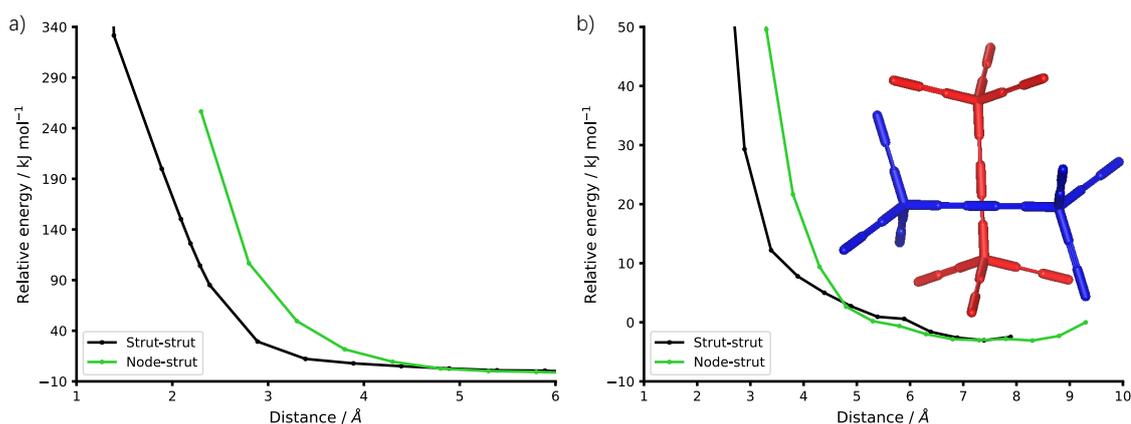


Figure 5.23: Energy profile of two stacked dimers as a function of the distance between one central *sp* carbon atom in each dimer (black, strut-strut), and between a central *sp* carbon atom of one dimer and an *sp*³ carbon node of the other (green, node-strut). a) Profile assessing the region at smaller distances. b) Profile assessing the region at larger distances. An image of the two-dimer setup is given as an inset to b), showing each dimer as a different colour for clarity. The energies were calculated relative to the energy of the two dimers separated at the modelled largest distance of 9.3 Å. Figure adapted from reference 95, upon which this chapter is based.

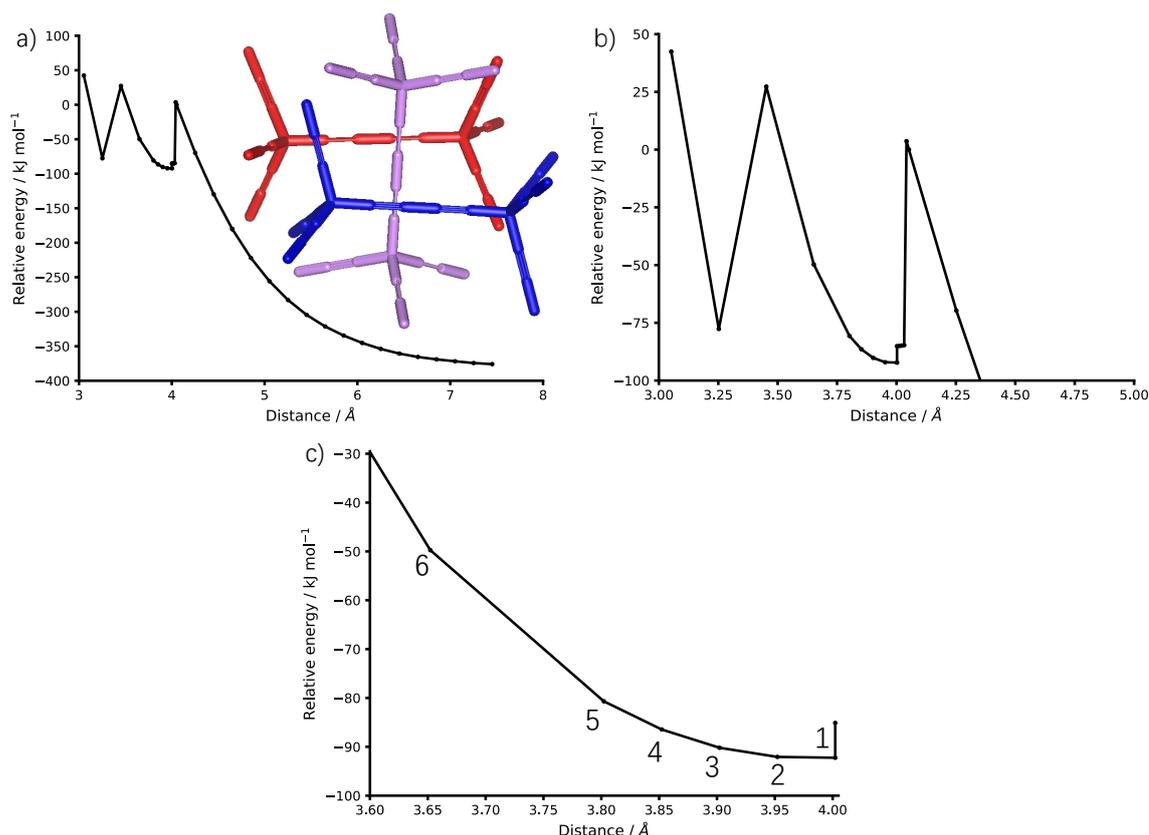


Figure 5.24: Energy profile of three stacked dimers as a function of the distance between one central *sp* carbon atom of each outer dimer. a) Full profile. b) Profile assessing the region at smaller distances. c) Profile assessing the structures in the energy well at 3.75–4.00 Å. An image of the three-dimer setup, which contains structures 1–6, is given as an inset to a), showing each dimer as a different colour for clarity. The energies were calculated relative to a distance of 4.05 Å. Figure adapted from reference 95, upon which this chapter is based.

The NMR spectra of the structures labelled 1–6, corresponding to those in the energy well of the three-dimer stack system at 4 Å, are given in Figure 5.25.⁹⁵ It can be seen that compared to the NMR spectra of the Gaussian fragment models, such as the single dimer, the peaks of structures 1–6 have broadened to wider chemical shift ranges, indicating a larger number of potential environments within the structures, along with shifting of the peaks to larger chemical shift values, giving a better match to the experimental ¹³C ssNMR spectrum of OSPC-1 as a result of the increasing interactions between chains.⁹⁵

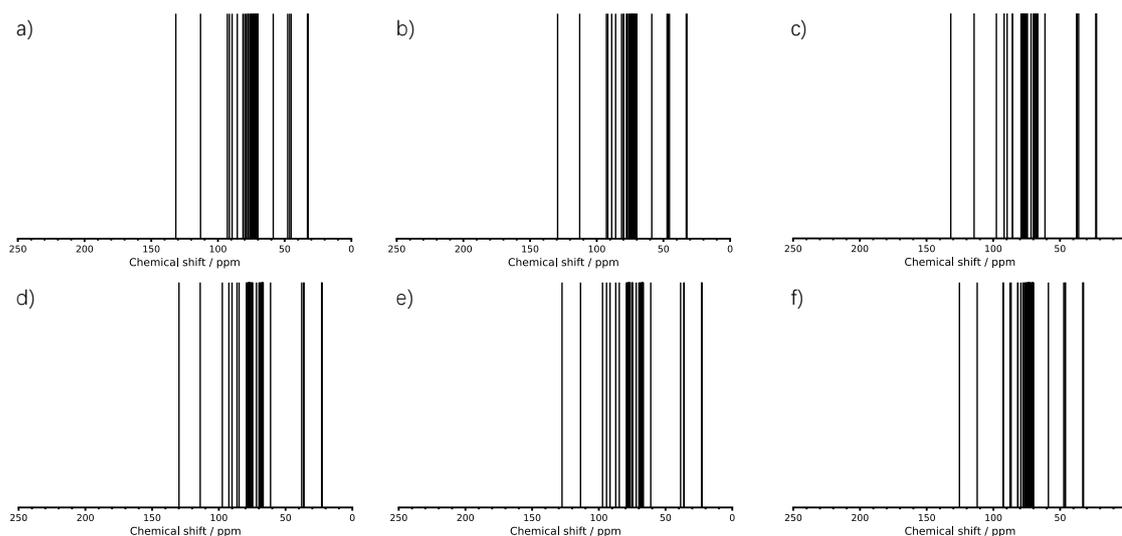


Figure 5.25: NMR spectra of structures 1–6, calculated using the B3LYP/6-311G model chemistry in Gaussian 09.^{79,85,90,91,212,380} a) structure 1, b) structure 2, c) structure 3, d) structure 4, e) structure 5, f) structure 6. Figure adapted from reference 95, upon which this chapter is based.

As seen from previous modelling, a range of structures may exist within the minimum energy well of the amorphous OSPC-1 system.^{16,325,363} Therefore, a combination of the one-dimensional periodic chains within this energy well, structures MF ($y = 3.12 \text{ \AA}$, $z = 6.92 \text{ \AA}$) to RT ($y = 3.10 \text{ \AA}$, $z = 6.60 \text{ \AA}$), was assessed to rationalise the experimental ^{13}C ssNMR spectrum of OSPC-1.⁹⁵ As in the spectra of structures 1–6, the NMR spectra of MF to RT (Figure 5.26) showed a broadened range of peaks, shifted to higher chemical shifts compared to the fragment models and less dense one-dimensional ethynyl end groups periodic structures.⁹⁵

When each of the peaks from MF to RT were combined into one spectrum (Figure 5.26g), four regions of the NMR spectrum could be identified.⁹⁵ Peak α corresponded to the range of peaks originating from the sp carbon atoms in the saturated OSPC-1 linker, appearing in the range of 80–140 ppm. Peak β corresponded to the narrow range of peaks originating from the sp^3 carbon atoms in the OSPC-1 nodes, appearing in the range of 50–56 ppm. Peaks χ and δ corresponded to the broad range of peaks originating from the sp carbon atoms in the unreacted alkynes. As these atoms appeared in two distinct regions of the spectrum, they were classified into two peaks, χ and δ , appearing in the ranges of 165–180 ppm and 80–110 ppm, respectively. Each of peaks α – δ can be identified within the experimental ^{13}C ssNMR spectrum of OSPC-1a (Figure 5.26i).⁹⁵

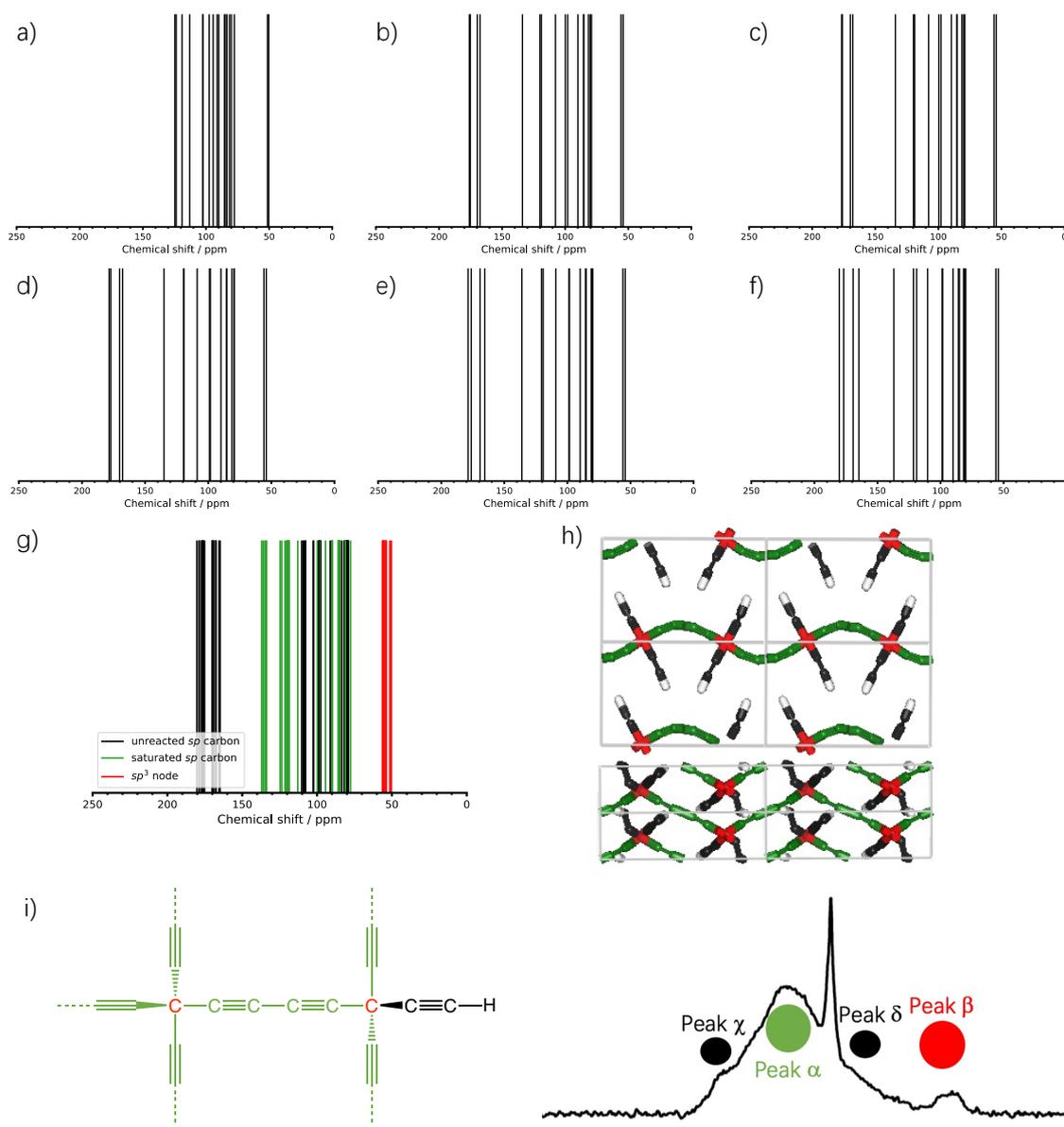


Figure 5.26: NMR spectra of structures MF to RT, calculated using CASTEP 8.0.^{66,68,86,97,213,215,216,218,220,222} a) structure MF ($y = 3.12 \text{ \AA}$, $z = 6.92 \text{ \AA}$), b) structure MR ($y = 3.10 \text{ \AA}$, $z = 6.92 \text{ \AA}$), c) structure RQ ($y = 3.10 \text{ \AA}$, $z = 6.90 \text{ \AA}$), d) structure RR ($y = 3.10 \text{ \AA}$, $z = 6.80 \text{ \AA}$), e) structure RS ($y = 3.10 \text{ \AA}$, $z = 6.70 \text{ \AA}$), f) structure RT ($y = 3.10 \text{ \AA}$, $z = 6.60 \text{ \AA}$). g) combined NMR spectrum of MF to RT. h) example model structure (top and side view of model MR). i) allocation of peaks α – δ in the experimental ^{13}C ssNMR spectrum of OSPC-1a. Key for g)–i): sp^3 carbon – red, sp carbon in saturated OSPC-1 linker – green, sp carbon in unreacted alkyne – black. Figure adapted from reference 95, upon which this chapter is based. Experimental ^{13}C ssNMR spectrum of OSPC-1a reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Whilst the ethynyl end group models were sufficient to model the NMR spectrum of OSPC-1a,¹⁶ OSPC-1b contains additional potential end groups within the structure, such as bromine atoms and TMS-capped butadiynyl groups.⁹⁵ To fully rationalise the experimental ^{13}C ssNMR spectrum of OSPC-1b,³²³ it

was therefore necessary to include these within the modelling approach. A small set of one-dimensional periodic chains were generated with two bromine atoms, and one bromine atom and one TMS-capped butadiynyl group, as the end groups, referred to as the bromine end group and the TMS-Br end group models, respectively. These models, which were obtained following the same protocols as the ethynyl end group models, were collected by Adam Rowling.⁹⁵

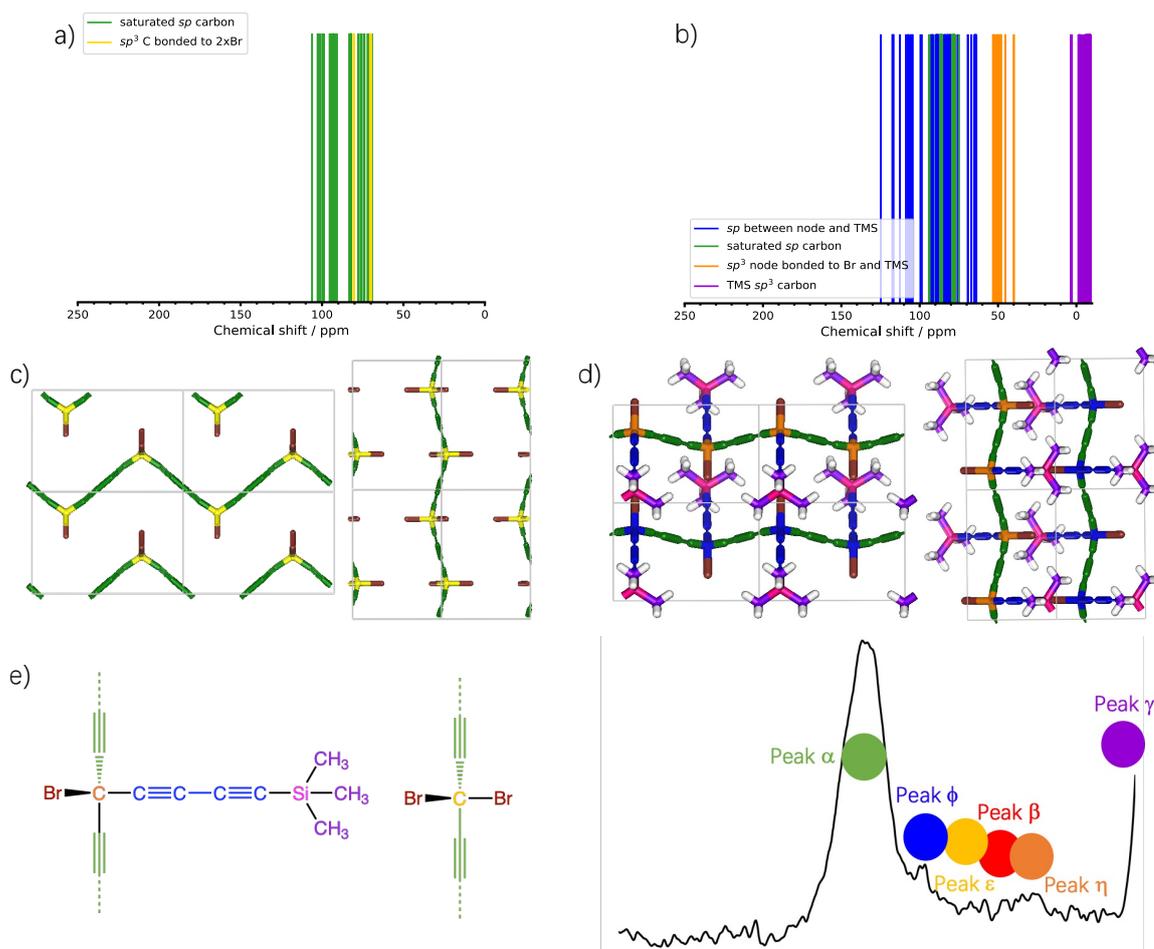


Figure 5.27: NMR spectra of the bromine end group and TMS-Br end group models, calculated using CASTEP 8.0.^{66,68,86,97,213,215,216,218,220,222} a) combined NMR spectrum of the bromine end group models, b) combined NMR spectrum of the TMS-Br end group models, c) example model structure of the bromine end group models (top and side view of model BB). d) example model structure of the TMS-Br end group models (top and side view of model BO). e) allocation of peaks α , β , ϵ – η in the experimental ^{13}C ssNMR spectrum of OSPC-1b. Key: sp carbon in saturated OSPC-1 linker – green, sp^3 carbon in OSPC-1 node bonded to two bromine end groups – yellow, sp carbon in butadiynyl chain between OSPC-1 node and TMS end group – blue, sp^3 carbon in TMS end group – purple, sp^3 carbon in OSPC-1 node bonded to a bromine end group and a TMS-capped butadiynyl group – orange, Br – burgundy, Si – pink. Figure adapted from reference 95, upon which this chapter is based. Experimental ^{13}C ssNMR spectrum of OSPC-1b obtained from reference 323.

The peaks arising from each model were combined into one distinct NMR spectrum per end group set as for the ethynyl end group models (Figure 5.27a-d). Following this, one overall spectrum considering all of the potential end groups within the OSPC-1b structure was hypothesised (Figure 5.27e). Four extra peak environments were observed, in addition to peaks α and β described above, due to the lower symmetry within the OSPC-1b structure.⁹⁵ Peak ϵ occurred due to the sp^3 carbon atoms bonded to two bromine end groups, occurring in the range of 70-83 ppm. Peak ϕ occurred due to the sp carbon atoms in the butadiynyl chain between the sp^3 OSPC-1 carbon node and the silicon atom of the TMS end group, occurring in the range of 60-125 ppm. Peak γ occurred due to the sp^3 carbon atoms within the TMS end groups, occurring in the range of -10-5 ppm. Finally, peak η occurred due to the sp^3 OSPC-1 carbon nodes bonded to one bromine end group and one TMS-capped butadiynyl end group, occurring in the range of 40-53 ppm.⁹⁵

5.4.5 Full NMR spectrum modelling

It was not possible to calculate the intensities of each simulated chemical shielding environment within the CASTEP models.⁹⁵ To circumvent this, a full simulated NMR spectrum was calculated using Gaussian functions (Equation 5.4)⁹⁵ to convolute and broaden each of peaks α - δ to model the OSPC-1a spectrum and each of peaks α , β , ϵ - η to model the OSPC-1b spectrum.⁹⁵ The DFT calculations of models MF-RT for OSPC-1a and the entire set of bromine end group and TMS-Br end group models for OSPC-1b were utilised to assign peaks α - η .⁹⁵

$$f(x, A, \mu, \sigma) = \frac{A}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (5.4)$$

Equation 5.4 defines the Gaussian convolution of each peak as a function of continuous variable x , peak height (Equation 5.5, the peak height is a constant value for each peak), centre of the peak μ and standard deviation σ . Once each peak was assigned, the μ value was obtained as the centre of the peak in the DFT calculations.⁹⁵ The overall convoluted spectrum (Figure 5.29) was then obtained by summing the contributions from each peak in the system. The parameters used for the convolution of peaks α - η are given in Table 5.5. The peak heights and standard deviation values were modified to give the best fit to experiment.^{16,95,323}

$$\text{Peak height} = \frac{A}{\sigma\sqrt{2\pi}} \quad (5.5)$$

The ratios of respective environments within each structure were calculated using the area of each peak in the overall spectrum, approximated using the

trapezium rule. The area of each peak is approximated by calculating the area of a series of trapezia underneath each curve (Equation 5.6).⁹⁵

$$\text{Peak area} = \sum_{i=1}^n \frac{f(x_i) + f(x_{i+1})}{2} \cdot (x_{i+1} - x_i) \quad (5.6)$$

The area of each trapezium was calculated by multiplying the average intensity across two consecutive chemical shift values (the trapezium ‘height’) by the increment in chemical shift value (the trapezium ‘width’, a value of one). The overall area was then obtained by summing the area of each trapezium.⁹⁵

Figure 5.28 shows a schematic of how the trapezium rule was used to approximate the peak areas, using the example of peak α , which has a bell-shaped curve. It can be seen that using a relatively large increment width, as illustrated by the black trapezia, corresponds to overestimations of the peak area at the outer regions of the peak (~ 80 – 100 ppm and ~ 160 – 180 ppm). The central region of peak α , in the region of 120 – 140 ppm, is underestimated. In contrast, using smaller trapezia widths, such as the yellow trapezia, gives a better approximation of the overall area by minimising the over- and underestimation that occurs when calculating the area over a large increment width. Figure 5.28b shows how the area of peak α , calculated across the peak width of 80 – 200 ppm, converges as a function of increment width. The shape of the curve for peak α means that as the relative intensity of the outer regions is small in comparison to that of the central region, the overestimation of the outer regions is of smaller magnitude than the underestimation of the central region. This corresponds to a decrease in the overall peak area with increasing increment width. Figure 5.28b also shows that the chosen increment width of 1 ppm was sufficiently small to yield a good approximation for the overall peak areas.

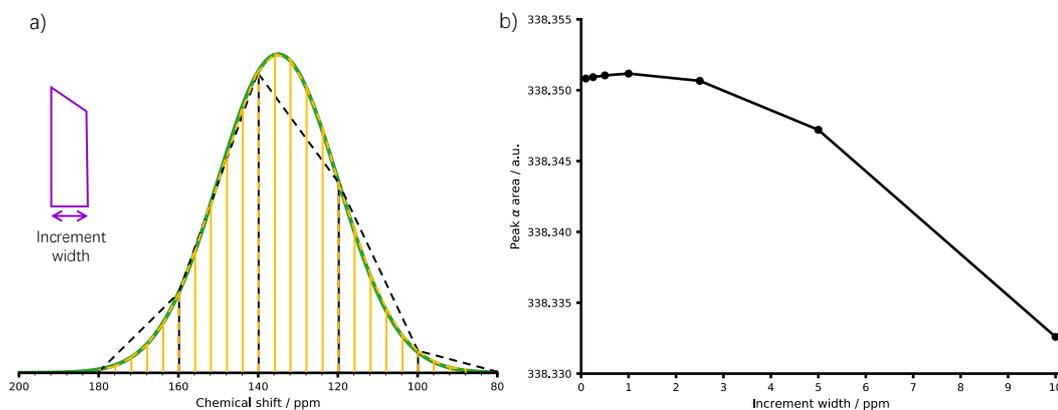


Figure 5.28: a) Schematic showing how the area of peak α (green) can be approximated using large trapezia of width 20 ppm (black) or smaller trapezia of width 4 ppm (yellow). b) Plot showing how the area of peak α converged with respect to the increment width for the chemical shift range of 80 – 200 ppm.

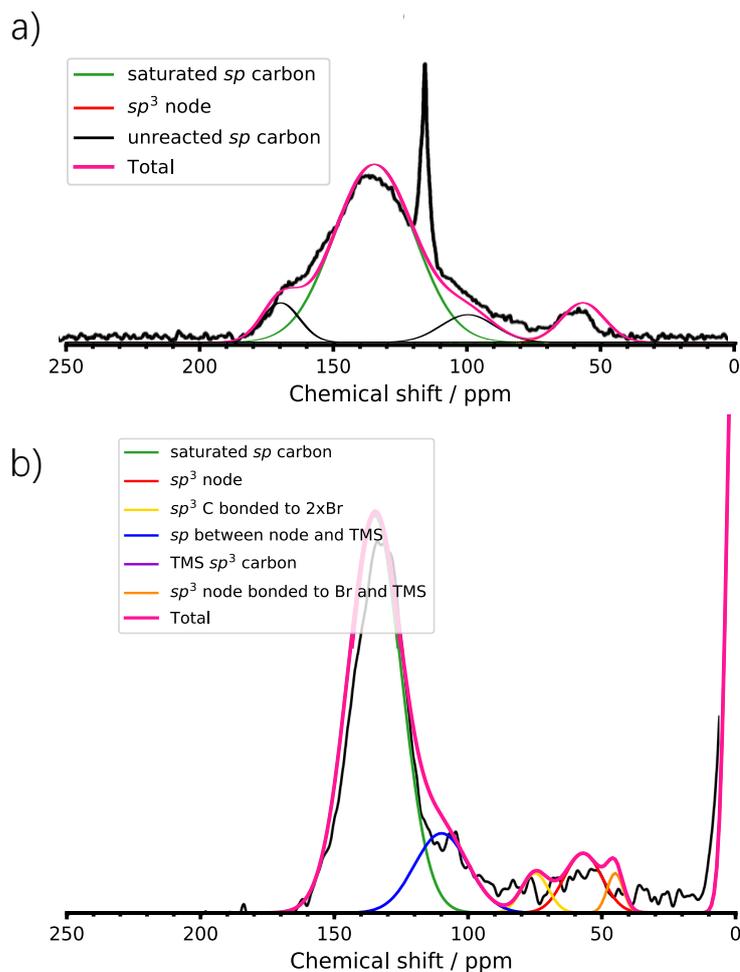


Figure 5.29: NMR spectrum of a) OSPC-1a and b) OSPC-1b, calculated using CASTEP 8.0,^{66,68,86,97,213,215,216,218,220,222} convoluted using Gaussian broadening functions, and overlaid onto the relevant experimental ^{13}C ssNMR spectrum. The parameters used for the convolution of peaks α – η are given in Table 5.5. Key: sp^3 carbon in saturated OSPC-1 network – red, sp carbon in saturated OSPC-1 linker – green, sp carbon in unreacted alkyne – black, sp^3 carbon in OSPC-1 node bonded to two bromine end groups – yellow, sp carbon in butadiynyl chain between OSPC-1 node and TMS end group – dark blue, sp^3 carbon in TMS end group – purple, sp^3 carbon in OSPC-1 node bonded to a bromine end group and a TMS-capped butadiynyl group – orange, overall simulated spectrum – pink. Figure adapted from reference 95, upon which this chapter is based. Experimental ^{13}C ssNMR spectrum of OSPC-1a reproduced (adapted) from reference 16 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. Experimental ^{13}C ssNMR spectrum of OSPC-1b obtained from reference 323.

Table 5.5: The parameters used for the convolution of the simulated a) OSPC-1a and b) OSPC-1b spectra given in Figure 5.29. Key: α - sp carbon in saturated OSPC-1 linker, β - sp^3 carbon in saturated OSPC-1 network, χ and δ - sp carbon in unreacted alkyne, ϵ - sp^3 carbon in OSPC-1 node bonded to two bromine end groups, ϕ - sp carbon in butadiynyl chain between OSPC-1 node and TMS end group, γ - sp^3 carbon in TMS end group, η - sp^3 carbon in OSPC-1 node bonded to a bromine end group and a TMS-capped butadiynyl group. Table adapted from reference 95, upon which this chapter is based.

Peak	μ	Height	σ
α	135	9	15
β	57	2	8
χ	170	2	7
δ	100	1.4	10

Peak	μ	Height	σ
α	135	20	10
β	57	3	7
ϵ	75	2	5
ϕ	110	4	10
γ	0	35	3
η	45	2	3

From Figure 5.29, it can be seen that the convoluted NMR spectra of OSPC-1a and OSPC-1b gave very good matches to the experimental for each material.^{16,95,323} Additionally, the peak ratios were able to offer insight into the structure. For OSPC-1a (Table 5.6a), the simpler case due to the single end group modelled (the ethynyl end groups), the ratio of peaks $\alpha : \beta$, corresponding to saturated sp carbon atoms in the OSPC-1 linkers, and sp^3 carbon nodes, respectively, was 8.4 : 1, comparing well to the idealised ratio of 8 : 1.^{16,95} The ratio of saturated sp carbon atoms in the OSPC-1 linkers to the unreacted ethynyl end groups (peaks $\alpha : \chi + \delta$) of 8.4 : 2 suggested that approximately 25% of the ethynyl groups within the OSPC-1a structure were unreacted end groups, however, this is open to flexibility depending on the other alternative end groups that could appear within the Eglinton homocoupling mechanism,³⁵⁰ such as radicals, anions, and oxygenated species which would change the chemical shift values and peak intensities.^{95,405,406}

For OSPC-1b (Table 5.6b), the ratios of peaks $\alpha : \beta$, and peaks $\beta : \epsilon : \eta$ (where ϵ and η corresponded to the sp^3 carbon atoms bonded to two bromine end groups, and one bromine end group and one TMS-capped butadiynyl end group, respectively) were 9.5 : 1 and 3.5 : 2 : 1, respectively.⁹⁵ This indicated that approximately half of the sp^3 hybridised OSPC-1 nodes were bonded to the sp linkers and two ethynyl end groups, whilst the other half were bonded to the sp linkers and alternative end groups possible within the OSPC-1b structure.⁹⁵

Table 5.6: Ratios of the peak environments within a) OSPC-1a and b) OSPC-1b. Data obtained from reference 95, upon which this chapter is based.

a)

Peak	Environment	Area	~Ratio (All)
α	OSPC sp	338.4	8.4
β	OSPC sp^3	40.1	1
χ	Ethynyl end group	35.1	1
δ	Ethynyl end group	35.1	1

b)

Peak	Environment	Area	~Ratio (All)	~Ratio (α : β)	~Ratio (β : ϵ : η)
α	OSPC sp	501.3	33	9.5	
β	OSPC sp^3	52.6	3.5	1	3.5
ϵ	C-Br	25.1	2		2
ϕ	Butadiynyl chain (OSPC-TMS end group)	100.3	7		
γ	TMS sp^3	97.5	6.5		
η	sp^3 C (TMS and Br end groups)	15.05	1		1

The rationalisation of the ^{13}C ssNMR spectra of OSPC-1a¹⁶ and OSPC-1b³²³ using the approach described here required the OSPC-1 chains to remain within close proximity to maximise interactions within the framework, leading to a dense polymer with little to no porosity.⁹⁵ This contradicted the reported experimental porosities of OSPC-1a and OSPC-1b ($766\text{ m}^2\text{ g}^{-1}$ and $909\text{ m}^2\text{ g}^{-1}$, respectively),^{16,323} and it was therefore important to rationalise the porosity of these materials.⁹⁵ One explanation would be that the materials have little to no microporosity due to the dense structure but a large quantity of mesoporosity due to inefficient packing of the polymer in the solid state. However, both OSPC-1a and OSPC-1b demonstrate type II isotherms, and peaks within the micropore regions of the PSDs, indicating a large quantity of microporosity throughout the structures.^{16,323} Alternatively, the porosity may arise due to multiple variations of the OSPC-1 structure within the material, with the dense, non-porous structural form approximated by this modelling approach being most visible in the experimental ^{13}C ssNMR spectrum.⁹⁵

Dipole-dipole relaxation effects due to chemical shift anisotropy (CSA) are a key feature in ^{13}C NMR spectra,⁴⁰⁷ with longer relaxation times occurring within carbon atoms with a greater ability to rotate.⁹⁵ A similar molecule to the OSPC-1 dimer, 1,4-diphenyl-1,3-butadiyne,⁴⁰⁸ which has the same linker but different nodes, demonstrated a large range of relaxation times depend-

ing on the chemical environment of each atom (Figure 5.30a). The linker atoms have the longest relaxation times within the molecule, showing the greatest degree of rotation. This indicated that an open and highly porous OSPC-1 structure, with a large ability to rotate, would show far longer relaxation times in comparison to dense, non-porous structures with the linkers held in place by interactions with nearby chains.⁹⁵ This would suggest that the open and porous structural form of OSPC-1 would not be visible within the normal collection times for an ^{13}C ssNMR experiment, such as those used in the original reports of OSPC-1a and OSPC-1b.^{16,95,323}

This hypothesis was confirmed by analysis of an additional ^{13}C ssNMR spectrum of OSPC-1b with a long relaxation time of 300 s (Figure 5.30b), compared to the initial relaxation time of 60 s (collected by Adam Rowling).⁹⁵ The spectrum did not show any significant differences to the initial, which could suggest that there was only one structural form of OSPC-1 present, contradicting the reported porosity, or it could indicate that an open, porous form of OSPC-1 was present in very small quantities and was masked in the experimental ^{13}C ssNMR spectrum.^{95,409-411} This was entirely plausible, as unlike the vast majority of organic molecules, OSPC-1 does not contain hydrogen atoms, which are typically exploited to enhance the signal to noise ratio of ^{13}C ssNMR spectra. This would mean that the signals of the entire spectrum would be reduced in intensity, as the spectrum could only be acquired through direct excitation of the ^{13}C nuclei themselves.⁹⁵ Consequently, the less intense peaks arising from small quantities of the open, porous form of OSPC-1, which would contribute a very small percentage of the overall material mass, could easily become indistinguishable from the background noise.⁹⁵

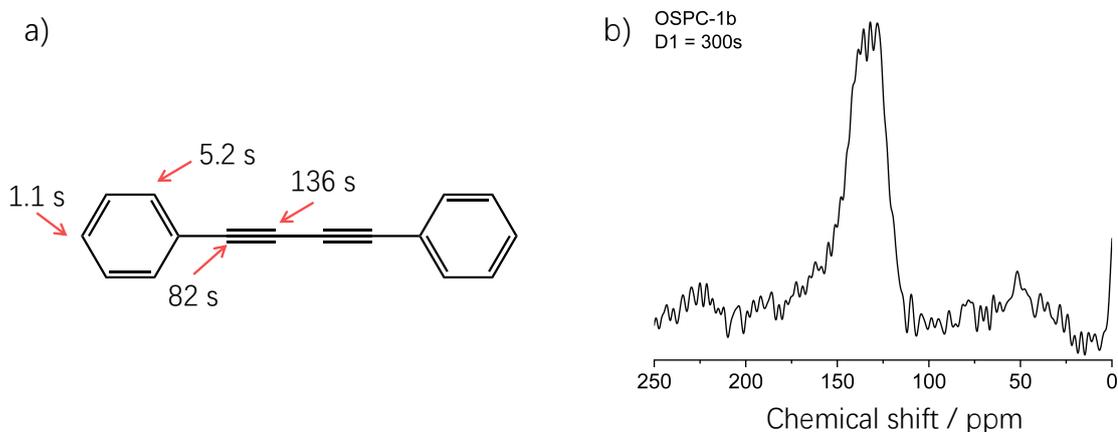


Figure 5.30: Relaxation times in ^{13}C ssNMR. a) ^{13}C ssNMR relaxation times of 1,4-diphenyl-1,3-butadiyne.⁴⁰⁸ b) ^{13}C ssNMR spectrum of OSPC-1b, collected with a relaxation time of 300 s. Figure adapted from reference 95, upon which this chapter is based.

5.4.6 Structure rationalisation

From analysis of the isotherms and PSDs of OSPC-1a and OSPC-1b, it is clear that both materials contain micro- and mesoporosity.^{16,323} For OSPC-1b, the maximum reported surface area of 909 m² g⁻¹ comprised of 617 m² g⁻¹ microporosity and 292 m² g⁻¹ mesoporosity, as judged from the experimental isotherm.³²³ The PSD of OSPC-1b is given in Figure 5.31.

This showed a large distribution of peaks with intensities below 0.50 nm, a shoulder of peaks in the range of 0.60–0.80 nm, a relatively large peak centred at 1.25 nm, and a smaller peak at 1.70 nm.^{95,323} The structure and properties of conjugated microporous polymer materials, which are very similar to OSPC-1, have been previously rationalised by the formation of spherical particles. This rationalisation involves each particle comprising of a dense, non-porous core, surrounded by gradually increasing porosity and decreasing density moving from the centre of the spherical particle outwards towards the outer edges.^{19,38,261} As described in Chapter 3 section 3.3, the reaction conditions play a vital role in the composition of the framework and ratios of non-porous central core to outer edges.⁹⁵ Following this hypothesis, it was plausible that the OSPC-1 framework comprised of a dense, non-porous core with a large quantity of network interpenetration, rationalised by the calculated periodic models plotted in Figure 5.29, surrounded by a thin shell of highly porous, less dense and open OSPC-1 structure to rationalise the experimental porosity.^{16,95,323}

To consider this further, a 100 nm sphere could be imagined, composed of dense, non-porous OSPC-1 framework such as periodic model MF ($y = 3.12 \text{ \AA}$, $z = 6.92 \text{ \AA}$, ethynyl end groups) to rationalise the experimental ¹³C ssNMR spectrum, surrounded by a shell made from OSPC-1 in the *dia* topology to rationalise the experimental porosity. A surface area of 643 m² g⁻¹, which is within the large error margin associated with calculating BET surface areas of microporous materials of the experimental OSPC-1 surface areas,^{16,36,323,412} could be obtained with a *dia* shell thickness of only 2 nm of the overall sphere, corresponding to 10% of the sphere volume.⁹⁵

This theory was validated by the construction of a sphere of OSPC-1 with ethynyl end groups (Figure 5.32a) using the Ambuild code.^{19,95} This was done by seeding a tetraethynylmethane monomer into the centre of a cuboidal unit cell. The four terminal carbon atoms acted as end group atoms, which form part of the new bonds, and the hydrogen atoms acted as cap atoms, which define the vector that the new bond can form along. The network formed *via* a loop consisting of a *growBlocks* step to bond a new building block to a free end group within the simulation cell, followed by geometry optimisation and a *zipBlocks* step, which relaxed the original bond length and bond angle

margins defined to assess whether any additional bonds could form.⁹⁵ To give a representative OSPC-1 framework, relatively tight margins of 2 Å and 60° were utilised during the *zipBlocks* steps. This was followed by geometry optimisation and NVT (constant number of molecules, cell volume and temperature) molecular dynamics within HOOMD-blue (Figures C.1 and C.2).^{193,194} The loop continued until 442 tetraethynylmethane monomer building blocks had been added into the unit cell, at which point the *growBlocks* step was no longer included within the protocol, continuing only with the *zipBlocks*, geometry optimisation and molecular dynamics steps to allow any remaining bonds to form within the specified criteria. This yielded a structure with a diameter of approximately 7 nm, with ~60% of unreacted ethynyl end groups, all of which were located around the edges of the polymer cluster.⁹⁵

This structure was comparable to the proposed outer shell of the OSPC-1 structure, with large porous regions around the edges of the polymer cluster. The PSD of the simulated sphere is given in Figure 5.32b, calculated using Poreblazer 4.0.¹⁶² The simulated PSD matched the experimental fairly well in the region of 0.7–2.0 nm, although it does contain sharper peaks compared to the experimental due to the smaller system size, and a better match would be expected with a larger model size.⁹⁵ The PSD functions were also calculated for OSPC-1 in the *dia*, *dia-c*, *pts-c* and *unc-c* topologies (Figure 5.32c), as these maintained the OSPC-1 bonding motif after the CASTEP geometry optimisation.^{66,67,86,97,213,219–221} OSPC-1 in the *dia-c* topology was non-porous with pores smaller than 0.5 nm so the PSD was not plotted for this model.⁹⁵ The PSD of OSPC-1 in the *dia*, *pts-c* and *unc-c* topologies each have one peak, centred at 1.15 nm, 0.59 nm and 0.77 nm, respectively.⁹⁵

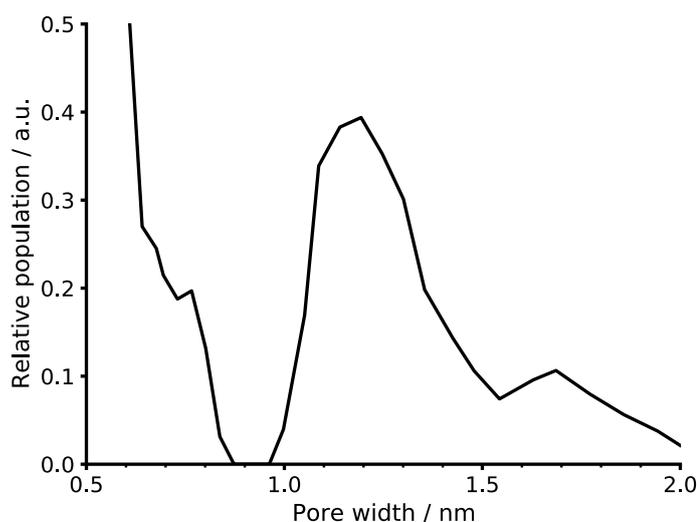


Figure 5.31: PSD of OSPC-1b. Figure adapted from references 95, upon which this chapter is based, and 323.

The PSDs of the OSPC-1 in the crystalline topologies show an incredibly good match to the experimental distribution, supporting the hypothesis of a dense, non-porous OSPC-1 core surrounded by a highly porous shell composed from OSPC-1 in the *dia*, *pts-c* and *unc-c* topologies. This was further confirmed by the high-resolution transmission electron microscopy (TEM) of OSPC-1a (Figure 5.32d), which shows a comparable surface topology to the hypothesised model.⁹⁵

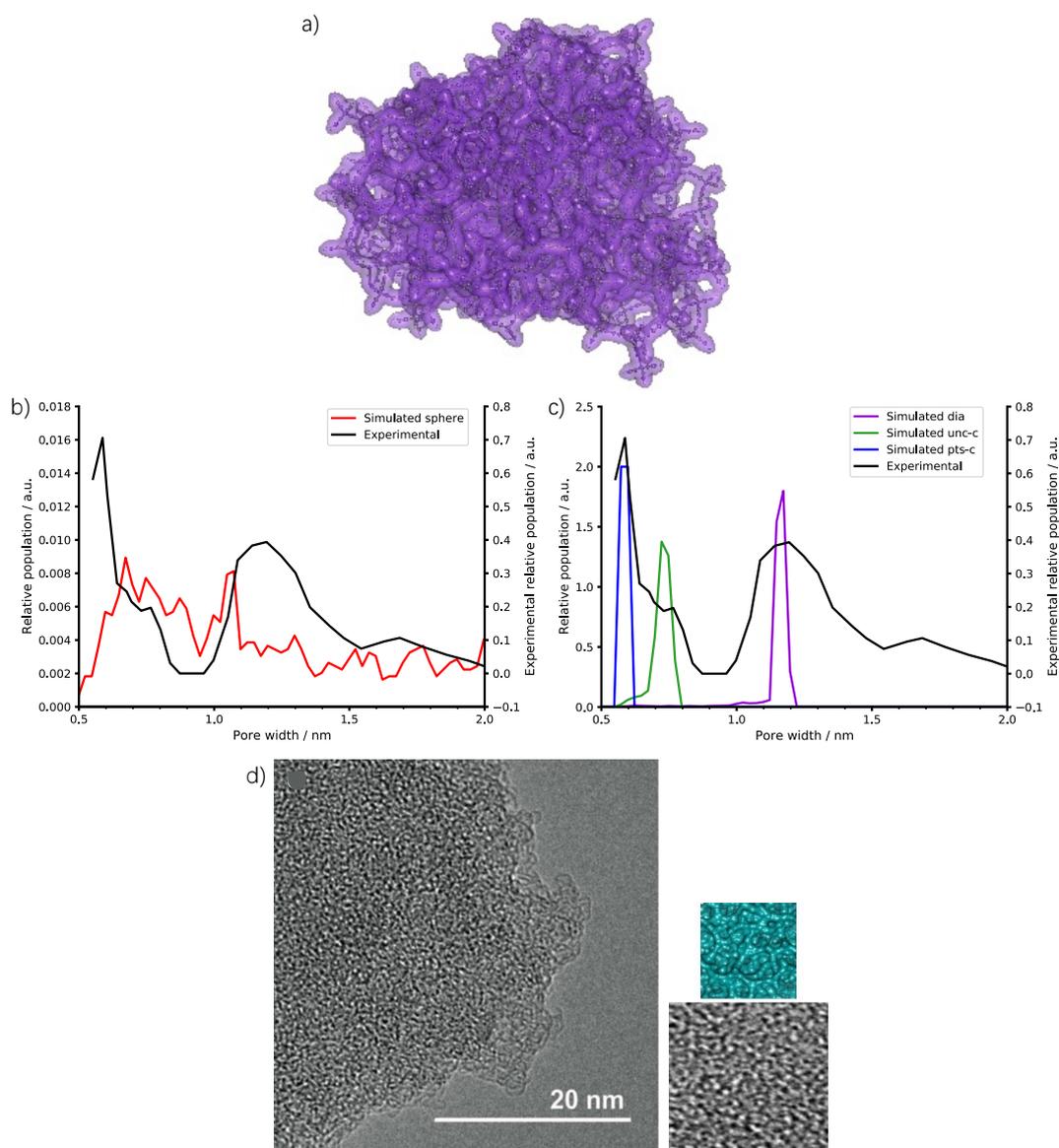


Figure 5.32: a) Proposed sphere model of OSPC-1, simulated using Ambuild¹⁹ and visualised with a van der Waals surface. b) PSD of OSPC-1b – black, and simulated sphere – red. c) PSD of OSPC-1b – black, and OSPC-1 in the *dia* (purple), *pts-c* (blue) and *unc-c* (green) topologies. A section of the simulated sphere model with a highlighted pore of diameter 1.2 nm is given in the inset. The PSDs of the sphere and crystalline models were calculated using Poreblazer 4.0.¹⁶² d) High-resolution TEM image of OSPC-1a, with a zoomed in section to show a similarly sized structure to the simulated sphere model. Figure adapted from reference 95, upon which this chapter is based, and high-resolution TEM image taken from reference 16.

5.5 Conclusions

Organically synthesised porous carbon (OSPC)-1 is an intriguing potential battery anode of the future, displaying over twice the capacity of graphite, the current commercial standard.⁸⁷ Yet despite being synthesised *via* two distinct synthetic methods,^{87,323} the second of which shows no possibility for the formation of sp^2 hybridised carbon,³²³ the question still arose as to why, at first glance, the ^{13}C solid-state nuclear magnetic resonance spectra of OSPC-1a and OSPC-1b indicate that the materials are composed of sp^2 and sp^3 hybridised carbon atoms.^{16,323}

This led to the work described within this chapter, where a series of small fragments of the proposed OSPC-1 structure, crystalline topologies, and periodic models were undertaken and analysed, considering the alternative end groups within the OSPC-1a and OSPC-1b mechanisms to rationalise the structure.⁹⁵

The fragments of the OSPC-1 structure did not rationalise the experimental ^{13}C solid-state nuclear magnetic resonance spectra well, and neither did OSPC-1 in the various crystalline topologies studied.⁹⁵ It was established that the only way to rationalise the NMR spectrum was through a series of dense, interpenetrated OSPC-1 chains, where interactions between neighbouring linkers could occur. However, this was unable to rationalise the experimental porosity, leading to the rationalisation of the structure *via* a spherical model of OSPC-1.⁹⁵ This was dominated by the dense, non-porous and highly interpenetrated model that was able to rationalise the experimental ^{13}C solid-state nuclear magnetic resonance spectrum, with a thin outer shell based on OSPC-1 in the open, porous *dia*, *pts-c* and *unc-c* topologies.⁹⁵ It is anticipated that this approach could be used to rationalise the NMR spectra of other amorphous materials in the future.

Chapter 6

Artificial synthesis of a covalent triazine framework as a case study for rationalising the FT-IR spectra of amorphous CTF materials

Common sense is not so common.

Voltaire

6.1 Summary

Covalent triazine-based frameworks (CTFs) are a sub-class of microporous organic polymers with varying properties depending on the synthetic route employed. When synthesised under kinetic control, an amorphous framework is generated, giving rise to additional peaks within the experimental FT-IR spectrum compared to those of the crystalline analogue. The origin of these additional peaks was previously unknown, which led to the artificial synthesis of the amorphous framework, described in this chapter and presented by Mollart and co-workers.⁴⁵ The aim of this chapter was to rationalise these unidentified peaks by modelling the full acid-catalysed cyclotrimerisation mechanism and identifying local structural features that had previously not been considered.

This was undertaken by generating a representative structural model of an amorphous CTF material using the Ambuild code, mimicking the full acid-

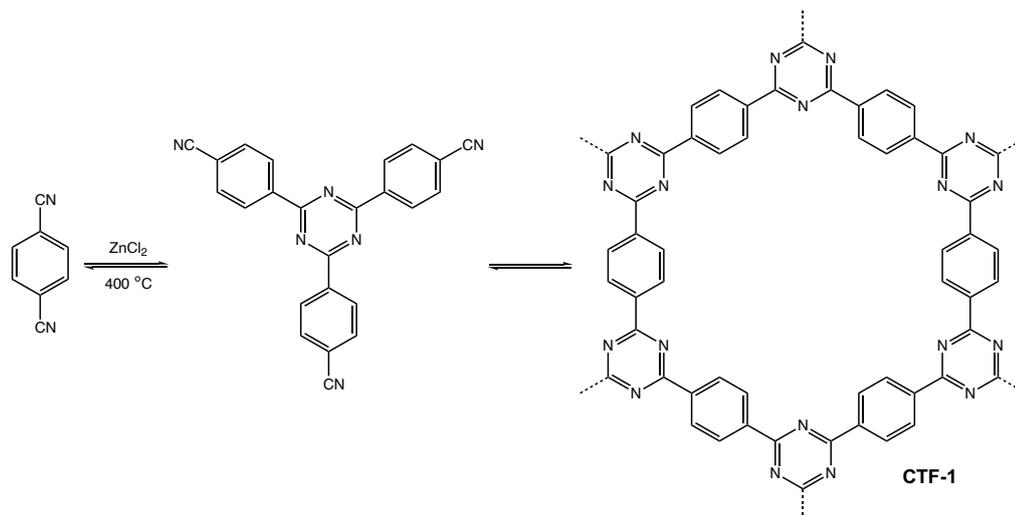
catalysed cyclotrimerisation reaction. This model showed a range of structural features as well as the expected triazine ring product. The FT-IR spectrum of each feature was then calculated, and the peaks compared to the experimental spectrum. The experimental FT-IR spectrum was then rationalised using the range of structural features observed and absorbed guests within the pore structure.⁴⁵

6.2 Introduction

CTFs have applications in fields such as drug delivery,³⁴ bandgap engineering,⁴¹³ photocatalytic hydrogen production,⁴¹⁴ heterogeneous catalysis,⁴¹⁵ electrochemical energy storage,⁴¹⁶ and gas uptake and separations.⁴¹³ These can be considered as either a sub-class of covalent organic frameworks (COFs) or conjugated microporous polymers (CMPs), due to their covalent, π -conjugated and microporous nature.²⁰ These materials exhibit varying degrees of crystallinity within the framework depending on the synthetic route employed, where a fully crystalline CTF appears more COF-like, whilst an amorphous CTF more closely resembles a CMP.^{5,14,20,413} Like many porous organic polymers, they can be formed from inexpensive, commercially available starting materials.¹⁴

CTFs were pioneered by Thomas and Antonietti and first reported in 2008.¹⁴ The first synthetic method involved an ionothermal synthesis of 1,4-dicyanobenzene using one equivalent of molten ZnCl_2 as solvent and catalyst at high temperatures (400 °C, Scheme 6.1).¹⁴ After 40 h, this was judged by FT-IR to be an almost complete conversion to form the expected black CTF-1 product, with product yields in the range of 90% and a Brunauer–Emmett–Teller (BET) surface area of 791 $\text{m}^2 \text{g}^{-1}$.^{14,417,418} The resulting CTF was thermally and chemically stable, and crystalline, with an analogous layered structure to COF-1.^{11,14,419}

ZnCl_2 was chosen as the ionic salt melt as it is a Lewis acid within which the monomers are soluble, allowing it to catalyse the cyclotrimerisation reaction, and is a stable solvent at the high temperature range required for the ionothermal synthesis.⁴¹⁸ However, the drawback is that it is difficult to remove all of the salt post-synthesis, leaving up to 5 wt.% residue.⁴¹³ This led the researchers to come up with new strategies to minimise the amount of residual salt, such as by grinding up the product at the end of the synthesis, followed by washing in water and dilute hydrochloric acid.⁴¹⁷



Scheme 6.1: Reaction scheme to form CTF-1 from 1,4-dicyanobenzene via an ionothermal synthesis catalysed by ZnCl₂ at 400 °C. Figure reprinted (adapted) with permission from P. Kuhn, A. Forget, D. Su, A. Thomas and M. Antonietti, *J. Am. Chem. Soc.*, 2008, **130**, 13333–13337. Copyright 2008 American Chemical Society.

Modifying the synthetic procedure can change the crystallinity and porosity of the CTF framework produced. Firstly, increasing the quantity of catalyst relative to the monomers yields an amorphous CTF material with an enhanced surface area.^{14,419,420} Elevating the temperature (Figure 6.1)⁴¹⁷ and reaction time also increases the surface area and total pore volume of the CTF, with the BET surface area of CTF-1 increasing from 920 to 3270 m² g⁻¹ when using five equivalents of ZnCl₂ salt and a two-step heating procedure of 400 °C for 20 h, followed by 600 °C for 96 h, rather than 400 °C for 40 h.^{5,417} The increase in surface area with increased temperature was suggested as being due to the decomposition and subsequent rearrangement of the polymer to a less-dense amorphous material with more mesopores in addition to the existing microporosity, which was confirmed by nitrogen sorption isotherm. Only the polymer synthesised at 400 °C gave rise to a type I isotherm, indicative of micropores only. All the others had type IV isotherms with a characteristic hysteresis to suggest mesopore formation in addition to the existing microporous network observed within the crystalline equivalent.^{417,420} Whilst the reaction was judged to be reversible at 400 °C, further increasing the temperature limited the reversibility.^{224,417,418} Increasing the reaction time led to an increased surface area by making the reversibility of the triazine ring formation more facile, leading to better thermodynamic control of the reaction.^{419,421}

Judicious choice of the monomers was required to only consider those that would be stable under the reaction conditions. However, many are thermally stable to temperatures well above 400 °C, maintaining a large library of start-

ing materials to choose from.¹⁴ Changing the monomer size and/or geometry does not appear to show a direct correlation with surface area and pore size, leading the researchers to suggest that the structures interpenetrate in three-dimensional space.⁴¹⁸ Applying more ZnCl_2 salt melt to increase the dilution of the polymer, and/or using much larger monomers with increased nodal dimensionality both help to decrease interpenetration in the network, giving an enhanced surface area. For example, replacing the CTF-1 monomer with a tetrahedral 1,3,5,7-tetrakis(4-cyanophenyl)adamantane (TCA, Figure 6.2) monomer at 400 °C for 40 h with 4 mol equiv. of salt gives BET surface areas as high as $1618 \text{ m}^2 \text{ g}^{-1}$ compared to CTF-1 at 400 °C for 40 h with 5 mol equiv. of salt reporting a surface area of $920 \text{ m}^2 \text{ g}^{-1}$.^{43,418}

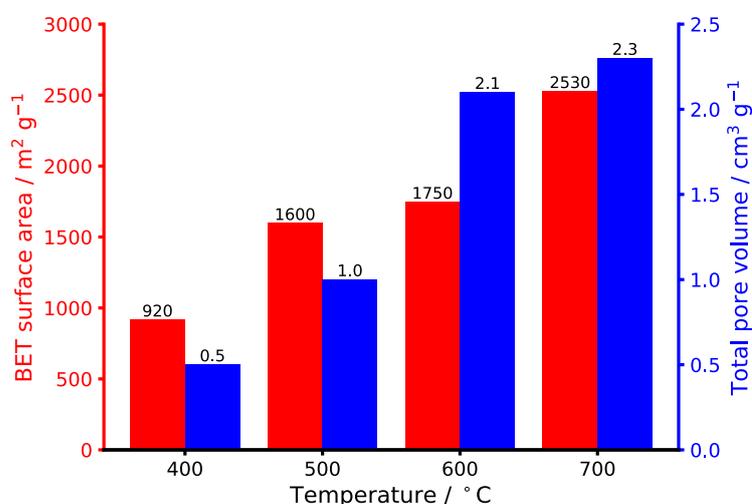


Figure 6.1: Influence of temperature on the BET surface area (red) and total pore volume (blue) of CTF-1. Data reprinted (adapted) with permission from P. Kuhn, A. Thomas and M. Antonietti, *Macromolecules*, 2009, **42**, 319–326. Copyright 2009 American Chemical Society. Pore volumes have been rounded to one decimal place.

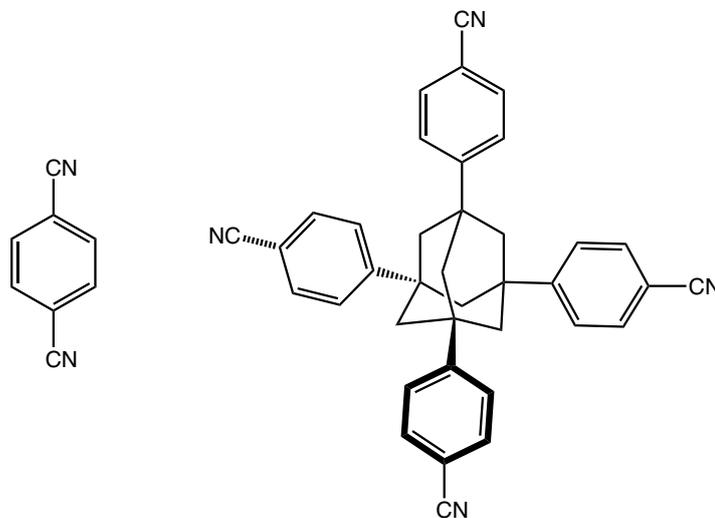


Figure 6.2: Structures of the 1,4-dicyanobenzene (left) and 1,3,5,7-tetrakis(4-cyanophenyl)adamantane (TCA, right) monomers.⁴¹⁸



Figure 6.3: Reaction scheme to synthesise amorphous CTF materials from the cyclotrimerisation and subsequent decomposition and rearrangement of tetracyanoquinodimethane (TCNQ). Figure reproduced (adapted) from reference 416 with permission from John Wiley and Sons. Copyright 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

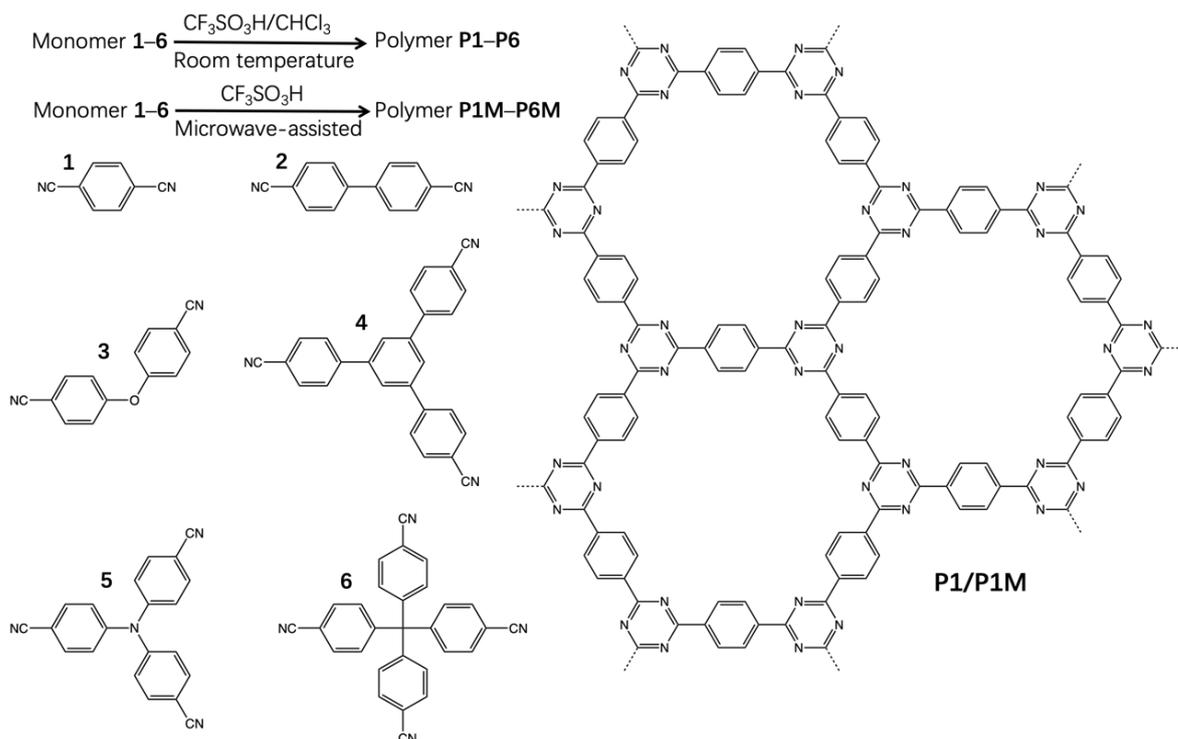
Molten ZnCl_2 at 400 °C was also used to make tetracyanoquinodimethane (TCNQ)-CTFs with varying nitrogen content for energy storage and/or supercapacitor applications.⁴¹⁶ The reaction scheme is given in Figure 6.3. Ideal supercapacitor materials require both a high nitrogen content to raise the electrochemical reactivity, yielding a better cycling stability and a higher specific capacity and rate performance,²⁶⁸ and high surface areas to have an increased contact with the electrolyte and more ion diffusion.⁴¹⁶

Once the cyclotrimerisation was complete, the temperature was elevated to varying levels up to 900 °C, which led to decomposition and rearrangement of the structure to an amorphous TCNQ-CTF material in all cases, as determined by PXRD.⁴¹⁶ The BET surface area increased with increasing temperature, but the nitrogen content decreased, especially on going from 600 °C (17.39 wt.% nitrogen content) to 700 °C (8.54 wt.% nitrogen content). The researchers postulated that the loss of nitrogen gave rise to the additional mesopore formation.⁴¹⁶ TCNQ-CTF synthesised with the elevated temperature of 800 °C (TCNQ-CTF-800) was judged to be the 'best' of the reported materials due to having a high BET surface area of 3663 $\text{m}^2 \text{g}^{-1}$ and a relatively high nitrogen content of 8.13 wt.%. This almost halved on increasing the temperature to 900 °C, suggesting that whilst the surface area increased, the reduction in nitrogen content would decrease the potential for energy applications.⁴¹⁶ TCNQ-CTF-800 showed promise for use in both an alkaline and ionic-liquid system, with a high capacitance (383 F g^{-1}) and commendable energy density (42.8 W h kg^{-1}), respectively, and good cycling stability in both systems.⁴¹⁶

Whilst the original method to make CTFs required a high-temperature ionothermal synthesis with a long reaction time (40 h was used to synthesise CTF-1 in the original report),¹⁴ other methods have since been developed. These include an AlCl_3 -catalysed Friedel-Crafts alkylation of cyanuric acid with an aromatic co-monomer *via* a solvent-free mechanochemical synthesis,^{223,422} and a time- and energy-saving microwave-assisted synthesis in the presence of a trifluoromethanesulfonic (TFMS) acid or anhydride catalyst.^{223,413,423} Of these, perhaps the most well-utilised catalyst is TFMS acid, which was first

used to synthesise CTFs in 2012 by Ren and co-workers, acting as a Brønsted acid to produce amorphous CTFs P1–P6 at room temperature, and P1M–P6M when synthesised under microwave conditions (Scheme 6.2).⁴¹³

The TFMS acid-catalysed cyclotrimerisation mechanism used to form these CTFs is given in Figure 6.4.⁴⁵ Firstly, one of the cyano groups on the monomer attacks a nearby proton, forming protonated intermediate **1**. Then, a cyano group on a second equivalent of monomer attacks the carbocation of protonated intermediate **1**, forming protonated intermediate **2**. This step occurs again, this time with a third equivalent of monomer attacking the carbocation of protonated intermediate **2** to form protonated intermediate **3**, which simply needs to cyclise to form the expected triazine ring product and regenerate the acid catalyst.⁴⁵ In the case of TFMS acid, it is likely that due to the strength of this ‘superacid’ ($\text{pK}_a \sim -14.7 \pm 2.0$ in water),⁴²⁴ arising due to the many possible resonance structures available between the sulphur and oxygen atoms, along with the electron-withdrawing fluorine groups, it will be fully dissociated in solution.



Scheme 6.2: Reaction scheme to prepare amorphous CTF materials P1–P6 *via* a trifluoromethanesulfonic (TFMS) acid-catalysed room temperature synthesis, and P1M–P6M *via* a TFMS acid-catalysed microwave-assisted synthesis. The idealised polymer structure of P1/P1M is also indicated. Scheme reproduced (adapted) from reference 413 with permission from John Wiley and Sons. Copyright 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The key advantage of preparing CTFs using TFMS acid over the conventional ionothermal ZnCl_2 synthetic route is that the former allows the reaction to occur at far reduced (room, in the case of P1–P6) temperatures and reaction times, meaning that some monomers that were previously too unstable to be considered under the harsh ionothermal conditions could be utilised in CTF formation with the milder TFMS acid method.⁴¹³ There is also less risk of competing side reactions and thermal decomposition occurring when using the TFMS acid method.⁴¹⁹ Additionally, this route had a better agreement in the experimental and theoretical elemental analysis compared to the ZnCl_2 ionothermal synthesis, suggesting that the materials possessed fewer defects; and it does not have the same issues with separating the excess catalyst from the product upon completion of the reaction.^{413,419} However, the TFMS acid itself does exhibit some safety concerns due to the strength of the acid.⁴²⁴

P1–P6 and P1M–P6M, which were thermally and chemically stable, like the original CTFs formed,^{14,413,417,418} had enhanced BET surface areas of up to $1152 \text{ m}^2 \text{ g}^{-1}$, and unlike the original black CTF-1 material, which has the same monomer as P1/P1M, had absorption and photoluminescence behaviour that could be tuned by the choice of monomers employed.^{223,413} Whilst P1–P6 were totally amorphous, P1M, P2M and P4M showed preferential ordering within the structure with some evidence of crystallinity, which was attributed solely to the microwave synthesis as all of the other variables were the same as the conventional (non-microwave) TFMS acid-catalysed synthesis.⁴¹³

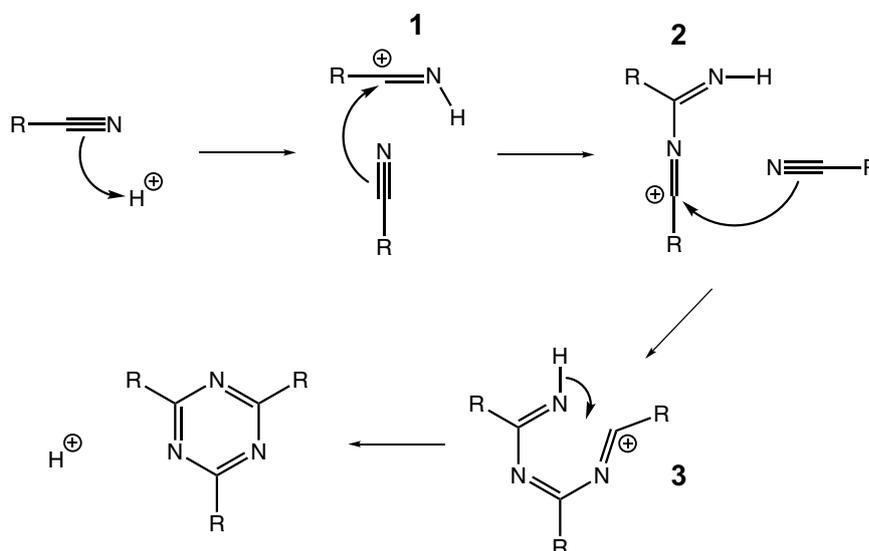


Figure 6.4: Trifluoromethanesulfonic (TFMS) acid-catalysed ring formation mechanism to synthesise the amorphous CTFs first reported in reference 413. In the case of P1, the R group is a phenyl ring with a para-substituted cyano group. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

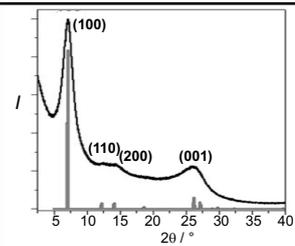
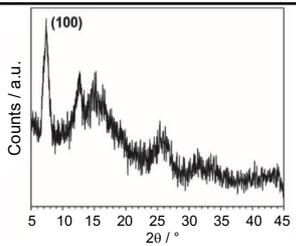
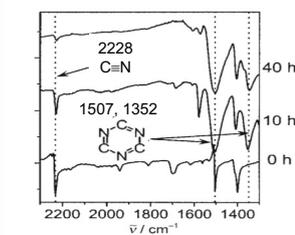
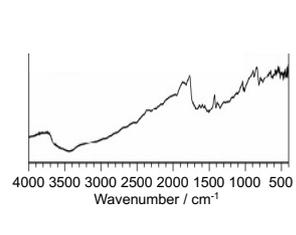
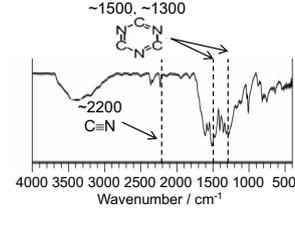
It was postulated that the increased pressure employed during the microwave synthesis enabled easier making and breaking of the triazine rings within the polymer, giving a more thermodynamic product with limited ordered domains within the framework compared to the totally amorphous room temperature reaction.⁴¹³ In general, the extra crystallinity and higher density of the microwave-synthesised materials gives them a lower porosity than the more amorphous, lower density room-temperature ones, as reported in the original CTF-1 paper, however, P1 and P1M are both non-porous.^{14,413}

Table 6.1 compares the Fourier-transform infrared (FT-IR) and powder x-ray diffraction (PXRD) patterns of the CTFs synthesised from a 1,4-dicyanobenzene monomer with the different synthetic approaches: CTF-1 (ionothermal ZnCl_2 -catalysed synthesis) is crystalline with permanent porosity,¹⁴ P1M (microwave-assisted TFMS acid-catalysed synthesis) shows preferential ordering with some crystallinity and no porosity, and P1 (room temperature TFMS acid-catalysed synthesis) is completely amorphous with no porosity.⁴¹³ P1 is therefore arguably the only CMP derivative, as CMPs are formally defined as amorphous frameworks.²⁰ There are two possibilities for why CTF-1 is porous and P1/P1M have no permanent porosity: firstly, the ZnCl_2 ionic salt may be acting as a template, as hypothesised by Kuhn and co-workers⁴¹⁸ and secondly, the elevated temperature, which gives rise to a more defect-rich structure as shown by elemental analysis.^{413,419}

As has been demonstrated throughout this thesis, simulating representative structural models of polymer systems is not trivial. Some research groups have taken the approach of using a kinetic Monte Carlo (KMC) method to assemble the product.^{11,425} However, whilst the KMC approach is appropriate to model COF-5, which is synthesised *via* a one-step condensation reaction, it cannot currently be used to model amorphous CTFs such as P1 due to the more complex multi-step mechanism required to make this material.^{45,425}

CTFs have been widely developed since the original report in 2008,¹⁴ with a number of different methods now available to synthesise them as discussed above to give a small number of crystalline CTF materials and a vast majority of amorphous CTFs.^{14,413,422,423} However, almost all of the computational modelling of CTFs has taken the assumption that every monomer will fully react to form the expected triazine ring product, meaning that small, density functional theory (DFT)-optimised fragments could be used as building blocks to create the overall polymer network, adding each fragment as a rigid-body and reacting all available end groups in the addition of each layer to the structure. Upon adding each layer, the polymer structure was optimised using the Polymer Consistent Forcefield (PCFF),¹¹⁵ and once all layers had been added, the structure was fully optimised.⁴³

Table 6.1: Comparison of the PXRD patterns and FT-IR spectra of the CTF synthesised from a 1,4-dicyanobenzene monomer using the different synthetic approaches. CTF-1 spectra reproduced (adapted) from reference 14 with permission from John Wiley and Sons. Copyright 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. P1/P1M spectra reproduced (adapted) from reference 413 with permission from John Wiley and Sons. Copyright 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Name	CTF-1	P1M	P1
Synthetic route	<ul style="list-style-type: none"> High-temperature ionothermal synthesis. ZnCl₂ catalyst. 	<ul style="list-style-type: none"> Microwave-assisted synthesis. TFMS acid catalyst. 	<ul style="list-style-type: none"> Room temperature synthesis. TFMS acid catalyst.
Degree of crystallinity	<ul style="list-style-type: none"> Crystalline. 	<ul style="list-style-type: none"> Some preferential ordering. 	<ul style="list-style-type: none"> Amorphous.
PXRD spectrum			<ul style="list-style-type: none"> Not reported due to the amorphous nature of the material.
FT-IR spectrum			

This gave rise to structural models of varying sizes that were able to rationalise the differences in porosity of the P1-P5 and P1M-P5M set of CTFs,⁴¹³ showing that the porosity of the CTF was highly dependent on the dimensionality and flexibility of the node and linker chosen. The conclusion drawn was that in order to synthesise a high-porosity CTF material from a two-dimensional node and a one-dimensional linker, the best combination would be to have high nodal dimensionality and short, rigid linkers, in order to prevent interpenetration of the structure and maintain porosity throughout.⁴³

However, whilst the assumption that every monomer would fully react to form part of a triazine ring was a reasonable first assumption to make, based on the difficulty in modelling the full reaction mechanism, the fact that many CTF materials are amorphous gives a strong likelihood that this assumption will not be true in all cases, and that there will be additional structural diversity that will form under the kinetically controlled reaction.⁴⁵ This is supported by the experimental FT-IR spectrum of P1,⁴¹³ which has extra features compared to the CTF-1 spectrum¹⁴ that cannot be fully rationalised when simply considering the expected triazine product.

In this chapter, the artificial synthesis of the fully amorphous P1 material,⁴⁵ first reported by Ren and co-workers,⁴¹³ is shown, following the full synthetic

pathway and TFMS acid-catalysed cyclotrimerisation mechanism to generate a representative structural model using the Ambuild code.¹⁹ Following this, the FT-IR spectra and porosity properties of this representative model, calculated using Gaussian B3LYP/6-31G* and Poreblazer,^{79,85,90,91,93,94,162,163,212} respectively, are compared to previously reported experimental data.^{45,413} The simulated model showed evidence of a range of additional structural features, including alternative rings, pre-rings, and neutral intermediates. It was found that the only way to rationalise the experimental P1 FT-IR spectrum was to incorporate the additional structural features and absorbed guest molecules such as water, hydronium ions, and carbon dioxide.⁴⁵

6.3 Methodology

In the first step of the reaction mechanism, shown in Figure 6.4 above,⁴⁵ the nitrogen atoms of the cyano groups on the monomer are protonated. This was mimicked by making the assumption that in the presence of a ‘superacid’ such as TFMS, the acid would have fully dissociated in solution and the nitrogen atoms would therefore already be protonated.⁴⁵ Also, a hydrogen atom was added to the carbon atom adjacent to the nitrogen atom of each cyano group to add cap atoms to direct the Ambuild bonding, and the C≡N cyano (triple) bond was converted to a C=N imine (double) bond in order to establish the correct PCFF typing in the product.^{45,115}

This created the Ambuild building block for P1 (Figure 6.5a), which was equivalent to di-protonated intermediate **1** in the reaction mechanism (Figure 6.5b), with the carbon and nitrogen atoms of the imine acting as end groups, and the adjacent hydrogen atoms as cap atoms. Bonding rules were then specified that allowed an unreacted carbon end group to react with an unreacted nitrogen end group, where new bonds were able to form as part of a *zipBlocks* test if unreacted carbon and nitrogen end group atoms were within a length of 3.2 Å and an angle of 120°.⁴⁵ Please refer to Chapter 2 section 2.4.1 for a full explanation of the parameters used within the Ambuild code. As an initial test of this building block and the forcefield used in the artificial synthesis, the first protocol used a simplified building block, where only one of the cyano groups on the monomer was able to react and form bonds (Figure 6.5c).⁴⁵ This made assessing the new bonds, angles, and dihedrals formed within the intermediate structures as part of the test network generation easier by slowing the rate of the acid-catalysed ring formation reaction.

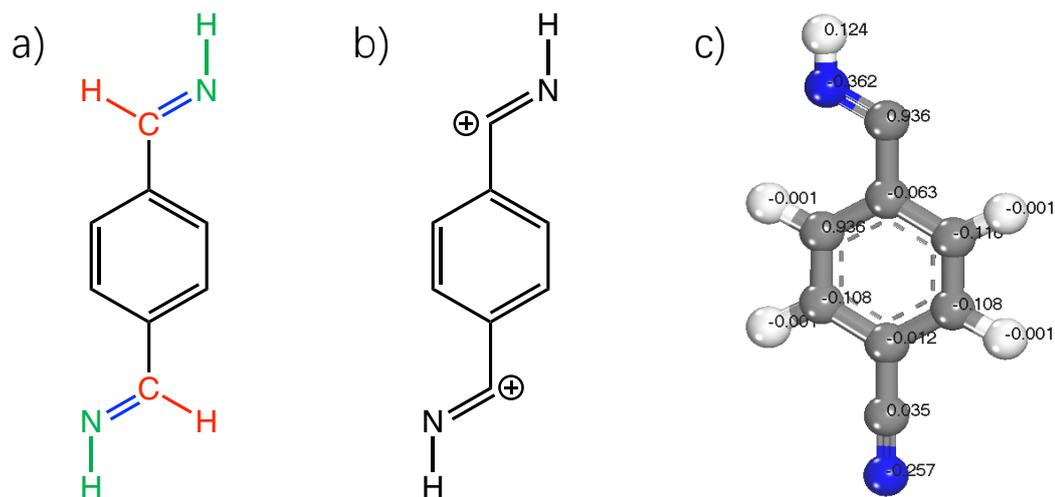


Figure 6.5: The Ambuild building blocks used in the artificial synthesis of P1. a) The P1 monomer building block is created by protonating the nitrogen atoms of the cyano groups (green), adding an Ambuild hydrogen cap atom to the carbon atoms of the cyano groups (red), and replacing the $C\equiv N$ cyano bond with a $C=N$ imine bond to give the correct PCFF typing in the product (blue). b) The resulting monomer building block is equivalent to the di-protonated 1,4-dicyanobenzene intermediate. c) The simplified test building block used to set up the Ambuild system for P1 showing the Gasteiger charges on each atom.¹⁹⁸ Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

Throughout the test protocol, the structures of the Ambuild intermediates and triazine ring product were compared to those of the equivalent structures optimised within the Forcite package of Materials Studio 5.0,¹⁵⁸ with forces calculated using the Gasteiger approach and optimisation using the PCFF forcefield (Figures 6.6–6.9),^{115,198} which was used as the starting point for the bonds, angles and dihedrals used within the Ambuild process.⁴⁵ Following this, the bonds, angles, and dihedrals within Ambuild were adjusted to better match those from the full PCFF-optimised structures,¹¹⁵ giving the final parameters reported in Table 6.2. The final force constants used within the full Ambuild methodology were large in order to maintain the rigidity and planarity of the structure throughout the HOOMD-blue geometry optimisation (Table 6.2).^{45,193,194} Once the forcefield was parametrised, the more complex building block given in Figure 6.5a was used to generate the full P1 structure, with an image of the input file given in Appendix D Figure D.1 and full details given after Table 6.2.⁴⁵ Please refer to Chapter 2 section 2.4.1 for a full explanation of the parameters used within the Ambuild code.

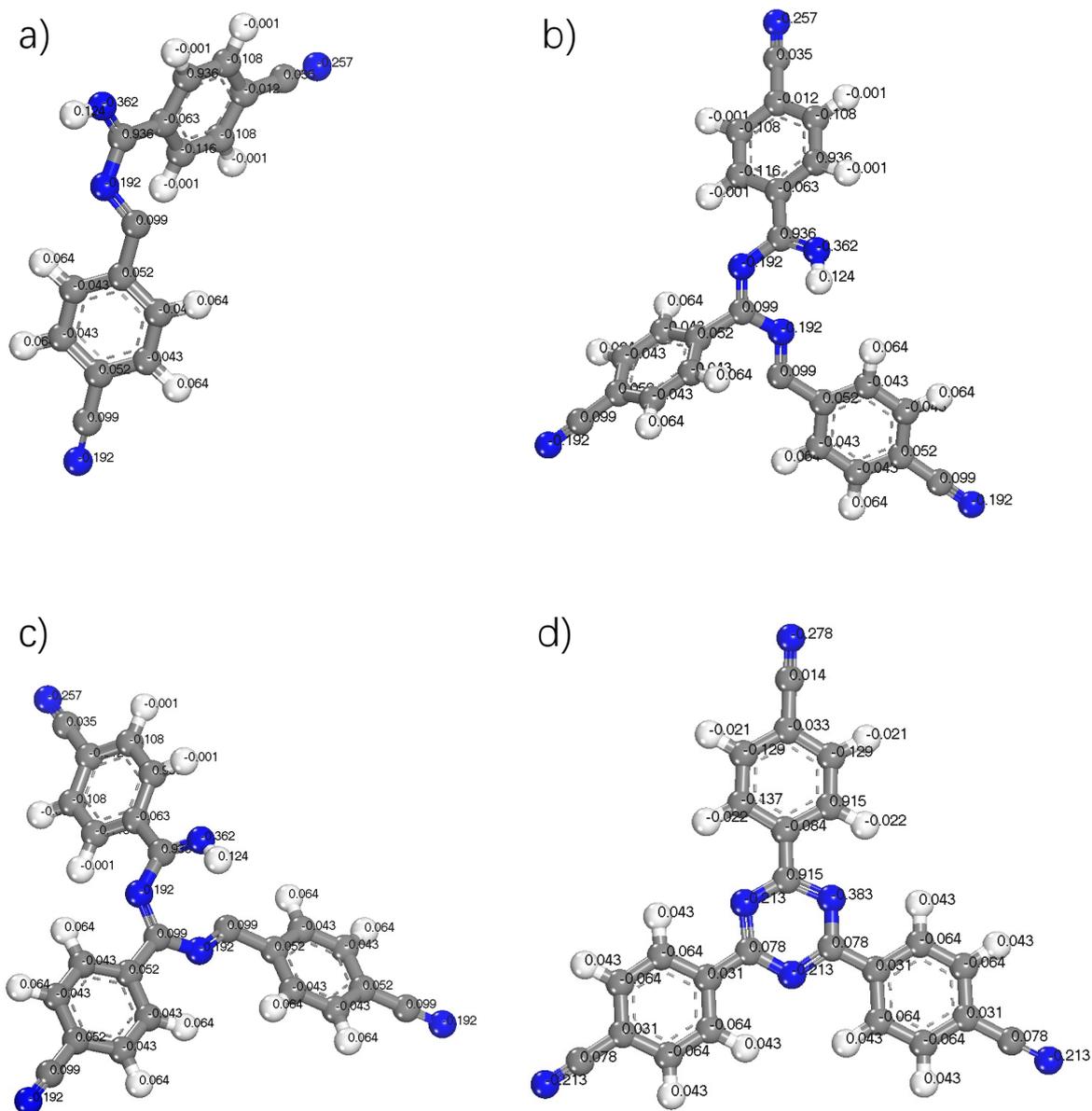


Figure 6.6: PCFF-optimised intermediates and triazine ring formed from the test building blocks (Figure 6.5 c), with the charges calculated using the Gasteiger approach shown on each atom.^{115,198} a) Monoprotonated intermediate formed from two test building blocks (intermediate 2 in Figure 6.4), b) monoprotonated intermediate formed from three test building blocks, c) monoprotonated intermediate formed from three test building blocks, arranged in the correct orientation to form the expected triazine ring product (intermediate 3 in Figure 6.4), d) triazine ring product formed from three test reactants. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

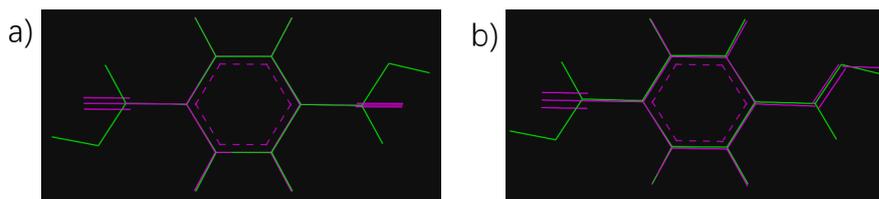


Figure 6.7: Comparison of the final P1 Ambuild building block (Figure 6.5a), optimised using the parameters reported in Table 6.2 within HOOMD-blue (green) and the 1,4-dicyanobenzene monomer (a) and protonated 1,4-dicyanobenzene monomer (b) optimised using the PCFF (pink).^{115,193,194} Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

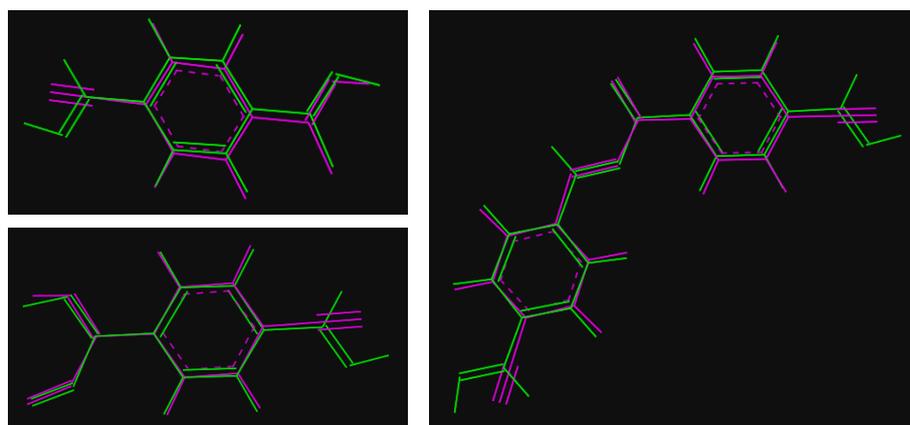


Figure 6.8: Comparison of the intermediate **3** structure formed from the final P1 Ambuild building block (Figure 6.5a), optimised using the parameters reported in Table 6.2 within HOOMD-blue (green) and the equivalent intermediate optimised using the PCFF (pink).^{115,193,194} Intermediate **3** is chiral in three-dimensional space and has been shown in multiple images to differentiate between the two distinct conformers that are equally viable. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

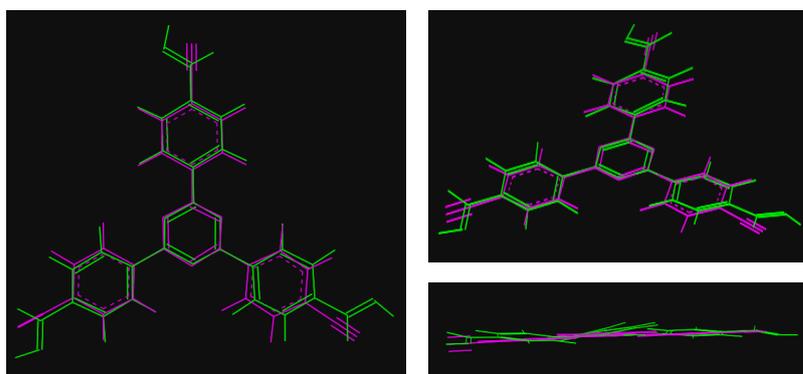


Figure 6.9: Comparison of the triazine ring formed from the final P1 Ambuild building block (Figure 6.5a), optimised using the parameters reported in Table 6.2 within HOOMD-blue (green) and the equivalent ring structure optimised using the PCFF (pink).^{115,193,194} Multiple views are given for visualisation purposes. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

Table 6.2: Forcefield parameters used in the Ambuild artificial synthesis of P1. The starting values used in the HOOMD-blue optimisation are based upon the Polymer Consistent Forcefield (PCFF).^{115,193,194} *k* – force constant, Angle – equilibrium angle, *d* – sign factor, *n* – angle multiplicity factor. Table reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

Starting values based upon PCFF			Ambuild parameter as used			
Angle	<i>k</i> / kcal mol ⁻¹	Angle / °	<i>k</i> / kcal mol ⁻¹	Angle / °		
n=c=1-cp	57	123.67	140	120		
c=1-n=c=1	67.6	125.1	2000	120		
Dihedral	<i>k</i> / kcal mol ⁻¹	<i>d</i>	<i>n</i>	<i>k</i> / kcal mol ⁻¹	<i>d</i>	<i>n</i>
n=c=1-cp-cp	10	3	1	100	-1	1
c=1-n=c=1-hc	10	3	1	500	-1	1
hc-cp-cp-c=	14	1	2	10	-1	1
cp-c=1-n=c=1	4	3	1	100	1	1
n=c=1-n=c=1	4	1	3	500	-1	1
n=c=1-n=hn	4	1	3	500	-1	1

One monomer building block, one triflate counter-ion and one chloroform solvent molecule were randomly seeded into an Ambuild simulation cell of length 50 Å.⁴⁵ A loop was undertaken where all of the triflate and chloroform blocks in the cell were deleted, and a new monomer building block was bonded to an available end group using a *growBlocks* step (simulating the formation of protonated intermediates **2** and **3** in the reaction mechanism). The structure was optimised using the adapted PCFF forcefield within HOOMD-blue,^{115,193,194} with a van der Waals cut-off distance of 10 Å. A sub-loop was then entered to seed in ten triflate counter-ions and ten chloroform molecules, which increased the structural sampling of the amorphous network by randomly re-seeding the building blocks into different configurations from the previous seed. The geometry was re-optimised to minimise the potential energy of the structure.⁴⁵ Then, a *zipBlocks* test was used to determine whether any of the existing end groups within the cell were within the pre-specified margins of a length of 3.2 Å and an angle of 120°. If so, a new bond formed between the two end groups (simulating the final step in the reaction mechanism to form the triazine ring). If any new bonds were able to form as part of the *zipBlocks* test, the structure was re-optimised.⁴⁵ The network generation continued until no new bonds had formed during the last twenty *zipBlocks* tests, at which point the network generation was complete and the structure was desolvated by removing any chloroform solvent molecules and triflate counter-ions within the simulation cell, giving the final amorphous P1 structure.⁴⁵

6.4 Results and Discussion

6.4.1 Network growth

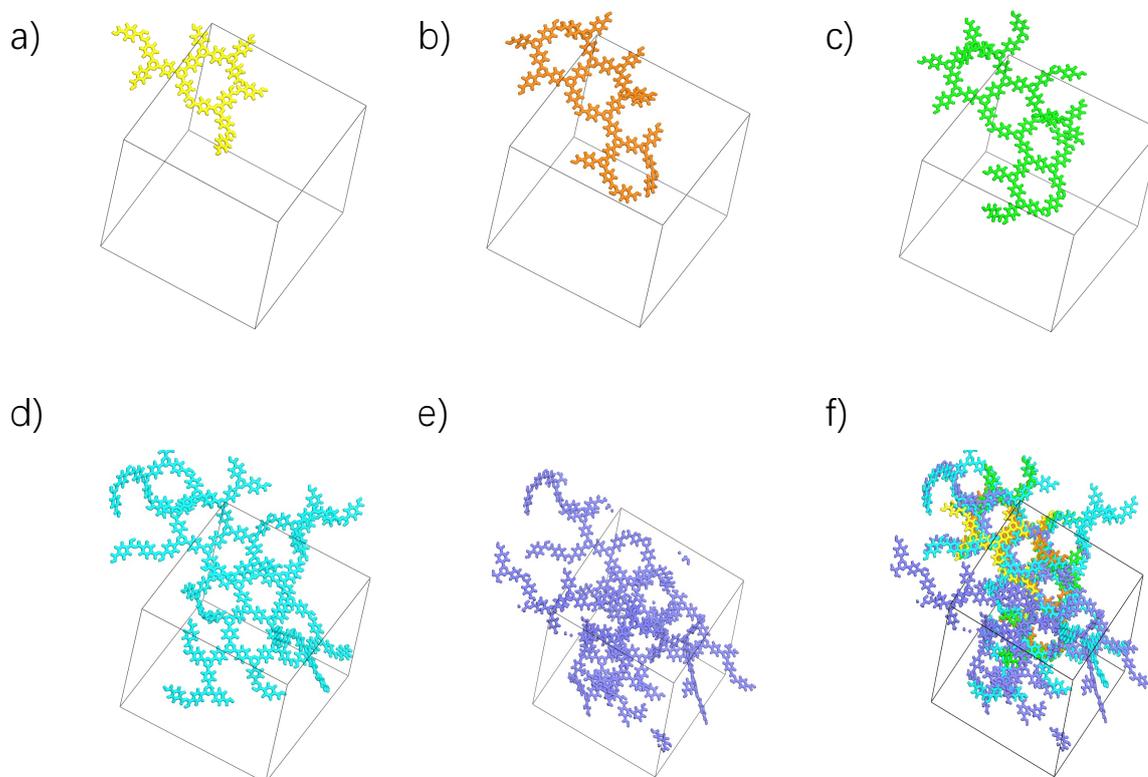


Figure 6.10: Images of the simulated P1 structure at different timesteps in the network generation. a) step 15. b) step 27. c) step 36. d) step 83. e) step 109. f) overlay of the structures in a)–e). Key: step 15 – yellow, step 27 – orange, step 36 – green, step 83 – blue, step 109 – purple. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

The structure was monitored throughout the network generation (Figure 6.10) and from very early stages in the reaction (step 15, Figure 6.10a), it could be seen that the expected 2-dimensional sheet-like structure of P1 was forming. There was some observed interruption to the planarity at longer timescales due to the solvent and triflate counter-ions present and from network interpenetration between polymer sheets to increase the stability of the structure through additional van der Waals dispersion interactions.⁴⁵

In addition to the expected triazine rings, other structural features were observed from the simulation.⁴⁵ These included alternative ring structures such as four- and eight-membered rings, as well as intermediate ‘pre-ring’ structures that could either cyclise to form a ring or react with an additional monomer to form a larger intermediate pre-ring structure. For example, a pre-4-ring structure could either cyclise to form a 4-membered ring or could react with an additional building block to form a pre-triazine feature.

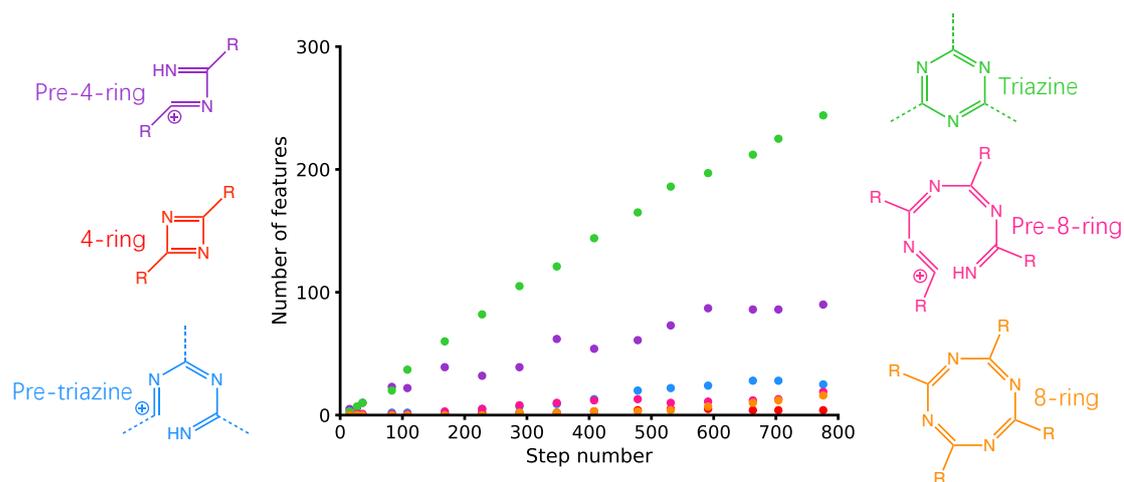


Figure 6.11: Prevalence of the different structural features observed within the simulated P1 structure as a function of the network generation step number. Idealised images of each structural feature are also included. Key: pre-4-ring – purple, 4-ring – red, pre-triazine ring – blue, triazine ring – green, pre-8-ring – pink, 8-ring – orange. Data reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

The prevalence of these alternative features was also tracked during the network generation (Figure 6.11). As expected, the most common ring feature observed at the end of the reaction was the triazine product.⁴⁵ This was followed by pre-4-rings and pre-triazines, which is also logical as these ring features are intermediates in the triazine ring formation process and could therefore react to form a triazine ring over a longer timescale. It is proposed that there was a larger percentage of pre-4-rings observed compared to the pre-triazines due to the rapid cyclisation step to form a triazine from the pre-triazine, which would mean that once the pre-triazine formed, the cyclisation to form a triazine ring would occur readily, whilst the pre-4-rings require an extra step to form the triazine product (reacting with a third equivalent of monomer to form the pre-triazine). The other ring features observed, namely the 4-rings, pre-8-rings and 8-rings, occurred much less prevalently, which is likely due to the nature of these rings meaning that they cannot react to form the expected triazine product due to being either too large, as in the case of the pre-8-rings and 8-rings, or to cyclising irreversibly before a third monomer building block could react with the ring structure in order to form a pre-triazine feature.⁴⁵

These additional ring features are a consequence of the amorphous conditions used in the synthesis⁴¹³ and replicated during the Ambuild simulation,⁴⁵ which place the reaction under kinetic control and ensure that any bond formation occurring is irreversible. This leads to the formation of alternative ring structures that would likely not be seen as part of the final product prepared

using the thermodynamic conditions employed in the synthesis of the crystalline CTF-1 material,¹⁴ which allow for the repeated making-and-breaking of chemical bonds in order to form the most stable, minimum energy framework structure.⁴⁵ An additional example of an ‘irregularity’ within the amorphous simulated P1 structure (Figures 6.12a and c) is a spiral-type macrocycle made up of a chain of P1 that continues to grow rather than cyclising to form the idealised planar sheet, as seen in Figures 6.12b and d. These additional features highlight the importance of being able to simulate as high a degree of statistical sampling as possible, either by increasing the simulation cell size, by generating multiple repeat structures, or by incorporating delete and re-seed steps in order to visualise different ‘snapshots’ of the experimental material.⁴⁵ Whilst some regions of the simulated P1 structure appear to look ‘idealised’ as in the case of the fragment given in Figures 6.12b and d, giving a first glance at a nicely ordered structure, others contain additional features such as the variety of ring structures observed (Figure 6.11) and the spiral-type macrocycle given in Figures 6.12a and c, and it is important to consider all of this structural diversity when modelling structures that aim to be representative of the experiment.⁴⁵

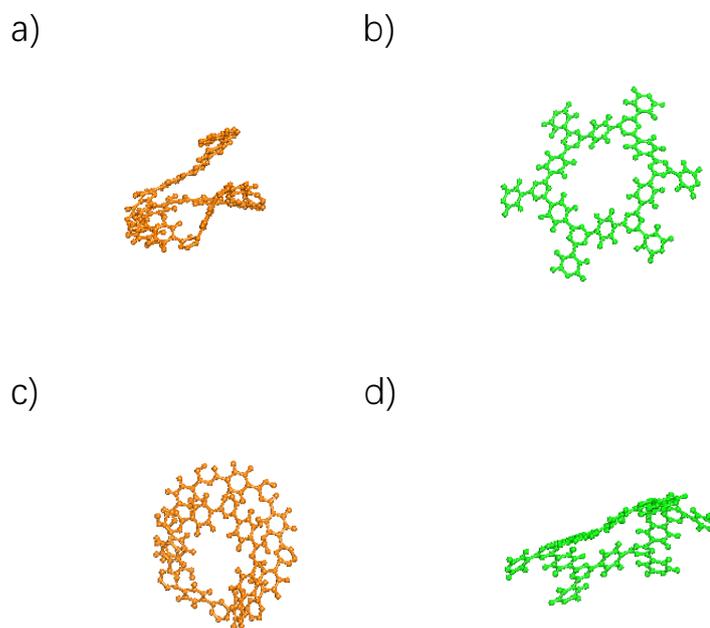


Figure 6.12: Additional structural features observed within the simulated P1 network. a) and c) top and side view, respectively, of a spiral-type macrocycle of P1, b) and d) top and side view, respectively, of an idealised sheet-like fragment of P1. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

6.4.2 Structure analysis and comparison to experiment

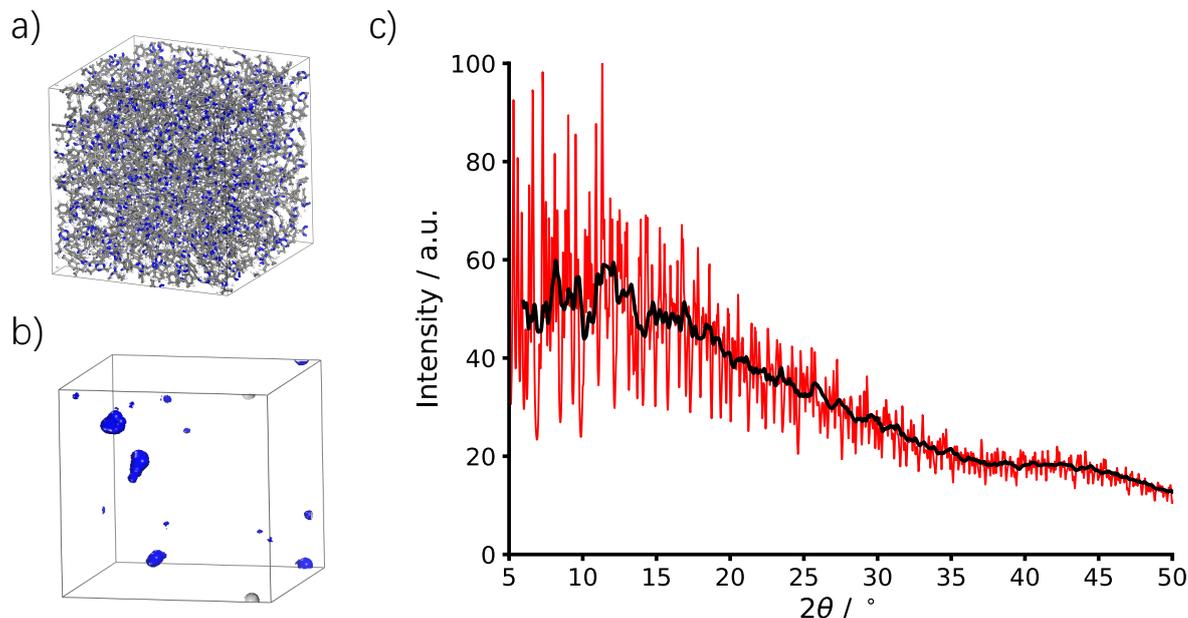


Figure 6.13: The final simulated P1 framework. a) Structure. b) Connolly surface (blue), obtained within Materials Studio 5.0 using a probe radius of 1.82 Å, appropriate for nitrogen sorption.¹⁵⁸ c) Simulated PXRD pattern (red) and rolling average (black), obtained using Mercury with a wavelength of 1.54056 Å.¹⁸⁹ Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

The final simulated P1 network (Figure 6.13a), with a combination of the structural features discussed previously, had a zero solvent-accessible surface area as a result of small, non-interconnected regions of porosity within the framework (Figure 6.13b).⁴⁵ Additionally, the structure was amorphous as determined by a simulated PXRD pattern (Figure 6.13c) that showed no clear peaks that evidenced any degree of crystallinity in the material. It is worth noting that whilst at first glance the simulated PXRD pattern does appear to contain relatively sharp peaks, these are an artefact of the system setup by making the assumption that a periodic unit cell could be used to model the full PXRD pattern of the amorphous P1 system.⁴⁵ It would be expected that increasing the simulation size or averaging the data over a large number of repeat structures, would give additional disorder within the system, leading to a broader PXRD pattern that more closely mimics the rolling average. These findings matched those reported experimentally, which added confidence to the simulated structure.^{45,413}

The FT-IR spectrum of P1 compared to CTF-1 contains additional peaks as well as those arising due to unreacted cyano groups ($\sim 2200\text{ cm}^{-1}$) and triazine rings ($\sim 1300\text{ cm}^{-1}$ and $\sim 1500\text{ cm}^{-1}$, Table 6.1), which occur in the regions of $\sim 1000\text{ cm}^{-1}$, 1600 cm^{-1} , 2350 cm^{-1} , and a broad peak in the

range of ~ 2900 – 3700 cm^{-1} , with contributions in the regions of $\sim 3200\text{ cm}^{-1}$, $\sim 3400\text{ cm}^{-1}$ and $\sim 3600\text{ cm}^{-1}$.^{14,45,413} Whilst some of these peak environments could be classified by considering the protonated intermediates within the structure made possible by the TFMS acid-catalysed mechanism,⁴²⁶ there were still some peaks that were of unknown origin.⁴⁵

It was postulated that these unknown peaks arose due to some of the additional structural features observed in the simulated model of P1, such as alternative ring structures and pre-rings, along with protonated intermediates and the expected triazine product.⁴⁵ Based on this hypothesis, a subset of fragments of the P1 structure containing these features were analysed. Each fragment in Figures 6.14 and 6.15 was optimised to a ground-state geometry using the B3LYP/6-31G* model chemistry within Gaussian 09,^{79,85,90,91,93,94,212} confirmed by harmonic wavenumber analysis to ensure that no imaginary vibrational modes were present.⁴⁵

The vibrational wavenumbers were scaled by applying a literature scaling factor of 0.961 to account for the systematic overestimation of the vibrational wavenumbers calculated when using this model chemistry to collect IR information,⁴²⁷ and to ensure that each vibration can be described using a harmonic approximation.⁴⁵ Unless otherwise specified, the IR intensities were plotted directly without scaling, and a Gaussian convolution (Equation 6.1) was applied to each peak in turn using a full-width half-maximum (FWHM) value of 20 cm^{-1} to mimic experimental peak broadening.⁴⁵ The value of 20 cm^{-1} was used as it is reflective of the meaningful resolution that is possible for the FT-IR spectrum of amorphous materials. A range of FWHM values were tested and 20 cm^{-1} was most appropriate to model the broadening.

$$f(x, A, \mu, \sigma) = \frac{A}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (6.1)$$

Equation 6.1 defines the Gaussian convolution as a function of continuous variable x , peak height (Equation 6.2, A is a constant), centre of the peak μ and standard deviation σ (Equation 6.3), that is applied to broaden each IR peak.⁴⁵

$$\text{Peak height} = \frac{A}{\sigma\sqrt{2\pi}} \quad (6.2)$$

$$\sigma = \frac{\text{FWHM}}{\sqrt{8\ln 2}} \quad (6.3)$$

The overall spectrum is then obtained by summing the contributions from each peak in the system.⁴⁵ Figures 6.14 and 6.15 show the calculated IR

spectra for each fragment.⁴⁵ When analysing these spectra, it could be seen that the charged fragments shown in Figure 6.14b, d and f do not match the experimental P1 spectrum very well, with more intense peaks in the higher wavenumber regions (4000–2000 cm^{-1}) compared to the equivalent neutral fragments in Figure 6.14c, e and g.⁴⁵ When these were discounted from the set of fragments, a number of the remaining ‘unknown’ peaks could be identified due to a good match between the calculated fragment models and experiment, namely, the peak at 3200 cm^{-1} , which features contributions from the 4-ring, pre-8-ring, and 8-ring systems, and the peaks at 1000 cm^{-1} and 1600 cm^{-1} , which arise from the summation of IR intensities from all of the remaining fragment models within this wavenumber range.⁴⁵

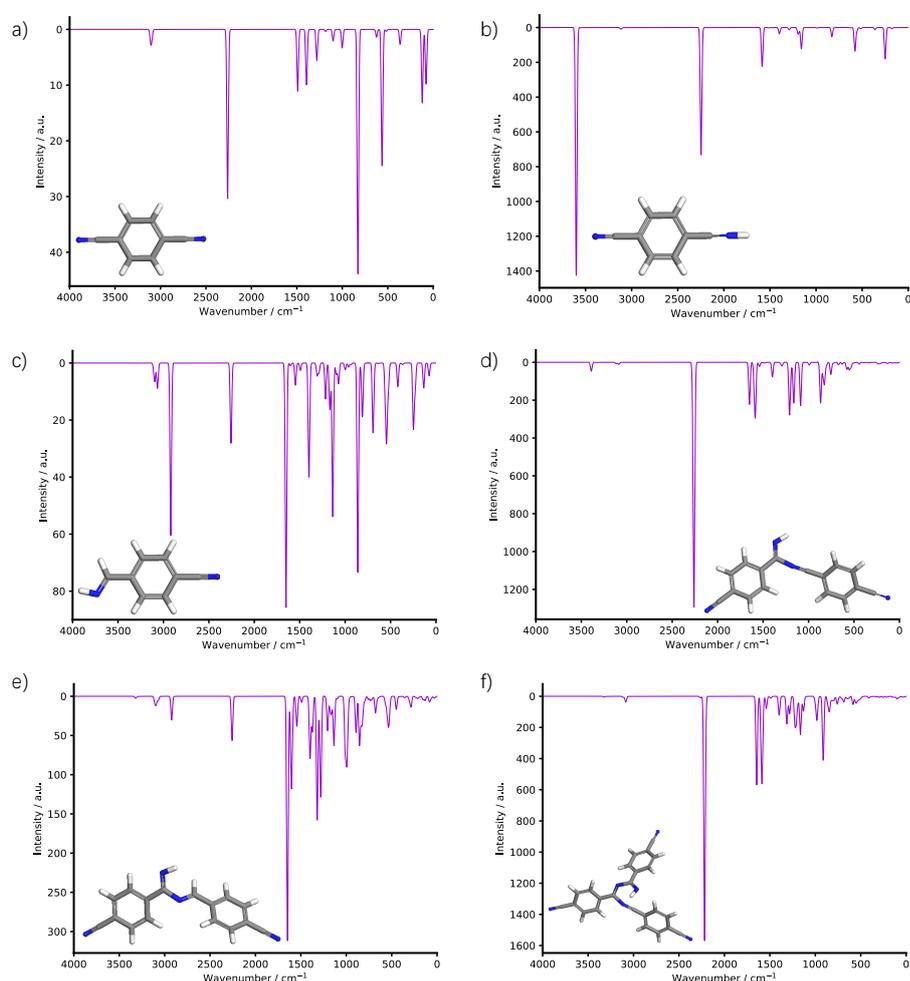


Figure 6.14: IR spectra of the fragments analysed by simulated FT-IR, with the structures of each fragment given as insets. a) 1,4-dicyanobenzene monomer. b) Protonated P1 monomer, +1 charge. c) Protonated P1 monomer with an additional hydrogen atom on one cyano carbon to mimic the simplified test building block, neutral charge. d) Dimer of two reacted P1 monomers to mimic protonated intermediate **2** in the mechanism, +1 charge. e) Dimer of two reacted P1 monomers with an additional hydrogen atom on one cyano carbon, neutral charge. f) Trimer of three reacted P1 monomers to mimic protonated intermediate **3** in the mechanism, +1 charge. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

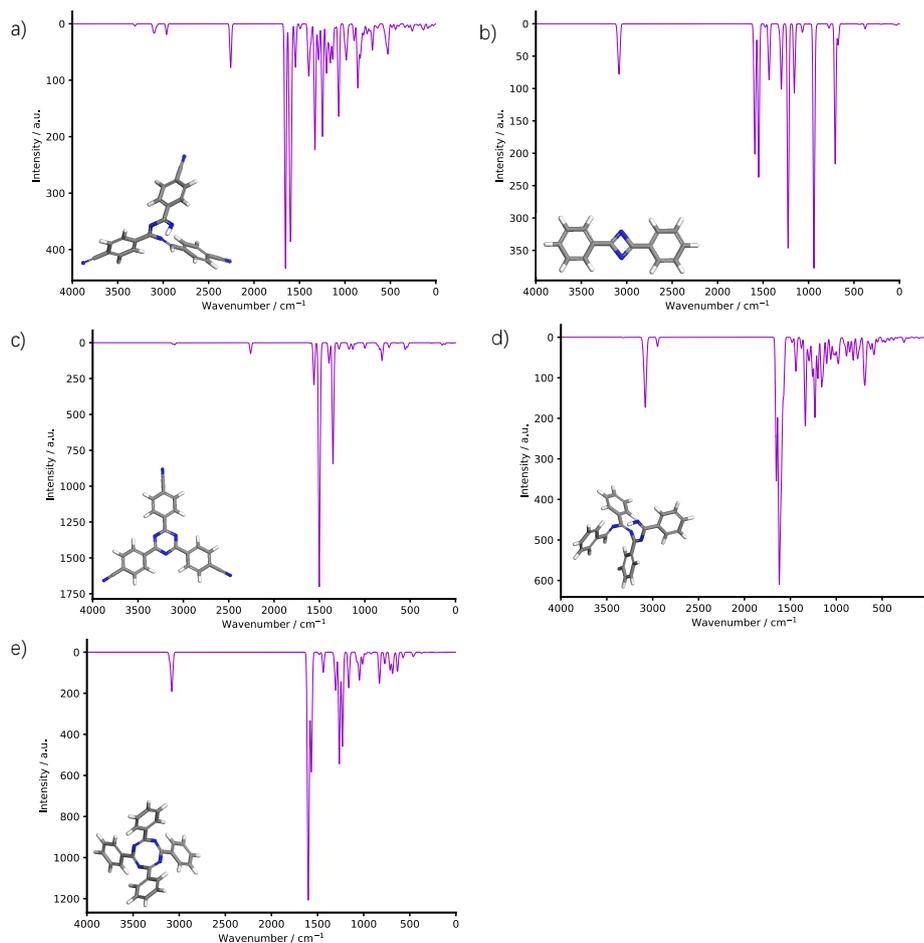


Figure 6.15: IR spectra of the fragments analysed by simulated FT-IR, with the structures of each fragment given as insets. a) Trimer of three reacted P1 monomers with an additional hydrogen atom on one cyano carbon, neutral charge. b) Dimer of two P1 monomers that have reacted to form a 4-ring. c) Expected triazine ring product. d) Four P1 monomers that have reacted to form a pre-8-ring. e) Four P1 monomers that have reacted to form an 8-ring. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

Figure 6.16 shows an overlay of the FT-IR spectra of all of the neutral fragments given in Figures 6.14 and 6.15, along with the FT-IR spectrum achieved by summing the intensities of each fragment at each wavenumber into one spectrum, which is appropriate here considering that the simulated P1 structure contains each of these features.⁴⁵ The peak at 3600 cm^{-1} was partly accounted for by the presence of the protonated and positively charged monomer that is not included in the spectra displayed in Figure 6.16, however, the charge may be spread across the whole system, rather than being localised within a single monomer unit. This could occur *via* the presence of water in the system, which is possible as microporous polymers like CTFs can readily absorb water into the pore structure of the material.^{45,428}

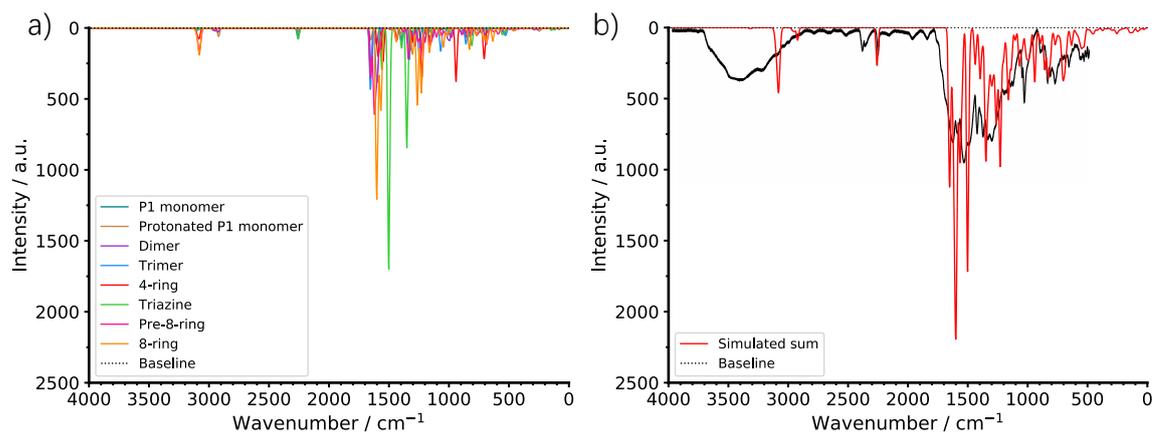


Figure 6.16: Simulated FT-IR spectra of the neutral P1 fragments analysed. a) Individual contributions from each neutral fragment. Key: 1,4-dicyanobenzene monomer – teal, protonated P1 monomer with an additional hydrogen atom on one cyano carbon to mimic the simplified test building block – brown, dimer of two reacted P1 monomers with an additional hydrogen atom on one cyano carbon – purple, trimer of three reacted P1 monomers with an additional hydrogen atom on one cyano carbon – blue, dimer of two P1 monomers that have reacted to form a 4-ring – red, expected triazine ring product – green, four P1 monomers that have reacted to form a pre-8-ring – pink, four P1 monomers that have reacted to form an 8-ring – orange. b) Sum of the contributions from the fragments listed in a) – red, experimental P1 FT-IR spectrum⁴¹³ – black. The black dotted baseline at the top of each plot is included for added clarity. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies. Experimental FT-IR spectrum reproduced (adapted) from reference 413 with permission from John Wiley and Sons. Copyright 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The calculated FT-IR spectrum of a single water molecule in the presence of a standard polarisable continuum solvation model in Gaussian 09²¹² (mimicking the dielectric effects of bulk water) showed peaks in the regions of $\sim 3600\text{ cm}^{-1}$ and $\sim 1700\text{ cm}^{-1}$ (a broad experimental peak appears at 3400 cm^{-1} for a single water molecule),⁴²⁹ which assigned one of the remaining unidentified peaks in the P1 spectrum.⁴⁵ Due to the acidic conditions and presence of atmospheric water, it was also possible that hydronium (H_3O^+) ions may be present within the hydrogen-bonded water network. The FT-IR spectrum of a single hydronium ion was therefore calculated following the same protocol as the water molecule, and this gave peaks in the regions of $\sim 3500\text{ cm}^{-1}$, $\sim 1700\text{ cm}^{-1}$ and $\sim 1000\text{ cm}^{-1}$. The resulting FT-IR spectra of the water molecule and hydronium ion were broadened by a larger FWHM value of 50 cm^{-1} to better mimic the experimental peak broadening and account for the additional environments of these due to diffusion through the extended water hydrogen-bonding network.⁴⁵

Including the presence of water and hydronium ions gave a more consistent match to experiment⁴¹³ and corresponded to a better match than by including the charged fragments given in Figure 6.14b, d and f, but there was still

one remaining peak to be identified in the region of 2350 cm^{-1} .⁴⁵ This region of the FT-IR spectrum is associated with matrix-isolated carbon dioxide molecules,⁴³⁰ suggesting that the peak may arise due to carbon dioxide absorbed within the P1 structure. To test this, the FT-IR spectrum was calculated for a single carbon dioxide molecule and it could be seen that this was able to rationalise the remaining peak at 2350 cm^{-1} .⁴⁵ CTFs have been well-reported for CO_2 -related applications, including for carbon dioxide fixation, photoreduction and capture.^{27-29,32,431} This is due to the CO_2 molecules having a very high affinity for the nitrogen-rich triazine rings, giving CTFs a high carbon dioxide uptake.^{45,428}

This means that it was a reasonable assumption that the experimental P1 structure contained absorbed carbon dioxide, water, and hydronium ions. To support this, it is worth noting that unexplained peaks within these regions have been previously reported in other CTF materials.^{27-29,32,45,431} Figure 6.17 presents the simulated P1 FT-IR spectrum given in Figure 6.16, along with contributions from absorbed carbon dioxide, water, and hydronium ions, with unscaled intensities for each contribution and the intensities of each P1 fragment summed at each wavenumber.

Figure 6.17 gives a much better match to experiment on adding the contributions from the absorbed guest molecules.⁴⁵ However, due to the extended hydrogen-bonding network within the absorbed water molecules and the ability of the water to interact with the P1 framework, it was a sensible approximation to scale the intensities of the peaks within the calculated FT-IR spectrum for the single water molecule. This gives the final simulated FT-IR spectrum for P1, with the assignments of each peak above 1700 cm^{-1} presented in Table 6.3 and the spectrum given in Figure 6.18.⁴⁵ This spectrum includes contributions from the summed neutral fragments considered, absorbed carbon dioxide, water, and hydronium ions. The intensities of the water molecule were scaled by a factor of 10 to better take into account the extended interactions with the framework. As can be seen from Table 6.3, all of the neutral P1 fragments and guests considered contributing to the set of peaks below 1700 cm^{-1} . It is clear that this spectrum gives a better, and overall, very good match to experiment, considering that due to the amorphous nature of this material, an exact match will never be possible without a simulation containing the exact same chemical composition as the experimental.⁴⁵ This indicates that it was a reasonable approach to scale the intensities of the water molecules in the simulation protocol.⁴⁵

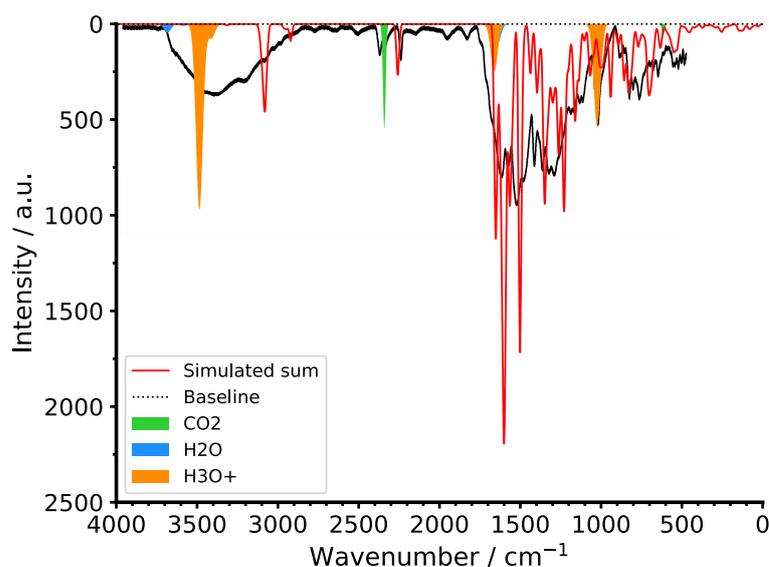


Figure 6.17: Simulated FT-IR spectrum of the neutral P1 fragments analysed and plotted in Figure 6.16, with the intensities of each fragment summed at each wavenumber. Peaks from absorbed carbon dioxide, water, and hydronium ions are also included in the spectrum. Intensity values for all of the simulated data are unscaled. Key: Sum of the contributions from the neutral fragments – red, simulated FT-IR of a single CO₂ molecule – green, simulated FT-IR spectra of a single water molecule – blue, simulated FT-IR of a single hydronium (H₃O⁺) ion – orange, experimental P1 FT-IR spectrum⁴¹³ – black. The black dotted baseline at the top of the plot is included for added clarity. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies. Experimental FT-IR spectrum reproduced (adapted) from reference 413 with permission from John Wiley and Sons. Copyright 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Table 6.3: The assignments of each peak above 1700 cm⁻¹ in Figure 6.18. Table reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies.

Wavenumber / cm ⁻¹	Assignment
500–1700	All neutral models and guests
~1300, ~1500	Triazine ring
~2200	Unreacted cyano groups
~2350	Absorbed guest: CO ₂ molecule
~3200	4-ring, pre-8-ring, 8-ring
~3400	Absorbed guest: H ₃ O ⁺ ion
~3600	Absorbed guest: H ₂ O molecule

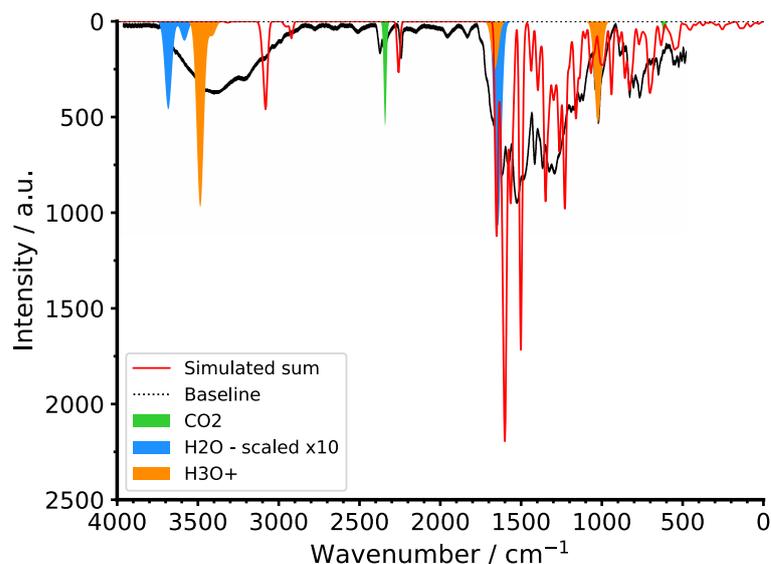


Figure 6.18: Final simulated FT-IR spectrum of the neutral P1 fragments analysed and plotted in Figure 6.16, with the intensities of each fragment summed at each wavenumber. Peaks from absorbed carbon dioxide, water, and hydronium ions are also included in the spectrum. Intensity values for the water molecule are scaled by a factor of 10, all other intensities are unscaled. Key: Sum of the contributions from the neutral fragments – red, simulated FT-IR of a single CO₂ molecule – green, simulated FT-IR spectra of a single water molecule – blue, simulated FT-IR of a single hydronium (H₃O⁺) ion – orange, experimental P1 FT-IR spectrum⁴¹³ – black. The black dotted baseline at the top of the plot is included for added clarity. Figure reproduced (adapted) from reference 45, upon which this chapter is based, with permission from the PCCP Owner Societies. Experimental FT-IR spectrum reproduced (adapted) from reference 413 with permission from John Wiley and Sons. Copyright 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

6.5 Conclusions

Covalent triazine frameworks are an intriguing class of microporous framework materials with good chemical and thermal stability due to their π -conjugated polymer skeleton, giving them applications in a wide variety of fields.^{34,413–416}

An interesting feature of these materials is their ability to display varying properties depending on the synthetic route employed. For example, a highly porous, crystalline material with the same layered, sheet-like topology as COF-1 can be made using a high-temperature conventional ionothermal synthesis with a 1,4-dicyanobenzene monomer.^{14,417,418} Using a trifluoromethanesulfonic acid-catalysed synthesis under microwave-assisted or room temperature conditions with the same starting material gives rise to a framework with preferential ordering from some crystalline domains within the overall amorphous framework, or a totally amorphous material, respectively, with both of these latter materials exhibiting little to no porosity.⁴¹³

Previous modelling of CTF materials such as P1, the fully amorphous analogue of the CTF prepared from a 1,4-dicyanobenzene monomer, assumed that all monomers would fully react to form part of the expected triazine product, with no intermediates and no templating effects, using building blocks with this triazine feature 'built in'.⁴³ This was a reasonable first assumption to make, and is likely to be true in the case of the crystalline CTF-1 material, formed under thermodynamic control with a high temperature and long reaction time.^{14,417,418} However, in the case of the amorphous P1 structure,⁴¹³ formed under kinetic control, the FT-IR spectrum of the experimental material could not be fully rationalised by taking this approach.⁴⁵

Therefore, the work reported in this chapter involved the artificial synthesis of the P1 structure using the Ambuild code,¹⁹ where the full TFMS acid-catalysed cyclotrimerisation reaction was modelled, and solvent and triflate counter-ions also added into the reaction mixture to better mimic the real-world synthesis.⁴⁵ This structure exhibited the expected amorphous structure, as shown by a simulated PXRD pattern, and a zero solvent-accessible surface area, as observed experimentally.^{45,413}

However, the simulated structure not only contained the expected triazine rings, but also featured a large range of structural diversity, including larger and smaller ring features resulting from network errors (8- and 4-rings), and ring intermediates (pre-rings) that could either cyclise to form a ring, or react with an additional monomer to form a larger pre-ring. It was noted that due to the kinetic control employed, any bond formation will be irreversible.⁴⁵

It was found that, in order to fully rationalise the experimental FT-IR spectrum of P1,⁴¹³ the simulated framework must consider this structural diversity, including the presence of alternative ring features, pre-rings and neutral intermediate structures, as well as the expected triazine rings and adsorbed guests within the micropore structure to include water, carbon dioxide, and hydronium ions.⁴⁵ This is proposed as a simple approach to rationalise the properties of novel CTF materials synthesised under amorphous conditions, and demonstrates that the Ambuild code can be reliably used to model a representative CTF structure that gives rise to a comparable FT-IR spectrum to the experimental.^{45,413}

Chapter 7

Conclusions

The whole of science is nothing more than a refinement of everyday thinking.

Albert Einstein

Amorphous microporous organic polymers (MOPs) are an intriguing class of materials with applications in a range of areas.^{5,223,270} They often exhibit a superior chemical and thermal stability in comparison to their crystalline counterparts due to hypercrosslinking within the polymer structure.¹⁹ Yet despite the established applications of these materials, the current limiting factor to materials design and discovery of MOP frameworks is their amorphous nature. This makes simulation of the structures at the atomic scale crucial to understanding the molecular properties, paving the way to design strategies for novel materials.^{19,38}

This thesis displays a range of four case study materials, all of which had previously been synthesised in the ‘real-world’. Each material was simulated, following the full synthetic mechanism used during the experimental synthesis, termed the ‘artificial synthesis’ protocol, before the resulting properties were compared to the experimental.

Chapter 3 describes the artificial synthesis of the first reported conjugated microporous polymer (CMP), CMP-1.^{13,38} Since 2010, the differing porosity properties of CMP materials synthesised using different reaction solvents have been known, and an empirical trend of ‘good’ and ‘bad’ solvents to yield CMPs with high surface areas was formulated.²³⁵ The artificial synthesis described here, which modelled the CMP-1 structure using the same four solvents used in the original 2010 report,^{38,235} demonstrated that the solvent can influence the ratios of micro- and mesoporosity comprising the total framework porosity. This is due to the miscibility of the molecular building blocks

within solvents of different polarities, which affects the formation of the growing polymer network and resulting pore structure.³⁸

Chapter 4 describes a hypercrosslinked polymer (HCP) synthesised from waste polystyrene and loaded with azobenzene *via* physical steeping.^{24,26} This material exhibited differing pore size distributions and gas uptakes depending on the presence and isomer of azobenzene, with a far enhanced carbon dioxide uptake in the presence of *cis*-azobenzene.²⁶ This was rationalised by the increased quantity of micropores in the 11–13 Å region of the pore size distribution and the more favourable interaction between the dipole of *cis*-azobenzene and quadrupole of carbon dioxide, rather than the ability of the carbon dioxide to diffuse throughout the HCP pore structure.²⁶

Chapter 5 rationalises the structure of organically synthesised porous carbon, OSPC-1.^{16,95} The structure of this material was determined by x-ray photoelectron spectroscopy as an sp^3 hybridised carbon node connected by sp hybridised carbon linkers, yet displayed evidence of an alternative structure on first inspection of the solid-state nuclear magnetic resonance (NMR) spectrum.⁹⁵ NMR spectra of a series of fragment and periodic OSPC-1 models which replicate the chemistry of the synthesis and potential end groups were generated to compare to the experimental.⁹⁵ From these, it was established that the OSPC-1 framework was composed of a dense, non-porous and highly interpenetrated structure, surrounded by a thin outer shell based on OSPC-1 in the open, highly porous *dia*, *pts-c* and *unc-c* topologies.⁹⁵

Finally, Chapter 6 rationalises the Fourier-transform infrared (FT-IR) spectrum of the first covalent triazine framework, CTF-1, later synthesised using a room temperature methodology to yield a fully amorphous framework named P1.^{14,45,413} In contrast to the material synthesised using a high-temperature synthesis, which gave a crystalline material formed from the idealised triazine ring product,¹⁴ this synthesis allowed the incorporation of alternative ring features and intermediates into the structure.^{45,413} The resulting P1 structure, which incorporates these features, was able to replicate the experimental porosity and a rolling average XRD spectrum, expected from the experimental material.⁴⁵ To fully rationalise the experimental FT-IR spectrum, this structural diversity, along with the presence of absorbed guest molecules, was required alongside the idealised triazine rings.⁴⁵

Overall, this thesis has shown how simulations can be used to rationalise various properties of amorphous microporous polymers. A combination of computational approaches was utilised to best describe the structures of the studied materials. This typically involved using the Ambuild artificial synthesis procedure to generate the structures, followed by analysis of local structural features using quantum methods such as density functional theory. The artificial

synthesis approach has been shown throughout to yield materials with comparable properties to the experimental.^{19,26,38,45,95,104} This demonstrates the importance of considering the full synthetic conditions and catalytic mechanism within the modelling approach to rationalise these amorphous materials. However, it must be noted that in order to generate representative structural models, multiple repeat structures must be generated in as large a simulation cell as possible, each an individual snapshot of the experimental material. Following these guidelines, the methodologies and understanding developed within this thesis can be applied to the development and design of new materials in the future.

7.1 Future work

Following on from the successful use of simulations to rationalise the experimental properties of established materials, future work would involve extending this approach for use within materials design and discovery of new amorphous microporous polymers.

Firstly, the coarse grain approach, which was explored to model CMP frameworks by Mollart and co-workers,¹⁰⁴ should be expanded to allow for the simulation of larger scale phenomena such as phase separation and gelation. The current system setup utilised hybrid coarse-grained building blocks, which retained the reactive sites as fully atomistic and fully coarse-grained toluene solvent, could be further developed to allow for other solvents to be coarse-grained in a similar manner. This would also allow for a larger number of repeat models to be generated in the same timescale, crucial for simulation of amorphous materials, where each simulated model is a unique ‘snapshot’ of the experimental material.

Additionally, the structure generation process can be further automated to allow for multiple models to be obtained within a reasonable timescale with minimal user input. This would permit not only the seasoned computational chemist, but any scientist interested in amorphous MOPs, to explore potential candidates to be synthesised in the real-world. This could take the form of a library of different chemical mechanisms to choose from, feeding in the structures of molecular building blocks as inputs and then allowing the structures to be artificially synthesised in a high-throughput manner.

Appendix A

```
#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory, cell dimensions, and create the cell:
paramsDir = 'Parameters'
boxDim = [200,200,200]
mycell = ab_cell.Cell(boxDim, atomMargin=0.1, bondMargin=0.5, bondAngleMargin=5, paramsDir=paramsDir )

#Add building block fragments to the library
mycell.libraryAddFragment( filename = 'Blocks/DBB.car', fragmentType = 'DBB', markBonded = True )
mycell.libraryAddFragment( filename = 'TEB_Cu.car', fragmentType = 'TEB', markBonded = True )
mycell.libraryAddFragment( filename = 'Blocks/Pd_cat.car', fragmentType = 'Catalyst', markBonded = True, catalyst = True )
mycell.libraryAddFragment( filename = 'Blocks/TEA.car', fragmentType = 'TEA' )
mycell.libraryAddFragment( filename = 'Blocks/Solvent.car', fragmentType = 'Solvent' )

#Specify bonding rules: either monomer can bond to the catalyst, and monomers already bonded to the catalyst can form a new bond
#and break their respective bonds to the catalyst.
mycell.addBondType( 'Catalyst:a-DBB:a' )
mycell.addBondType( 'Catalyst:a-TEB:a' )
mycell.addBondType( 'DBB:a*-TEB:a*' )

#Seed the monomers and catalyst into the cell, optimise the geometry, seed the solvent and TEA, and give an output step.
mycell.seed( 4, fragmentType = 'Catalyst', zone = [x1, x2, y1, y2, z1, z2] )
mycell.seed( 100, fragmentType = 'TEB', zone = [x1, x2, y1, y2, z1, z2] )
mycell.seed( 100, fragmentType = 'DBB', zone = [x1, x2, y1, y2, z1, z2] )
mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
mycell.seed( 1076, fragmentType = 'TEA' )
mycell.seed( 3229, fragmentType = 'Solvent' )
mycell.dump()

for i in range(100):
    #Remove unreacted monomers, run NVT MD.
    for j in range(10):
        mycell.deleteBlocks( fragmentTypes = 'DBB' )
        mycell.deleteBlocks( fragmentTypes = 'TEB' )
        mycell.runMD( doDihedral = True, dt = 0.0001, mdCycles = 1000000, rCut = 10, T = 55.0, integrator = 'nvt' )
        #Re-seed the monomers and optimise the geometry.
        for k in range(10):
            mycell.seed( 10, fragmentType = 'TEB', zone = [x1, x2, y1, y2, z1, z2] )
            mycell.seed( 10, fragmentType = 'DBB', zone = [x1, x2, y1, y2, z1, z2] )
            mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
        #zipBlocks test, then geometry optimisation. If any new bonds can form, allow them to do so and re-optimize the geometry.
        if mycell.zipBlocks( bondMargin = 10.0, bondAngleMargin = 90, selfBond = False, clashCheck = False ) > 0:
            mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
            mycell.cat1Paf2( ['DBB', 'TEB'] )
        mycell.dump()
```

Figure A.1: Ambuild input file for the network generation procedure to generate the models of CMP-1 using toluene, THF, 1,4-dioxane and DMF as the reaction solvent choices. The number of solvent building blocks can be varied to reflect the correct molar ratio in each case.

```
#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory, cell dimensions, and create the cell:
paramsDir = 'Parameters'
boxDim = [200,200,200]
mycell = ab_cell.Cell(boxDim, atomMargin=0.1, bondMargin=0.5, bondAngleMargin=5, paramsDir=paramsDir )

#Add building block fragments to the library
mycell.libraryAddFragment( filename = 'Blocks/DBB.car', fragmentType = 'DBB', markBonded = True )
mycell.libraryAddFragment( filename = 'TEB_Cu.car', fragmentType = 'TEB', markBonded = True )
mycell.libraryAddFragment( filename = 'Blocks/Pd_cat.car', fragmentType = 'Catalyst', markBonded = True, catalyst = True )
mycell.libraryAddFragment( filename = 'Blocks/TEA.car', fragmentType = 'TEA' )
mycell.libraryAddFragment( filename = 'Blocks/Solvent.car', fragmentType = 'Solvent' )

#Seed the monomers and catalyst into the cell, optimise the geometry, seed the solvent and TEA, and give an output step.
mycell.seed( 4, fragmentType = 'Catalyst', zone = [x1, x2, y1, y2, z1, z2] )
mycell.seed( 100, fragmentType = 'TEB', zone = [x1, x2, y1, y2, z1, z2] )
mycell.seed( 100, fragmentType = 'DBB', zone = [x1, x2, y1, y2, z1, z2] )
mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
mycell.seed( 1076, fragmentType = 'TEA' )
mycell.seed( 3229, fragmentType = 'Solvent' )
mycell.dump()

for i in range(10000):
    #Run a loop of NPT MD followed by geometry optimisation. Give an output after each three loops through.
    for j in range(3):
        mycell.runMD( doDihedral = True, dt = 0.0001, mdCycles = 1000000, rCut = 10, T = 55.0, P = 0.0625, integrator = 'npt' )
        mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
    mycell.dump()
```

Figure A.2: Ambuild input file for the diffusion of the Phase Separated Systems and Solvent/Monomer Phase Interface System models of CMP-1 using toluene, THF, 1,4-dioxane and DMF as the reaction solvent choices. The number of solvent building blocks can be varied to reflect the correct molar ratio in each case.

Table A.1: Porosity data acquired using Poreblazer 4.0¹⁶² for the fully mixed CMP-1 models generated using 1%, 33%, 67% and 100% of the experimental stoichiometry of toluene and TEA, averaged across repeat models 1–4. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

% Solvent	Desolvation scheme	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
100	1	3.50	6.65	0.368	724	1.079	20.75
	2	2.74	4.90	0.196	0	1.177	15.23
	3	2.45	4.47	0.146	0	1.242	18.93
	4	3.48	6.80	0.429	886	1.091	20.95
	5	4.16	7.64	0.324	517	1.139	24.82
	6	2.94	6.08	0.194	227	1.203	21.41
67	1	6.17	10.21	0.920	2365	0.640	27.05
	2	Empty cell after desolvation.					
	3	15.06	18.29	5.708	2216	0.159	27.03
	4	7.23	10.86	0.920	2406	0.640	27.04
	5	5.38	9.82	0.584	1318	0.845	27.06
	6	2.75	6.07	0.213	184	1.123	22.47
33	1	7.13	10.57	0.918	2224	0.638	32.52
	2	10.01	13.57	2.351	5562	0.344	32.54
	3	9.28	12.86	2.040	4818	0.383	32.55
	4	6.31	9.98	0.925	2385	0.638	32.52
	5	6.36	11.22	0.334	761	1.143	45.05
	6	3.89	8.14	0.256	429	1.184	36.82
1	1	6.11	11.89	0.874	2030	0.640	100.00
	2	6.10	12.47	1.068	2562	0.562	99.60
	3	6.43	12.67	1.044	2477	0.566	99.80
	4	5.34	11.91	0.889	2011	0.631	100.00
	5	2.64	7.24	0.227	48	1.049	100.00
	6	4.07	9.51	0.433	612	0.896	100.00

Table A.2: Porosity data acquired using Poreblazer 4.0¹⁶² for the fully mixed CMP-1 models generated using 1%, 33%, 67% and 100% of the experimental stoichiometry of THF and TEA, averaged across repeat models 1–4. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

% Solvent	Desolvation scheme	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
100	1	2.18	4.65	0.236	0	1.038	14.21
	2			Empty cell after desolvation.			
	3			Empty cell after desolvation.			
	4	2.40	4.30	0.233	0	1.038	14.21
	5	7.30	11.39	1.169	2716	0.640	25.68
	6	2.60	4.25	0.273	0	1.035	14.22
67	1	15.71	21.19	6.452	4398	0.142	30.82
	2			Empty cell after desolvation.			
	3			Empty cell after desolvation.			
	4	16.02	20.41	6.452	4331	0.142	30.82
	5	8.40	12.81	1.413	3552	0.487	30.69
	6	2.55	5.34	0.320	0	0.953	16.34
33	1	6.92	11.60	1.078	2878	0.580	37.24
	2			Empty cell after desolvation.			
	3	14.33	18.42	5.715	2093	0.158	37.21
	4	7.54	10.95	1.069	2721	0.580	37.24
	5	7.43	11.16	1.050	2709	0.591	37.23
	6	2.57	5.16	0.155	27	1.213	29.13
1	1	6.05	12.41	0.999	2286	0.570	93.27
	2	6.12	12.37	1.082	2540	0.539	91.33
	3	6.20	12.90	1.085	2542	0.538	92.09
	4	6.64	12.72	1.000	2308	0.570	93.28
	5	3.33	8.90	0.337	296	0.920	100.00
	6	5.65	12.54	0.600	1218	0.778	98.89

Table A.3: Porosity data acquired using Poreblazer 4.0¹⁶² for the fully mixed CMP-1 models generated using 1%, 33%, 67% and 100% of the experimental stoichiometry of 1,4-dioxane and TEA, averaged across repeat models 1–4. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

% Solvent	Desolvation strategy	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
100	1	1.93	4.27	0.109	0	1.287	17.15
	2			Empty cell after desolvation.			
	3			Empty cell after desolvation.			
	4	2.42	4.72	0.137	74	1.267	17.28
	5	1.97	4.63	0.121	0	1.270	19.01
	6	2.63	4.79	0.174	82	1.199	17.56
67	1	7.40	11.15	1.026	2678	0.594	27.75
	2			Empty cell after desolvation.			
	3			Empty cell after desolvation.			
	4	7.38	10.77	1.026	2630	0.595	27.74
	5	7.34	11.46	0.977	2489	0.616	27.75
	6	2.47	4.76	0.152	0	1.218	21.86
33	1	6.60	10.62	0.876	2179	0.657	34.16
	2	14.17	20.70	5.173	1898	0.173	34.11
	3	14.05	16.04	5.180	2167	0.173	34.11
	4	6.78	10.32	0.880	2403	0.658	34.13
	5	6.46	9.55	0.626	1622	0.821	34.14
	6	3.05	5.45	0.238	137	1.079	28.98
1	1	5.99	12.33	0.958	2155	0.581	89.54
	2	6.17	12.16	1.012	2281	0.554	87.34
	3	6.06	12.06	1.011	2329	0.555	89.18
	4	6.16	12.48	0.953	2091	0.581	89.53
	5	3.47	8.91	0.373	430	0.895	100.00
	6	4.84	11.13	0.475	737	0.847	95.18

Table A.4: Porosity data acquired using Poreblazer 4.0¹⁶² for the fully mixed CMP-1 models generated using 1%, 33%, 67% and 100% of the experimental stoichiometry of DMF and TEA, averaged across repeat models 1–4. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

% Solvent	Desolvation strategy	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
100	1			Empty cell after desolvation.			
	2			Empty cell after desolvation.			
	3			Empty cell after desolvation.			
	4			Empty cell after desolvation.			
	5	13.64	24.24	1.221	2285	0.519	63.74
	6	7.743	15.12	0.728	1646	0.732	50.31
67	1	15.33	20.47	6.311	2169	0.145	29.73
	2			Empty cell after desolvation.			
	3			Empty cell after desolvation.			
	4	16.11	21.48	6.309	2162	0.145	29.73
	5	12.91	20.16	1.264	2576	0.513	62.06
	6	7.38	12.45	1.249	2922	0.649	48.57
33	1	7.54	10.26	0.974	2571	0.625	35.92
	2			Empty cell after desolvation.			
	3	15.29	21.46	5.753	1882	0.158	20.87
	4	6.98	10.57	0.978	2592	0.624	20.84
	5	6.41	10.00	0.843	1967	0.711	20.84
	6	2.59	5.24	0.180	171	1.173	16.95
1	1	6.60	12.98	0.835	1907	0.688	92.57
	2	6.59	12.08	0.874	2006	0.679	88.93
	3	6.23	12.62	0.854	1976	0.684	88.34
	4	7.11	11.51	0.823	1846	0.700	91.64
	5	3.25	8.09	0.281	189	1.000	100.00
	6	5.05	11.33	0.467	826	0.873	98.28

Table A.5: Porosity data acquired using Poreblazer 4.0¹⁶² for the CMP-1 models generated using the fully mixed large cell configuration with 100% of the experimental stoichiometry of solvent and TEA. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

Solvent	Model	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
Toluene	1	18.37	25.46	8.311	11275	0.111	32.55
	2	15.44	23.55	6.067	10461	0.148	29.77
	3	15.38	21.29	5.543	10038	0.160	28.62
	4	19.80	29.41	10.962	12109	0.086	35.36
	Average	17.25	24.93	7.721	10971	0.126	31.57
THF	1	17.78	29.98	8.951	11648	0.104	33.15
	2	13.84	20.73	4.788	9229	0.182	27.52
	3	17.26	26.48	9.146	12060	0.102	33.48
	4	23.05	33.49	12.822	12818	0.074	36.90
	Average	17.98	27.67	8.927	11439	0.116	32.77
Dioxane	1	21.88	29.30	11.832	12883	0.080	35.94
	2	21.78	29.94	11.837	12458	0.080	36.06
	3	20.16	26.74	8.768	11821	0.106	32.91
	4	13.51	21.30	5.040	9382	0.174	28.03
	Average	19.33	26.82	9.369	11636	0.110	33.24
DMF	1	18.29	27.90	9.280	11549	0.100	33.64
	2	19.99	29.75	9.069	11849	0.103	33.28
	3	20.18	29.32	10.553	11970	0.089	34.83
	4	20.65	32.62	11.926	12345	0.079	36.02
	Average	19.78	29.90	10.207	11928	0.093	34.44

Table A.6: Porosity data acquired using Poreblazer 4.0¹⁶² for the CMP-1 models generated using the four clusters configuration with 100% of the experimental stoichiometry of solvent and TEA. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

Solvent	Model	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
Toluene	1	4.67	9.65	0.769	1286	0.616	18.65
	2	3.45	7.37	0.500	306	0.728	17.48
	3	6.84	12.01	1.145	2658	0.514	19.96
	4	18.92	25.69	4.866	7349	0.179	27.86
	Average	8.47	13.68	1.820	2900	0.509	20.99
THF	1	6.74	11.43	1.233	3039	0.493	20.48
	2	4.01	9.67	0.675	999	0.652	18.41
	3	8.26	14.59	1.843	4689	0.382	21.87
	4	14.24	22.87	3.065	6354	0.265	24.50
	Average	8.31	14.64	1.704	3770	0.448	21.31
Dioxane	1	14.41	22.50	2.528	5247	0.305	23.75
	2	16.45	24.90	4.166	7667	0.205	26.73
	3	19.11	33.35	5.774	8755	0.155	30.16
	4	4.42	9.81	0.887	1792	0.582	18.85
	Average	13.60	22.64	3.339	5865	0.312	24.87
DMF	1	9.59	20.98	2.438	5517	0.315	23.35
	2	12.59	23.70	3.539	7280	0.234	25.43
	3	10.60	19.98	2.167	5024	0.343	22.50
	4	4.42	9.29	0.647	1110	0.672	18.60
	Average	9.30	18.49	2.198	4733	0.391	22.47

Table A.7: Porosity data acquired using Poreblazer 4.0¹⁶² for the CMP-1 models generated using the two clusters configuration with 100% of the experimental stoichiometry of solvent and TEA. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

Solvent	Model	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
Toluene	1	8.95	12.91	1.724	4551	0.405	20.96
	2	21.48	35.76	6.829	7499	0.133	30.30
	3	8.56	16.36	1.748	4327	0.401	20.81
	4	6.49	11.46	0.991	2340	0.561	18.07
	Average	11.37	19.12	2.823	4679	0.375	22.54
THF	1	11.30	19.90	2.277	4980	0.331	22.79
	2	18.15	23.16	3.384	5321	0.243	24.73
	3	6.42	11.41	0.767	1488	0.632	18.09
	4	13.58	24.26	3.540	6487	0.235	24.44
	Average	12.36	19.68	2.492	4569	0.360	22.51
Dioxane	1	9.36	16.60	2.186	5261	0.343	21.95
	2	5.36	10.41	0.908	1903	0.581	18.48
	3	17.86	25.62	4.241	6904	0.201	26.58
	4	5.04	8.80	0.533	787	0.726	17.48
	Average	9.41	15.36	1.967	3714	0.463	21.12
DMF	1	24.45	37.88	7.917	7173	0.116	31.57
	2	16.47	23.27	3.758	6415	0.223	25.23
	3	2.85	6.09	0.308	0	0.832	16.59
	4	11.84	17.55	2.365	5731	0.323	22.55
	Average	13.90	21.20	3.587	4829	0.374	23.98

Table A.8: Porosity data acquired using Poreblazer 4.0¹⁶² for the CMP-1 models generated using the one small cluster configuration with 100% of the experimental stoichiometry of solvent and TEA. PLD – pore limiting diameter, MPD – maximum pore diameter, He volume – pore volume based on a helium probe, accessible SA – network-accessible surface area, % size of initial cell – the final simulation cell length as a percentage of the initial. Table replicated from reference 38, upon which Chapter 3 is based.

Solvent	Model	PLD / Å	MPD / Å	He Volume / cm ³ g ⁻¹	Accessible SA / m ² g ⁻¹	Density / g cm ⁻³	% Size of Initial Cell
Toluene	1	6.95	12.07	1.194	2682	0.505	19.08
	2	13.89	20.81	2.123	4358	0.348	21.48
	3	15.32	23.35	4.118	6554	0.206	26.35
	4	10.45	18.57	2.033	4906	0.358	21.77
	Average	11.65	18.70	2.367	4625	0.354	22.17
THF	1	11.22	17.73	2.463	5546	0.314	23.05
	2	21.94	30.22	5.407	6862	0.164	28.84
	3	14.26	21.76	2.964	5983	0.271	24.57
	4	7.45	12.43	1.203	2852	0.501	19.28
	Average	13.72	20.54	3.009	5311	0.313	23.94
Dioxane	1	7.30	14.06	1.512	3902	0.438	20.66
	2	23.67	35.26	5.717	6376	0.155	29.02
	3	20.88	34.50	4.618	5797	0.186	26.30
	4	15.53	22.77	3.973	7029	0.214	25.25
	Average	16.85	26.65	3.955	5776	0.248	25.31
DMF	1	19.24	24.43	3.263	5015	0.248	24.58
	2	26.43	36.17	7.600	7030	0.120	30.93
	3	6.01	11.49	1.004	2334	0.552	19.01
	4	18.95	35.46	4.932	6751	0.178	27.84
	Average	17.66	26.89	4.200	5282	0.275	25.59

```

import sys
import re
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib.ticker import FormatStrFormatter
import numpy as np

#To edit-----
#Need to change this if you want to keep something other than oxygen.
contains_oxygen_regex = re.compile('^(\d)\s+(-?\d+\.\d+)\s+(-?\d+\.\d+)\s+(-?\d+\.\d+)\$')

#Specify number of bins and number of solvent molecules in the unit cell.
number_of_bins = 10
number_of_solvent = 3229

#Collecting fractional coordinates to analyse-----
input_file = sys.argv[1]
input_file_split = input_file.split(sep='.')[0]

'''
Collect all of the lines containing oxygen atoms and add to a list of tuples (called oxygen_string_list), then concatenate these tuples to make
a string for each line, and add each line to the oxygen_data.csv file.
'''
def collect_oxygen_data(input_file):
    oxygen_string_list = []
    with open(input_file, 'r') as save_oxygen_data:
        for line in save_oxygen_data.readlines():
            match = contains_oxygen_regex.match(line)
            if match is not None:
                oxygen_string_list.append(match.groups())

    with open(str(input_file_split)+'_oxygen_data.csv', 'w') as oxygen_data:
        oxygen_data.write('Atom,x,y,z\n')
        for oxygen_string_tuple in oxygen_string_list:
            oxygen_linestring = ','.join(oxygen_string_tuple)
            oxygen_data.write(f'{oxygen_linestring}\n')
    collect_oxygen_data(input_file)

'''
Opens the .xyz file specified in the command to run this script, calls it 'input_file', strips any whitespace from it, then searches line by
line for 'Axes'. If it's there, the line will be split with ':' delimiters and saved as the variable 'parts'. Then, the 'cell_length' is saved
by taking the 1st index (second section) of the 'parts' variable.
'''
def collect_cell_length(input_file):
    global cell_length
    with open(input_file, 'r') as input_file:
        for line in input_file:
            line = line.strip(' \n,')
            if 'Axes' in line:
                parts = line.split(':')
                cell_length = float(parts[1].strip(' '))
    collect_cell_length(input_file)
'''

```

Figure A.3: The Python script utilised to split the CMP-1 Solvent/Monomer Phase Interface Systems into sections to generate concentration profiles of each model. The example is shown for Slice A, and to slice the system in the other two directions (Slices B and C), the ranges of x , y and z can be modified. Otherwise, the script is the same for Slices B and C. This script was utilised to count the number of DMF, THF and toluene molecules within the simulation cell. The number of 1,4-dioxane molecules were counted by considering the profiles generated for each of the two oxygen atoms per molecule separately, and then averaging the result. The number of solvent molecules can be varied to reflect the correct molar ratio in each case.

```

'''
This converts the raw oxygen data to fractional coordinates by dividing the values by the cell length obtained above. This will make it
easier to run this code on multiple cells.
'''
df_raw_oxygen = pd.read_csv(str(input_file_split)+'_oxygen_data.csv')
df_raw_oxygen['x'] = (df_raw_oxygen['x'] / cell_length)
df_raw_oxygen['y'] = (df_raw_oxygen['y'] / cell_length)
df_raw_oxygen['z'] = (df_raw_oxygen['z'] / cell_length)
df_raw_oxygen.to_csv(str(input_file_split)+'_fractional_oxygen_data.csv', index=False)

#Slice A-----
#Specify ranges of x, y and z, and the increment.
x1 = 0.45
x2 = 0.55
z1 = 0.45
z2 = 0.55
y_range = 1
increment = (y_range / number_of_bins) - 0.0000000001
centre_of_bin = 0.5 * (y_range / number_of_bins)

#This works out which values from the fractional_oxygen_data.csv dataset are within the x range, and adds them to a new .csv file.
df_fractional_oxygen = pd.read_csv(str(input_file_split)+'_fractional_oxygen_data.csv')
in_x_range = df_fractional_oxygen[df_fractional_oxygen['x'].between(x1, x2)]
in_x_range.to_csv(str(input_file_split)+'_x_data.csv', index=False)

#This works out which values from the x_data.csv dataset are within the z range, and adds them to a new .csv file.
df_x = pd.read_csv(str(input_file_split)+'_x_data.csv')
in_x_and_z_range = df_x[df_x['z'].between(z1, z2)]
in_x_and_z_range.to_csv(str(input_file_split)+'_x_and_z_data.csv', index=False)

'''
#This creates a loop to take each fragment (size specified by number_of_bins above) and count the number of oxygen atoms in each
(as there's one per solvent molecule). Starting y values given first - will change these throughout the loop.
'''
y1 = 0
y2 = y1 + increment

with open(str(input_file_split)+'_slice_A_to_plot.csv', 'w') as output_file:
    output_file.write('Centre of bin,Raw count,Scaled count\n')

for i in range(number_of_bins):
    df_x_and_z = pd.read_csv(str(input_file_split)+'_x_and_z_data.csv')
    in_fragment_range = df_x_and_z[df_x_and_z['y'].between(y1,y2)]
    in_fragment_range.to_csv(str(input_file_split)+'_slice_A_fragment'+str(i)+'_df.csv', index=False)
    y1 = y2
    y2 = y1 + increment
#Count number of oxygen atoms in each fragment file.
count_number_of_lines = in_fragment_range.shape[0]
'''
Scaled count done by working out the count as a percentage of the total number of solvent molecules, then multiplying by 10 to give a
whole number to plot. The axes are the normalised y values x 10^-1.
'''
scaled_count = 100 * (count_number_of_lines / number_of_solvent) * 10
with open(str(input_file_split)+'_slice_A_to_plot.csv', 'a') as output_file:
    to_write = str((i / number_of_bins) + centre_of_bin) + ',' + str(count_number_of_lines) + ',' + str(scaled_count) + '\n'
    output_file.write(to_write)

```

Figure A.4: Continuation of the Python script utilised to split the CMP-1 Solvent/Monomer Phase Interface Systems into sections to generate concentration profiles of each model. The example is shown for Slice A, and to slice the system in the other two directions (Slices B and C), the ranges of x , y and z can be modified. Otherwise, the script is the same for Slices B and C. This script was utilised to count the number of DMF, THF and toluene molecules within the simulation cell. The number of 1,4-dioxane molecules were counted by considering the profiles generated for each of the two oxygen atoms per molecule separately, and then averaging the result.

Table A.9: Porosity data acquired after the network generation using Materials Studio 5.0¹⁵⁸ for the CMP-1 Solvent/Monomer Phase Interface Systems, averaged across slices A and B. Connolly - Connolly surface area, SSA - smoothed solvent surface area, SASA - smoothed solvent-accessible surface area. Table replicated from reference 38, upon which Chapter 3 is based.

Solvent	Surface Area / m ² g ⁻¹	Fragment ID						
		0	1	2	3	4	5	6
Toluene	Connolly	9104	9058	4732	4139	4120	4521	4577
	SSA	21857	19096	5789	4338	3888	4478	4603
	SASA	21857	19095	5776	4330	3868	4466	5750
THF	Connolly	8668	8300	4494	4514	4059	4128	4503
	SSA	20886	20316	5464	4522	3885	4116	4868
	SASA	20879	20309	5453	4497	3855	4108	4851
Dioxane	Connolly	7711	7659	4487	3794	4372	3932	4206
	SSA	16430	18593	5279	3362	4472	3694	3661
	SASA	16428	18592	5275	3351	4458	3684	3654
DMF	Connolly	8403	9110	4666	4709	4180	4415	4403
	SSA	19313	24633	5270	4503	3589	4846	5192
	SASA	19309	24631	5264	4488	3570	4837	5181

Table A.10: Porosity data acquired after the network generation, desolvation, homocoupling, workup and cell equilibration using Materials Studio 5.0¹⁵⁸ for the CMP-1 Solvent/Monomer Phase Interface Systems, averaged across slices A and B. Connolly - Connolly surface area, SSA - smoothed solvent surface area, SASA - smoothed solvent-accessible surface area. Table replicated from reference 38, upon which Chapter 3 is based.

Solvent	Surface Area / m ² g ⁻¹	Fragment ID						
		0	1	2	3	4	5	6
Toluene	Connolly	3840	3078	3516	5347	2683	3467	3792
	SSA	567	397	92	606	738	248	107
	SASA	560	390	12	338	737	163	102
THF	Connolly	2477	1892	1101	1252	3244	2712	3094
	SSA	183	93	123	57	1070	64	1615
	SASA	177	47	122	52	1067	26	1609
Dioxane	Connolly	2052	2323	2465	4220	2117	3397	3126
	SSA	219	121	53	1413	243	247	39
	SASA	219	111	45	1368	201	192	2
DMF	Connolly	2644	3830	2599	2413	3509	2900	1297
	SSA	264	469	251	96	572	443	241
	SASA	234	446	212	88	534	380	241

Appendix B

```
#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory, cell dimensions, and create the cell:
paramsDir = 'Parameters'
boxDim = [100,100,100]
mycell = ab_cell.Cell(boxDim, atomMargin=0.1, bondMargin=0.5, bondAngleMargin=5, paramsDir=paramsDir )

#Add building block fragments to the library
mycell.libraryAddFragment( filename = 'Blocks/Styrene.car', fragmentType = 'Styrene', markBonded = True )

#Specify bonding rules: the styrene monomers may bond to each other.
mycell.addBondType( 'Styrene:a-Styrene:a' )

#Seed one monomer into the centre then five more into the central region to prevent formation of a linear polymer.
mycell.seed( 1, fragmentType='Styrene', center=True)
mycell.seed(5, fragmentType='Styrene', zone=[40,60,40,60,40,60])

for i in range(200):
    #Run a loop to grow ten monomers onto the existing structure, run NVT MD and geometry optimisation.
    for j in range(10):
        mycell.growBlocks( 10, cellEndGroups = ['Styrene:a'], libraryEndGroups = ['Styrene:a'], maxTries = 500 )
        mycell.runMD( doDihedral = True, dt = 0.0001, mdCycles = 1000000, rCut = 10, T = 55.0, integrator = 'nvt' )
        mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
        mycell.dump()
```

Figure B.1: Ambuild input file for the polystyrene formation process.

```

#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory and load in the polymer structure:
paramsDir = 'Parameters'
mycell = ab_util.cellFromPickle("step_x.pkl.gz", paramsDir = paramsDir )

#Add building block fragments to the library
mycell.libraryAddFragmant( filename='Blocks/FDA.car', fragmentType='FDA', markBonded=True )

#Specify bonding rules: the styrene may bond to the FDA, and the FDA can bond to other FDA building blocks.
mycell.addBondType( 'FDA:a-FDA:b' )
mycell.addBondType( 'Styrene:b-FDA:a' )
mycell.addBondType( 'Styrene:b-FDA:b' )

for i in range(100):
    #Grow FDA onto the unreacted Styrene end groups within the structure.
    mycell.growBlocks( 1, cellEndGroups = ['Styrene:b'], libraryEndGroups = ['FDA:a','FDA:b'], maxTries = 500 )
    #Run a loop of optimisation and NVT MD, followed by zipBlocks tests.
    for j in range(2):
        mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
        mycell.runMD( doDihedral = True, dt = 0.0001, mdCycles = 1000000, rCut = 10, T = 1.0, integrator = 'nvt' )
        mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
        mycell.dump()
        if mycell.zipBlocks(bondMargin = 10.0, bondAngleMargin = 90, selfBond = False, clashCheck = False) > 0:
            mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
            mycell.dump()

```

Figure B.2: Ambuild input file for the crosslinking process with NVT ensemble.

```

#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory and load in the polymer structure:
paramsDir = 'Parameters'
mycell = ab_util.cellFromPickle("step_x.pkl.gz", paramsDir = paramsDir )

for _ in range(200):
    #Grow FDA onto the unreacted Styrene end groups within the structure.
    mycell.growBlocks( 1, cellEndGroups = ['Styrene:b'], libraryEndGroups = ['FDA:a','FDA:b'], maxTries = 500 )
    #Run a loop of optimisation and NVT MD, followed by zipBlocks tests.
    for _ in range(2):
        mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
        mycell.runMD( doDihedral = True, dt = 0.00000005, mdCycles = 50000, rCut = 10, T = 1.0, P = 0.0625, integrator = 'npt' )
        mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
        mycell.dump()
        if mycell.zipBlocks(bondMargin = 4.5, bondAngleMargin = 60, selfBond = False, clashCheck = False) > 0:
            mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
            mycell.dump()

```

Figure B.3: Ambuild input file for the crosslinking process with NPT ensemble.

```
#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory and load in the polymer structure:
paramsDir = 'Parameters'
mycell = ab_util.cellFromPickle("step_x.pkl.gz", paramsDir = paramsDir )

#Add building block fragments to the library
mycell.libraryAddFragment( filename = 'Blocks/azobenzene.car', fragmentType = 'azobenzene' )

#Seed 24 azobenzene building blocks at random, optimise the geometry and run NVT MD.
mycell.seed( 24, fragmentType='azobenzene')
mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
mycell.runMD( doDihedral = True, dt = 0.0001, mdCycles = 10000, rCut = 10, T = 55.0, integrator = 'nvt' )
mycell.optimiseGeometry( quiet = True, dt = 0.0001, optCycles = 1000000, rCut = 10 )
mycell.dump()
```

Figure B.4: Ambuild input file for the azobenzene loading process.

Appendix C

Table C.1: Reported chemical shifts for various sp^2 ,^{236,365,367-371} and sp hybridised carbon materials.^{236,365,369,372} Table adapted from reference 95, upon which Chapter 5 is based.

C Hybridisation	Chemical Environment	Peak (ppm)	Reference
sp	Alkyne (CMP)	82.2–95.9	377
sp	Alkyne (CMP)	90.2–91.4	198
sp	Acetylenic	89.8	380
sp	Acetylenic - Pt	127–130	380
sp	Glaser carbon acetylenic	50	373
sp	Glaser carbon Allenic carbyne	18	373
sp	Glaser phenyl cumulenenic carbyne	187	373
sp	Glaser phenyl carbyne	40 and 85	373
sp^2	Phenyl (CMP)	124–139	377
sp^2	Graphite oxide (GO)	130	378
sp^2	Fullerene shell	111	376
sp^2	Activated carbon fibres	131	376
sp^2	Graphite	119	376
sp^2	Pyrene (CMP)	124–138	379
sp^2	Phenyl (CMP)	124–141	198
sp^2	Phenyl (PAF)	120–150	375
sp^2	Glaser carbon	127	373
sp^2	Glaser phenyl	127	373

Table C.2: Continuation of the reported chemical shifts for various sp^3 ,^{365–370} sp^2 ,^{236,365,367–371} and sp hybridised carbon materials.^{236,365,369,372} Anions, ketyls and radicals are also included.^{373–379} Table adapted from reference 95, upon which Chapter 5 is based.

sp^3	Alkyl (CMP)	114.8–119.2	377
sp^3	Nanodiamond (ND)	34.9	374
sp^3	Aliphatic (ND)	45	374
sp^3	Alcohol (ND)	75	374
sp^3	C-OH (GO)	60	378
sp^3	C-O-C (GO)	70	378
sp^3	Diamond phase	35	376
sp^3	Ultra-nanocrystalline diamond	35	376
sp^3	C-OH	73	376
sp^3	Alkyne	25–55	376
sp^3	Methyl (PAF)	65	375
sp^3	Glaser carbon	25	373
sp^3	Glaser phenyl	28	373
sp anion	Acetylide (Pt)	83–92, and 101	382
sp^2 anion	Ph-C(-)-N (Na ⁺)	161	383
sp^3 anion	Methanide	73.6	384
sp^2 ketyl	Ketyl (metal coordinated ligands)	210.9	385
sp^2 ketyl	Ketyl (metal coordinated ligands)	236 and 242	386
sp^2 ketene	Ketene	160	381
sp^2 radicals	Phenalenyl	115–14 and 140–150	387
sp^3 radicals	3,4-dimethylenethiophene intermediate (terminal carbon)	105	387
sp^3 radicals	3,4-dimethylenethiophene	65	387

Table C.3: Table of the models within Set 2, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FJ	10.9196	3.5000	10.9196
IL	10.9196	3.5000	10.5000
IM	10.9196	3.5000	9.9196
IN	10.9196	3.5000	9.5000
IO	10.9196	3.5000	8.9196
IP	10.9196	3.5000	8.5000
IQ	10.9196	3.5000	7.9196
IR	10.9196	3.5000	7.5000
IS	10.9196	3.5000	6.9196
IT	10.9196	3.5000	6.5000
IU	10.9196	3.5000	6.0000
IV	10.9196	3.5000	5.5000

Table C.4: Table of the models within Set 3, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FM	10.9196	3.4000	10.9196
JE	10.9196	3.4000	10.5000
JF	10.9196	3.4000	9.9196
JG	10.9196	3.4000	9.5000
JH	10.9196	3.4000	8.9196
JI	10.9196	3.4000	8.5000
JJ	10.9196	3.4000	7.9196
JK	10.9196	3.4000	7.5000
JL	10.9196	3.4000	6.9196
JM	10.9196	3.4000	6.5000
JN	10.9196	3.4000	6.0000
JO	10.9196	3.4000	5.5000

Table C.5: Table of the models within Set 4, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FN	10.9196	3.3000	10.9196
JQ	10.9196	3.3000	10.5000
JR	10.9196	3.3000	9.9196
JS	10.9196	3.3000	9.5000
JT	10.9196	3.3000	8.9196
JU	10.9196	3.3000	8.5000
JV	10.9196	3.3000	7.9196
JW	10.9196	3.3000	7.5000
JX	10.9196	3.3000	6.9196
JY	10.9196	3.3000	6.5000
JZ	10.9196	3.3000	6.0000

Table C.6: Table of the models within Set 5, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FO	10.9196	3.2000	10.9196
KC	10.9196	3.2000	10.5000
KD	10.9196	3.2000	9.9196
KE	10.9196	3.2000	9.5000
KF	10.9196	3.2000	8.9196
KG	10.9196	3.2000	8.5000
KH	10.9196	3.2000	7.9196
KI	10.9196	3.2000	7.5000
KJ	10.9196	3.2000	6.9196
KK	10.9196	3.2000	6.5000
KL	10.9196	3.2000	6.0000

Table C.7: Table of the models within Set 6, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FS	10.9196	3.1800	10.9196
KO	10.9196	3.1800	10.5000
KP	10.9196	3.1800	9.9196
KQ	10.9196	3.1800	9.5000
KR	10.9196	3.1800	8.9196
KS	10.9196	3.1800	8.5000
KT	10.9196	3.1800	7.9196
KU	10.9196	3.1800	7.5000
KV	10.9196	3.1800	6.9196
KW	10.9196	3.1800	6.5000
KX	10.9196	3.1800	6.0000

Table C.8: Table of the models within Set 7, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FU	10.9196	3.1600	10.9196
LA	10.9196	3.1600	10.5000
LB	10.9196	3.1600	9.9196
LC	10.9196	3.1600	9.5000
LD	10.9196	3.1600	8.9196
LE	10.9196	3.1600	8.5000
LF	10.9196	3.1600	7.9196
LG	10.9196	3.1600	7.5000
LH	10.9196	3.1600	6.9196
LI	10.9196	3.1600	6.5000
LJ	10.9196	3.1600	6.0000

Table C.9: Table of the models within Set 8, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FW	10.9196	3.1400	10.9196
LM	10.9196	3.1400	10.5000
LN	10.9196	3.1400	9.9196
LO	10.9196	3.1400	9.5000
LP	10.9196	3.1400	8.9196
LQ	10.9196	3.1400	8.5000
LR	10.9196	3.1400	7.9196
LS	10.9196	3.1400	7.5000
LT	10.9196	3.1400	6.9196
LU	10.9196	3.1400	6.5000
LV	10.9196	3.1400	6.0000

Table C.10: Table of the models within Set 9, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å		
	x	y	z
FY	10.9196	3.1200	10.9196
LY	10.9196	3.1200	10.5000
LZ	10.9196	3.1200	9.9196
MA	10.9196	3.1200	9.5000
MB	10.9196	3.1200	8.9196
MC	10.9196	3.1200	8.5000
MD	10.9196	3.1200	7.9196
ME	10.9196	3.1200	7.5000
MF	10.9196	3.1200	6.9196
MG	10.9196	3.1200	6.5000
MH	10.9196	3.1200	6.0000

Table C.11: Table of the models within Set 10, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å			Structure	Cell length / Å		
	x	y	z		x	y	z
FP	10.9196	3.1000	10.9196	RR	10.9196	3.1000	6.8000
MK	10.9196	3.1000	10.5000	RS	10.9196	3.1000	6.7000
ML	10.9196	3.1000	9.9196	RT	10.9196	3.1000	6.6000
MM	10.9196	3.1000	9.5000	MS	10.9196	3.1000	6.5000
MN	10.9196	3.1000	8.9196	RU	10.9196	3.1000	6.4000
MO	10.9196	3.1000	8.5000	RV	10.9196	3.1000	6.3000
MP	10.9196	3.1000	7.9196	RW	10.9196	3.1000	6.2000
MQ	10.9196	3.1000	7.5000	RX	10.9196	3.1000	6.1000
MR	10.9196	3.1000	6.9196	MT	10.9196	3.1000	6.0000
RQ	10.9196	3.1000	6.9000				

Table C.12: Table of the models within Set 11, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å			Structure	Cell length / Å		
	x	y	z		x	y	z
MH	10.9196	3.1200	6.0000	QE	10.9196	2.7000	6.0000
PP	10.9196	3.1160	6.0000	QF	10.9196	2.6000	6.0000
PQ	10.9196	3.1120	6.0000	QG	10.9196	2.5000	6.0000
PR	10.9196	3.1080	6.0000	QO	10.9196	2.4000	6.0000
PS	10.9196	3.1040	6.0000	QP	10.9196	2.3000	6.0000
MT	10.9196	3.1000	6.0000	QQ	10.9196	2.2000	6.0000
QB	10.9196	3.0000	6.0000	QR	10.9196	2.1000	6.0000
QC	10.9196	2.9000	6.0000	QS	10.9196	2.0000	6.0000
QD	10.9196	2.8000	6.0000				

Table C.13: Table of the models within Set 12, along with the cell dimensions of each. Figure adapted from reference 95, upon which this chapter is based.

Structure	Cell length / Å			Structure	Cell length / Å		
	x	y	z		x	y	z
NM	21.8392	21.8392	21.8392	OJ	21.8392	10.4196	21.8392
NN	21.8392	21.4196	21.8392	OR	21.8392	9.8392	21.8392
NO	21.8392	20.8392	21.8392	OS	21.8392	9.4196	21.8392
NP	21.8392	20.4196	21.8392	OT	21.8392	8.8392	21.8392
NQ	21.8392	19.8392	21.8392	OU	21.8392	8.4196	21.8392
NR	21.8392	19.4196	21.8392	OV	21.8392	7.8392	21.8392
NS	21.8392	18.8392	21.8392	OW	21.8392	7.4196	21.8392
NT	21.8392	18.4196	21.8392	OX	21.8392	6.8392	21.8392
NU	21.8392	17.8392	21.8392	OY	21.8392	6.4196	21.8392
NV	21.8392	17.4196	21.8392	OZ	21.8392	5.8392	21.8392
NW	21.8392	16.8392	21.8392	PA	21.8392	5.4196	21.8392
NX	21.8392	16.4196	21.8392	PB	21.8392	4.8392	21.8392
NY	21.8392	15.8392	21.8392	PD	21.8392	4.7343	21.8392
NZ	21.8392	15.4196	21.8392	PF	21.8392	4.7243	21.8392
OA	21.8392	14.8392	21.8392	PG	21.8392	4.7143	21.8392
OB	21.8392	14.4196	21.8392	PH	21.8392	4.7043	21.8392
OC	21.8392	13.8392	21.8392	PI	21.8392	4.6943	21.8392
OD	21.8392	13.4196	21.8392	PJ	21.8392	4.6843	21.8392
OE	21.8392	12.8392	21.8392	PK	21.8392	4.6743	21.8392
OF	21.8392	12.4196	21.8392	PL	21.8392	4.6643	21.8392
OG	21.8392	11.8392	21.8392	PM	21.8392	4.6543	21.8392
OH	21.8392	11.4196	21.8392	PN	21.8392	4.6443	21.8392
OI	21.8392	10.8392	21.8392				

```

#!/usr/bin/env python3
from ambuild import ab_util
from ambuild import ab_cell

#Specify the parameters directory, cell dimensions, and create the cell:
paramsDir = 'Parameters'
boxDim=[70,70,200]
mycell = ab_cell.Cell(boxDim, atomMargin=0.1, bondMargin=0.5, bondAngleMargin=5, paramsDir=paramsDir)

#Add building block fragments to the library.
mycell.libraryAddFragment(filename='Blocks/tetra_A.car', fragmentType='tetra_A')
mycell.libraryAddFragment(filename='Blocks/tetra_B.car', fragmentType='tetra_B')
mycell.libraryAddFragment(filename='Blocks/tetra_C.car', fragmentType='tetra_C')

#Specify bonding rules: each building block type can bond to all other types.
mycell.addBondType('tetra_A:a-tetra_A:a')
mycell.addBondType('tetra_B:a-tetra_B:a')
mycell.addBondType('tetra_C:a-tetra_C:a')
mycell.addBondType('tetra_A:a-tetra_B:a')
mycell.addBondType('tetra_B:a-tetra_C:a')
mycell.addBondType('tetra_A:a-tetra_C:a')

#Seed one block in the first (central) region A. Grow more, optimise the geometry, grow more.
mycell.seed(1, fragmentType='tetra_A', zone=[34,35,34,35,99,101])
mycell.growBlocks(4, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_A:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(12, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_A:a')
#zipBlocks test followed by geometry optimisation, growBlocks and geometry re-optimisation.
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(36, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_A:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation, growBlocks and geometry re-optimisation.
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(48, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_A:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#Run NVT MD, zipBlocks test followed by geometry optimisation. Give an output step after the loop.
for i in range (10):
    mycell.runMD(doDihedral=True, rCut=10, mdCycles=1000, T=255.0, dt=0.0001, integrator = 'nvt')
    mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
    mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.dump()

#Seed in the second (middle) regions B
mycell.growBlocks(100, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_B:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation, growBlocks and geometry re-optimisation.
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(100, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_B:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation, growBlocks and geometry re-optimisation.
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(100, cellEndGroups='tetra_A:a', libraryEndGroups='tetra_B:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation.
mycell.zipBlocks(bondMargin=1.8, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#Run NVT MD, zipBlocks test followed by geometry optimisation. Give an output step after the loop.
for i in range (10):
    mycell.runMD(doDihedral=True, rCut=10, mdCycles=1000, T=255.0, dt=0.0001, integrator = 'nvt')
    mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
    mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.dump()

```

Figure C.1: Ambuild input file for the formation of a sphere of OSPC-1. Building blocks tetra_A, tetra_B and tetra_C are identical and are simply used to arrange the sphere. The cell is cuboidal in order to incorporate solvent at a later point in the reaction.

```
#Seed in the third (outer) regions C
mycell.growBlocks(100, cellEndGroups='tetra_B:a', libraryEndGroups='tetra_C:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation, growBlocks and geometry re-optimisation.
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(100, cellEndGroups='tetra_B:a', libraryEndGroups='tetra_C:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation, growBlocks and geometry re-optimisation.
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.growBlocks(100, cellEndGroups='tetra_B:a', libraryEndGroups='tetra_C:a')
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
#zipBlocks test followed by geometry optimisation
mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)

#Run NVT MD, zipBlocks test followed by geometry optimisation. Give an output step after the loop.
for i in range(100):
    mycell.runMD(doDihedral=True, rCut=10, mdCycles=1000, T=255.0, dt=0.0001, integrator='nvt')
    mycell.zipBlocks(bondMargin=2.0, bondAngleMargin=60, clashCheck=False)
    mycell.optimiseGeometry(quiet=True, rCut=10, dt=0.0001, optCycles=1000000)
mycell.dump()
```

Figure C.2: Continuation of the Amsbuild input file for the formation of a sphere of OSPC-1. Building blocks tetra_A, tetra_B and tetra_C are identical and are simply used to arrange the sphere. The cell is cuboidal in order to incorporate solvent at a later point in the reaction.

Appendix D

```
#!/usr/bin/env python3
from ambuild import ab_cell
from ambuild import ab_util

#Specify the parameters directory, cell dimensions, and create the cell:
paramsDir = 'Parameters'
boxDim=[50,50,50]
mycell = ab_cell.Cell( boxDim, atomMargin=0.1, bondMargin=0.5, bondAngleMargin=5, paramsDir=paramsDir )

#Add building block fragments to the library
mycell.libraryAddFragment( filename='Blocks/1_4-Dicyanobenzene-capped.car', fragmentType='TCNQ', markBonded=True )
mycell.libraryAddFragment( filename='Blocks/triflic_acid.car', fragmentType='Triflate', markBonded=True )
mycell.libraryAddFragment( filename='Blocks/chloroform.car', fragmentType='Chloroform', markBonded=True )

#Specify bonding rules: the nitrogen atoms in the di-protonated imine can bond to the carbon atoms.
mycell.addBondType( 'TCNQ:a-TCNQ:b' )

#Seed one monomer, catalyst counter-ion, and solvent into the cell.
mycell.seed( 1, fragmentType='TCNQ' )
mycell.seed( 1, fragmentType='Triflate' )
mycell.seed( 1, fragmentType='Chloroform' )

#Remove solvent and triflate, grow a second monomer onto the first, optimise the geometry.
for i in range(100):
    mycell.deleteBlocks( fragmentTypes='Triflate' )
    mycell.deleteBlocks( fragmentTypes='Chloroform' )
    mycell.growBlocks( 1, cellEndGroups=['TCNQ:a','TCNQ:b'], libraryEndGroups=['TCNQ:a','TCNQ:b'], dihedral=0 )
    mycell.optimiseGeometry( doDihedral=True, rCut=10, dump=False, dt=0.0001, optCycles=1000000 )
    #Re-seed solvent and optimise the geometry.
    for j in range(100):
        mycell.seed( 10, fragmentType='Triflate' )
        mycell.seed( 10, fragmentType='Chloroform' )
    mycell.optimiseGeometry( doDihedral=True, rCut=10, dump=False )
    #zipBlocks test. If any new bonds can form, allow them to do so and re-optimize the geometry. Then give an output step.
    for j in range(1):
        if mycell.zipBlocks( bondMargin=3.2, bondAngleMargin=120, selfBond=True, clashCheck=False ) > 0:
            mycell.cat1Paf2( ['TCNQ','TCNQ'] )
            mycell.optimiseGeometry( doDihedral=True, rCut=10, dump=False )
    mycell.dump()
```

Figure D.1: Ambuild input file for the CTF-1 formation process.

Bibliography

1. J. Rouquerol, D. Avnir, C. W. Fairbridge, D. H. Everett, J. M. Haynes, N. Pernicone, J. D. F. Ramsay, K. S. W. Sing and K. K. Unger, *Pure Appl. Chem.*, 1994, **66**, 1739-1758.
2. K. S. W. Sing, D. H. Everett, R. A. W. Haul, L. Moscou, R. A. Pierotti, J. Rouquerol and T. Siemieniewska, *Pure Appl. Chem.*, 1985, **57**, 603-619.
3. R. Brodschneider and K. Crailsheim, *Apidologie*, 2010, **41**, 278-294.
4. N. G. Patel, M. H. Haydak and T. A. Gochnauer, *Nature*, 1960, **186**, 633-634.
5. S. Das, P. Heasman, T. Ben and S. Qiu, *Chem. Rev.*, 2017, **117**, 1515-1563.
6. L. Wang, Y. Zhong, C. Qian, D. Yang, J. Nie and G. Ma, *Acta Biomater.*, 2020, **114**, 193-205.
7. M. Moshoeshe, M. S. Nadiye-Tabbiruka and V. Obuseng, *Am. J. Mater. Sci.*, 2017, **7**, 196-221.
8. M. Król, *Crystals*, 2020, **10**, 622.
9. T. D. Bennett, Y. Yue, P. Li, A. Qiao, H. Tao, N. G. Greaves, T. Richards, G. I. Lampronti, S. A. T. Redfern, F. Blanc, O. K. Farha, J. T. Hupp, A. K. Cheetham and D. A. Keen, *J. Am. Chem. Soc.*, 2016, **138**, 3484-3492.
10. H. Furukawa, N. Ko, Y. B. Go, N. Aratani, S. B. Choi, E. Choi, A. Ö. Yazaydin, R. Q. Snurr, M. O’Keeffe, J. Kim and O. M. Yaghi, *Science*, 2010, **329**, 424-428.
11. A. P. Côté, A. I. Benin, N. W. Ockwig, M. O’Keeffe, A. J. Matzger and O. M. Yaghi, *Science*, 2005, **310**, 1166-1170.
12. US Pat., 3729457, 1971.
13. J.-X. Jiang, F. Su, A. Trewin, C. D. Wood, N. L. Campbell, H. Niu, C. Dickinson, A. Y. Ganin, M. J. Rosseinsky, Y. Z. Khimyak and A. I. Cooper, *Angew. Chem. - Int. Ed.*, 2007, **46**, 8574-8578.

14. P. Kuhn, M. Antonietti and A. Thomas, *Angew. Chem. - Int. Ed.*, 2008, **47**, 3450–3453.
15. T. Ben, H. Ren, S. Ma, D. Cao, J. Lan, X. Jing, W. Wang, J. Xu, F. Deng, J. M. Simmons, S. Qiu and G. Zhu, *Angew. Chem. - Int. Ed.*, 2009, **48**, 9457–9460.
16. Z. Zhao, S. Das, G. Xing, P. Fayon, P. Heasman, M. Jay, S. Bailey, C. Lambert, H. Yamada, T. Wakihara, A. Trewin, T. Ben, S. Qiu and V. Valtchev, *Angew. Chem. - Int. Ed.*, 2018, **57**, 11952–11956.
17. N. B. McKeown, *Polymer*, 2020, **202**, 122736.
18. D. Taylor, S. J. Dalgarno and F. Vilela, in *Concepts and Design of Materials Nanoarchitectonics*, ed. O. Azzaroni and K. Ariga, The Royal Society of Chemistry, Cambridge, 2022, ch. 11, pp. 226–246.
19. J. M. H. Thomas, C. Mollart, L. Turner, P. Heasman, P. Fayon and A. Trewin, *J. Phys. Chem. B*, 2020, **124**, 7318–7326.
20. R. Dawson and A. Trewin, in *Porous Polymers: Design, Synthesis and Applications*, ed. S. Qiu and T. Ben, The Royal Society of Chemistry, Cambridge, 2015, ch. 7, pp. 155–185.
21. M. O’Keeffe, *Chem. Soc. Rev.*, 2009, **38**, 1215–1217.
22. Y. Kou, Y. Xu, Z. Guo and D. Jiang, *Angew. Chem. - Int. Ed.*, 2011, **50**, 8753–8757.
23. P. Fayon, J. M. H. Thomas and A. Trewin, *J. Phys. Chem. C*, 2016, **120**, 25880–25891.
24. X. Dong, A. Akram, B. Comesaña-Gándara, X. Dong, Q. Ge, K. Wang, S.-P. Sun, B. Jin and C. H. Lau, *ACS Appl. Polym. Mater.*, 2020, **2**, 2586–2593.
25. Y. Zhu and W. Zhang, *Chem. Sci.*, 2014, **5**, 4957–4961.
26. A. Liu, C. Mollart, A. Trewin, X. Fan and C. H. Lau, *ChemSusChem*, 2023, **16**, e202300019.
27. Q.-Q. Dang, C.-Y. Liu, X.-M. Wang and X.-M. Zhang, *ACS Appl. Mater. Interfaces*, 2018, **10**, 27972–27978.
28. J. Li, P. Liu, H. Huang, Y. Li, Y. Tang, D. Mei and C. Zhong, *ACS Sustain. Chem. Eng.*, 2020, **8**, 5175–5183.
29. Y.-M. Li, L. Yang, L. Sun, L. Ma, W.-Q. Deng and Z. Li, *J. Mater. Chem. A*, 2019, **7**, 26071–26076.
30. J.-H. Zhang, S.-M. Xie, L. Chen, B.-J. Wang, P.-G. He and L.-M. Yuan, *Anal. Chem.*, 2015, **87**, 7817–7824.

31. K. Zhang, D. Kopetzki, P. H. Seeberger, M. Antonietti and F. Vilela, *Angew. Chem. - Int. Ed.*, 2013, **52**, 1432-1436.
32. J. Roeser, K. Kailasam and A. Thomas, *ChemSusChem*, 2012, **5**, 1793-1799.
33. J. Zou, Y. Zhao, C. Mollart, M. J. G. Peach, P. Fayon, P. Heasman, P. A. T. J. Fletcher, J. Xu, W. Liang, A. Trewin and T. Ben, *Small*, 2024, **20**, 2308677.
34. A. Rengaraj, P. Puthiaraj, Y. Haldorai, N. S. Heo, S.-K. Hwang, Y.-K. Han, S. Kwon, W.-S. Ahn and Y. S. Huh, *ACS Appl. Mater. Interfaces*, 2016, **8**, 8947-8955.
35. S. Brunauer, P. H. Emmett and E. Teller, *J. Am. Chem. Soc.*, 1938, **60**, 309-319.
36. M. F. De Lange, T. J. H. Vlugt, J. Gascon and F. Kapteijn, *Microporous Mesoporous Mater.*, 2014, **200**, 199-215.
37. T. Düren, F. Millange, G. Férey, K. S. Walton and R. Q. Snurr, *J. Phys. Chem. C*, 2007, **111**, 15350-15356.
38. C. Mollart and A. Trewin, *J. Mater. Chem. A*, 2024, **12**, 4159-4168.
39. L. J. Abbott, K. E. Hart and C. M. Colina, *Theor. Chem. Acc.*, 2013, **132**, 1334.
40. D. W. Lewis, D. J. Willock, C. R. A. Catlow, J. M. Thomas and G. J. Hutchings, *Nature*, 1996, **382**, 604-606.
41. A. Wang, R. Tan, C. Breakwell, X. Wei, Z. Fan, C. Ye, R. Malpass-Evans, T. Liu, M. A. Zwijnenburg, K. E. Jelfs, N. B. McKeown, J. Chen and Q. Song, *J. Am. Chem. Soc.*, 2022, **144**, 17198-17208.
42. L. J. Abbott and C. M. Colina, *Macromolecules*, 2014, **47**, 5409-5415.
43. C. Reece, D. J. Willock and A. Trewin, *Phys. Chem. Chem. Phys.*, 2015, **17**, 817-823.
44. A. Trewin, D. J. Willock and A. I. Cooper, *J. Phys. Chem. C*, 2008, **112**, 20549-20559.
45. C. Mollart, S. Holcroft, M. J. G. Peach, A. Rowling and A. Trewin, *Phys. Chem. Chem. Phys.*, 2022, **24**, 20025-20029.
46. E. Schrödinger, *Ann. Phys.*, 1926, **384**, 361-376.
47. E. Schrödinger, *Ann. Phys.*, 1926, **384**, 489-527.
48. E. Schrödinger, *Naturwissenschaften*, 1926, **14**, 664-666.
49. E. Schrödinger, *Ann. Phys.*, 1926, **384**, 734-756.
50. E. Schrödinger, *Ann. Phys.*, 1926, **385**, 437-490.
51. E. Schrödinger, *Ann. Phys.*, 1926, **386**, 109-139.

52. B. Zwiebach, *Mastering Quantum Mechanics*, The MIT Press, Cambridge, Massachusetts, 2022.
53. J. Harvey, *Computational Chemistry*, Oxford University Press, Oxford, 2018.
54. R. Shankar, *Principles of Quantum Mechanics*, Springer, New York, 1994.
55. M. Born and R. Oppenheimer, *Ann. Phys.*, 1927, **389**, 457–484.
56. D. R. Hartree, *Proc. Camb. Phil. Soc.*, 1928, **24**, 89–110.
57. D. R. Hartree, *Proc. Camb. Phil. Soc.*, 1928, **24**, 111–132.
58. D. R. Hartree, *Proc. Camb. Phil. Soc.*, 1928, **24**, 426–437.
59. V. Fock, *Z. Phys.*, 1930, **61**, 126–148.
60. V. Fock, *Z. Phys.*, 1930, **62**, 795–805.
61. T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory*, John Wiley & Sons, Ltd., Chichester, 2012.
62. J. C. Slater, *Phys. Rev.*, 1929, **34**, 1293–1322.
63. J. C. Slater, *Phys. Rev.*, 1930, **35**, 509–529.
64. W. Pauli, *Z. Phys.*, 1927, **43**, 601–623.
65. *Molecular Mechanics Tools*, <https://education.molssi.org/mm-tools/01-introduction/index.html>, (accessed 04.03.2024).
66. P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
67. W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **137**, A1697–A1705.
68. W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
69. L. J. Sham and W. Kohn, *Phys. Rev.*, 1966, **145**, 561–567.
70. S. Armstrong, PhD thesis, Lancaster University, 2022.
71. P. A. M. Dirac, *Proc. R. Soc. Lond. A*, 1930, **126**, 360–365.
72. L. H. Thomas, *Math. Proc. Cam. Philos. Soc.*, 1927, **23**, 542–548.
73. E. Fermi, *Rend. Lincei*, 1927, **6**, 602–607.
74. J. P. Perdew and K. Schmidt, *AIP Conf. Proc.*, 2001, **577**, 1–20.
75. R. V. Leeuwen and E. J. Baerends, *Int. J. Quantum Chem.*, 1994, **52**, 711–730.
76. M. M. Gugelchuk and Y. Cui, *J. Mol. Struct.: THEOCHEM*, 1996, **365**, 111–117.
77. J. P. Perdew and A. Zunger, *Phys. Rev. B*, 1981, **23**, 5048–5079.
78. J. C. Slater and J. C. Phillips, *Phys. Today*, 1974, **27**, 49–50.
79. S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200–1211.

80. *Density Functional Theory for Experimentalists: A Four-Lecture Primer*, <https://epubs.stfc.ac.uk/manifestation/11251048/RAL-TR-2014-002.pdf>, (accessed 04.03.2024).
81. D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, 1980, **45**, 566–569.
82. J. P. Perdew, *MRS Bull.*, 2013, **38**, 743–750.
83. A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
84. J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822–8824.
85. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
86. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
87. Y. Zhao and D. G. Truhlar, *J. Chem. Phys.*, 2006, **125**, 194101.
88. J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
89. J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.*, 1996, **105**, 9982–9985.
90. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
91. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
92. J. P. Perdew and Y. Wang, *Phys. Rev. B*, 1992, **45**, 13244–13249.
93. P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213–222.
94. M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654–3665.
95. C. Mollart, A. Rowling, E. Sherrett, M. J. G. Peach and A. Trewin, unpublished work.
96. T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
97. M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, *Rev. Mod. Phys.*, 1992, **64**, 1045–1097.
98. *CASTEP background theory*, <https://www.tcm.phy.cam.ac.uk/castep/documentation/WebHelp/content/modules/castep/thcastepbackground.htm>, (accessed 06.03.2024).
99. *Plane wave basis set*, <https://www.tcm.phy.cam.ac.uk/castep/documentation/WebHelp/content/modules/castep/thcastepplanebasis.htm>, (accessed 07.03.2024).
100. *Introduction to first-principles modelling and CASTEP*, https://www-users.york.ac.uk/~pjh503/DFT/intro_DFT_CASTEP.pdf, (accessed 06.03.2024).

101. *Reciprocal space*, http://goodwin.chem.ox.ac.uk/goodwin/TEACHING_files/-11_handout.pdf, (accessed 05.07.2024).
102. *Brillouin Zone Construction*, https://www.doitpoms.ac.uk/tlplib/brillouin_zones/zone_construction.php, (accessed 18.04.2024).
103. *Molecular Mechanics (MM)*, <https://www.compchems.com/molecular-mechanics-mm/#why-molecular-mechanics>, (accessed 12.03.2024).
104. C. Mollart, B. Ciborowski and A. Trewin, *Mol. Syst. Des. Eng.*, 2023, **8**, 1456–1461.
105. *Molecular modelling*, http://chem.winthrop.edu/molecular_modeling/molecular_mechanics.html, (accessed 12.03.2024).
106. *HOOMD-blue md.dihedral*, <https://hoomd-blue.readthedocs.io/en/v4.5.0/module-md-dihedral.html>, (accessed 13.03.2024).
107. *HOOMD-blue md.bond*, <https://hoomd-blue.readthedocs.io/en/v4.5.0/module-md-bond.html>, (accessed 13.03.2024).
108. *HOOMD-blue md.angle*, <https://hoomd-blue.readthedocs.io/en/v4.5.0/module-md-angle.html>, (accessed 13.03.2024).
109. *angle_style quartic command*, https://docs.lammps.org/angle_quartic.html, (accessed 18.03.2024).
110. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard III and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
111. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
112. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.
113. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell Jr., *J. Comput. Chem.*, 2010, **31**, 671–690.
114. H. Sun, *J. Phys. Chem. B*, 1998, **102**, 7338–7364.
115. H. Sun, *Macromolecules*, 1995, **28**, 701–712.
116. *Classification of forcefield used in molecular dynamics*, <https://www.j-octa.com/about-j-octa/force-fields-for-MD-calculation.html>, (accessed 06.07.2024).
117. *CHARMM force fields, parameterization strategies and future/ongoing force field developments*, http://www.mmts.org/workshops/mmts-ctbp_2006/Lecture_pdf/params_alex_jul06a.pdf, (accessed 06.07.2024).
118. S. Zhang, M. Z. Makoś, R. B. Jadrich, E. Kraka, K. Barros, B. T. Nebgen, S. Tretiak, O. Isayev, N. Lubbers, R. A. Messerly and J. S. Smith, *Nat. Chem.*, 2024, **16**, 727–734.

119. T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama and A. C. T. van Duin, *npj Comput. Mater.*, 2016, **2**, 15011.
120. A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, *J. Phys. Chem. A*, 2001, **105**, 9396–9409.
121. *Scientists develop new machine learning method for modeling chemical reactions*, <https://phys.org/news/2024-03-scientists-machine-method-chemical-reactions.html>, (accessed 20.03.2024).
122. J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev and A. E. Roitberg, *J. Chem. Phys.*, 2018, **148**, 241733.
123. C. Devereux, J. S. Smith, K. K. Huddleston, K. Barros, R. Zubatyuk, O. Isayev and A. E. Roitberg, *J. Chem. Theory Comput.*, 2020, **16**, 4192–4202.
124. P. Ren, Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, B. B. Gupta, X. Chen and X. Wang, *ACM Comput. Surv.*, 2021, **54**, 1–40.
125. L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande and T. J. Martínez, *Nat. Chem.*, 2014, **6**, 1044–1048.
126. L.-P. Wang, in *Computational Approaches for Chemistry Under Extreme Conditions*, ed. N. Goldman, Springer Cham, Switzerland, 1st edn., 2019, ch. 6, pp. 127–159.
127. R. DeVane, M. L. Klein, C.-c. Chiu, S. O. Nielsen, W. Shinoda and P. B. Moore, *J. Phys. Chem. B*, 2010, **114**, 6386–6393.
128. S. J. Marrink, A. H. de Vries and A. E. Mark, *J. Phys. Chem. B*, 2004, **108**, 750–760.
129. L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S.-J. Marrink, *J. Chem. Theory Comput.*, 2008, **4**, 819–834.
130. V. V. Pisarev and S. A. Zakharov, *J. Phys. Conf.*, 2018, **946**, 012100.
131. G. D. Smith and W. Paul, *J. Phys. Chem. A*, 1998, **102**, 1200–1208.
132. J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller and A. Tkatchenko, *Chem. Rev.*, 2021, **121**, 9816–9872.
133. S. A. Hollingsworth and R. O. Dror, *Neuron*, 2018, **99**, 1129–1143.
134. R. Iftimie, P. Minary and M. E. Tuckerman, *Proc. Natl. Acad. Sci.*, 2005, **102**, 6654–6659.
135. *Molecular Dynamics: Equations of motion*, <https://www.compchems.com/-molecular-dynamics-equations-of-motion>, (accessed 15.03.2024).

136. *G. Molecular Dynamics*, [https://bio.libretexts.org/Under_Construction/Purgatory/Book%3A_Biochemistry_Online_\(Jakubowski\)/01%3A_LIPID_/1.3-%3A_Dynamics_of_Membrane_Lipids/Molecular_Mechanics_and_Dynamics/G_Molecular_Dynamics](https://bio.libretexts.org/Under_Construction/Purgatory/Book%3A_Biochemistry_Online_(Jakubowski)/01%3A_LIPID_/1.3-%3A_Dynamics_of_Membrane_Lipids/Molecular_Mechanics_and_Dynamics/G_Molecular_Dynamics), (accessed 15.03.2024).
137. H. Grad, *Commun. Pure Appl. Math.*, 1949, **2**, 331-407.
138. L. Verlet, *Phys. Rev.*, 1967, **159**, 98-103.
139. N. S. Martys and R. D. Mountain, *Phys. Rev. E*, 1999, **59**, 3733-3736.
140. R. W. Hockney, *Methods Comput. Phys.*, 1970, **9**, 135-211.
141. *Molecular Dynamics: Ensembles*, <https://www.compchems.com/molecular-dynamics-ensembles>, (accessed 18.03.2024).
142. *Thermostats in Molecular Dynamics*, <https://www.compchems.com/thermostats-in-molecular-dynamics>, (accessed 18.04.2024).
143. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
144. H. C. Andersen, *J. Chem. Phys.*, 1980, **72**, 2384-2393.
145. H. A. Posch, W. G. Hoover and F. J. Vesely, *Phys. Rev. A*, 1986, **33**, 4253-4265.
146. G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
147. *Barostats in Molecular Dynamics*, <https://www.compchems.com/barostats-in-molecular-dynamics>, (accessed 18.04.2024).
148. M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182-7190.
149. M. Parrinello and A. Rahman, *J. Chem. Phys.*, 1982, **76**, 2662-2666.
150. M. Parrinello and A. Rahman, *Phys. Rev. Lett.*, 1980, **45**, 1196-1199.
151. *Mulliken Populations*, [https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Book%3A_Quantum_States_of_Atoms_and_Molecules_\(Zielinski_et_al\)/10%3A_Theories_of_Electronic_Molecular_Structure/10.07%3A_Mulliken_Populations](https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Book%3A_Quantum_States_of_Atoms_and_Molecules_(Zielinski_et_al)/10%3A_Theories_of_Electronic_Molecular_Structure/10.07%3A_Mulliken_Populations), (accessed 17.04.2024).
152. R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833-1840.
153. K. V. Kumar, S. Gadipelli, B. Wood, K. A. Ramisetty, A. A. Stewart, C. A. Howard, D. J. L. Brett and F. Rodriguez-Reinoso, *J. Mater. Chem. A*, 2019, **7**, 10104-10137.
154. *Hysteresis in adsorption and desorption isotherm*, <https://www.microtrac.com/applications/knowledge-base/hysteresis-in-adsorption>, (accessed 02.04.2024).
155. J. Wang and X. Guo, *Chemosphere*, 2020, **258**, 127279.

156. M. G. Martin, *Mol. Simul.*, 2013, **39**, 1212–1222.
157. A. Gupta, S. Chempath, M. J. Sanborn, L. A. Clark and R. Q. Snurr, *Mol. Simul.*, 2003, **29**, 29–46.
158. BIOVIA, Accelrys Materials Studio 5.0, Dassault Systèmes, San Diego, CA, 2009.
159. I. Langmuir, *J. Am. Chem. Soc.*, 1918, **40**, 1361–1403.
160. S. Lowell, J. E. Shields, M. A. Thomas and M. Thommes, in *Characterization of Porous Solids and Powders: Surface Area, Pore Size and Density*, ed. S. Lowell, J. E. Shields, M. A. Thomas and M. Thommes, Springer Netherlands, Dordrecht, 2004, ch. 5, pp. 58–81.
161. H. Swenson and N. P. Stadie, *Langmuir*, 2019, **35**, 5409–5426.
162. L. Sarkisov, R. Bueno-Perez, M. Sutharson and D. Fairen-Jimenez, *Chem. Mater.*, 2020, **32**, 9849–9867.
163. L. Sarkisov and A. Harrison, *Mol. Simul.*, 2011, **37**, 1248–1257.
164. T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza and M. Haranczyk, *Microporous Mesoporous Mater.*, 2012, **149**, 134–141.
165. M. L. Connolly, *Science*, 1983, **221**, 709–713.
166. D. Ongari, P. G. Boyd, S. Barthel, M. Witman, M. Haranczyk and B. Smit, *Langmuir*, 2017, **33**, 14529–14538.
167. *An Introduction to NLDFT Models for Porosity Characterisation*, https://micromeritics.com/Repository/Files/WP_An_Introduction_to_NLDFT.pdf, (accessed 06.07.2024).
168. A. V. Neimark, Y. Lin, P. I. Ravikovitch and M. Thommes, *Carbon*, 2009, **47**, 1617–1628.
169. G. Horváth and K. Kawazoe, *J. Chem. Eng. Jpn.*, 1983, **16**, 470–475.
170. E. P. Barrett, L. G. Joyner and P. P. Halenda, *J. Am. Chem. Soc.*, 1951, **73**, 373–380.
171. *Solubility*, <https://www.britannica.com/science/solubility-chemistry>, (accessed 02.04.2024).
172. L. R. Snyder, *J. Chromatogr. A*, 1974, **92**, 223–230.
173. L. R. Snyder, *J. Chromatogr. Sci.*, 1978, **16**, 223–234.
174. P.-G. Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, New York, 1979.
175. W. Burchard, in *Chemistry and Technology of Water-Soluble Polymers*, ed. C. A. Finch, Springer Science+Business Media, New York, 1st edn., 1983, ch. 7, pp. 125–142.

176. F. Franks, in *Chemistry and Technology of Water-Soluble Polymers*, ed. C. A. Finch, Springer Science+Business Media, New York, 1st edn., 1983, ch. 9, pp. 157–178.
177. *Solubility Parameters: Theory and Applications*, <https://cool.cultural-heritage.org/coolaic/sg/bpg/annual/v03/bp03-04.html>, (accessed 02.04.2024).
178. J. H. Hildebrand and R. S. Scott, *J. Phys. Chem.*, 1951, **55**, 619–620.
179. A. F. M. Barton, *Handbook of Solubility Parameters and Other Cohesion Parameters*, CRC Press, Boca Raton, 2nd edn., 1991.
180. J. Chen, T. Qiu, W. Yan and C. F. J. Faul, *J. Mater. Chem. A*, 2020, **8**, 22657–22665.
181. J. F. Moulder, W. F. Stickle, P. E. Sobol and K. D. Bomben, *Handbook of X-ray Photoelectron Spectroscopy*, Perkin-Elmer Corporation, Minnesota, 1992.
182. P. J. Hore, *Nuclear Magnetic Resonance*, Oxford University Press, Oxford, 1995.
183. S. P. Brown and L. Emsley, in *Handbook of Spectroscopy*, ed. G. Gauglitz and T. Vo-Dinh, Wiley, Weinheim, 2003, vol. 1, ch. 9, pp. 269–326.
184. *Spectroscopy in a Suitcase*, <https://edu.rsc.org/resources/spectroscopy-resource-packs/280.article>, (accessed 28.03.2024).
185. B. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*, Wiley, Chichester, 2008.
186. *Single Crystal and Powder X-ray Diffraction*, <https://www.jove.com/v/10-462/single-crystal-and-powder-x-ray-diffraction>, (accessed 28.03.2024).
187. *Bragg's Law*, [https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_\(Analytical_Chemistry\)/Instrumentation_and_Analysis/Diffraction_Scattering_Techniques/Bragg's_Law](https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Instrumentation_and_Analysis/Diffraction_Scattering_Techniques/Bragg's_Law), (accessed 28.03.2024).
188. M. C. Rowe and B. J. Brewer, *Comput. Geosci.*, 2018, **120**, 21–31.
189. C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler and P. A. Wood, *J. Appl. Cryst.*, 2020, **53**, 226–235.
190. K. Momma and F. Izumi, *J. Appl. Cryst.*, 2008, **41**, 653–658.
191. *Transmission Electron Microscopy vs Scanning Electron Microscopy*, <https://www.thermofisher.com/uk/en/home/materials-science/learning-center/applications/sem-tem-difference>, (accessed 17.04.2024).
192. *Ambuild wiki*, <https://github.com/linucks/ambuild/wiki>, (accessed 04.04.2024).

193. J. A. Anderson, C. D. Lorenz and A. Travesset, *J. Comput. Phys.*, 2008, **227**, 5342–5359.
194. J. Glaser, T. D. Nguyen, J. A. Anderson, P. Lui, F. Spiga, J. A. Millan, D. C. Morse and S. C. Glotzer, *Comput. Phys. Commun.*, 2015, **192**, 97–107.
195. X. Zhao, J. Gong, P. Alam, C. Ma, Y. Wang, J. Guo, Z. Zeng, Z. He, H. H. Y. Sung, I. D. Williams, K. S. Wong, S. Chen, J. W. Y. Lam, Z. Zhao and B. Z. Tang, *CCS Chem.*, 2021, **4**, 1912–1920.
196. J. M. H. Thomas and A. Trewin, *J. Phys. Chem. C*, 2014, **118**, 19712–19722.
197. J. Lenhard, S. Stephan and H. Hasse, *Ann. Phys.*, 2024, **536**, 2400115.
198. J. Gasteiger and M. Marsili, *Tetrahedron*, 1980, **36**, 3219–3228.
199. *molcharge - Partial charges*, <https://docs.eyesopen.com/toolkits/python/quacpactk/molchargetheory.html>, (accessed 19.03.2024).
200. *Openeye code*, <https://github.com/choderalab/openmoltools/blob/master/openmoltools/openeye.py>, (accessed 19.03.2024).
201. D. Ongari, P. G. Boyd, O. Kadioglu, A. K. Mace, S. Keskin and B. Smit, *J. Chem. Theory Comput.*, 2019, **15**, 382–401.
202. A. K. Rappe and W. A. Goddard III, *J. Phys. Chem.*, 1991, **95**, 3358–3363.
203. E. Bitzek, P. Koskinen, F. Gähler, M. Moseler and P. Gumbsch, *Phys. Rev. Lett.*, 2006, **97**, 170201.
204. *md.minimize*, <https://hoomd-blue.readthedocs.io/en/v3.1.0/module-md-minimize.html>, (accessed 07.07.2024).
205. *hoomd.md*, <https://hoomd-blue.readthedocs.io/en/v3.1.0/package-md.html#module-hoomd.md>, (accessed 07.07.2024).
206. G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, 1994, **101**, 4177–4189.
207. G. Torrie and J. Valleau, *J. Comput. Phys.*, 1977, **23**, 187–199.
208. I. T. Todorov, W. Smith, K. Trachenko and M. T. Dove, *J. Mater. Chem.*, 2006, **16**, 1911–1918.
209. *07_Geometry_Optimisation_and_MD*, https://github.com/linucks/ambuild/wiki/07_Geometry_Optimisation_and_MD, (accessed 18.03.2024).
210. J. Hoshen and R. Kopelman, *Phys. Rev. B*, 1976, **14**, 3438–3445.
211. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

212. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision E.01*, Gaussian Inc., Wallingford CT, 2015.
213. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson and M. C. Payne, *Z. Kristallogr. - Cryst. Mater.*, 2005, **220**, 567–570.
214. A. P. Bartók and J. R. Yates, *J. Chem. Phys.*, 2019, **150**, 161101.
215. C. Bonhomme, C. Gervais, F. Babonneau, C. Coelho, F. Pourpoint, T. Azaïs, S. E. Ashbrook, J. M. Griffin, J. R. Yates, F. Mauri and C. J. Pickard, *Chem. Rev.*, 2012, **112**, 5733–5779.
216. C. J. Pickard and F. Mauri, *Phys. Rev. B*, 2001, **63**, 245101.
217. C. G. Van de Walle and P. E. Blöchl, *Phys. Rev. B*, 1993, **47**, 4244–4255.
218. J. R. Yates, C. J. Pickard and F. Mauri, *Phys. Rev. B*, 2007, **76**, 024401.
219. G. P. Francis and M. C. Payne, *J. Phys.: Condens. Matter*, 1990, **2**, 4395–4404.
220. H. J. Monkhorst and J. D. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.
221. B. G. Pfrommer, M. Côté, S. G. Louie and M. L. Cohen, *J. Comput. Phys.*, 1997, **131**, 233–240.
222. M. Profeta, F. Mauri and C. J. Pickard, *J. Am. Chem. Soc.*, 2003, **125**, 541–548.
223. J.-S. M. Lee and A. I. Cooper, *Chem. Rev.*, 2020, **120**, 2171–2214.
224. Y. Xu, S. Jin, H. Xu, A. Nagai and D. Jiang, *Chem. Soc. Rev.*, 2013, **42**, 8012–8031.
225. W. Wei, G. Chang, Y. Xu and L. Yang, *J. Mater. Chem. A*, 2018, **6**, 18794–18798.

226. J.-X. Jiang, A. Trewin, D. J. Adams and A. I. Cooper, *Chem. Sci.*, 2011, **2**, 1777–1781.
227. B. Liang, H. Wang, X. Shi, B. Shen, X. He, Z. A. Ghazi, N. A. Khan, H. Sin, A. M. Khattak, L. Li and Z. Tang, *Nat. Chem.*, 2018, **10**, 961–967.
228. T. Chen, W. Zhang, B. Li, W. Huang, C. Lin, Y. Wu, S. Chen and H. Ma, *ACS Appl. Mater. Interfaces*, 2020, **12**, 56385–56392.
229. Y. Xie, T.-T. Wang, X.-H. Liu, K. Zou and W.-Q. Deng, *Nat. Commun.*, 2013, **4**, 1960.
230. L. Chen, Y. Yang and D. Jiang, *J. Am. Chem. Soc.*, 2010, **132**, 9138–9143.
231. X. Liu, Y. Xu and D. Jiang, *J. Am. Chem. Soc.*, 2012, **134**, 8738–8741.
232. J. Tan, J. Wan, J. Guo and C. Wang, *Chem. Commun.*, 2015, **51**, 17394–17397.
233. *Conjugated microporous polyaryleneethynylene copolymer*, <https://www-sigmaaldrich.com/GB/en/product/aldrich/799491>, accessed (18.12.2023).
234. R. Dawson, A. Laybourn, R. Clowes, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Macromolecules*, 2009, **42**, 8809–8816.
235. R. Dawson, A. Laybourn, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Macromolecules*, 2010, **43**, 8524–8530.
236. J.-X. Jiang, F. Su, A. Trewin, C. D. Wood, H. Niu, J. T. A. Jones, Y. Z. Khimyak and A. I. Cooper, *J. Am. Chem. Soc.*, 2008, **130**, 7710–7720.
237. A. Laybourn, R. Dawson, R. Clowes, T. Hasell, A. I. Cooper, Y. Z. Khimyak and D. J. Adams, *Polym. Chem.*, 2014, **5**, 6325–6333.
238. A. I. Cooper, *Adv. Mater.*, 2009, **21**, 1291–1295.
239. S. Yuan, B. Dorney, D. White, S. Kirklin, P. Zapol, L. Yu and D.-J. Liu, *Chem. Commun.*, 2010, **46**, 4547–4549.
240. J. Schmidt, M. Werner and A. Thomas, *Macromolecules*, 2009, **42**, 4426–4429.
241. J. Chen, W. Yan, E. J. Townsend, J. Feng, L. Pan, V. Del Angel Hernandez and C. F. J. Faul, *Angew. Chem. - Int. Ed.*, 2019, **58**, 11715–11719.
242. R. S. Sprick, J.-X. Jiang, B. Bonillo, S. Ren, T. Ratvijitvech, P. Guiglion, M. A. Zwijnenburg, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2015, **137**, 3265–3270.
243. J. Weber and A. Thomas, *J. Am. Chem. Soc.*, 2008, **130**, 6334–6335.
244. B. Kiskan and J. Weber, *ACS Macro Lett.*, 2012, **1**, 37–40.
245. Y. Lei, Z. Tian, H. Sun, F. Liu, Z. Zhu, W. Liang and A. Li, *ACS Appl. Mater. Interfaces*, 2021, **13**, 5823–5833.

246. H. Ma, F. Li, P. Li, H. Wang, M. Zhang, G. Zhang, M. Baumgarten and K. Müllen, *Adv. Funct. Mater.*, 2016, **26**, 2025–2031.
247. Y. Su, F. Wang, S. Wu, Y. Fan, W. Bai, S. Wang, H. Sun, Z. Zhu, W. Liang and A. Li, *Sep. Purif. Technol.*, 2021, **259**, 118203.
248. A. Li, H.-X. Sun, D.-Z. Tan, W.-J. Fan, S.-H. Wen, X.-J. Qing, G.-X. Li, S.-Y. Li and W.-Q. Deng, *Energy Environ. Sci.*, 2011, **4**, 2062–2065.
249. L. Xu, Y. Zang, J. Xiao, Y. Wu, Y. Pan, T. Wu, Y. Tang, J. Cui, H. Jia and F. Miao, *Sep. Purif. Technol.*, 2021, **261**, 118291.
250. M. Zhang, T. Zhao, J. Dou, Z. Xu, W. Zhang, X. Chen, X. Wang and B. Zhou, *ChemElectroChem*, 2019, **6**, 5946–5950.
251. K. Sonogashira, Y. Tohda and N. Hagihara, *Tetrahedron Letters*, 1975, **16**, 4467–4470.
252. H. Doucet and J.-C. Hierso, *Angew. Chem. - Int. Ed.*, 2007, **46**, 834–871.
253. M. Trunk, A. Herrmann, H. Bildirir, A. Yassin, J. Schmidt and A. Thomas, *Chem. Eur. J.*, 2016, **22**, 7179–7183.
254. J.-X. Jiang, F. Su, H. Niu, C. D. Wood, N. L. Campbell, Y. Z. Khimyak and A. I. Cooper, *Chem. Commun.*, 2008, 486–488.
255. M. Carril, A. Correa and C. Bolm, *Angew. Chem. - Int. Ed.*, 2008, **47**, 4862–4865.
256. K. S. Sindhu, A. P. Thankachan, A. M. Thomas and G. Anilkumar, *ChemistrySelect*, 2016, **1**, 556–559.
257. O. Vechorkin, D. Barmaz, V. Proust and X. Hu, *J. Am. Chem. Soc.*, 2009, **131**, 12078–12079.
258. D.-L. Zhu, R. Xu, Q. Wu, H.-Y. Li, J.-P. Lang and H.-X. Li, *J. Org. Chem.*, 2020, **85**, 9201–9212.
259. R. Chinchilla and C. Nájera, *Chem. Rev.*, 2007, **107**, 874–922.
260. D. Tan, W. Fan, W. Xiong, H. Sun, Y. Cheng, X. Liu, C. Meng, A. Li and W.-Q. Deng, *Macromol. Chem. Phys.*, 2012, **213**, 1435–1440.
261. C. Mollart and A. Trewin, *Phys. Chem. Chem. Phys.*, 2020, **22**, 21642–21645.
262. R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724–728.
263. W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 2003, **56**, 2257–2261.
264. T. Yamamoto, A. Morita, Y. Miyazaki, T. Maruyama, H. Wakayama, Z. H. Zhou, Y. Nakamura, T. Kanbara, S. Sasaki and K. Kubota, *Macromolecules*, 1992, **25**, 1214–1223.

265. L. Tan and B. Tan, in *Porous Polymers: Design, Synthesis and Applications*, ed. S. Qiu and T. Ben, The Royal Society of Chemistry, Cambridge, 2015, ch. 4, pp. 66-94.
266. L. Tan and B. Tan, *Chem. Soc. Rev.*, 2017, **46**, 3322-3356.
267. K. Amin, N. Ashraf, L. Mao, C. F. J. Faul and Z. Wei, *Nano Energy*, 2021, **85**, 105958.
268. J. Xie, P. Gu and Q. Zhang, *ACS Energy Lett.*, 2017, **2**, 1985-1996.
269. R. Joseph, W. T. Ford, S. Zhang, M. P. Tsyurupa, A. V. Pastukhov and V. A. Davankov, *J. Polym. Sci. Part A: Polym. Chem.*, 1997, **35**, 695-701.
270. S. Xu, Y. Luo and B. Tan, *Macromol. Rapid Commun.*, 2013, **34**, 471-484.
271. USSR Pat., 299165, 1969.
272. M. P. Tsyurupa and V. A. Davankov, *React. Funct. Polym.*, 2006, **66**, 768-779.
273. C. D. Wood, B. Tan, A. Trewin, H. Niu, D. Bradshaw, M. J. Rosseinsky, Y. Z. Khimiyak, N. L. Campbell, R. Kirk, E. Stöckel and A. I. Cooper, *Chem. Mater.*, 2007, **19**, 2034-2048.
274. B. Li, R. Gong, W. Wang, X. Huang, W. Zhang, H. Li, C. Hu and B. Tan, *Macromolecules*, 2011, **44**, 2410-2414.
275. J. Germain, J. M. J. Fréchet and F. Svec, *J. Mater. Chem.*, 2007, **17**, 4989-4997.
276. V. A. Davankov, M. M. Ilyin, M. P. Tsyurupa, G. I. Timofeeva and L. V. Dubrovina, *Macromolecules*, 1996, **29**, 8398-8403.
277. A. V. Pastukhov, M. P. Tsyurupa and V. A. Davankov, *J. Polym. Sci. Part B: Polym. Phys.*, 1999, **37**, 2324-2333.
278. M. P. Tsyurupa, Z. K. Blinnikova, Y. A. Davidovich, S. E. Lyubimov, A. V. Naumkin and V. A. Davankov, *React. Funct. Polym.*, 2012, **72**, 973-982.
279. J. Hradil and E. Králová, *Polymer*, 1998, **39**, 6041-6048.
280. D. Zhang, L. Tao, J. Ju, Y. Wang, Q. Wang and T. Wang, *Polymer*, 2015, **60**, 234-240.
281. V. A. Davankov, S. V. Rogoshin and M. P. Tsyurupa, *J. Polym. Sci.: Polym. Symp.*, 1974, **47**, 95-101.
282. B. Li, R. Gong, Y. Luo and B. Tan, *Soft Matter*, 2011, **7**, 10910-10916.
283. J.-H. Ahn, J.-E. Jang, C.-G. Oh, S.-K. Ihm, J. Cortez and D. C. Sherrington, *Macromolecules*, 2006, **39**, 627-632.
284. M. P. Tsyurupa and V. A. Davankov, *React. Funct. Polym.*, 2002, **53**, 193-203.

285. D. Chen, S. Gu, Y. Fu, Y. Zhu, C. Liu, G. Li, G. Yu and C. Pan, *Polym. Chem.*, 2016, **7**, 3416–3422.
286. Y. Yang, Q. Zhang, S. Zhang and S. Li, *Polymer*, 2013, **54**, 5698–5702.
287. M. G. Schwab, A. Lennert, J. Pahnke, G. Jonschker, M. Koch, I. Senkovska, M. Rehahn and S. Kaskel, *J. Mater. Chem.*, 2011, **21**, 2131–2135.
288. Y. Luo, S. Zhang, Y. Ma, W. Wang and B. Tan, *Polym. Chem.*, 2013, **4**, 1126–1131.
289. B. Li, Z. Guan, X. Yang, W. D. Wang, W. Wang, I. Hussain, K. Song, B. Tan and T. Li, *J. Mater. Chem. A*, 2014, **2**, 11930–11939.
290. P. Cui, X.-F. Jing, Y. Yuan and G.-S. Zhu, *Chin. Chem. Lett.*, 2016, **27**, 1479–1484.
291. M. Errahali, G. Gatti, L. Tei, G. Paul, G. A. Rolla, L. Canti, A. Fraccarollo, M. Cossi, A. Comotti, P. Sozzani and L. Marchese, *J. Phys. Chem. C*, 2014, **118**, 28699–28710.
292. Y. Luo, B. Li, W. Wang, K. Wu and B. Tan, *Adv. Mater.*, 2012, **24**, 5703–5707.
293. A. Modak, Y. Maegawa, Y. Goto and S. Inagaki, *Polym. Chem.*, 2016, **7**, 1290–1296.
294. P. Puthiaraj and W.-S. Ahn, *Ind. Eng. Chem. Res.*, 2016, **55**, 7917–7923.
295. M. Saleh, H. M. Lee, K. C. Kemp and K. S. Kim, *ACS Appl. Mater. Interfaces*, 2014, **6**, 7325–7333.
296. X. Yang, M. Yu, Y. Zhao, C. Zhang, X. Wang and J.-X. Jiang, *J. Mater. Chem. A*, 2014, **2**, 15139–15145.
297. S. Yao, X. Yang, M. Yu, Y. Zhang and J.-X. Jiang, *J. Mater. Chem. A*, 2014, **2**, 8054–8059.
298. C. Zhang, P.-C. Zhu, L. Tan, J.-M. Liu, B. Tan, X.-L. Yang and H.-B. Xu, *Macromolecules*, 2015, **48**, 8509–8514.
299. T. Ratvijitvech, M. Barrow, A. I. Cooper and D. J. Adams, *Polym. Chem.*, 2015, **6**, 7280–7285.
300. Y. Yang, B. Tan and C. D. Wood, *J. Mater. Chem. A*, 2016, **4**, 15072–15080.
301. A. Toutianoush and B. Tieke, *Macromol. Rapid Commun.*, 1998, **19**, 591–595.
302. R. Ou, H. Zhang, V. X. Truong, L. Zhang, H. M. Hegab, L. Han, J. Hou, X. Zhang, A. Deletic, L. Jiang, G. P. Simon and H. Wang, *Nat. Sustain.*, 2020, **3**, 1052–1058.

303. K. Griffiths, N. R. Halcovitch and J. M. Griffin, *New J. Chem.*, 2022, **46**, 4057–4061.
304. H. K. Cammenga, V. N. Emel'yanenko and S. P. Verevkin, *Ind. Eng. Chem. Res.*, 2009, **48**, 10120–10128.
305. A. R. Dias, M. E. Minas Da Piedade, J. A. Martinho Simões, J. A. Simoni, C. Teixeira, H. P. Diogo, Y. Meng-Yan and G. Pilcher, *J. Chem. Thermodyn.*, 1992, **24**, 439–447.
306. E. Durgun and J. C. Grossman, *J. Phys. Chem. Lett.*, 2013, **4**, 854–860.
307. G. D. Han, S. S. Park, Y. Liu, D. Zhitomirsky, E. Cho, M. Dincă and J. C. Grossman, *J. Mater. Chem. A*, 2016, **4**, 16157–16165.
308. T. J. Kucharski, N. Ferralis, A. M. Kolpak, J. O. Zheng, D. G. Nocera and J. C. Grossman, *Nat. Chem.*, 2014, **6**, 441–447.
309. K. Masutani, M.-a. Morikawa and N. Kimizuka, *Chem. Commun.*, 2014, **50**, 15803–15806.
310. G. S. Hartley, *J. Chem. Soc.*, 1938, 633–642.
311. G. S. Hartley and R. J. W. Le Fèvre, *J. Chem. Soc.*, 1939, 531–535.
312. K. Tsuda, G. C. Dol, T. Gensch, J. Hofkens, L. Latterini, Weener, Meijer and D. Schryver, *J. Am. Chem. Soc.*, 2000, **122**, 3445–3452.
313. N. Wazzan, PhD thesis, The University of Edinburgh, 2009.
314. N. Biswas and S. Umapathy, *J. Phys. Chem. A*, 1997, **101**, 5555–5566.
315. P. C. Chen and Y. C. Chieh, *J. Mol. Struct.: THEOCHEM*, 2003, **624**, 191–200.
316. C. R. Crecca and A. E. Roitberg, *J. Phys. Chem. A*, 2006, **110**, 8188–8203.
317. H. Fliegl, A. Köhn, C. Hättig and R. Ahlrichs, *J. Am. Chem. Soc.*, 2003, **125**, 9821–9827.
318. N. Kurita, T. Ikegami and Y. Ishikawa, *Chem. Phys. Lett.*, 2002, **360**, 349–354.
319. N. Kurita, S. Tanaka and S. Itoh, *J. Phys. Chem. A*, 2000, **104**, 8114–8120.
320. H. Rau, in *Photochromism: molecules and systems*, ed. H. Dürr and H. Bouas-Laurent, Elsevier, Amsterdam, 2nd edn., 2003, ch. 4, pp. 165–188.
321. T. H. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007–1023.
322. S. Grimme, *WIREs Comput. Mol. Sci.*, 2011, **1**, 211–228.
323. A. Rowling, J. Doulcet, R. Dawson, N. Tapia-Ruiz and A. Trewin, *ACS Appl. Mater. Interfaces*, under review.

324. *Energy Trends: December 2023*, <https://www.gov.uk/government/statistics/energy-trends-december-2023>, (accessed 15.02.2024).
325. P. Heasman and A. Trewin, *J. Phys. Chem. C*, 2019, **123**, 25603–25610.
326. H. O. Pierson, *Handbook of Carbon, Graphite, Diamonds and Fullerenes: Processing, Properties and Applications*, Elsevier, Amsterdam, 1994.
327. M. T. Weller, T. Overton, J. Rourke, F. A. Armstrong and P. W. Atkins, *Inorg. Chem.* Oxford University Press, Oxford, 6th ed., 2014.
328. Y. S. Chung, S. H. Yoo and C. K. Kim, *Ind. Eng. Chem. Res.*, 2009, **48**, 4346–4351.
329. Q. Li, J. Chen, L. Fan, X. Kong and Y. Lu, *Green Energy Environ.*, 2016, **1**, 18–42.
330. W. M. Haynes, *CRC Handbook of Chemistry and Physics*, CRC Press, Boca Raton, 2016.
331. *Centers for Disease Control and Prevention: Cobalt*, <https://www.cdc.gov/niosh/topics/cobalt/default.html>, (accessed 09.02.2024).
332. R. D. Shannon, *Acta Cryst.*, 1976, **A32**, 751–767.
333. G. Çakmak and T. Öztürk, *Diam. Relat. Mater.*, 2019, **96**, 134–139.
334. Y. Xiong, Y. Liu, L. Chen, S. Zhang, X. Zhu, T. Shen, D. Ren, X. He, J. Qiu, L. Wang, Q. Hu and H. Zhang, *Energy Environ. Mater.*, 2022, **5**, 872–876.
335. Y.-S. Hu, P. Adelhelm, B. M. Smarsly, S. Hore, M. Antonietti and J. Maier, *Adv. Funct. Mater.*, 2007, **17**, 1873–1878.
336. Y. J. Hwang, S. K. Jeong, K. S. Nahm, J. S. Shin and A. Manuel Stephan, *J. Phys. Chem. Solids*, 2007, **68**, 182–188.
337. W. Long, B. Fang, A. Ignaszak, Z. Wu, Y.-J. Wang and D. Wilkinson, *Chem. Soc. Rev.*, 2017, **46**, 7176–7190.
338. A. M. Stephan, T. P. Kumar, R. Ramesh, S. Thomas, S. K. Jeong and K. S. Nahm, *Mater. Sci. Eng. A*, 2006, **430**, 132–137.
339. S.-H. Yeon, K.-N. Jung, S. Yoon, K.-H. Shin and C.-S. Jin, *J. Phys. Chem. Solids*, 2013, **74**, 1045–1055.
340. B. Sakintuna and Y. Yürüm, *Ind. Eng. Chem. Res.*, 2005, **44**, 2893–2902.
341. M. R. Benzigar, S. N. Talapaneni, S. Joseph, K. Ramadass, G. Singh, J. Scaranto, U. Ravon, K. Al-Bahily and A. Vinu, *Chem. Soc. Rev.*, 2018, **47**, 2680–2721.
342. A. D. Roberts, X. Li and H. Zhang, *Chem. Soc. Rev.*, 2014, **43**, 4341–4356.
343. P. Kalyani and A. Anitha, *Int. J. Hydrogen Energy*, 2013, **38**, 4034–4045.

344. M.-M. Titirici and M. Antonietti, *Chem. Soc. Rev.*, 2010, **39**, 103–116.
345. Z. Ma, T. Kyotani and A. Tomita, *Chem. Commun.*, 2000, 2365–2366.
346. H. Nishihara and T. Kyotani, *Adv. Mater.*, 2012, **24**, 4473–4498.
347. R. Ryoo, S. H. Joo and S. Jun, *J. Phys. Chem. B*, 1999, **103**, 7743–7746.
348. V. Presser, M. Heon and Y. Gogotsi, *Adv. Funct. Mater.*, 2011, **21**, 810–833.
349. G. N. Yushin, E. N. Hoffman, A. Nikitin, H. Ye, M. W. Barsoum and Y. Gogotsi, *Carbon*, 2005, **43**, 2075–2082.
350. G. Eglinton and A. R. Galbraith, *J. Chem. Soc.*, 1959, 889–896.
351. Y. X. Shao, Y. H. Cai, D. Dong, S. Wang, S. G. Ang and G. Q. Xu, *Chem. Phys. Lett.*, 2009, **482**, 77–80.
352. F. Tao, M. H. Qiao, Z. H. Li, L. Yang, Y. J. Dai, H. G. Huang and G. Q. Xu, *Phys. Rev. B*, 2003, **67**, 115334.
353. P. K. Chu and L. Li, *Mater. Chem. Phys.*, 2006, **96**, 253–277.
354. A. Dato, V. Radmilovic, Z. Lee, J. Phillips and M. Frenklach, *Nano Lett.*, 2008, **8**, 2012–2016.
355. A. Duarte-Moller, F. Espinosa-Magaña, R. Martínez-Sánchez, M. Avalos-Borja, G. A. Hirata and L. Cota-Araiza, *J. Electron Spectrosc. Relat. Phenom.*, 1999, **104**, 61–66.
356. A. P. Hitchcock, S. Johnston, T. Tyliczszak, C. C. Turci, M. Barbatti, A. B. Rocha and C. E. Bielschowsky, *J. Electron Spectrosc. Relat. Phenom.*, 2002, **123**, 303–314.
357. V. Mortet, L. Zhang, M. Eckert, J. D’Haen, A. Soltani, M. Moreau, D. Troadec, E. Neyts, J.-C. De Jaeger, J. Verbeeck, A. Bogaerts, G. Van Tendeloo, K. Haenen and P. Wagner, *Phys. Status Solidi A*, 2012, **209**, 1675–1682.
358. M. Häussler, R. Zheng, J. W. Y. Lam, H. Tong, H. Dong and B. Z. Tang, *J. Phys. Chem. B*, 2004, **108**, 10645–10650.
359. L. Chen, N. Li, X. Yu, S. Zhang, C. Liu, Y. Song, Z. Li, S. Han, W. Wang, P. Yang, N. Hong, S. Ali and Z. Wang, *Chem. Eng. J.*, 2023, **462**, 142139.
360. Z. Yang, Y. Song, C. Zhang, J. He, X. Li, X. Wang, N. Wang, Y. Li and C. Huang, *Adv. Energy Mater.*, 2021, **11**, 2101197.
361. Z. Yang, X. Ren, Y. Song, X. Li, C. Zhang, X. Hu, J. He, J. Li and C. Huang, *Energy Environ. Mater.*, 2023, **6**, e12269.
362. C. Dong, J. Chu, L. Cao, F. Cui, S. Liang, X. Zhang, X. Tao, H.-G. Wang and G. Zhu, *CCS Chem.*, 2022, **5**, 607–615.

363. P. Heasman, E. Varley and A. Trewin, *Energy Fuels*, 2022, **36**, 6560–6568.
364. J. Wang, X. Fu, N. Yan and Y. Zhang, *ChemSusChem*, 2021, **14**, 3806–3809.
365. F. Cataldo and D. Capitani, *Mater. Chem. Phys.*, 1999, **59**, 225–231.
366. M. Dubois, K. Guérin, E. Petit, N. Batische, A. Hamwi, N. Komatsu, J. Giraudet, P. Pirotte and F. Masin, *J. Phys. Chem. C*, 2009, **113**, 10371–10378.
367. E. Karakhanov, A. Maximov, M. Terenina, V. Vinokurov, L. Kulikov, D. Makeeva and A. Glotov, *Catal. Today*, 2020, **357**, 176–184.
368. A. M. Panich, *Diam. Relat. Mater.*, 2007, **16**, 2044–2049.
369. S. Ren, R. Dawson, A. Laybourn, J.-x. Jiang, Y. Khimiyak, D. J. Adams and A. I. Cooper, *Polym. Chem.*, 2012, **3**, 928–934.
370. H. He, T. Riedl, A. Lerf and J. Klinowski, *J. Phys. Chem.*, 1996, **100**, 19954–19958.
371. N. J. Brownbill, R. S. Sprick, B. Bonillo, S. Pawsey, F. Aussenac, A. J. Fielding, A. I. Cooper and F. Blanc, *Macromolecules*, 2018, **51**, 3088–3096.
372. K. J. Harris, G. M. Bernard, C. McDonald, R. McDonald, M. J. Ferguson and R. E. Wasylshen, *Inorg. Chem.*, 2006, **45**, 2461–2473.
373. L. George, K.-P. Netsch, G. Penn, G. Kollenz and C. Wentrup, *Org. Biomol. Chem.*, 2006, **4**, 558–564.
374. T. M. Cooper, D. M. Krein, A. R. Burke, D. G. McLean, J. E. Rogers, J. E. Slagle and P. A. Fleitz, *J. Phys. Chem. A*, 2006, **110**, 4369–4375.
375. S. Gu, S. Wu, L. Cao, M. Li, N. Qin, J. Zhu, Z. Wang, Y. Li, Z. Li, J. Chen and Z. Lu, *J. Am. Chem. Soc.*, 2019, **141**, 9623–9628.
376. F. Breher, J. Grunenberg, S. C. Lawrence, P. Mountford and H. Rügger, *Angew. Chem. - Int. Ed.*, 2004, **43**, 2521–2524.
377. T.-K. Huang, Y. Chi, S.-M. Peng, G.-H. Lee, S.-L. Wang and F.-L. Liao, *Organometallics*, 1995, **14**, 2164–2166.
378. M. H. Chisholm, D. Ho, J. C. Huffman and N. S. Marchant, *Organometallics*, 1989, **8**, 1626–1636.
379. K. Fukui, K. Sato, D. Shiomi, T. Takui, K. Itoh, K. Gotoh, T. Kubo, K. Yamamoto, K. Nakasuji and A. Naito, *Synth. Met.*, 1999, **103**, 2257–2258.
380. R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 2008, **72**, 650–654.

381. University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, TURBOMOLE v7.5, TURBOMOLE GmbH, Karlsruhe, 2017.
382. J. R. Cheeseman, G. W. Trucks, T. A. Keith and M. J. Frisch, *J. Chem. Phys.*, 1996, **104**, 5497–5509.
383. T. A. Keith and R. F. W. Bader, *Chem. Phys. Lett.*, 1992, **194**, 1–8.
384. T. A. Keith and R. F. W. Bader, *Chem. Phys. Lett.*, 1993, **210**, 223–231.
385. W. Deng, J. R. Cheeseman and M. J. Frisch, *J. Chem. Theory Comput.*, 2006, **2**, 1028–1037.
386. C. Mollart, P. Heasman, P. A. T. J. Fletcher, P. Fayon, J. M. H. Thomas, V. Franckevičius, M. J. G. Peach and A. Trewin, *Adv. Mater.*, submitted.
387. M. C. Bowman, A. D. Burke, J. M. Turney and H. F. I. Schaefer, *J. Phys. Chem. A*, 2018, **122**, 9498–9511.
388. T. H. Dunning Jr. and P. J. Hay, in *Modern Theoretical Chemistry*, ed. H. F. Schaefer III, Plenum, New York, 1977, vol. 3, ch. 1, pp. 1–28.
389. P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 270–283.
390. P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 299–310.
391. W. R. Wadt and P. J. Hay, *J. Chem. Phys.*, 1985, **82**, 284–298.
392. *dia*, <https://rcsr.anu.edu.au/nets/dia>, (accessed 18.07.2024).
393. *dia-c*, <https://rcsr.anu.edu.au/nets/dia-c>, (accessed 18.07.2024).
394. *dia-c3*, <https://rcsr.anu.edu.au/nets/dia-c3>, (accessed 18.07.2024).
395. *dia-c4*, <https://rcsr.anu.edu.au/nets/dia-c4>, (accessed 18.07.2024).
396. *dia-c6*, <https://rcsr.anu.edu.au/nets/dia-c6>, (accessed 18.07.2024).
397. *pts-c*, <https://rcsr.anu.edu.au/nets/pts-c>, (accessed 18.07.2024).
398. *unc-c*, <https://rcsr.anu.edu.au/nets/unc-c>, (accessed 18.07.2024).
399. O. Delgado Friedrichs, M. O’Keeffe and O. M. Yaghi, *Acta Cryst. A*, 2003, **59**, 22–27.
400. C. Bonneau, O. Delgado-Friedrichs, M. O’Keeffe and O. M. Yaghi, *Acta Cryst. A*, 2004, **60**, 517–520.
401. C. Bonneau and M. O’Keeffe, *Acta Cryst. A*, 2015, **71**, 82–91.
402. Z. Guo, Z. Zhang and J. Sun, *Adv. Mater.*, 2024, **36**, 2312889.
403. *Reticular Chemistry Structure Resource*, <https://rcsr.anu.edu.au/>, (accessed 18.07.2024).
404. *What is Crystal Structure Prediction? And why is it so difficult?*, <https://www.ccdc.cam.ac.uk/discover/blog/what-is-crystal-structure-prediction-csp>, (accessed 19.02.2024).

405. R. E. Botto, R. Wilson and R. E. Winans, *Energy Fuels*, 1987, **1**, 173–181.
406. G. N. La Mar, *J. Am. Chem. Soc.*, 1971, **93**, 1040–1041.
407. K. Ohgo and K. K. Kumashiro, *J. Magn. Reson.*, 2018, **297**, 161–171.
408. J. W. Akitt and B. E. Mann, *NMR and Chemistry: An introduction to modern NMR spectroscopy*, CRC Press, Boca Raton, 2000.
409. I. C. Felli, R. Pierattelli, S. J. Glaser and B. Luy, *J. Biomol. NMR*, 2009, **43**, 187–196.
410. J. A. Franz, R. Garcia, J. C. Linehan, G. D. Love and C. E. Snape, *Energy Fuels*, 1992, **6**, 598–602.
411. J. C. C. Freitas, T. J. Bonagamba and F. G. Emmerich, *Energy Fuels*, 1999, **13**, 53–59.
412. J. W. M. Osterrieth, J. Rampersad, D. Madden, N. Rampal, L. Skoric, B. Connolly, M. D. Allendorf, V. Stavila, J. L. Snider, R. Ameloot, J. Marreiros, C. Ania, D. Azevedo, E. Vilarrasa-Garcia, B. F. Santos, X.-H. Bu, Z. Chang, H. Bunzen, N. R. Champness, S. L. Griffin, B. Chen, R.-B. Lin, B. Coasne, S. Cohen, J. C. Moreton, Y. J. Colón, L. Chen, R. Clowes, F.-X. Coudert, Y. Cui, B. Hou, D. M. D’Alessandro, P. W. Doheny, M. Dincă, C. Sun, C. Doonan, M. T. Huxley, J. D. Evans, P. Falcaro, R. Ricco, O. Farha, K. B. Idrees, T. Islamoglu, P. Feng, H. Yang, R. S. Forgan, D. Bara, S. Furukawa, E. Sanchez, J. Gascon, S. Telalović, S. K. Ghosh, S. Mukherjee, M. R. Hill, M. M. Sadiq, P. Horcajada, P. Salcedo-Abraira, K. Kaneko, R. Kukobat, J. Kenvin, S. Keskin, S. Kitagawa, K.-i. Otake, R. P. Lively, S. J. A. DeWitt, P. Llewellyn, B. V. Lotsch, S. T. Emmerling, A. M. Pütz, C. Martí-Gastaldo, N. M. Padial, J. García-Martínez, N. Linares, D. MasPOCH, J. A. Suárez del Pino, P. Moghadam, R. Oktavian, R. E. Morris, P. S. Wheatley, J. Navarro, C. Petit, D. Danaci, M. J. Rosseinsky, A. P. Katsoulidis, M. Schröder, X. Han, S. Yang, C. Serre, G. Mouchaham, D. S. Sholl, R. Thyagarajan, D. Siderius, R. Q. Snurr, R. B. Goncalves, S. Telfer, S. J. Lee, V. P. Ting, J. L. Rowlandson, T. Uemura, T. Iiyuka, M. A. van der Veen, D. Rega, V. Van Speybroeck, S. M. J. Rogge, A. Lemaire, K. S. Walton, L. W. Bingel, S. Wuttke, J. Andreato, O. Yaghi, B. Zhang, C. T. Yavuz, T. S. Nguyen, F. Zamora, C. Montoro, H. Zhou, A. Kirchon and D. Fairen-Jimenez, *Adv. Mater.*, 2022, **34**, 2201502.
413. S. Ren, M. J. Bojdys, R. Dawson, A. Laybourn, Y. Z. Khimyak, D. J. Adams and A. I. Cooper, *Adv. Mater.*, 2012, **24**, 2357–2361.
414. W. Huang, Q. He, Y. Hu and Y. Li, *Angew. Chem. - Int. Ed.*, 2019, **58**, 8676–8680.

415. R. Palkovits, M. Antonietti, P. Kuhn, A. Thomas and F. Schüth, *Angew. Chem. - Int. Ed.*, 2009, **48**, 6909–6912.
416. Y. Li, S. Zheng, X. Liu, P. Li, L. Sun, R. Yang, S. Wang, Z.-S. Wu, X. Bao and W.-Q. Deng, *Angew. Chem. - Int. Ed.*, 2018, **57**, 7992–7996.
417. P. Kuhn, A. Forget, D. Su, A. Thomas and M. Antonietti, *J. Am. Chem. Soc.*, 2008, **130**, 13333–13337.
418. P. Kuhn, A. Thomas and M. Antonietti, *Macromolecules*, 2009, **42**, 319–326.
419. H. Wang, D. Jiang, D. Huang, G. Zeng, P. Xu, C. Lai, M. Chen, M. Cheng, C. Zhang and Z. Wang, *J. Mater. Chem. A*, 2019, **7**, 22848–22870.
420. M. J. Bojdys, J. Jeromenok, A. Thomas and M. Antonietti, *Adv. Mater.*, 2010, **22**, 2202–2205.
421. P. Katekomol, J. Roeser, M. Bojdys, J. Weber and A. Thomas, *Chem. Mater.*, 2013, **25**, 1542–1548.
422. E. Troschke, S. Grätz, T. Lübken and L. Borchardt, *Angew. Chem. - Int. Ed.*, 2017, **56**, 6859–6863.
423. A. Herrera, A. Riaño, R. Moreno, B. Caso, Z. D. Pardo, I. Fernández, E. Sáez, D. Molero, A. Sánchez-Vázquez and R. Martínez-Alvarez, *J. Org. Chem.*, 2014, **79**, 7012–7024.
424. A. Trummal, L. Lipping, I. Kaljurand, I. A. Koppel and I. Leito, *J. Phys. Chem. A*, 2016, **120**, 3663–3669.
425. H. Li, A. D. Chavez, H. Li, H. Li, W. R. Dichtel and J.-L. Bredas, *J. Am. Chem. Soc.*, 2017, **139**, 16310–16318.
426. A. N. Marimuthu, F. Huis in't Veld, S. Thorwirth, B. Redlich and S. Brünken, *J. Mol. Spectrosc.*, 2021, **379**, 111477.
427. *Precomputed vibrational scaling factors*, <https://cccbdb.nist.gov/vibscale-just.asp>, (accessed 19.09.2023).
428. R. Dawson, L. A. Stevens, T. C. Drage, C. E. Snape, M. W. Smith, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2012, **134**, 10741–10744.
429. Y. Maréchal, *J. Mol. Struct.*, 2003, **648**, 27–47.
430. D. F. Dinu, M. Podewitz, H. Grothe, T. Loerting and K. R. Liedl, *Phys. Chem. Chem. Phys.*, 2020, **22**, 17932–17947.
431. W. Yu, S. Gu, Y. Fu, S. Xiong, C. Pan, Y. Liu and G. Yu, *J. Catal.*, 2018, **362**, 1–9.