

# MODELS FOR THE UNDERSTANDING OF MOLECULARLY IMPRINTED POLYMERS

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Todd Cowen BSc, MSc (Leicester)

Department of Chemistry

University of Leicester

September 2018

## Abstract

Polymer synthesis can be performed in the presence of template molecule to produce a corresponding cavity, giving a molecularly imprinted polymer (MIP) with affinity for the template. A brief review of the predominant techniques in MIP computational analysis will be given and contrasted with a molecular mechanics/molecular dynamics (MM/MD) alternative, including some examples of how this can be applied. Automated synthesis of MIPs *in silico* will then be demonstrated as a further development of the MM/MD method, producing polymers by mimicking radical polymerization atomistically. Comparative analysis in the design of a synthetic ephedrine receptor demonstrates that the new method can effectively identify affinity trends and binding site selectivities where analysis of template-monomer and template-solution systems cannot.

Studies of polymer nanoparticle dimensions were then pursued and found to correlate with polymer solubility when expressed as the Flory parameter  $\chi_{s,p}$ . A modified Flory-Huggins based thermodynamic model was then developed for the analysis of the hydrodynamic diameters, and the absolute size of polymer nanoparticles was found to be predictable by varying the polymerization conditions that influence  $\chi_{s,p}$ . The position of the spinodal, associated with a given  $\chi_{s,p}$  equivalent, allows an absolute value,  $\Delta\chi_{spinodal}$ , to be calculated. The hydrodynamic diameter,  $D$ , of nanoparticles at the primarily observed fraction was then found to be dependent on  $D$  (nm) =  $-74\Delta\chi_{spinodal} + 367$  nm, where  $\Delta\chi_{spinodal}$  must be positive for successful separation. The polymerization algorithm was then applied to the prepolymerization system in an attempt to improve understanding and prediction of  $\Delta\chi_{spinodal}$ . From these studies it is concluded that MIP synthesis occurs by a binodal-character phase separation, having implications for the synthetic mechanism. The polymerization algorithm, thermodynamic model, synthetic mechanism and described relationship between polymer diameter and solubility should therefore be useful additions to future analysis.

## Acknowledgment

Obviously I am indebted to all the people that I have interacted with over the last few years (Bashar and Georgie for example, who mentioned me in their acknowledgements), and through the previous portion of my life which now seems so irrelevant. I will otherwise limit this however to those who have had a direct impact on this work.

Thanks to all my collaborators, notably Antonio Guerreiro, Elena Piletska, Julie Settapani, Anca Florea, Cem Esen, Sabrina Di Masi, Daniel Lopez and Hasim Munawar; to all my students, especially Matthew Young, Alistair Watson, Hilary Wu and Emily Kohler; and for additional help from Mirko Busato, Andrew Myers, Shengfu Yang and Nikolai Brilliantov. Apologies for anyone I have forgotten or ignored.

I am particularly grateful to my two supervisors. Firstly Kal Karim; always helpful, always interesting, always ready to let me use as much money as I want. Secondly Sergey Piletsky, who has probably had a greater impact on my thinking than any other person I have met so far, for better or worse.

Finally I will reserve my strongest gratitude to Joanna Czulak: forever resilient, though rarely quietly. For your tolerance both in and out of the lab I am eternally grateful.

Dedicated to the memory of Sarah,  
my condolences to her friends and family.

# Table of Contents

Abstract .....	i
Acknowledgment.....	ii
List of tables and figures .....	vi
Tables .....	vi
Figures.....	viii
Symbols, abbreviations and terminology .....	xiii
Publications.....	xviii
1. Introduction .....	1
1.1. <i>Molecularly imprinted polymers</i> .....	4
1.2. <i>Computational approaches in the design of MIPs</i> .....	10
1.3. <i>Quantum mechanical methods overview and self-consistent field</i> .....	15
1.4. <i>Semi-empirical methods</i> .....	19
1.5. <i>Density functional theory</i> .....	22
1.6. <i>Examples of QM in imprinted polymer analysis</i> .....	24
1.7. <i>Conclusions</i> .....	30
2. Molecular mechanics and molecular dynamics .....	31
2.1. <i>Molecular mechanics</i> .....	31
2.2. <i>Molecular dynamics</i> .....	38
2.3. <i>Force fields</i> .....	39
2.4. <i>Simple molecular mechanics</i> .....	41
2.5. <i>Automated monomer screening with standard monomer database</i> .....	43
2.6. <i>Automated screening with electropolymerizable monomer database</i> .....	47
2.7. <i>Further applications of molecular mechanics modeling</i> .....	54
2.8. <i>Conclusions</i> .....	57
3. In silico synthesis of synthetic receptors.....	58
3.1. <i>Problems with existing computational approaches</i> .....	58
3.2. <i>Molecular models of polymers</i> .....	62
3.3. <i>Quantum mechanical polymerization</i> .....	66
3.4. <i>Manual and automated polymerization</i> .....	69
3.5. <i>Development of a polymerization algorithm</i> .....	73
3.6. <i>Volume requirements</i> .....	78

3.7.	<i>Experiments with ephedrine imprinted polymers</i> .....	79
4.	Size and solubility of polymer nanoparticles .....	85
4.1.	<i>Limits of atomistic modeling</i> .....	85
4.2.	<i>Introduction to solubility</i> .....	87
4.3.	<i>Nanoparticles in vivo</i> .....	88
4.4.	<i>Investigation of solubility parameters</i> .....	90
4.5.	<i>Gibbs energy of mixing calculation</i> .....	95
4.6.	<i>Spinodal and binodal phase separation</i> .....	98
4.7.	<i>The predictive capacity of <math>\Delta\chi</math></i> .....	107
4.8.	<i>Conclusions</i> .....	113
5.	Molecular mechanics of polymer solubility .....	114
5.1.	<i>Modeling solubility</i> .....	114
5.2.	<i>Models of imprinted polymer synthesis</i> .....	115
5.3.	<i>Determining the Gibbs energy of solvation</i> .....	121
5.4.	<i>Preliminary analysis of <math>\Delta G_{solv}</math> calculation</i> .....	123
5.5.	<i>Calculation of <math>\Delta G_{solv}</math> for polymers</i> .....	125
5.6.	<i>Conclusions</i> .....	127
6.	Conclusion.....	128
	Appendix 1. Docking of drugs with 4-aminobenzoic acid oligomer .....	131
	Appendix 2. Algorithms.....	132
	2.1. <i>Polymerization algorithm</i> .....	132
	2.2. <i>System preparation</i> .....	143
	2.3. <i>Intermolecular potential</i> .....	145
	2.4. <i>Analysis of <math>\Delta G_{solv}</math></i> .....	147
	Appendix 3. Experimental section for Chapter 4.....	150
	3.1. <i>Materials and methods</i> .....	150
	3.2. <i>Primary experiments</i> .....	150
	3.3. <i>Preliminary experiments</i> .....	151
	References.....	152

## List of tables and figures

### Tables

*Table 1.1: Compilation of the use of computational approaches in MIP design 2013-2016, covering work published after the last major review of this area.<sup>54</sup> References are given according to the target selected, the computational technique applied in the modeling, and the manner in which the models were used. This table was previously published in *Analytica Chimica Acta* (Cowen et al., 2016).<sup>51</sup>*

*Table 2.1: Mean bond lengths and potential energies of NAPMA copper complexes.*

*Table 2.2: Mean bond lengths and potential energies of NAPMA and acrylate copper complexes.*

*Table 2.3a: The results of the automated screening of aminoparathion with the Leapfrog algorithm.*

*Table 2.3b: The results of the automated screening of methyl parathion with the Leapfrog algorithm.*

*Table 2.4: Binding energies obtained from the Leapfrog screening of multiple templates (drugs represented in charged and uncharged form) with a virtual database of electropolymerizable monomers.*

*Table 2.5: Binding scores of cocaine and several interferent drugs with 4-aminobenzoic acid in neutral and charged forms.*

*Table 2.6: Results of the oligomer-drug docking studies.*

*Table 3.1: Empirical affinity results and theoretical predictions with three MIP compositions. Empirical values and LeapFrog relative binding score were retrieved from Piletsky et al., and pre-polymerization and polymer affinity scores were obtained using Equation 3.4. The abbreviations in parentheses are the functional monomers used in the syntheses: methacrylic acid (MAA), hydroxyethyl methacrylate (HEM), 2-vinylpyridine (2-VP).*

*Table 3.2: Comparison of empirical selectivity, theoretical selectivity based on the pre-polymerization system and theoretical selectivity based on the model polymer.*

*Table 4.1: Relationship between solubility parameter components,  $R_{s,p}$ ,  $\chi_{s,p}$ , and observed diameter of polymer nanoparticles. Only  $\chi_{s,p}$  is seen to effectively predict the initial reduction in polymer diameter with increasing THF content.*

*Table 5.1: Calculated and empirical values of  $\Delta G_{solv}$  for various compounds.<sup>317</sup> The calculated values adjusted for residual analyte energy can be observed to be more accurate in the first three rows.*

## Figures

*Figure 1.1: The standard schematic diagram giving an overview of molecularly imprinted polymers: a) a template molecule (gray) forms a complex in solution with functional monomers via strong intermolecular interactions (shown as colored shapes fitting the cavities of the template); b) polymerization is induced with the complex retaining its strong interactions with the functional groups; c) On template release a cavity remains with a geometry and functionality suitable for rebinding the template.*

*Figure 1.2: Synthesis of Wulff and Sarhan's D-glyceric acid imprinted polymer. The target is covalently bound to the monomer and eventual polymer before being removed, leaving a specific binding site. This method is referred to as covalent imprinting.*

*Figure 1.3: a) Slater 1s basis function (Slater type orbital, STO), a representation of the one electron wave function; b) STO-3G 1s basis function, an approximation of the STO using a combination of Gaussian functions.*

*Figure 1.4: Stewart's illustrations of water dimers, demonstrating the development of the Parametric Methods (a-c) and a comparison with 6-31G\* (reproduced with permission).<sup>173</sup>*

*Figure 1.5: The organometallic binding site in the nicotine imprinted polymer developed by Huynh et al. (reproduced with permission).*

*Figure 1.6: Structures of the target amino acid ester, its hydrolytic transition state, and the template used to produce a synthetic enzyme for this reaction.*

*Figure 2.1: Harmonic oscillator and Morse potential, both functions that may be used to describe the bond stretching (1,2) interactions between atoms. Harmonic oscillators are simpler, can be applied more broadly with minimal modification, and less computationally expensive, but the Morse potential more accurately represents the natural behavior of bonds.*

*Figure 2.2: The potential energy observed in rotation about the 2-3 bond of butane.*

*Figure 2.3: The Leapfrog automated screening in progress, with a template (vancomycin, center) being screened against a database of monomers. Points of maximal and minimal charge at optimal distance are given by blue and red points around the template, with neutral areas given in yellow.*

*Figure 2.4: Ochratoxin A (A) and two structural analogies (B and C) in their minimum energy configurations.*

*Figure 2.5: 4-aminobenzoic acid based oligomer used to dock cocaine and several interferent drugs.*

*Figure 2.6: The interactions observed between the linear molecularly imprinted polymer fragments and the quorum sensing signaling peptide (Motib, 2017). Initial predictions of the relative positions of the fragments were originally made with Leapfrog, screening the peptide with the different fragments, then refined with the Surfex docking software.*

*Figure 3.1: Comparison of monomer-template models and monomer-template-cross-linker interactions vs empirical binding (data from Sobiech et al. 2014, original graph). The data for the model including cross-linker molecules can be observed to correlate with the empirical values obtained, while the values for the monomer and template alone are less reliable.*

*Figure 3.2: An early example of a simple approach to approximating polymer-template/target interactions by modifying the monomer to better represent the equivalent polymer unit.*

*Figure 3.3: Polymers constructed from 50 individual monomers and observed for interaction with theophylline. The figure on the left shows a polymer composed of MAA alone, and the right a copolymer of MAA and MMA (Monti et al., 2006).<sup>233</sup>*

*Figure 3.4: The polymerization simulation applied by Schauerl and Lewis, which incorporates a zeolite modeling tool to grow a polymer via an evolutionary mechanism.*

*Figure 3.5: Manually produced methacrylic acid based polymer with ephedrine (red). Solvent chloroform is explicitly present but not displayed.*

*Figure 3.6: Explanation of the functioning of the polymerization algorithm. This scheme and the following images were originally published in Macromolecular Rapid Communications (Cowen, 2016).<sup>245</sup>*

*Figure 3.7: a) the reactive atom of the assigned molecule scans the surroundings in concentric spheres until an appropriately reactive atom is found with which to form a bond. Bond formation is favored with closer atoms, if multiple atoms are found within the same interval the more reactive is favored; b) reactions with sterically strained structures occur in the absence of suitable alkenes. Strained structures will not spontaneously break or react with each other.*

*Figure 3.8: a 104 kDa macromolecule synthesized in silico using the polymerization algorithm, with an ephedrine molecule for comparison.*

*Figure 4.1: Preliminary experiments using a variety of particle synthesis techniques and reaction environments gave a correlation between particle diameter solubility (expressed as the reciprocal of the Flory parameter  $1/\chi_{s,p}$ ).*

Figure 4.2: In idealized diagram for the theoretical basis of the phase boundaries. The binodal curve is built from the Gibbs free energy extrema at different temperatures or Flory parameter values. The spinodal occurs within the binodal, and is found from the inflection points. The critical point is the meeting point of the binodal and spinodal.

Figure 4.3: The relationship between free energy and the phase boundaries (spinodal in red, binodal in blue) in systems where individual polymers have a higher volume ( $x$ , degree of polymerization) than the solvent clusters. The dashed vertical line and  $\phi_1$  are for illustrative purposes and explained in the text. This diagram too is exaggerated, in reality the phase boundaries stretch from near zero to near total polymer content.

Figure 4.4: The positions of the various positions highlighted in Figure 4.3 at the point  $\phi_1$ . The second derivative is negative from within the spinodal condition until the spinodal boundary at the minimum; from the spinodal the second derivative is positive, representing the metastable region between the spinodal and binodal.

Figure 4.5: Gibbs energy of mixing over several polymer fractions with different Flory parameters;  $\partial\Delta G_m / \partial xN$  is observed to increase with greater  $\chi_{s,p}$  as the lower approaches the inflection. This graph and the following were originally published in *Polymer Chemistry* (Cowen, 2018).<sup>308</sup>

Figure 4.6: The relationship between particle diameter and  $\Delta\chi_{spinodal}$ , where  $\Delta\chi_{spinodal} = \chi_{s,p} - \chi_{spinodal}$ . At this  $xN$ ,  $D(nm) = -74\Delta\chi_{spinodal} + 367$  nm.

Figure 4.7: A possible explanation for the observations, showing insolubility of developing particles as a function of time,  $x/N$  and final particle diameter. The particle grows until it reaches  $\chi_{s,p}$ , the plateau of the sigmoid curve of insolubility. Separation occurs at the spinodal or from the binodal. The shape of the sigmoid curve, and thus the nanoparticle diameter, depends on  $\Delta\chi$  (the position of  $\chi_{s,p}$  relative to the spinodal or binodal).

*Figure 4.8: The relationship between particle diameter, fraction (expressed in values of  $xN$ ) and  $\Delta\chi_{spinodal}$ . The trend line shown is given by the equation in the Figure, but can be approximated to  $D(\text{nm}) = 173\ln[(xN)^2 10^{-36}/\Delta\chi_{spinodal}] - 193$  nm.*

*Figure 4.9: Simple diagram explaining the observed relationship between polymer MSPU solubility and nanoparticle diameter.*

*Figure 5.1: States described by the current research; a) relatively low affinity monomers; b) the oligomer in equilibrium between template and solvent; c) the final imprinted polymer.*

*Figure 5.2: Hypothetical mechanism of imprinted polymer synthesis (cross-linkers are omitted for schematic simplicity): a) Monomers are semi-randomly arranged in solution, and interactions between monomer and template are short lived; b) On polymerization the monomer units are arranged through the polymer largely without consideration for the template; c) As the polymer develops the affinity for the solvent is reduced and template binding occurs more frequently, with the random polymer being arranged into the most favorable arrangement for template binding.*

*Figure 5.3: Values for the Gibbs energy of solvation as calculated with the script given in Appendix 3.4, and empirical values taken from the literature.<sup>317</sup>*

## Symbols, abbreviations and terminology

#	Cross-linker
$\delta$	Solubility parameter; a value describing the chemical properties of a material for the study of solubility. This can be divided into parameters of dispersive forces, $\delta_d$ , polarity, $\delta_p$ , and hydrogen bonding, $\delta_h$ .
$\Delta$	Delta is used in various contexts to refer to the difference between two numerical values, giving another single value.
$\Delta\delta$	The difference in solubility parameter, inversely proportional to the solubility of the relevant materials in each other.
$\Delta\chi$	$\Delta\chi_{spinodal} = \chi_{s,p} - \chi_{spinodal}$ $\Delta\chi_{binodal} = \chi_{s,p} - \chi_{binodal}$
$\Delta G_m$	Gibbs free energy of mixing
$\Delta H_m$	Enthalpy of mixing
$\Delta S_m$	Entropy of mixing
$\varepsilon$	Maximal internal interaction energy between two molecules in the Lennard-Jones potential.
$\sigma$	The distance at which $U_{1,2} = 0$ and $U = U_\infty$ in the intermolecular potential calculation, as the short range interactions cross from attractive to repulsive.
$\chi_{binodal}$	The Flory parameter corresponding to the binodal at a given polymer fraction
$\chi_{s,p}$	The Flory parameter for a solvent and polymer
$\chi_{spinodal}$	The Flory parameter corresponding to the spinodal at a given polymer fraction
AAc	Acrylic acid

<i>Ab initio</i>	“From the beginning”; techniques which start from fundamental principles of quantum mechanics. Typically refers to Hartree-Fock and related levels of theory in which individual atomic orbitals are calculated and built into molecular orbitals.
AIM	Atoms in molecules theory, or the quantum theory of atoms in molecules (QTAIM); structure can be determined from the electron density of the molecule or molecules.
APS	Ammonium persulfate
APTMS	(3-Aminopropyl)trimethoxysilane
ATRP	Atom-transfer radical polymerization
B3LYP	Becke 3-parameter Lee-Yang-Parr; commonly used hybrid functional combining Becke’s three-parameter exchange functional with Lee, Yang and Parr’s nonlocal correlation functional . These functions relate to more accurate representations of the spin interactions of electrons.
Basis set	A set of functions used to create a representation of the molecular electron distribution. Those used in the papers reviewed are all split valence or ‘Pople’ basis sets (see Section 1.3).
Bis	<i>N,N'</i> -methylenebis(acrylamide)
Cross-linker	Monomers used to give greater rigidity to a MIP. Screening of cross-linkers is relatively unusual compared to that of functional monomers. Often also given as ‘crosslinker’ in the literature.
DDS	Drug delivery system
DFT	Density functional theory. DFT treats molecular orbitals as continuous bands and treats electron density as fundamental property from which calculations can be performed, as opposed to HF which must account for all particles.
EGDMA	Ethylene glycol dimethacrylate, a common cross-linker in MIP synthesis.

FHT	Flory-Huggins Theory
Force field	A set of parameters used to describe and refine the energy of an arrangement of atoms in molecular mechanics and molecular dynamics.
Functional monomer	Monomers used to maximize the electrostatic interactions between the target molecule and the MIP binding site. Screening of functional monomers is a typical use of computational methods in MIP design.
GAFF	General Amber force field; a molecular mechanics force field available primarily in the Amber programs.
<i>H</i>	Enthalpy
HF	Hartree-Fock; an <i>ab initio</i> technique based on the self-consistent field method of repeating calculations with improving values for orbital energies and their coefficients until self-consistent.
<i>In silico</i>	'In silicon'; performed via computer simulation.
Leapfrog	A program available within the Sybyl software used to automatically screen a library of functional monomers. Often stylized as 'LeapFrog' or 'LEAPFROG'.
M	Functional monomer
Macroscopic	Here this term will apply to observations on the scale of millimeters or larger.
MD	Molecular dynamics; a method of simulating the change of a system through time based on MM, allowing the consideration of thermodynamic and related physical effects.
Microscopic	Here this term will apply to observations on the scale of nanometers (typically 0.1-10 nm).
Mesoscopic	Here this term will apply to observations on the scale of high nanometers to micrometers (At a low of approximately 10 nm, typically from 100 nm).

MIP	Molecularly imprinted polymer; a typically organic polymer synthesized with a template molecule resulting in a selective 'imprint' binding site.
MM	Molecular mechanics; empirically based atomistic models used to predict the energies of different molecules, allowing structural predictions, interaction energies, etc.
NAPMA	N-(3-Aminopropyl)methacrylamide
NiPAm	<i>N</i> -isopropylacrylamide
Particle	The word 'particle' is used throughout, and often refers to very different things. In the discussion of QM this will more commonly refer to subatomic particles, and in MD refers to molecules. In most cases 'particle' will refer to macroscopic materials, in most cases can be used interchangeably with 'nanoparticles'.
PCM	Polarizable continuum model; the application of a dielectric constant (or similar) across the observed system to replicate solvent effects.
NBO	Natural Bond Orbital – NBO theory is an approach to calculating electron density. NBO charge transfer has been used to predict the strength of electrostatic interactions.
NP	Nanoparticle
QM	Quantum mechanics, i.e. those computational approaches which use principles of quantum chemistry in calculating molecular geometries, electron density distribution, etc.
$r_{1,2}$	Intermolecular distance
RDF	Radial distribution function; used within MD simulations, RDF gives the relative length of time that a particular distance (the density) was observed between one atom and another atom or group of atoms.
RHF	Restricted Hartree-Fock
$S$	Entropy

Semi-empirical	QM techniques which use <i>ab initio</i> methods to replicate valence electronic structure and empirical parameters for the core electrons. This has benefits in computational expense but has recently fallen out of favor.
T	Template/Target
TBAm	<i>N-tert</i> -butylacrylamide
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
Template/ Target	The template is the molecule used to form the MIP imprint, the target is that for which the MIP is being prepared. This is usually but not necessarily the same compound.
$U$	Internal Energy, equal to the sum of potential and kinetic energy. In many applications of molecular mechanics kinetic energy is zero, and internal and potential energy are equivalent.
$U_{1,2}$	Internal interaction energy between two molecules
$U_{\infty}$	Internal interaction energy between two molecules at infinite distance, therefore the total intramolecular potential energy of the two molecules.

## Publications

1. T. Cowen, K. Karim, S. A. Piletsky, 'Solubility and size of polymer nanoparticles', *Polymer Chemistry*, 2018, **9**, 4566-4573.
2. J. Settipani, K. Karim, A. Chauvin, S. M. Ibnou-Ali, F. Paille-Barrere, E. Mirkes, A. Gorban, L. Larcombe, M. J. Whitcombe, T. Cowen, S. A. Piletsky, 'Theoretical aspects of peptide imprinting: screening of MIP (virtual) binding sites for their interactions with amino acids, di- and tripeptides', *Journal of the Chinese Advanced Materials Society*, 2018, DOI: 10.1080/22243682.2018.1467279.
3. A. Florea, T. Cowen, S. Piletsky, K. De Wael, 'Polymer platforms for selective detection of cocaine in street samples adulterated with levamisole', *Talanta*, 2018, **186**, 362-367.
4. A. Motib, A. Guerreiro, F. Al-Bayati, E. Piletska, I. Manzoor, S. Shafeeq, A. Kadam, O. Kuipers, L. Hiller, T. Cowen, S. Piletsky, P. W. Andrew, H. Yesilkaya, 'Modulation of Quorum Sensing in a Gram-Positive Pathogen by Linear Molecularly Imprinted Polymers with Anti-infective Properties', *Angewandte Chemie International Edition*, 2017, **52**, 16555-16558.
5. K. Karim, T. Cowen, A. Guerreiro, E. Piletska, M. J. Whitcombe, S. A. Piletsky, 'A protocol for the computational design of high affinity molecularly imprinted polymer synthetic receptors', *Global Journal of Biotechnology and Biomaterial Science*, 2017, **3**, 1-7

6. T. Cowen, M. Busato, K. Karim, S. A. Piletsky, 'In silico synthesis of synthetic receptors; a polymerization algorithm', *Macromolecular Rapid Communications*, 2016, **37**, 2011-2016.
  
7. T. Cowen, K. Karim, S. Piletsky, 'Computational approaches in the design of synthetic receptors – a review', *Analytica Chimica Acta*, 2016, **936**, 62-74.
  
8. K. Karim, L. Giannoudi, E. Piletska, I. Chianella, O. Y. F. Henry, P. Laitenberger, S. A. Piletsky, T. Cowen, 'Development of MIP sensor for monitoring propofol in clinical procedures', *Journal of the Chinese Advanced Materials Society*, 2015, **3**, 149-160.

*Some of the contents of this thesis were presented at the 9<sup>th</sup> (oral, Chapter 3) and 10<sup>th</sup> (poster, Chapter 4) International Conference on Molecular Imprinting. Work related to that in Chapter 2 has been exhibited elsewhere.*

# 1. Introduction

*It would be so much better if we called most of them models rather than theories. Then we would not have to defend their truth but only their usefulness. Most of us do need models in order to think.*

- George Scatchard<sup>1</sup>

It is common in contemporary communication of science to refer to the information gained in scientific inquiry as truth, the suggestion being that a particular discovery uncovers another small piece of the reality hidden beneath ignorance. It is arguably more useful to consider scientific theories however simply as models, allowing the user of the model to predict outcomes and explain observations. Whether all science progresses via punctuated equilibrium style paradigm shifts as described by Thomas Kuhn or not, history suggests that a theory is initially produced based on existing evidence and developed through application, until the evidence no longer supports the theory and an alternative is proposed. While this could fairly be considered a process of working towards a greater understanding of underlying reality, and in some examples the minute developments suggest a close approximation to the actual natural structure, to consider a particular theory 'true' suggests that those preceding it are false. That a particular model of a natural phenomenon can potentially switch in some short time from being a useful, predictive and true representation to one that is false in the light of a new theory is unsupported in all but the most extreme events.

Examples can be given in the analysis of elemental mercury and the planet Mercury. The model of planetary orbit derived from Newtonian laws of motion allows predictions of all the major bodies of the solar system except Mercury, which required input from relativity for accurate description. Similarly, elemental mercury from quantum mechanical calculations would be predicted to be solid under ambient conditions, and only predicted to be liquid when core electron relativistic effects are accounted for (this is discussed in more detail in Section 1.4.). While such examples

demonstrate the limitations of a particular model, Newtonian mechanics and non-relativistic quantum mechanics cannot be considered incorrect, but only limited. For certain examples a theory may become so removed from the modern state of science that its inclusion causes more confusion than benefit (phlogiston theory, geocentrism, spontaneous generation, etc.) even if the antiquated model is effectively predictive in certain experiments. For many other examples however this is not the case, as demonstrated by the continued use of Newtonian mechanics where it has benefits over those models that may be viewed as superseding it.

This perspective may be helpful in contextualizing the two major research areas described in this thesis. The first, developed in Chapter 2 and Chapter 3, describes the application of a model to aid technological development and analysis. While this work is more of an application of science than a science in itself, it functions as a technique to potentially facilitate scientific advancement. While applying an existing model of nature (classical mechanics) the work describes the design of a technique which provides an original model of a polymer network, 'original model' in the sense that such a representation has not been produced before, and a method of synthesizing such a structure. This work is a response to the standard approach to molecular modelling, described from Section 1.2., and a development from those approaches described in Chapter 2.

This molecular modelling was performed explicitly to build more appropriate models of polymers for technological applications. The second research area however presents an attempt to describe an aspect of nature, while having some possible useful application as a convenient by-product. This research is described in Chapter 4, with some attempts at further development of the model being given in Chapter 5, where it is aided by the techniques described in Chapters 2 and 3. This work was in many ways more ambitious than the previous in that it represents a proposed contribution to polymer science as a whole, drawing on theory from various sub areas to describe the properties of polymer nanoparticles from their synthetic conditions.

Chapter 5 then attempts to tie together all the previously described work, thus achieving an end point to the project. The modeling techniques described in Chapter 2 will be applied with the polymerization algorithm of Chapter 3, with the goal of

improving the natural model of polymer nanoparticles explained in Chapter 4. Chapter 5 therefore demonstrates an attempt to combine the two major research areas by developing more accurate descriptions of synthesis conditions with molecular modeling, thus applying microscopic analysis to mesoscopic properties. This section also includes some discussion of the mechanisms of polymer synthesis implied by the previous chapters and their possible implications for future modeling.

Section 1.1 however begins with an explanation of the theory underlying molecularly imprinted polymers (MIPs), and leads to a discussion through the rest of Chapter 1 of the predominant (quantum mechanical) molecular models currently used in the design and analysis of imprinted polymers.

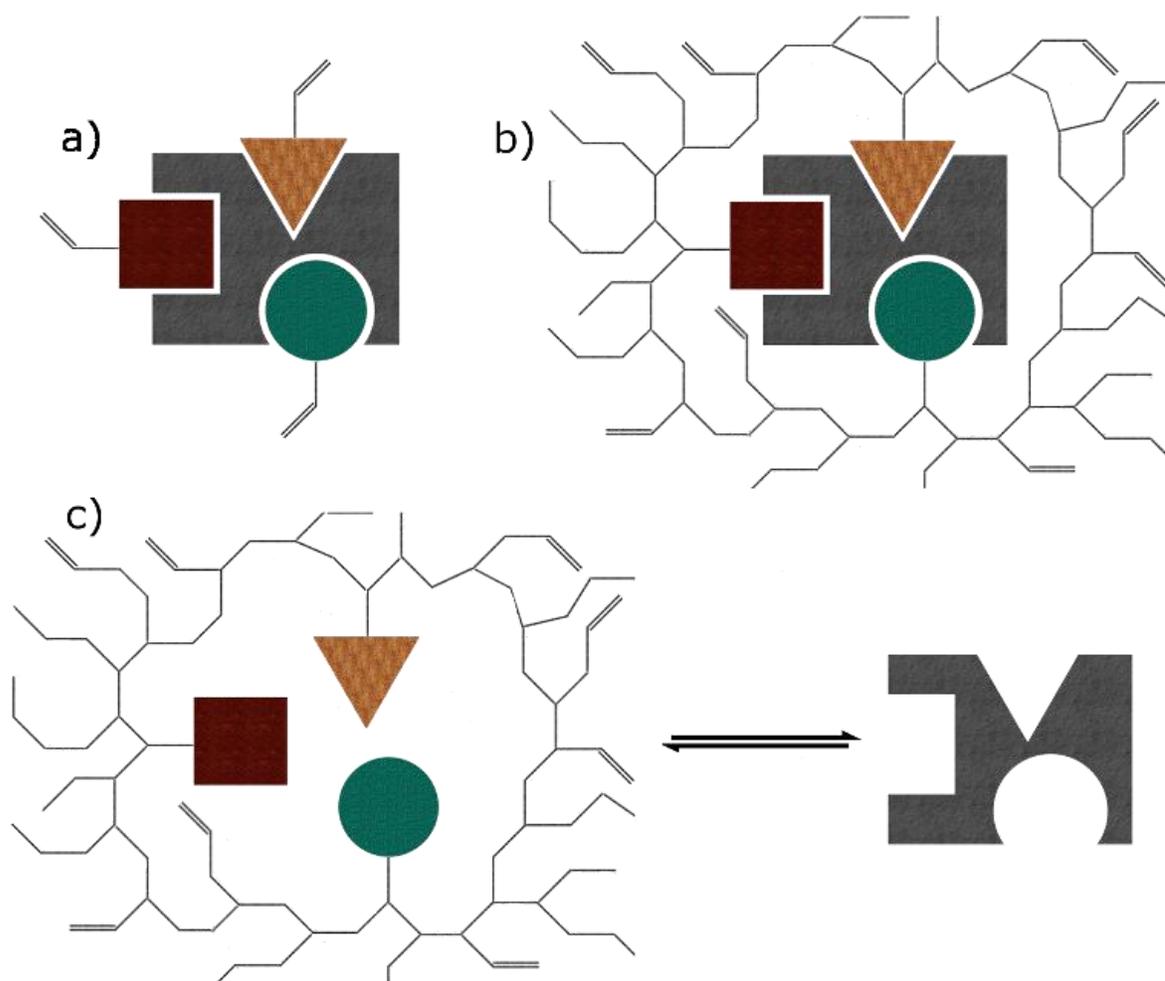
### *Objectives by chapter*

1. Provide background information on MIPs and the quantum mechanical methods used in their design and analysis.
2. Introduce molecular mechanics as an alternative, giving original examples of how they can be applied to MIPs.
3. Develop a polymerization algorithm as an improvement to the existing techniques in MIP analysis.
4. Describe the relationship between polymer solubility and polymer nanoparticle hydrodynamic diameter.
5. Investigate the feasibility of applying the polymerization algorithm to improve the description of polymer solubility.

## *1.1. Molecularly imprinted polymers*

Molecular imprinting is a process in which a polymer is synthesized in the presence of some template molecule, evidently leaving an imprint cavity within the polymer that will retain, or can readily revert to, a complementary chemical structure on removal of the template (Figure 1.1). If successfully formed the cavity and the template will inherently exhibit affinity as a result of the area and/or charges of interaction between the two surfaces. Synthesis may require the presence of a monomer-template complex, resulting from favorable monomer-template interactions, which can then be incorporated into the polymer as it develops. Specialized functional monomers are often chosen which will form strong non-covalent bonds with the template, resulting in a more stable complex and stronger interactions between the template and polymer. On removing the template the cavity remains, or can be induced, and so acts as a binding site for the molecule originally used as the template or some analogue of the template depending on the requirements.

The invention of molecular imprinting is commonly attributed to M. V. Polyakov in 1931, for his publication documenting the observed differences in rebinding simple aromatics (benzene, toluene or xylene) between silica dried in the presence of each of these compounds.<sup>2,3</sup> For various reasons, primarily that Polyakov published his work in Russian, the first research into MIPs has also been attributed to Frank Dickey in 1949.<sup>4</sup> Dickey's work similarly involved the preparation of silica gels in the presence of methyl orange and its homologues, the resultant specific adsorbents showing selectivity for their specific template and greater adsorption than equivalent control gels.<sup>5</sup> This work is of particular interest, aside from the relatively strong effect observed, due to the concluding remarks that these adsorbents could be used for the separation of optical isomers, and that the research could lead to the development of synthetic enzymes which catalyze the reactions of specific reagents.



*Figure 1.1: The standard schematic diagram giving an overview of molecularly imprinted polymers: a) a template molecule (gray) forms a complex in solution with functional monomers via strong intermolecular interactions (shown as colored shapes fitting the cavities of the template); b) polymerization is induced with the complex retaining its strong interactions with the functional groups; c) On template release a cavity remains with a geometry and functionality suitable for rebinding the template.*

Dickey based these presumptions on the work of Linus Pauling. Pauling's investigations into antibody production led him to conclude that the biomolecules are formed around an antigen (the antigen acting as template) thus giving the typical high affinity interactions, but also that, in contrast to previous hypotheses, that the amino acid sequence of the primary protein structure is the same in all antibodies, and that affinity derives from the functionalities of the peptide chain being arranged appropriately around the antigen.<sup>6</sup> Pauling's model was obviously attractive to Dickey as his own research suggested that high affinity adsorbents could be made for different target compounds by varying only the template.

The implementation of selected monomers with strong affinities arising from electrostatic interactions did not occur until much later, though the theory developed through the mid-twentieth century. In 1955, Haldeman and Emmett validated the previous proposals regarding the formation of micropores in silica suited to the properties of the template, and first used the word 'imprint' (interchangeably with 'footprint') in the context of polymeric adsorbents.<sup>4</sup> More typical organic molecular imprinting was developed independently by two different research groups with publications in 1972: Wulff and Sarhan, who produced polymerizable compounds containing the chiral target of interest, allowing racemic solutions to be easily separated (i.e. covalent imprinting, Figure 1.2),<sup>7</sup> and Takagishi and Klotz, who published their research with a methyl orange imprinted polymer, synthesized by cross-linking polyethyleneimine in either the presence or absence of the target compound,<sup>8</sup> the real significance of this work being the incorporation of hydrophobic groups into the otherwise hydrophilic polymer, leading to significantly enhanced affinity. Both works therefore center on the direct interactions between functional groups of the polymer and target, allowing greater control over the strength of interactions in the binding site.

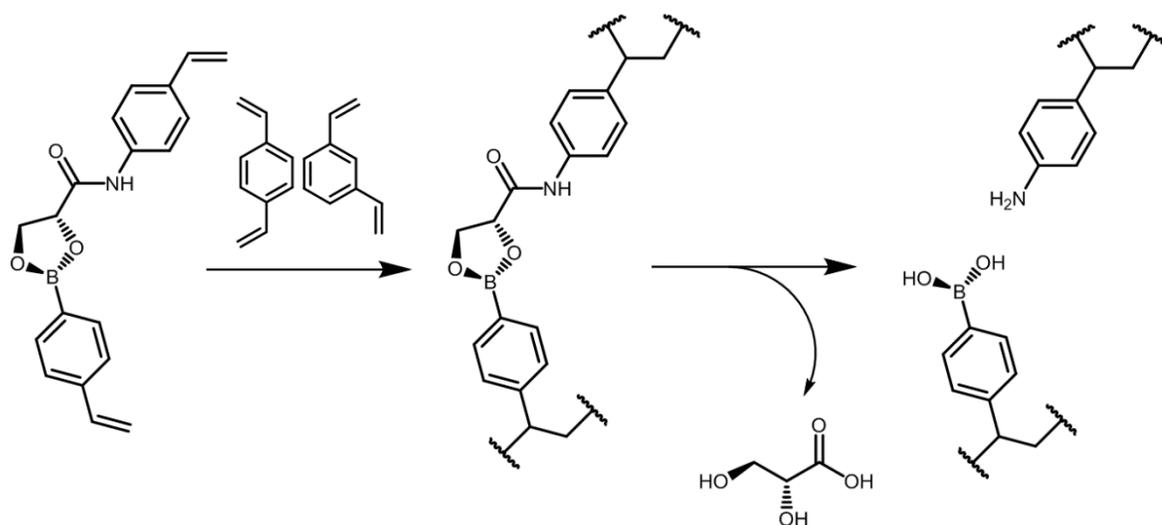


Figure 1.2.: Synthesis of Wulff and Sarhan's D-glyceric acid imprinted polymer. The target is covalently bound to the monomer and eventual polymer before being removed, leaving a specific binding site. This method is referred to as covalent imprinting.

Aside from the work of a small number of dedicated researchers notably Mosbach and his associates, who produced several articles documenting their progress using MIPs as sensors and for chromatographic separation (a technique which developed into molecularly imprinted solid-phase extraction, MISPE<sup>9</sup>),<sup>10-14</sup> relatively little attention was given to molecularly imprinted polymers until the latter half of the 1990s. The availability of advanced spectroscopic techniques, in particular NMR, allowed functional monomers selection and relative ratio to be used in the design of MIPs, with good results in terms of template specificity.<sup>15</sup> Polyakov concluded his original work by suggesting that the adsorbents he was investigating may be useful in studying the structure of the molecules used as templates; his proposed application was rendered obsolete by the same technologies that validated his discovery.<sup>3</sup>

The number of groups actively working on MIPs doubling between 1996 and 1999,<sup>16</sup> and this time period correspondingly coincides with the first MIP nanoparticles (NPs),<sup>17, 18</sup> and the first references to MIP 'plastic antibodies'.<sup>19</sup> Catalytic synthetic enzymes had been previously demonstrated,<sup>20</sup> and MIP-based assays (in place of antibodies or enzymes) were in the early stage of development.<sup>21, 22</sup> Much

contemporary research has focused on the creation of MIP-based ELISA (enzyme-linked immunosorbent assay) style analytical methods.<sup>23, 24</sup> This technique has been used to both select appropriate MIP nanoparticles (by having the target compound screened against a library of affinity particles) and analyze samples for a specific target using immobilized imprinted polymers. This technique has for some time been referred to as MIA (molecularly imprinted assay)<sup>25</sup> or more recently MINA (molecularly imprinted nanoparticle assay)<sup>26</sup> to emphasize that they exist as an alternative to and not simply a replica of an existing biomolecule-based technique.

Antibodies have been used to great effect in the treatment of a range of diseases and poisonings, as well as in immunoassays for biochemical diagnostics, but they are unstable outside of the environment to which they are adapted, expensive in time and resources to produce, and difficult to create for a specific target.<sup>27, 28</sup> The paratope (binding site) of an antibody is analogous to that of a MIP, leading to the potential for synthetic antibodies which exhibit greater diversity of application, with hugely reduced costs, via molecular imprinting technologies. Natural antibodies typically cost between \$100 and \$1000 per milligram, whilst equivalent MIPs could be available for \$0.1 per mg, and demonstrate stable functioning in organic solvents, high and low pH, temperatures above 150 °C and in extensive reuse, as well as being easily prepared compared to the relatively crude method of biological antibody production.<sup>29</sup> However, the real benefit of MIPs over antibodies will likely be in binding small organic molecules (typical of many drugs, poisons and explosives), for which antibodies can often not be produced.

This 'plastic antibody' moniker has led MIP developers to mimic their biological equivalents to various extremes, including the production of literal antibody replicas by first imprinting a natural antibody and then using the imprinted polymer as a stencil.<sup>30</sup> These plastic antibodies were selected to bind to human rhino virus 14 (HRV 14), and showed greater sensitivity than their natural equivalents. Related research into virus imprinted polymers spans a variety of harmful species, including the tobacco mosaic virus, bovine leukemia virus, dengue virus, and HIV.<sup>31-34</sup> Paradigmatic advances in these areas are likely to result from further research into cooperative binding, in

which the affinity of one target is dependent on the presence of another, mimicking the behavior of many biological receptors and enzymes.<sup>35</sup>

Speculative bionanotechnological applications of MIP NPs however only became part of the standard description from 2010, when Hoshino et al. demonstrated the successful use of MIP nanoparticles in a live mammalian model.<sup>36</sup> The polymers were imprinted with melittin (a 26 amino acid peptide and the main component of bee venom) and administered intravenously to mice. The MIP nanoparticles were found to be nontoxic and effectively neutralized lethal doses of melittin, with the venom-MIP complex later being cleared from the blood in the liver. While the availability of cheap, thermally stable antivenom has obvious value (effectively dealing with snakebites alone could save nearly 100,000 lives per year globally<sup>37</sup>) the demonstration of peptide sequestration has huge potential throughout the biological sciences.

Related applications of MIPs center on their potential as drug delivery systems. Simpler examples include the use of macroscopic objects which can be directly applied to the point of delivery with a controlled release, such as contact lenses imprinted with ophthalmic drugs.<sup>38</sup> More sophisticated design is required for the application of MIP NPs, where the polymer may be imprinted with some relevant agent to be released inside the body; examples include the suppression of tumors in mice using magnetic 5-fluorouracil imprinted NPs, where the drug is released after being guided to the cells with magnets.<sup>39</sup>

The particle may alternatively be imprinted using an epitope corresponding to a particular target, the MIP NP therefore requiring no external guidance.<sup>40</sup> This technique has previously been applied in cancer cell imaging, and could be used in a variety of other situations.<sup>41</sup> Successful binding to the target may have some biological effect in itself, but more common in the context of cancer suppression and similar research is the reliance on drug release on binding.<sup>42</sup> The polymer is preloaded via synthesis in a solution of the relevant agent and the release profile monitored. On target binding the agent may be released as a result of environmental changes (e.g. pH decrease in the presence of cancer cells) or as a direct result of binding. These interactions rely on a phenomenon referred to as the gate effect, where polymer-target binding induces structural changes in the polymer as a whole, increasing permeability and releasing the

drug.<sup>43</sup> This effect is also used in many MIP-based sensors, where a polymer coating an electrode displays variable permeability dependent on the quantity of bound target.

All these examples of MIP research rely heavily on successful synthesis, which is dependent on effective design. General polymerization conditions have been previously determined through experimental trial and error, resulting in standard protocols which can be broadly applied for MIP nanoparticle synthesis.<sup>44</sup> However, new synthetic techniques are always evolving,<sup>45</sup> and for many applications attempts must be made to ensure high affinity and selective binding sites are formed, which require more specific tailoring of the materials and conditions of synthesis.

## *1.2. Computational approaches in the design of MIPs*

The work of Wulff and other described previously demonstrated the significance of including functional groups into the polymer binding cavity, as opposed to earlier methods employing silica which rely on only hydrophobic siloxane and hydrophilic silanol interactions.<sup>46</sup> Constructing a polymer combining a cross-linked polymer matrix with functional groups with specific affinity for the target could increase both affinity and selectivity, with the use of functional monomers allowing systematic variation as required for the analyte. With this new protocol a number of different functional monomers were then available as options to be included in the prepolymerization mixture, which necessarily required some screening for selection of that which would be most appropriate. Monomer screening can be performed by a variety of methods, but the relative efficiency of molecular modeling has made it particularly popular for this purpose.

While the use of simple molecular modeling to visualize template-monomer interactions was a practice observed at the time,<sup>47</sup> adoption of computational methods explicitly for the purpose of MIP design first appeared in the late 1990s.<sup>48-50</sup> Studies of

the interaction between dibenzothiophene sulfone (a pollutant in crude oil) and six different monomers, with varying ratios of the latter, were presented at a conference in 1998.<sup>48</sup> This work was not formally published until 2001, but the basic methodology has remained largely unchanged to the present day in routine monomer screening, though the calculations have been somewhat simplified and the specific modeling technique (semi-empirical AM1) has fallen out of favor (see Table 1.1).<sup>49, 51</sup> The computational approach to rational design employing molecular modeling however was only widely acknowledged in 2001, with a publication describing the use of an automated screening technique using a relatively large library of functional monomers.<sup>52</sup> This paper demonstrated that good predictions could be generated without recourse to expensive empirical methods such as combinatorial screening, which, although highly efficient compared to traditional methods,<sup>53</sup> would still require an unreasonable amount of time and resources to replicate the data obtained from computational modeling.

While the number of researchers focused primarily on the theoretical element of MIP design remains relatively low, computational approaches have become well established in the years since this publication. Table 1.1 gives an overview of the molecular modeling techniques applied in imprinted polymer research. As the vast majority of this research involves applying quantum mechanical methods, an overview of these methods and the resulting applications will be given to further demonstrate the current state of the area. From there considerations will be given regarding the suitability of these models in designing and analyzing MIPs.

Table 1.1: Compilation of the use of computational approaches in MIP design 2013-2016, covering work published after the last major review of this area.<sup>54</sup> References are given according to the target selected, the computational technique applied in the modeling, and the manner in which the models were used. This table was previously published in *Analytica Chimica Acta* (Cowen et al., 2016).<sup>51</sup>

Application	Method	Targets
<i>Functional monomer screening</i> <i>(* using screening/docking program)</i>	MM/MD	*Curcumin, <sup>55</sup> *fenthion, <sup>56</sup> *methidathion, <sup>57</sup> *propofol, <sup>58</sup> *amlodipine, <sup>59</sup> *endotoxins, <sup>60</sup> *cocaine, <sup>61</sup> thymopentin, <sup>62-65</sup> diuron, <sup>66</sup> *metoprolol, <sup>67</sup> *paracetamol (modified), <sup>68</sup> iprodione, <sup>69</sup> 5-(3,5-Dichloro-2-hydroxybenzylamino)-2-hydroxybenzoic acid, <sup>70</sup> naproxen, <sup>71</sup> biotin, <sup>72</sup> hexazinone, <sup>73</sup> 1-(2,4-Difluorophenyl)-2-(1H-1,2,4-triazol-1-yl)ethanone. <sup>74</sup>
	Semi-empirical	Metaproterenol, <sup>75</sup> diuron, <sup>66</sup> ciprofloxacin, <sup>76</sup> baicalein, <sup>77</sup> Sulfamethizole, <sup>78</sup> 4-(2-Aminoethyl)aniline, <sup>79</sup> hexazinone, <sup>73</sup> theophylline, <sup>80</sup> cotinine. <sup>81</sup>
	Ab initio	Baicalein, <sup>82</sup> phenothiazine, <sup>83</sup> phenol, <sup>84</sup> acephate, <sup>85</sup> phenazopyridine, <sup>86</sup> fusaric acid, <sup>87</sup> uracil, <sup>88</sup> 5-fluorouracil, <sup>88</sup> metformin, <sup>89</sup> tanshinone IIA, <sup>90</sup> pantoprazole. <sup>91</sup>
	DFT	(S)-Warfarin, <sup>92</sup> hydrochlorothiazide, <sup>93</sup> Atrazine, <sup>94</sup> sulfanilamide, <sup>95</sup> carbofuran, <sup>96</sup> fenitrothion, <sup>97</sup> 2,3,7,8-Tetrachlorodibenzo-p-dioxin, <sup>98</sup> butylphthalide, <sup>99</sup> mesalamine, <sup>100</sup> enrofloxacin, <sup>101</sup>

		1,4-dihydroxyanthraquinone, <sup>102</sup> ractopamine, <sup>103</sup> acephate, <sup>85</sup> Sulfamethizole, <sup>78</sup> dopamine, <sup>104</sup> $\Delta^9$ -tetrahydrocannabinol, <sup>105</sup> 11-nor-9-carboxy- $\Delta^9$ -tetrahydrocannabinol, <sup>105</sup> tramado <sup>106</sup> 5-(3,5-Dichloro-2-hydroxybenzylamino)-2-hydroxybenzoic acid, <sup>70</sup> 6-thioguanine, <sup>107</sup> metformin, <sup>89</sup> melamine, <sup>108-110</sup> spermidine, <sup>111</sup> epinephrine, <sup>112</sup> quinoline, <sup>113</sup> triamterene. <sup>114</sup>
<i>Advanced structural analysis / ratio optimization</i>	MM/MD	Curcumin, <sup>55</sup> amlodipine, <sup>59</sup> phosmet, <sup>115</sup> estrone, <sup>115</sup> metolcarb, <sup>115</sup> enrofloxacin, <sup>115</sup> tetracycline, <sup>116</sup> tyramine, <sup>117</sup> octopamine, <sup>118</sup> paracetamol (modified), <sup>68</sup> butylated hydroxyanisole, <sup>119</sup> bupivacaine, <sup>120-122</sup> norfloxacin, <sup>123</sup> dibenzothiophene, <sup>124</sup> 4-nitrophenol, <sup>125</sup> bisphenol A, <sup>126</sup> phenylalanine, <sup>127</sup> 5-(3,5-Dichloro-2-hydroxybenzylamino)-2-hydroxybenzoic acid, <sup>70</sup> naproxen, <sup>71</sup> copper(II), <sup>128</sup> caffeine, <sup>129</sup> theophylline, <sup>129</sup> 1-(2,4-Difluorophenyl)-2-(1H-1,2,4-triazol-1-yl)ethanone, <sup>74</sup> bupivacaine, <sup>121, 122</sup> (S)-propranolol, <sup>130</sup> 1,2,3-trichlorobenzene. <sup>131</sup>
	Semi-empirical	Metaproterenol, <sup>75</sup> phosmet, <sup>115</sup> estrone, <sup>115</sup> metolcarb, <sup>115</sup> enrofloxacin, <sup>115</sup> tetracycline, <sup>116</sup> ciprofloxacin, <sup>76</sup> sulfamethizole, <sup>78</sup> uracil, <sup>88</sup> 5-fluorouracil, <sup>88</sup> aspartame, <sup>132</sup> theophylline, <sup>80</sup> erythromycin, <sup>133</sup> sulfadiazine. <sup>134</sup>
	Ab initio	Tetracycline, <sup>116</sup> phenol, <sup>84</sup> acephate, <sup>85</sup> metformin, <sup>89</sup> tanshinone IIA, <sup>90</sup> cotinine, <sup>81</sup> sulfadiazine. <sup>134</sup>
	DFT	Metaproterenol, <sup>75</sup> uric acid, <sup>135</sup> (S)-warfarin, <sup>92</sup> tyramine, <sup>117</sup> fenitrothion, <sup>97</sup> minoxidil, <sup>136</sup>

		<p>Famciclovir,<sup>137</sup> melamine,<sup>108-110</sup> cichoric acid,<sup>138</sup> 2,3,7,8-Tetrachlorodibenzo-p-dioxin,<sup>98</sup> deltamethrin,<sup>139</sup> enrofloxacin,<sup>101</sup> barbital,<sup>140</sup> ractopamine,<sup>103</sup> 5-fluorouracil,<sup>141</sup> acephate,<sup>85</sup> sulfamethizole,<sup>78</sup> carnosine,<sup>142</sup> butylated hydroxyanisole,<sup>119</sup> dopamine,<sup>104</sup> <math>\Delta^9</math>-tetrahydrocannabinol,<sup>105</sup> 11-nor-9-carboxy-<math>\Delta^9</math>-tetrahydrocannabinol,<sup>105</sup> 6-thioguanine,<sup>143</sup> flumequine,<sup>144</sup> benzothiophene sulfone,<sup>145, 146</sup> dibenzothiophene sulfone,<sup>145, 146</sup> 4,6-methylidibenzothiophene sulfone,<sup>145, 146</sup> erythromycin,<sup>147</sup> tramadol,<sup>106</sup> <math>\alpha</math>-amanitin,<sup>148</sup> propranolol,<sup>149</sup> salbutamol,<sup>150, 151</sup> cocaine,<sup>152</sup> palmitic acid,<sup>153</sup> oleic acid,<sup>153</sup> elaidic acid,<sup>153</sup> nicotine,<sup>154</sup> 5-(3,5-Dichloro-2-hydroxybenzylamino)-2-hydroxybenzoic acid,<sup>70</sup> naproxen,<sup>71</sup> copper(II),<sup>128</sup> atrazine,<sup>155</sup> quercetin,<sup>156</sup> metformin,<sup>89</sup> pyrene,<sup>157</sup> caffeine,<sup>158</sup> spermidine,<sup>111</sup> gallic acid,<sup>159</sup> ciprofloxacin,<sup>160</sup> quinoline,<sup>113</sup> sulfadiazine,<sup>134</sup> neopterin.<sup>161</sup></p>
<i>Analysis of interactions in dynamic system</i>	Molecular Dynamics	<p>Naproxen,<sup>71</sup> (S)-propranolol,<sup>130</sup> bupivacaine,<sup>120-122</sup> norfloxacin,<sup>123</sup> dibenzothiophene,<sup>124</sup> 4-nitrophenol,<sup>125</sup> bisphenol A,<sup>126</sup> cholesterol,<sup>162</sup> 1,2,3-trichlorobenzene.<sup>131</sup></p>
<i>Comparative binding site interaction energy/polymer simulation</i>	Various	<p>Tyramine,<sup>117</sup> octopamine,<sup>118</sup> 6-thioguanine,<sup>143</sup> cholesterol,<sup>162</sup> caffeine,<sup>129</sup> theophylline,<sup>129</sup> 1-(2,4-Difluorophenyl)-2-(1H-1,2,4-triazol-1-yl)ethanone.<sup>74</sup></p>

### 1.3. Quantum mechanical methods overview and self-consistent field

Quantum mechanical (QM) methods give the most accurate description of molecular structure and interactions of the currently available modeling techniques. QM has grown significantly in popularity in recent years due to the combined progression of computational processing power and the development of more sophisticated and efficient theoretical atomic and molecular models. The widespread adoption of QM techniques in MIP design has occurred relatively recently, with almost all of the small number of groups and individuals using QM in the first decade of the new millennium applying semi-empirical models (i.e. not full QM, see Section 1.4).<sup>163</sup> Likely as a result of progression in available computer power, QM methods have grown to dominate molecular modeling of MIPs, with density functional theory (DFT) alone being used in approximately 56 % of recently published examples (Table 1.1).<sup>51</sup>

QM is primarily employed in solving Equation 1.1:

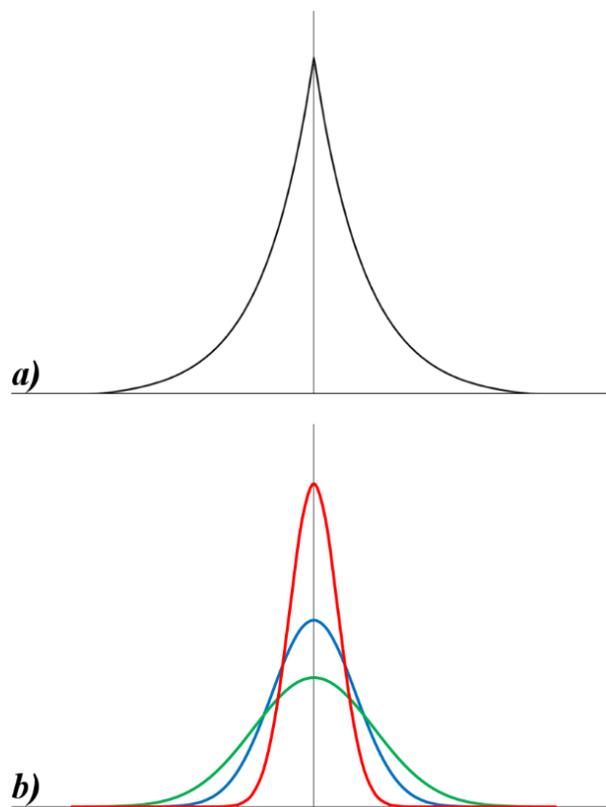
$$\Delta E = E_C - (E_T - \sum E_M) \quad \text{Eq. 1.1}$$

Which is also often the basis of methods in other techniques, if not explicitly stated. The energy,  $E$ , is typically potential energy or Gibbs free energy (by derivation), with  $E_C$  referring to that of the M-T complex,  $E_T$  that of the template, and  $E_M$  that of the monomer/s in isolation.  $\Delta E$  then gives a difference in energy between the bound complex and the individual components, i.e. an association energy, and comparison of different values of  $\Delta E$  provides a guide to monomer selection, appropriate component ratios, etc.

Calculating the energy of a molecule or complex starting from QM first principles (*ab initio*) is relatively complicated but can be explained in abbreviation to provide an understanding of the terms frequently encountered in the literature.

*Ab initio* techniques begin with one electron wave functions ( $\chi_i$ ). The one electron wave functions can be presented as Slater type orbitals (STOs, Figure 1.3a), which are good representations of a single electron in one dimension, but difficult to apply otherwise. Because of the difficulties in applying STOs they are often replicated by combining Gaussian functions, referred to as 'STO-nG basis sets', where the 'n' signifies the number of Gaussian orbitals added together to approximate the single STO (Figure

1.3b). A basis set is usually an atom centered function, resembling an atomic or molecular orbital, used to describe the electronic structure. These STO-nG basis sets are often referred to as the 'minimal basis', which is the basis set containing only the filled atomic orbitals (for example, for elements of period 2, groups 13 to 18, the minimal basis is just the 1s, 2s and 2p atomic orbitals).



*Figure 1.3: a) Slater 1s basis function (Slater type orbital, STO), a representation of the one electron wave function; b) STO-3G 1s basis function, an approximation of the STO using a combination of Gaussian functions.*

STO-nG basis sets were developed by John Pople, a Nobel laureate and major figure in computational chemistry. The minimal basis however is quite simple and does not accurately describe the orbital. For this reason, Pople then formed the split valence basis sets, which have the general notation N-XYG, where N is the number of Gaussian

functions describing the inner-shell orbitals, and X and Y the split valence orbitals (the hyphen is included to highlight the splitting of the valence shell). This is a double-zeta ( $\zeta$ ) basis set, as the valence orbitals are split into X and Y basis functions, which each denote a linear combination of Gaussian functions (there are also triple-zeta, N-XYZG, quadruple-zeta, etc., but these are less common). For example, 3-21G is a commonly used split valence basis set in which three primitive Gaussian functions are used to represent the inner orbitals and two different Gaussian functions, one a linear combination of two Gaussian functions, are used for the valence orbitals. Polarization terms, which describe the non-symmetrical electron distribution in approaching another atom, are denoted 3-21G\* or 3-21G(d) for an addition of some d character to p orbitals, or 3-21G\*\* (3-21G(d,p)) for the application of some p to d. There may also be diffuse functions which more accurately represent the portion of the orbitals far from the nucleus, shown by '+', for the exaggeration of p properties, or '++' for both p and s orbitals, e.g. 3-21+G\*. While various other kinds exist, Split-Valence (or Pople) basis sets are by far the most common in QM MIP design.

These basis sets describe atomic orbitals which must be brought together via the Linear Combination of Atomic Orbitals (LCAO) to find molecular orbitals. The molecular orbitals can then be used to calculate the molecular wave function,  $\Psi$ , when the overlap integrals are accounted for.

The molecular orbital can be represented in terms of a one-electron Fock operator,  $H\phi_i = \varepsilon_i\phi_i$ , where  $\varepsilon_i$  is the orbital energy (the combined kinetic and potential energy of an electron in orbital i), and therefore can be rewritten as:

$$H \sum_k c_{ik} \chi_k = \varepsilon_i \sum_k c_{ik} \chi_k \quad \text{Eq. 1.2a}$$

The molecular orbital being just the combination of atomic orbitals.

Integrating over all electronic coordinates and multiplying each side by  $\chi$  gives:

$$\sum_k c_{ik} \left( \int \chi_l H \chi_k dv \right) = \varepsilon_i \sum_k \left( \int \chi_l \chi_k dv \right) \quad \text{Eq. 1.2b}$$

which can be simplified with the overlap integral to:

$$\sum_k c_{ik} (H_{lk} - \varepsilon_i S_{lk}) = 0 \quad \text{Eq. 1.2c}$$

A determinantal equation is then required for the solution, maintaining the condition  $\det |H_{lk} - \varepsilon S_{lk}| = 0$ . Rearranging the equation gives values for the orbital energy

( $\epsilon$ ). As the coefficients ( $c$ ) are initially estimated, the original set of equations can be recalculated, giving more accurate values of  $\epsilon$  from the determinantal equation, which can be used to give more accurate values of  $c$ , and so on. This cycle is repeated until the results are self-consistent (results from one cycle match those previous). This process is the Self-Consistent Field (SCF) method, and it also forms an important part of the Hartree-Fock (HF) method. The purpose of this process is to achieve figures as close as possible to the Hartree-Fock limiting energy, the lowest predictable energy of a particular system.

Processing this and repeating for self-consistency requires high computational expense, but HF became very popular for producing results which conform highly with experimental data.<sup>164</sup> Despite its accuracy the pure form of HF has several problems however, and over-reliance on the method leads to difficulties in their accumulation. HF techniques neglect electron correlation, each electron being represented by an averaged charge, the 'mean field approximation'.

The pure form of the theory has gradually fallen out of favor for other reasons however, primarily due to the time typically required in running analysis on just atoms and simple molecules. Researchers working outside of areas that required highly precise electronic property calculations tended to prefer molecular mechanics or semi-empirical quantum methods for this reason, until it was essentially retired for all but the most niche areas of investigation with the refinement of Density Functional Theory. Hartree-Fock still lies at the foundations of quantum chemistry however and has been granted new dignity as a sizable contributor to hybrid functionals (see Section 1.5) which have been used extensively throughout chemical system simulations, including in the rational design of imprinted polymers.

## 1.4. *Semi-empirical methods*

*Ab initio* calculations have some flaws, the first and most significant error arising from the neglect of electron correlation. The energy of each electron is determined from its interaction with proximate nuclear and electronic charges, but the other electrons are considered to have the position probability of a lone electron, while in reality these particles are more likely to be distributed so as to increase the distance between them. For example, in an s orbital with two electrons, if one electron is on one side of the nucleus the other is more likely to be found on the opposite side, there is not an equal probability of finding the electron throughout the orbital. This creates a difference in the HF limitation energy and the true energy of the system referred to as correlation energy.

Additional deviation from the true energy may emerge as a result of relativistic effects being neglected by the Schrödinger equation, which result from the electrons close to the nucleus travelling at near the speed of light. An interesting example which demonstrates the importance of these considerations is elemental mercury, which would have a melting point approximately 105 K higher than its actual value, and therefore be solid under ambient conditions, if relativistic effects are excluded from the metallic system.<sup>165</sup>

Semi-empirical methods tend not to inherently overcome the issues described but do dramatically reduce the required computational power; even quite small molecules analyzed *ab initio* may require the processing of millions of integrals (the number of which typically increases with four orders of magnitude with an increase in basis functions). Semi-empirical methods simplify the required calculations by including empirically determined parameters to approximate results, resulting in much less computationally expensive analysis. For example, semi-empirical calculations typically use experimentally determined ionization potentials instead of the integrals used in *ab initio* methods to approximate the energy holding electrons within atoms.

J. J. P. Stewart, the creator of several semi-empirical methods, explained that while these techniques are significantly slower than molecular mechanics and less rigorous than *ab initio*, the lack of specialization makes them extremely versatile, providing

useful models in a broad range of applications.<sup>166</sup> Semi-empirical methods remain in routine use, often in combination with other techniques, and have found use in a wide variety of different areas, including the rational design of MIPs. The majority of the methods that are still used however, stem from the work of just two research groups.

Pople and co-workers developed the Complete Neglect of Differential Overlap (CNDO) method, a technique which applies Kronecker deltas (a function giving 1 if equal and 0 if not) to all valence electron functions so as to eliminate repulsive overlap calculations between inequivalent neighboring atoms. This allowed more complex systems to be analyzed without computational strain. The neglect of differential overlap style methods matured into MNDO (Modified Neglect of Differential Overlap), a method developed by Dewar and his team (which included Stewart), which showed obvious supremacy to MINDO/3 (Modified Intermediate Neglect of Differential Overlap, also created by Dewar), which was formed as an improvement upon CNDO, NDDO (Neglect of Diatomic Differential Overlap) and INDO (Intermediate Neglect of Differential Overlap, a refinement of CNDO), all of which were developed by the Pople group.<sup>167, 168</sup> However, various issues with MNDO, most importantly the absence of hydrogen bonds, prompted further development of the semi-empirical method, resulting in AM1 in the mid-1980s.<sup>169</sup> AM1 (Austin Model 1) resulted from a 'brute force approach', after gentle alteration failed, manifest in the addition of Gaussian terms to the core repulsion function.<sup>170</sup>

The rapidity with which these techniques emerged was due to the immediately obvious advantages of semi-empirical techniques. In 1985 it was shown that the newly developed MINDO/3, MNDO and AM1 methods gave results comparable to those obtained using *ab initio* calculations, but required less than a one hundredth the computer time for a single self-consistent field calculation, and one ten-thousandth the time required by 6-31G.<sup>171</sup> PM3 (Parametric Method 3, Figure 1.4) was revealed in a series of articles by J. J. P. Stewart in 1989 as MNDO-PM3, the first two parametric methods being MNDO and AM1.<sup>172, 173</sup> Further models have since been introduced, for example SAMI and PM5, but the wide-spread adoption of PM3, and to a lesser extent AM1, have left these relatively ignored, as researchers have further developed these original techniques.<sup>174</sup>

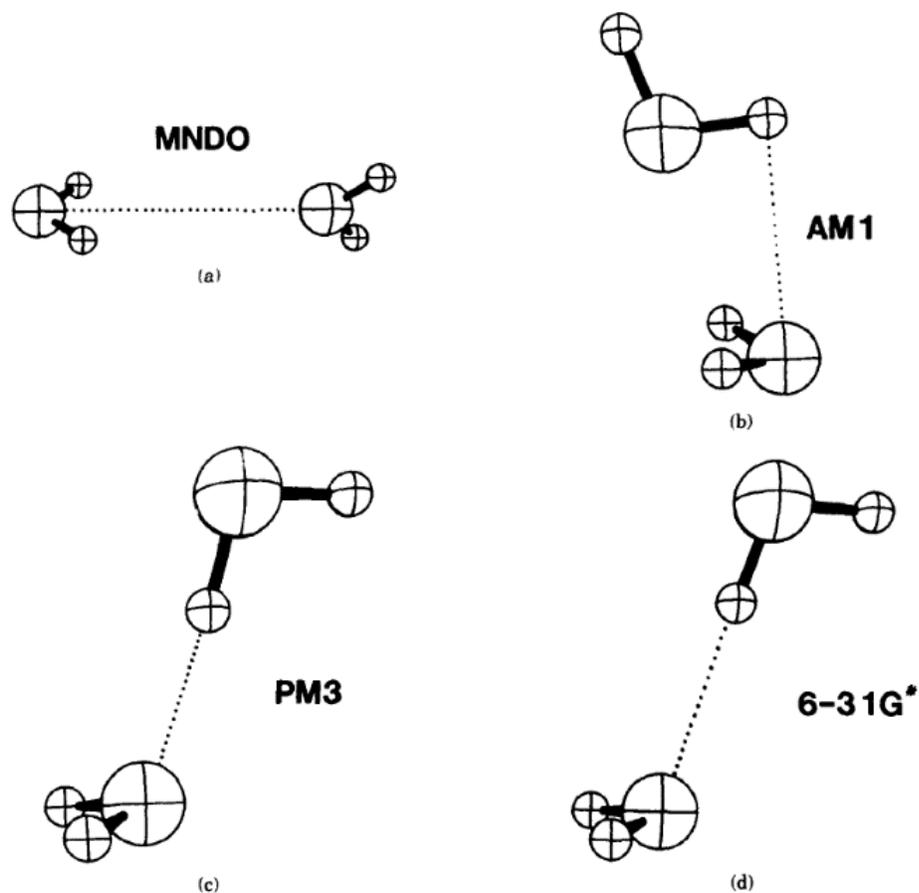


Figure 1.4: Stewart's illustrations of water dimers, demonstrating the development of the Parametric Methods (a-c) and a comparison with 6-31G\* (reproduced with permission).<sup>173</sup>

Despite the versatility of semi-empirical methods and advantages over many alternatives, these techniques no longer hold their high position in theoretical and computational chemistry. In MIP design and in the natural sciences generally, semi-empirical QM still has its role, but is no longer the only practical choice for highly detailed analysis.

## 1.5. Density functional theory

Density functional theory (DFT) has recently overtaken HF as the most widely used *ab initio* technique in theoretical chemistry. DFT considers only the electron density of a system and combines atomic orbitals into electronic bands, unlike Hartree-Fock which forms molecular orbitals, and must therefore attempt to include terms which account for the parameters for all particle interactions. Band theory regards extended areas of molecular orbitals as a continuum, allowing the same electronic structure to be represented in a much simpler fashion. In regarding the electron density (on which the functionals act; the electron density itself being a function of space and time) as a fundamental property, DFT is free to use calculations with only three or four variables instead of the coordinates of each atom, as required by HF theory. Based on the principles of the Hohenberg-Kohn theorem and its successors that all the ground-state properties of a system (including total energy, structures, vibrational frequencies, atomization and ionization energies, reaction paths, etc.) occur as functions of the electron density, DFT massively reduces the total computational expense relative to a similar scenario performed with Hartree-Fock.

DFT usually incorporates the Kohn-Sham potential as Kohn-Sham Density Functional Theory (KS DFT), which allows the free energy (as a function of electron density) of the theoretical system to be calculated exactly by more accurately describing the wave function. This too must be solved self-consistently as the effective potential depends on density distribution, but remains significantly more efficient than other *ab initio* methods. KS DFT allows precise system minimization but requires the inclusion of the Local Density Approximation (LDA) to express the exchange and correlation energies between electrons. However, while LDA can produce energy predictions within one percent of the true energy, this is not of a high enough standard for some applications. KS DFT may be further improved by introducing the Local Spin Density (LSD), which accounts for electrons more individually, or similar techniques, but it is now common to utilize a hybrid functional to overcome these problems.

Hybrid functionals combine Density Functional Theory with Hartree-Fock methods and other quantum mechanical techniques to better represent exchange and

correlation (which include Pauli exclusion effects), and other minor flaws in standard KS DFT. The first successful mixing of DFT (with LSD) and Hartree-Fock methods was described in 1993 by A. D. Becke, which improved the accuracy of property prediction whilst retaining the high efficiency of the traditional method.<sup>175</sup> In the same volume (the following year) Kim and Jordan presented the B3LYP functional, which combined Becke's three-parameter exchange functional (which describe relative contributions to this effect) with Lee, Yang and Parr's nonlocal correlation functional.<sup>176</sup> This LYP correlation calculation was used because it effectively converted the complex density matrix from Hartree-Fock into a more efficient DFT formula, allowing correlation energies to be determined from electron density, with minimal computational expense.<sup>177</sup> In a 2004 comparison of various techniques for the determination of the geometries and vibrational frequencies of transition-metal-olefin complexes, B3LYP was found to produce results closest to those that had been empirically determined, yielding better results than all other DFT based methods, which themselves performed better than both HF and MP2 (a post-HF *ab initio* method).<sup>178</sup> The authors also note that B3LYP and the DFT methods are significantly less computationally expensive, and show that MP2 'displays a more varied effect when comparing calculated results with experimental results'; in essence, that MP2 is comparatively both consistently and illogically inaccurate. B3LYP has become very common in MIP design for this reason, though studies in more direct application to the field have suggested that this technique may be less capable of reproducing real interactions than alternative, less common hybrid functionals.<sup>109, 151</sup>

Advances in both computational power and DFT have allowed much larger systems to be analyzed by non-empirical methods in recent years. Hung and Carter simulated over a million atoms using Orbital Free Density Functional Theory (OFDFT), a technique similar to Kohn-Sham Density Functional Theory (KSDF) but with one-electron wave functions removed.<sup>179</sup> This requires the kinetic energy to be approximated with Kinetic Energy Density Functionals (KEDFs), but creates a linear scaling of necessary processing in relation to system size, instead of the cubic scaling found in KSDF. Combined with the fact that *ab initio* methods have significant advantages over empirically based techniques, notably that the latter typically struggle

to describe excited states and transition states, there are few reasons to dismiss these methods on the basis of practicality.<sup>180</sup>

The emphasis that has been placed on the improvements of *ab initio* methods in relation to empirical and classical methods is not an attempt to discredit the latter two, but more to demonstrate that some of the critiques, mainly those surrounding efficiency, are increasingly unfounded. However, there are still some valid arguments for alternatives, which will be discussed after some examples are given of QM methods applied in MIP design and analysis.

### 1.6. *Examples of QM in imprinted polymer analysis*

QM techniques have been applied in the design of MIPs for countless applications, from the observation of everything from herbicide pollution to internal bleeding.<sup>181, 182</sup> Likely the most detailed example of QM methods applied in MIP design can be found in that of Khan *et al.* in their design of a 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) imprinted polymer.<sup>98</sup> In this research, the template TCDD, a library of 35 functional monomers, three different cross-linker molecules, and three commonly encountered template analogues were energetically minimized using a B3LYP/6-31+G(d,p) method. Using Equation 1.1, the functional monomers were screened against the template to determine a small number of likely candidates, which were then analyzed with a polarizable continuum model (PCM), which mimics the effects of a given solvent on the system components by the application of a dielectric constant in the surroundings. By the standard protocol this therefore allowed the researchers to select an appropriate monomer-solvent combination for successful complex formation. Cross-linkers were then also analyzed for their affinity for the template, selecting that with the lowest  $\Delta E$  value, therefore further increasing the relative affinity of the functional monomer selected. Analysis of the most appropriate monomers with the template analogues, under PCM approximations, indicated the MIPs likely selectivity for the chosen target over possible contaminants.

Khan *et al.* also applied some less common QM based techniques, notably determination of the HOMO-LUMO gap, which may be useful in predicting the relative stability of monomer-template complexes.<sup>98, 183</sup> This frontier orbital analysis can be used to calculate the interaction energy whilst also providing information on electron structure, which may be useful in considering further improvements to the system. Theoretical IR spectroscopy and molecular electrostatic potential (MEP) maps were also observed and considered useful by the group, but these techniques are less commonly applied.

The functional monomers used in MIP synthesis are typically small molecules (methacrylic acid being the most common<sup>184</sup>), but relatively large and complex structures can also be used, such as calixarenes (phenol based macrocycles) and DNA.<sup>185, 186</sup> An example of rational design with more unusual affinity structures can be observed in the research of Huynh *et al.* in their analysis of a nicotine imprinted polymers incorporating large organometallic structures.<sup>154</sup> Building on their earlier work with zinc porphyrins, the group developed a macromolecular zinc-phthalocyanine complex with bithienyl derived substituents and a similar hydroxyl containing structure. The nicotine-monomer complex was optimized with B3LYP/3-21G\* (B3LYP hybrid functional with 3-21G\* basis set), and the procedure was repeated with nicotine analogues. By freezing all monomer atoms not directly involved in target binding, the group managed to simulate the interactions which would be expected between nicotine, cotinine or myosmine with the final polymer in an aqueous medium (Figure 1.5).

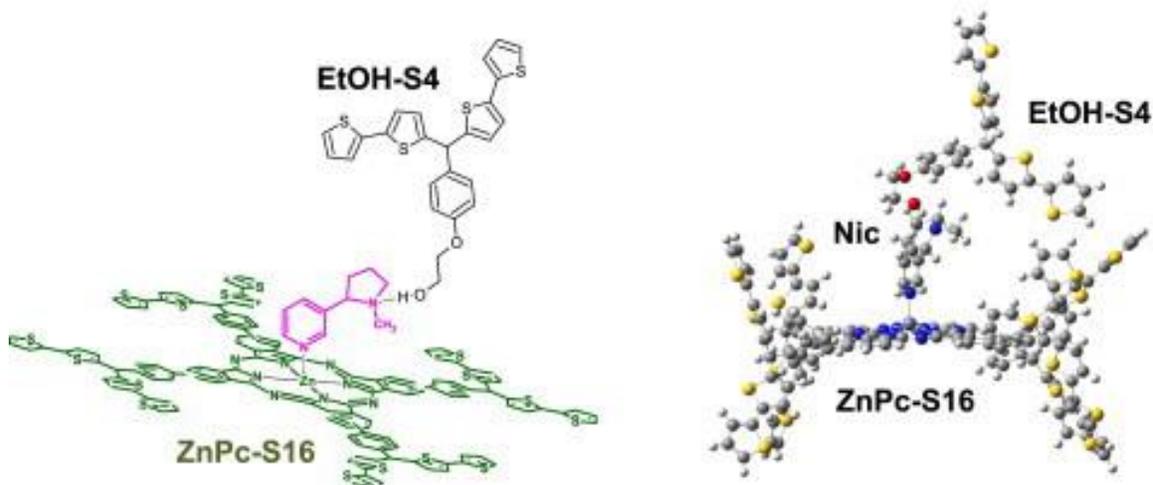


Figure 1.5.: The organometallic binding site in the nicotine imprinted polymer developed by Huynh *et al.* (reproduced with permission).

In this approach, which combined coordination of the target compound with traditional hydrogen bonding in the formation of the MIP, it was found that the latter interaction contributed only 20 % to the total binding energy. The complex was also found to be highly stable in synthesis, as predicted by the simulations, but with an unfortunate lack of specificity to nicotine. This is a development however from the earlier monomer structure in which the hydrogen bonding component of the system was covalently attached to a zinc porphyrin ring, creating a compound capable of detecting dinitrogen alkaloids.<sup>187</sup> The use of metal coordination in MIP design is therefore of potential interest, allowing polymer analysis easier in some respects (specifically fluorescence), but may not be entirely appropriate at this time due to the current lack of flexibility in design which leads to a relative loss of specificity.

Using more conventional functional monomers, Diñeiro *et al.* produced a homovanillic acid (a major dopamine metabolite, which is used as a marker in the screening for neuroblastoma, the most common form of infant cancer) imprinted polymer, similarly using B3LYP/6-31G\*.<sup>188</sup> The interaction energies were then improved by optimizing the geometries obtained from those already developed, using

a larger base set (B3LYP/6-311+G\*\*). During the design procedure, the monomer-template interactions were also simulated in a range of polymerization solvents, represented by continua of their relative dielectric permittivity with appropriate polarizations, to analyze the differences in binding interactions. They concluded that an MAA (methacrylic acid) MIP produced in a toluene porogenic solvent produced the best results, and the experimental evidence obtained after synthesis supported these claims. The group later reviewed the subject with an expanded library of functional monomers and found superior binding with trifluoromethacrylic acid.<sup>189</sup>

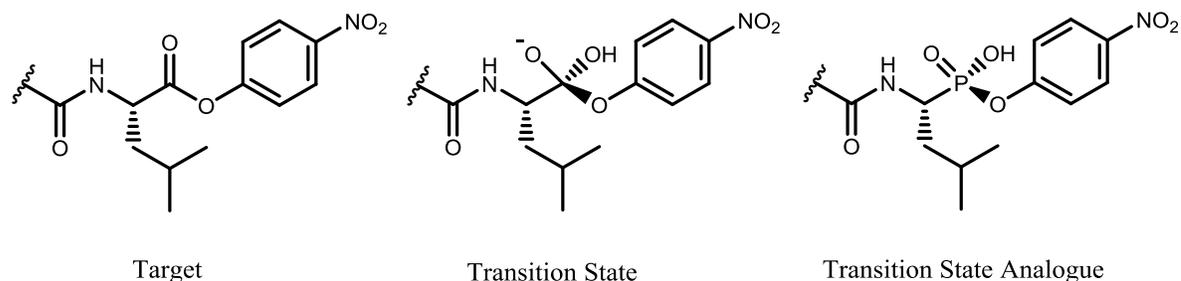
In terms of technological development, one area of particular interest both amongst developers and consumers will be the creation of a more effective method of glucose concentration monitoring for use by diabetics. These sensors usually rely on enzymatic binding, and for this reason are intolerant to environmental changes and have a short lifetime. For the realistic progression towards an artificial pancreas which can automatically deliver insulin based on glucose monitoring, MIP-based devices with the capacity to transduce binding into electrical signals are necessary.<sup>190</sup> Transduction for these devices can be achieved by utilizing the gate effect mentioned previously,<sup>191</sup> and several examples of glucose imprinted polymers can be found in the literature.<sup>192-194</sup> Difficulties arise however in producing polymers which distinguish between different sugars, and molecular modeling based design is necessary for effect discrimination.<sup>195</sup> Shariatinia *et al.* optimized a glucose complex with semi-empirical PM3, and then refined these results with HF/6-31G\* and B3LYP/6-31G\* on Gaussian 98.<sup>196</sup> These were then compared with MP2/6-311++G\*\* in these determination of the lowest binding interaction between  $\alpha$ -glucose or  $\beta$ -glucose with methacrylic acid (MAA), finding more favorable interactions with the  $\beta$ -form irrespective of which technique was used. This work demonstrated the importance of design, and that further research that is likely required before MIP-based blood sugar monitors can applied in practice.

Related research combining thorough theoretical analysis and bio-medical applications can be observed in the preparation of a pindolol imprinted polymer.<sup>197</sup> The template (a  $\beta$ -adrenoreceptor antagonist) and monomers were initially optimized using the Merck Molecular Force Field (MMFF), a set of parameters similar to MM3

designed explicitly for pharmaceutical applications, which emphasizes dynamics over mechanics (Sections 2.1 and 2.2).<sup>198</sup> The geometries of the monomers were then optimized by Hartree-Fock methods, and the energies calculated with B3LYP/6-31+G\*. Hartree-Fock was used instead of more advanced methods in an attempt to reduce computational expense, though a relatively large number (sixteen) of different functional monomers were observed, and this process would still be computationally expensive. The authors, somewhat unusually, also acknowledge the requirement to correct erroneous conclusions arising from Basis Set Super-position Error (BSSE), an effect resulting from overlap of the basis functions on groups involved in intermolecular interactions resulting in some, but incomplete, basis set mixing. This was resolved with the counterpoise method, in which mixed basis sets are used in all calculations to quantify the error.

Semi-empirical molecular orbital calculations were used to design a cinchonidine imprinted polymer to create a quinine receptor mimetic.<sup>199</sup> The purpose of this research was to find a method of screening bitterness-suppressing agents (such as arginine); quinine (a bitter substance) was used as a standard and monitored by spectrophotometry. By developing a MIP with lower specificity (cinchonidine lacks the 6-methoxy group of quinine) the retention time was reduced, making the method more suitable practically. The researchers were then able to observe the binding of quinine to the synthetic receptor and its inhibition by arginine. Demonstrating the effectiveness of arginine as a bitterness-suppressing agent has implications in drug manufacture due to the bitter taste of many pharmaceutical compounds and the resulting patient compliance. Geometries were initially predicted using MM2 molecular mechanical methods, then optimized using MOPAC, a semi-empirical QM program which combines MNDO, AM1, MINDO/3 and MNDO-PM3. The authors found that the MIP was less capable of recognizing the more complex interactions involved in bitterness perception than human subjects, but concluded that the use of MIPs as synthetic receptors in this area was highly efficacious. Similar biomimetic were developed by Sagawa *et al.*, in their production of a synthetic enzyme which could selectively hydrolyze *p*-nitrophenyl amino acid esters.<sup>200</sup> By optimizing a MIP with PM3 methods for a transition state analogue (phenyl-1-undecylcarbonylamino-3-methylbutyl phosphonate, Figure 1.6) of

the target molecule (N-dodecanoyl leucine-p-nitrophenyl) an effective biomimetic was produced which could catalyze the reaction in the same pseudo-first-order manner as that observed in natural enzymes.



*Figure 1.6: Structures of the target amino acid ester, its hydrolytic transition state, and the template used to produce a synthetic enzyme for this reaction.*

One major area of potential medicinal use for MIPs is in drug delivery systems. Though not a new idea, earlier proposals typically involved imprinting a polymer with some drug of interest and experimenting with the polymerization mixture until a MIP with appropriate dissociative properties in a relevant medium was discovered,<sup>201</sup> with more refined versions of this same method continuing into the 2010s.<sup>202</sup> Examples of such technology development including a rational design component include the thorough analysis of a MIP drug delivery system for naltrexone, an opioid receptor antagonist used in the treatment of alcohol and opioid dependence.<sup>203</sup> The template and five different functional monomers were first geometrically optimized with semi-empirical and HF methods before B3LYP/6-31+G\* was used to refine the complexes and find the relative energies, and the counterpoise method was used to correct the BSSE. The methacrylic acid and acrylic acid complexes were found to show similarly strong interactions so both were synthesized for comparison, resulting in the conclusion that the AA-based MIPs were superior, largely due to their greater loading capacity. Both MIP species showed controlled release profiles, with the relative release of total naltrexone from each being very similar, though less steady dissociation was found after reloading the polymers.

## 1.7. *Conclusions*

QM techniques, with DFT in particular, offer precise calculations of molecular structure and intermolecular binding interactions that are valuable in the design and analysis of MIPs. They are also applied frequently in the area and with varying level of detail, meaning a precedent that may be useful for repetition by different researchers with similar requirements. However, the high level of detail obtained using QM techniques come at a high computation cost, which creates significant limitations in terms of both model accuracy and depth of analysis. Monomer screening is restricted to a few molecules, with exceptions being notable due the time requirements, and all compounds must be relatively small. More significant for effective representation is the lack of accounting for environmental factors, primarily solvent but also the effects of time and temperature, which reduce the utility of QM for practical application. A more detailed models would require sacrificing some level of precision in individual molecular representation for the simultaneous simulation of a greater number of molecules.

## 2. Molecular mechanics and molecular dynamics

Molecular mechanics (MM) based techniques are relatively simple alternatives to QM, based on a set of functions which vary very little between the specific technique applied. Molecular dynamics (MD) techniques are historically considered a subset of MM (being essentially MM over time), though may often be referenced as a technique in itself (or occasionally as the superset of MM) due to its broad utility over raw MM. MM and MD currently represent approximately 36 % of the molecular modeling techniques used in MIP design, and many of these examples come from the work of a small number of research groups.<sup>51</sup> MM/MD however have many advantages over QM techniques and form the basis of much of the original research presented.

The aim of this chapter will be first to explain molecular mechanics and molecular dynamics, and to give some recent examples of how these methods have been applied through the course of this project. The first example describes a simple MM technique of minimization for energy and bond length analysis, the second demonstrates an automated screening technique using a standard set of monomers, the third example shows how the database of monomers can be modified for different applications, and the final example demonstrates molecular mechanics in combination with other techniques for more complex structural analysis.

### 2.1. *Molecular mechanics*

MM is the generally preferred method for modeling multi-component systems, largely due to its simplicity. Relatively simple equations describe energy as a function of bond length, bond angle, dihedral angle and non-bonding interactions. The interaction energies between individual atoms are described by parameters which are combined to build a force field (the force acting on each atom). Force fields can be used to determine the total energy of a system, accounting for all interactions, but the specifics of each vary depending on the form of the contributing parameters. While

some force fields may be objectively superior or inferior to others in broad application, the differences between particular force fields largely arise from their intended application. MM therefore often relies on different programs for different purposes, analysis of a small molecule generally requiring a different force field to that for a simulation of a large system. In general however, the force fields are all composed of the same basic parameter considerations.

MM force fields can be described by the same general equation:

$$E = E_{1,2} + E_{1,3} + E_{1,4} + E_{vdw} + E_{ele} + \dots \quad \text{Eq. 2.1}$$

Where  $E_{1,2}$  refers to the stretching of directly bound atoms,  $E_{1,3}$  to atoms separated by one atom and so refers to angle bending, etc. Additional terms are often included to create specialization or correct errors that may occur in their absence, for example an 'out of plane' energy or inversion barrier to promote appropriate geometry when not eliminated by the terms already present. The energy  $E$  refers to the total energy of the system, and each  $E_{1,2}$ ,  $E_{1,3}$ , etc. refers to the total energy of that kind of interaction.

The Tripos force field is described by relatively simple functions that can fit many applications and will be used as an example. The 1,2 stretching interaction is given by Equation 2.2:

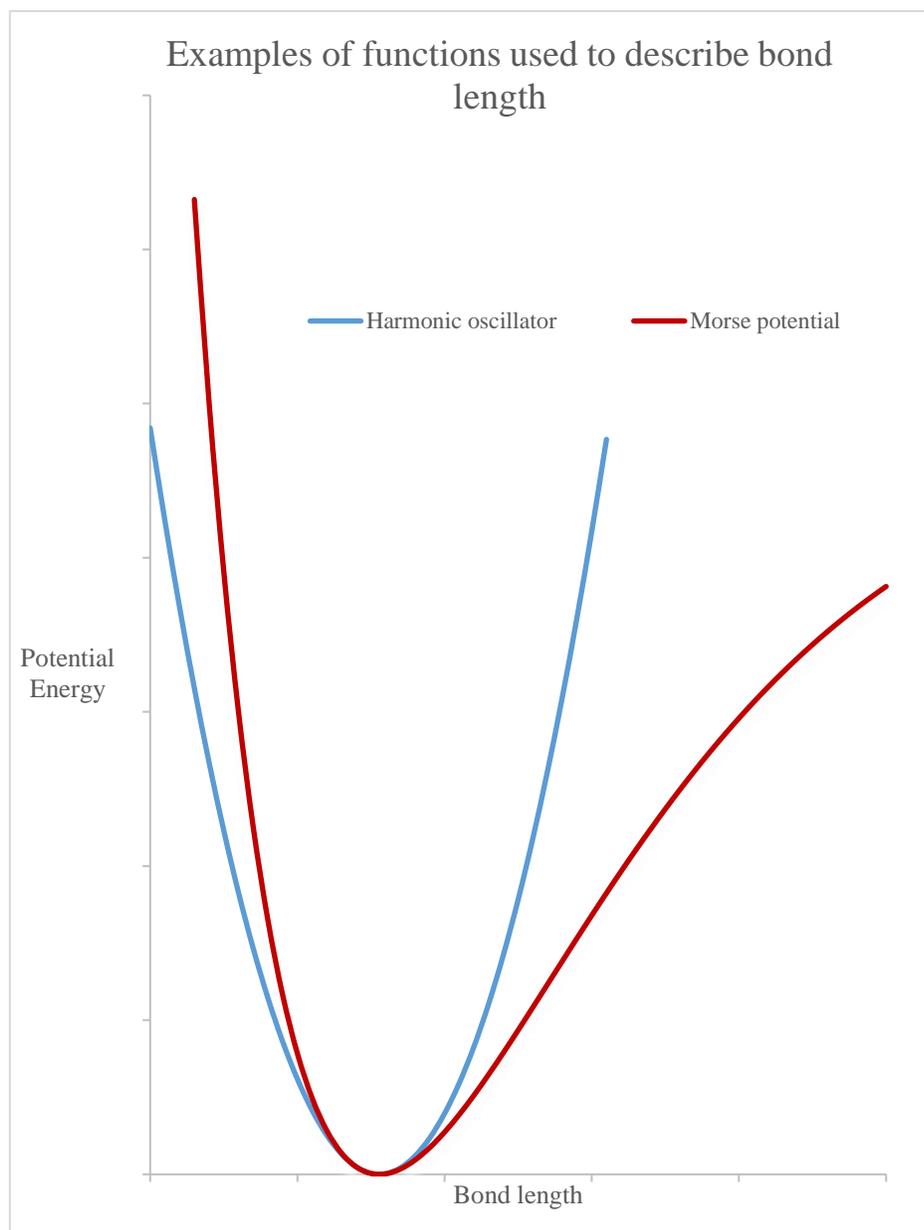
$$E_{1,2} = \sum_{i=1}^{N \text{ bonds}} \frac{1}{2} k_i^d (d_i - d_i^0)^2 \quad \text{Eq. 2.2}$$

Where  $d_i$  is the length of the  $i$ th bond,  $d_i^0$  is the equilibrium length for the  $i$ th bond, and  $k_i^d$  a bond stretching force constant. This harmonic oscillator model is a simplification but can be broadly applied as the conditions required for strong deviation from reality require values of  $(d_i - d_i^0)$  rarely encountered in typical simulations. When a more accurate model is required,  $E_{1,2}$  terms within which the energy converges towards a dissociation energy ( $D$ ) can be applied with a Morse potential (Equation 2.3):

$$E_{Morse} = \sum_{i=1}^{N \text{ bonds}} D \left( 1 - e^{\alpha(d_i - d_i^0)} \right)^2 \quad \alpha = \sqrt{\frac{k_i^d}{2D}} \quad \text{Eq. 2.3}$$

Examples of a harmonic oscillator and Morse potential are given in Figure 2.1. The Bond angles,  $E_{1,3}$ , are usually also represented by a basic harmonic function, though

very strained structures (e.g. three and four membered rings) can typically be described using additional terms.



*Figure 2.1: Harmonic oscillator and Morse potential, both functions that may be used to describe the bond stretching (1,2) interactions between atoms. Harmonic oscillators are simpler, can be applied more broadly with minimal modification, and less computationally expensive, but the Morse potential more accurately represents the natural behavior of bonds.*

Molecular mechanics is generally applied to polyatomic organic molecules, and dihedral angles must therefore also be accurately represented. The most common example used to explain these dihedral angles is butane (Figure 2.2), though something as simple as ethane must have appropriate parameters to describe the different energies of a staggered and eclipsed conformation. While the minima and maxima are equal in ethane, force fields for more complex molecules must account for local minima, such as that at  $60^\circ$  (gauche) in the butane structure.

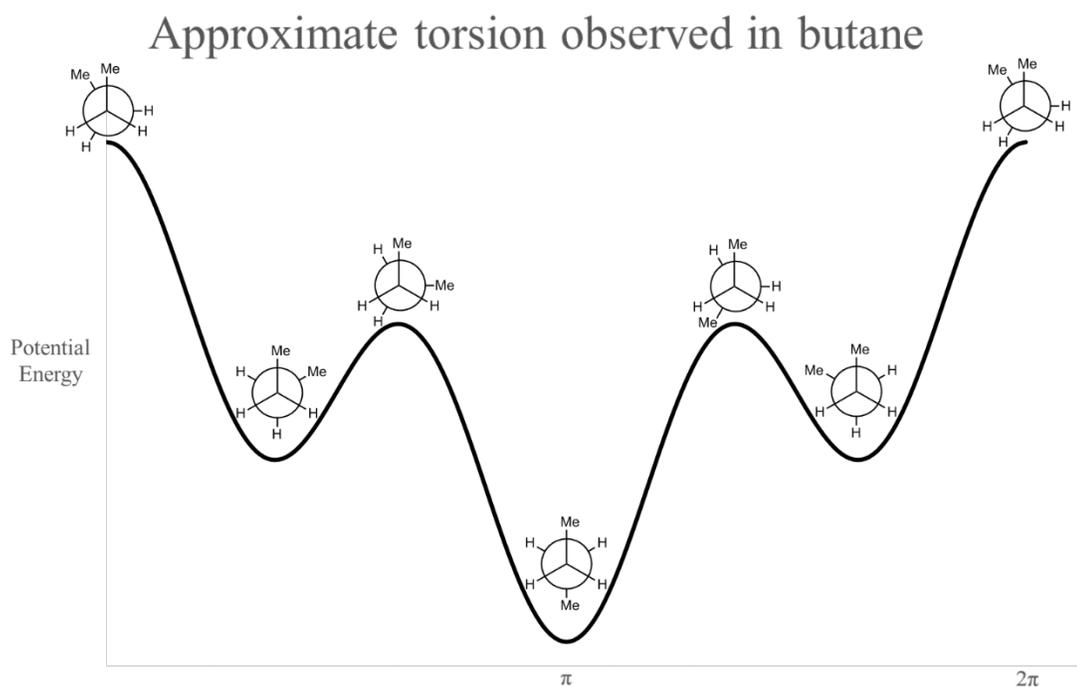


Figure 2.2: The potential energy observed in rotation about the 2-3 bond of butane.

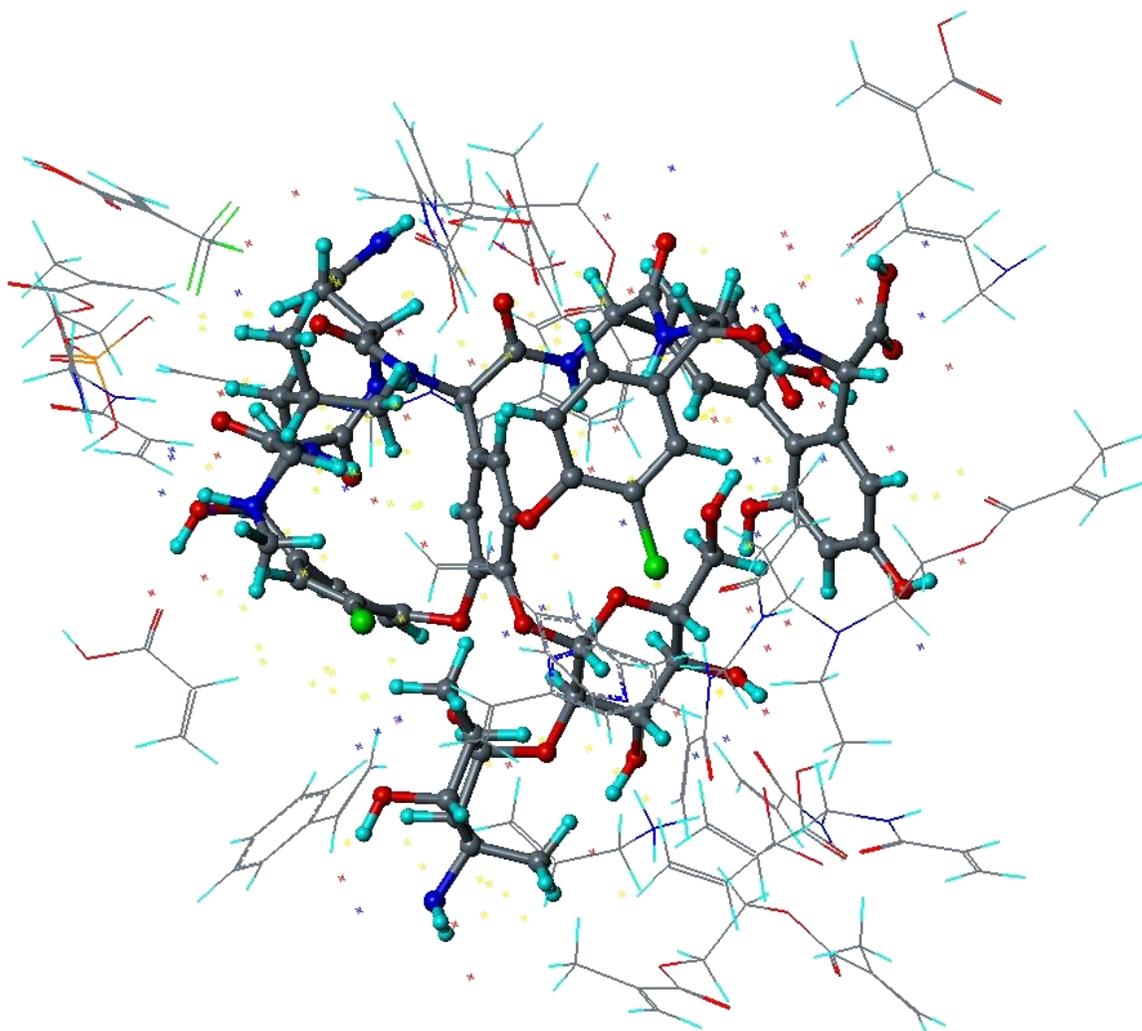
The 1-2, 1-3 and 1-4 interactions are all described using terms of potential energy increasing due to strain, and therefore contribute positively to the total energy of the system (if for example a molecule has all its bonds at their lowest energy lengths,  $E_{1,2} = 0$ ). Non-bonding interactions however may contribute to negative potential energy, equivalent to attractive interactions. Van der Waals interactions are given by a Lennard-Jones potential, Equation 2.4:

$$E_{vdW} = \sum_{i=1}^{N_{atoms}} \sum_{j>1} E_{ij} \left( \frac{1}{a_{ij}^{12}} - \frac{2}{a_{ij}^6} \right) \quad \text{Eq. 2.4}$$

Where  $E_{ij}$  is the van der Waals constant (dependent on the square of the multiplied energy of  $i$  and  $j$ ,  $E_{ij} = \sqrt{E_i E_j}$ ) and  $a_{ij}$  is equal to the interatomic distance,  $r_{ij}$ , divided by the sum of the van der Waals radii of the atoms;  $a_{ij} = r_{ij} / (R_i + R_j)$ . Hydrogen bonds are given by a variant of this with an additional constant in the  $a_{ij}$  term. Optional terms beyond this include electrostatic interactions, given by Coulomb's law, and specific additional functions to more accurately represent commonly encountered interactions, like carbon-carbon interactions at high and low interatomic distances.

Force fields are just the sum of their parameters, and are generally judged, especially when used for practical applications, by their ability to achieve results most similar to experimental data. However, the most reliable data usually comes from gas phase observations, as solutions and crystals include influences from the environment, and *ab initio* calculations.

The Tripos force field, previously described, is of particular importance in MIP design due to its use in the Leapfrog program, which is included within the Sybyl 7.3 software package.<sup>204</sup> MIPs are commonly designed by a screening process which involves placing a series of monomers adjacent to a template and solving  $\Delta E = E_c - (E_T + \sum E_M)$ , before comparing the resulting  $\Delta E$  values for the strongest complex. This approach forms the basis (and is a less efficient version in itself) of work presented in 2001 describing an automated screening process using Leapfrog.<sup>52</sup> The Leapfrog program first automatically finds the peaks and troughs of electron density surrounding the molecule to be screened as determined by the Tripos force field, represented on screen by colored points (as in Figure 2.3).



*Figure 2.3: The Leapfrog automated screening in progress, with a template (vancomycin, center) being screened against a database of monomers. Points of maximal and minimal charge at optimal distance are given by blue and red points around the template, with neutral areas given in yellow.*

A database, constructed from a number of previously effective functional monomers, is then screened against the molecule. Each monomer is placed in a position likely to give a strong interaction, based on the charge points, and the energy of the interaction is saved. The molecule is then moved slightly to note any change in interaction energy, and if stronger, overwrites the previous record. This process is repeated and recorded for every monomer of the database, until a final table of binding scores can be given for each monomer. The obvious advantage of this approach is, firstly, that the points of strongest interactions are found automatically, and secondly the large number of monomers that can be screened, again automatically. The original database included approximately 20 different monomers, in charged in neutral form, while later versions contain approximately 200.

The advantage of MM, both in general and in the design of screening and docking programs, is the efficiency of calculations. A Leapfrog screening involving thousands of energy calculations, structural modifications and data recording can be achieved in minutes using a standard desktop computer, while a similar process applied using QM software would be impossible with current technology.

## 2.2. *Molecular dynamics*

Molecular dynamics (MD) is the combination of MM and Newtonian laws of motion, allowing the simulation of atoms and molecules over time. By initial random assignment from a Maxwell-Boltzmann distribution this simulation can be performed at different temperatures, and the pressure can be controlled by modifying the available volume of the simulation, the molecules being confined to a 'box' of specified dimensions.

Dynamics simulations draw on observable systems and can be altered according to the required level of precision. For example, the observation of a molecule's properties in a particular solvent can be performed with a uniform dielectric constant providing the environment for the simulations. This is highly convenient for many situations, but the flaws in this can be readily observed in more complex molecules such as proteins, where the folding is not consistent with empirical data due to the more complicated solvent-solute interactions. Due to the relative simplicity of the MM calculations however, large numbers of molecules can be simulated simultaneously and in interaction with each other, permitting much more accurate analysis by including 'explicit' solvent. Experiments are usually performed either under the canonical ensemble, NVT (constant number of particles, volume and temperature) or, less commonly, the isothermal-isobaric ensemble, NPT (constant number of particles, pressure and temperature). Equilibration of an initially random arrangement of molecules can be performed with extended MD simulation at the required temperature, usually for a minimum of approximately of 3 ns but variable with system volume, or by simulated annealing (SA), in which the temperature is repeatedly elevated rapidly (usually to about 1000 K), allowing greater freedom of movement, and then gradually reduced to below ambient.

### 2.3. Force fields

While many different force fields and software packages have been produced for various purposes in computational chemistry, three particular examples should be briefly described due to their relevance to imprinted polymer design.

Amongst the most widely used parameter sets in general use are the MM2 family of force fields. MM2 was described by Allinger in 1977, following his relatively unsuccessful MM1 in 1973.<sup>205</sup> The differences between these values were changed little between MM1 and MM2 but the parameters were greatly improved by 'softening' the hydrogen atoms. This meant that the intermolecular forces between hydrogen and other atoms declined more gradually with distance than in the previous program, an alteration which dramatically improved the modeling of functional groups such as ethers and halides, and enhanced the reliability of organic molecules generally. MM2 was gradually improved in the years following its release, a process which continued after the release of MM3 in 1989.<sup>206, 207</sup> MM3 was developed in response to list of known or claimed issues with the MM2 program published in 1987. While MM2 was very good for most requirements it relied heavily on unjustified assumptions, prompting the release of MM4 in 1996, which demonstrated further improvements over both MM2 and MM3, particularly in its treatment of vibrational frequencies.<sup>208</sup> The MM2 family has repeatedly and consistently outperformed other molecular mechanics/dynamics techniques however, notably CFF (Consistent Force Field), CVFF (Consistent Valence Force Field) and UFF (Universal Force Field) which have been widely adopted in modeling packages.<sup>209</sup>

The Amber (Assisted Model Building with Energy Refinement) program, first published 1981, was design to simulate a range of differently sized molecules, unlike many of the other techniques available at the time, with a primary aim of modeling biological compounds and structures.<sup>210</sup> Amber quickly acquired additional capabilities, drawing on and collaborating with many individuals involved in computational science.<sup>211</sup> In a study of various computational methods' abilities to calculate the energies of different DNA fragments, Amber 4.1 was found to perform better than its MM competitors, including CFF95 and CHARMM, as well as the semi-

empirical methods AM1, PM3 and MNDO/M, and the simple *ab initio* SCF method.<sup>212</sup> The Amber software package also contains the Amber force fields in addition to other programs which are designed to co-operatively produce results greater than each alone would be capable of producing.<sup>213</sup> Amber has developed a reputation for quality and suitability for research, being both affordable and versatile. It is commonly used in the MIP MD modeling and has repeatedly demonstrated its efficacy, remaining relevant since its beginnings due to adaption to new situations and adoption of new methods to achieve excellent results.

The Tripos force field, sometimes used interchangeably with Sybyl (the supporting software package) was first publicized in 1989 as a good general tool for minimizing organic compounds.<sup>214</sup> In a comparative evaluation with other modeling techniques, Sybyl force fields was found to produce energy predictions in closer agreement with experimental data than the semi-empirical quantum methods, though it was out-performed by MM2.<sup>215</sup> In a more recent study, Tripos 5.2 force fields were found to produce results of approximately equal accuracy to those obtained with the General Amber Force Field (GAFF), and better than MMFF 94, CHARMM and Dreiding.<sup>216</sup> This is particularly impressive considering GAFF combines its predecessors' advanced protein force fields with further empirical data and information from *ab initio* MP2 hybrid functionals. While Tripos performs demonstrably better than many other force fields, its primary advantage is its exclusive integration into the Sybyl program, and is therefore tied to LeapFrog. This gives it a major advantage in MIP design, and can be shown to be effective in the design of materials for chemical extraction,<sup>217, 218</sup> sensors,<sup>219-222</sup> and biomedical analysis.<sup>223-225</sup>

## 2.4. Simple molecular mechanics

Examples of recent work involving MM analysis alone included the determination of the binding observed in a polymer designed for the analysis of an electrochemical sensor for the detection of aqueous copper. While the software and force fields are less suited to this work, the results appeared to conform to the experimental data. It was hypothesized initially that acrylic acid was required for the integrity of the sensor and involved in securing the polymer to the electrode. It was also observed however that a ratio of acrylic acid (Ac) and N-(3-Aminopropyl)methacrylamide (NAPMA) of 1:1 gave the most effective copper detection.

The structure of  $\text{Cu}(\text{H}_2\text{O})_6$  was constructed manually and energetically minimized using the Tripos force field with Gasteiger-Hückel charges with an applied dielectric constant of 1 (termination gradient  $0.01 \text{ kcal mol}^{-1}$ ). The total energy of the molecule was then recorded and the Cu-O bond lengths recorded. The water molecules were then systematically replaced with amine (datively) bound NAPMA ligands and the analytic process repeated. The results are given in Table 2.1.

Table 2.1: Mean bond lengths and potential energies of NAPMA copper complexes.

Formula	Mean H <sub>2</sub> O bond length, Å	Mean NAPMA bond length, Å	Total energy, kcal mol <sup>-1</sup>
$\text{Cu}(\text{H}_2\text{O})_6$	2.032	-	1265.333
$\text{Cu}(\text{H}_2\text{O})_5(\text{NAPMA})$	2.030	2.072	1216.704
$\text{Cu}(\text{H}_2\text{O})_4(\text{NAPMA})_2$	2.022	2.081	1161.064
$\text{Cu}(\text{H}_2\text{O})_3(\text{NAPMA})_3$	2.016	2.088	1109.566
$\text{Cu}(\text{H}_2\text{O})_2(\text{NAPMA})_4$	2.013	2.086	1067.215
$\text{Cu}(\text{H}_2\text{O})(\text{NAPMA})_5$	2.024	2.090	1044.768
$\text{Cu}(\text{NAPMA})_6$	-	2.098	1015.099

The complex energies suggest favorable interactions between (NAPMA) and the copper, with preferential binding over the solvated state. From the bond lengths it is

likely that a 2:4 arrangement is preferred with only NAPMA in aqueous solution, likely due to steric factors associated with greater numbers of ligand binding.

To account for the hypothesis that acrylate (Ac) was involved in the ion binding, this process was repeated with different monomer-water ratios (Table 2.2).

*Table 2.2: Mean bond lengths and potential energies of NAPMA and acrylate copper complexes.*

Formula	Mean H <sub>2</sub> O bond length, Å	Mean NAPMA bond length, Å	Mean Ac bond length, Å	Total energy, kcal mol <sup>-1</sup>
Cu(H <sub>2</sub> O) <sub>5</sub> (Ac)	2.000	-	1.918	790.407
Cu(H <sub>2</sub> O) <sub>4</sub> (Ac) <sub>2</sub>	1.963	-	1.893	410.902
Cu(H <sub>2</sub> O) <sub>3</sub> (Ac) <sub>3</sub>	1.933	-	1.898	143.809
Cu(H <sub>2</sub> O) <sub>2</sub> (Ac) <sub>3</sub> (NAPMA)	1.930	1.999	1.897	115.562
Cu(H <sub>2</sub> O)(Ac) <sub>3</sub> (NAPMA) <sub>2</sub>	1.907	2.004	1.902	115.432
Cu(Ac) <sub>3</sub> (NAPMA) <sub>3</sub>	-	2.020	1.905	150.899
Cu(H <sub>2</sub> O) <sub>3</sub> (Ac) <sub>2</sub> (NAPMA)	1.966	2.013	1.895	396.651
Cu(H <sub>2</sub> O) <sub>2</sub> (Ac) <sub>2</sub> (NAPMA) <sub>2</sub>	1.965	2.022	1.895	370.844
Cu(H <sub>2</sub> O)(Ac) <sub>2</sub> (NAPMA) <sub>3</sub>	1.971	2.038	1.895	394.170
Cu(Ac) <sub>2</sub> (NAPMA) <sub>4</sub>	-	2.053	1.915	397.362

Interactions between copper and acrylate were found to be very favorable. Substitution of water for NAPMA in the complex with three acrylate ligands gave greater stability with one equivalent, two NAPMA in place of water giving little difference or a small decrease in favorability. With two acrylate ligands in the complex binding of an additional two NAPMA with two water was found to be the most favorable arrangement. This suggests a 1:1 ratio of monomers would give the greatest binding of

Cu(II), conforming to the observed experimental data and confirming the interaction of acrylate in the copper binding.\*

## 2.5. *Automated monomer screening with standard monomer database*

The Leapfrog based automated monomer screening has been adopted as a standard component of the MIP design process by the Piletsky group. A protocol for the computational design of MIPs applying these techniques was recently published explaining the procedure in an open access journal, permitting easy access to this procedure by all interested researchers.<sup>204</sup>

An example of this process can be given with the design of a MIP-based magnetic assay for the detection of the pesticide methyl parathion.† In this example however the target was not used in the imprinting process, an analogue aminoparathion being used in place due to easier immobilization on the solid surface (a requirement in standard MIP NP synthesis). Both the aminoparathion, methyl parathion were initially drawn and energetically minimized with the Tripos force field and MMFF94 charges using the Sybyl 7.3 software package. Each structure was then screened against a database of 20 commonly used functional monomers (represented by their charged and uncharged forms where applicable) using the Sybyl Leapfrog algorithm, ranking the monomers by a relative binding energy and predicting the most stable arrangement of each monomer-target complex. The results are shown in Table 2.3a and 2.3b, and were used to determine the most suitable functional monomer for the prepolymerization mixture.

---

\* This research is currently being conducted by Sabrina Di Masi and will likely be published in early 2019.

† A manuscript is currently in review. This work will likely be published in late 2018, authored principally by Dr Joanna Czulak (Leicester Biotechnology Group/MIP diagnostics) and Prof Cem Esen (Aydın Adnan Menderes University).

*Table 2.3a: The results of the automated screening of aminoparathion with the Leapfrog algorithm.*

Monomer	Binding energy, kJ mol <sup>-1</sup>
Allyamine (+)	-193.803
Ethylene glycol methacrylate phosphate (2-)	-123.721
Acrylamido-2-methyl-1-propanesulfonic acid (-)	-119.286
Ethylene glycol dimethacrylate	-117.068
Ethylene glycol methacrylate phosphate	-107.361
Acrylamide	-99.244

*Table 2.3b: The results of the automated screening of methyl parathion with the Leapfrog algorithm.*

Monomer	Binding energy, kJ mol <sup>-1</sup>
Allyamine (+)	-288.947
N-phenylethylene diamine methacrylamide (+)	-184.891
N,N-diethylamino ethyl methacrylate (+)	-128.909
Ethylene glycol dimethacrylate	-122.717
Acrylamide	-114.140
Acrylamido-2-methyl-1-propanesulfonic acid (-)	-102.299

This standard database has been applied widely within the Leicester Biotechnology Group and in collaboration with other researchers. A recent example in a modified form involved the analysis of monomer binding, along with conformational analysis, in the preparation of an assay for the detection of Ochratoxin A (A). OTA and two structural analogues (shown in Figure 2.4) were constructed and minimized to a  $0.01 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$  gradient with the Tripos force field and MMFF94 charges using Sybyl 7.3 software. 5 cycles of 10 ps molecular dynamics simulation and minimization with the same parameters (*in vacuo*) were then performed, and the lowest potential energy structure recorded. These lowest energy structures were then solvated in water and a further 10 ps dynamic simulation was performed with explicit solvent at 300 K and analyzed for conformity with the *in vacuo* results.

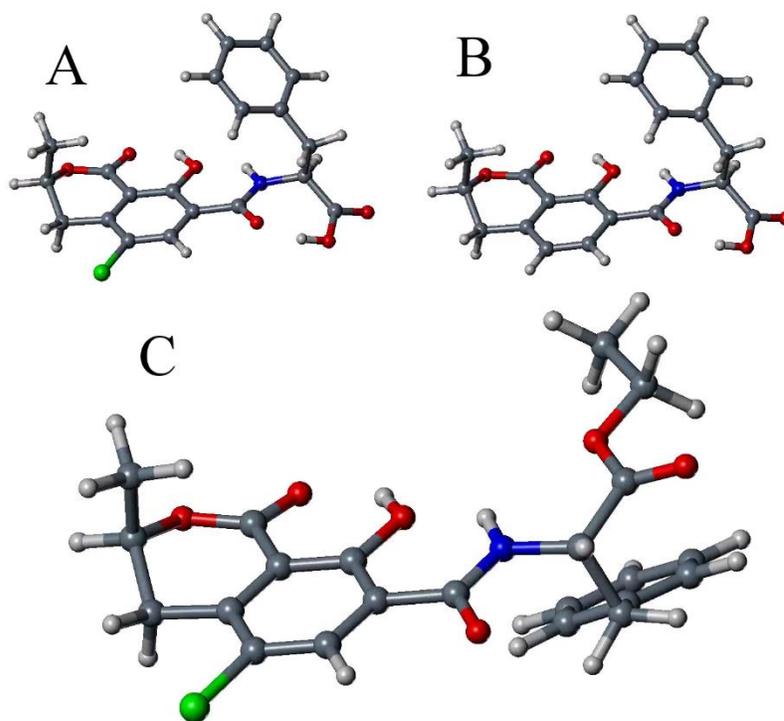


Figure 2.4: Ochratoxin A (A) and two structural analogues (B and C) in their minimum energy configurations.

Structure A obtained a minimal potential energy of  $-10.350 \text{ kcal mol}^{-1}$ , and structure B  $+4.754 \text{ kcal mol}^{-1}$ . The configurations of these two were almost identical, a very stable hydrogen bonding arrangement being found between carboxylic acid and amide oxygen, and also between amide nitrogen, hydroxyl and ring ester. The difference in energy between A and B presumably arises from a ring stabilizing effect from the chlorine for A.

The minimum potential energy observed for structure C was  $+2.145 \text{ kcal mol}^{-1}$ , this difference in energy between A and C arising from the lack of stabilizing carboxylic acid hydrogen bonding. The lowest energy configuration instead shows some attraction between the  $\text{sp}^3$  oxygen of the ester and the nitrogen of the amide, though this interaction is weak and the barrier to rotation breaking this bond is very low. Molecular dynamics in explicit solvent gave the same configuration as those determined *in vacuo* (Figure 2.4). Structures of A as would be required for MIP NP preparation were then screened against commonly used monomers for both the binding energy and the sites of interaction.\*

---

\* This research is currently being conducted by Daniel Lopez and will likely be published in early 2019.

## 2.6. *Automated screening with electropolymerizable monomer database*

This protocol can also be modified for alternative application. For example, in the development of MIP-based sensors for the detection of recreational drugs, in which the polymer was synthesized by electropolymerization onto the electrode. For this the monomers must be electropolymerizable, and the standard database was therefore unsuitable. A database of 19 commonly used electropolymerizable monomers was constructed and screened against Adipoyl-7-ADCA, and separately in analyzing a series of drugs, each of which was prepared in both charged and uncharged forms for analysis (cefadroxil, cefadroxil anion, cefadroxil cation, cefadroxil zwitterion, doxycycline, doxycycline cation, heroin and heroin cation).<sup>\*</sup> Each structure was energetically minimized using the Tripos force field with MMFF94 charges under a dielectric constant of 1. The minimized structure then underwent a 50 ps molecular dynamics simulation at 300 K, before being reminimized with the same parameters. The five highest binding scores achieved for each drug are given in Table 2.4.

---

<sup>\*</sup> This work will likely be published in 2019. Empirical analysis and instrument development is currently being conducted by Dr Anca Florea (University of Antwerp).

Table 2.4: Binding energies obtained from the Leapfrog screening of multiple templates (drugs represented in charged and uncharged form) with a virtual database of electropolymerizable monomers.

Template	Monomer	Binding, kJ mol <sup>-1</sup>
<i>Cefadroxil</i>	1H-pyrrole-2-carbohydrazide	-258.655
	Pyrrole-2-carboxylic acid	-205.602
	<i>o</i> -phenylenediamine	-168.783
	3-aminophenylboronic acid	-162.423
	4-aminothiophenol	-161.084
<i>Cefadroxil anion</i>	1H-pyrrole-2-carbohydrazide	-210.623
	Pyrrole-2-carboxylic acid	-193.719
	<i>o</i> -phenylenediamine	-123.805
	4-aminothiophenol	-114.725
	4-aminobenzoic acid	-69.413
<i>Cefadroxil cation</i>	Pyrrole-2-carboxylic acid	-171.167
	4-aminobenzoic acid	-161.753
	4-aminothiophenol	-126.315
	3-aminophenylboronic acid	-124.892
	1H-pyrrole-2-carbohydrazide	-122.508
<i>Cefaroxil zwitterion</i>	Pyrrole-2-carboxylic acid	-128.365
	1H-pyrrole-2-carbohydrazide	-121.629
	4-aminothiophenol	-116.817
	4-aminobenzoic acid	-108.617
	3-aminophenylboronic acid	-69.789

<i>Doxycycline</i>	<i>o</i> -phenylenediamine	-187.025
	4-aminobenzoic acid	-175.268
	1H-pyrrole-2-carbohydrazide	-158.239
	3-aminophenylboronic acid	-157.862
	4-aminothiophenol	-134.892
<i>Doxycycline cation</i>	<i>o</i> -phenylenediamine	-240.538
	Pyrrole-2-carboxylic acid	-171.460
	4-aminobenzoic acid	-171.042
	4,6-dihydroxybenzophenone	-165.979
	1H-pyrrole-2-carbohydrazide	-162.967
<i>Heroin</i>	Pyrrole-2-carboxylic acid	-117.612
	<i>o</i> -phenylenediamine	-108.240
	aniline	-103.721
	3-aminophenylboronic acid	-92.341
	1H-pyrrole-2-carbohydrazide	-89.959
<i>Heroin cation</i>	Pyrrole-2-carboxylic acid	-167.234
	aniline	-113.847
	4-aminothiophenol	-113.345
	<i>o</i> -phenylenediamine	-110.207
	1H-pyrrole-2-carbohydrazide	-108.784

This database was also applied in the design of a sensor with selective response for cocaine. Several common drugs (cocaine, morphine, levoamphetamine, dextroamphetamine, MDMA, cephalexin and nafcillin) were screened using Leapfrog with the electropolymerizable monomer database. From this screening it was determined that *o*-phenylenediamine and 4-aminobenzoic acid would produce the most selective MIP binding sites, a prediction which was later demonstrated empirically.<sup>226</sup>

Analysis of the polymers synthesized for cocaine however showed less selectivity for the target than would be necessary for real-world application, responses being observed for other drugs that suggested significant binding. To understand the interactions that were occurring on the MIP sensor, the affinity between the monomer 4-aminobenzoic acid and three interferent compounds was calculated.

Four monomer 4-aminobenzoic acid structures, neutral, anion, cation and zwitterion, were added to a simple database. This database was then screened against several drugs (cocaine, codeine, quinine, (R)-MDMA and (S)-MDMA) in both neutral and amine-protonated form. Each arrangement was then minimized using the Tripos force field and MMFF94 charges under a vacuum dielectric constant. The monomer-template complex potential energy was recorded, and each component was then removed and the energy recorded again in isolation, giving a total interaction energy  $\Delta E = E_{complex} - (E_{drug} + E_{monomer})$ . The results are given in Table 2.5.

Table 2.5: Binding scores of cocaine and several interferent drugs with 4-aminobenzoic acid in neutral and charged forms.

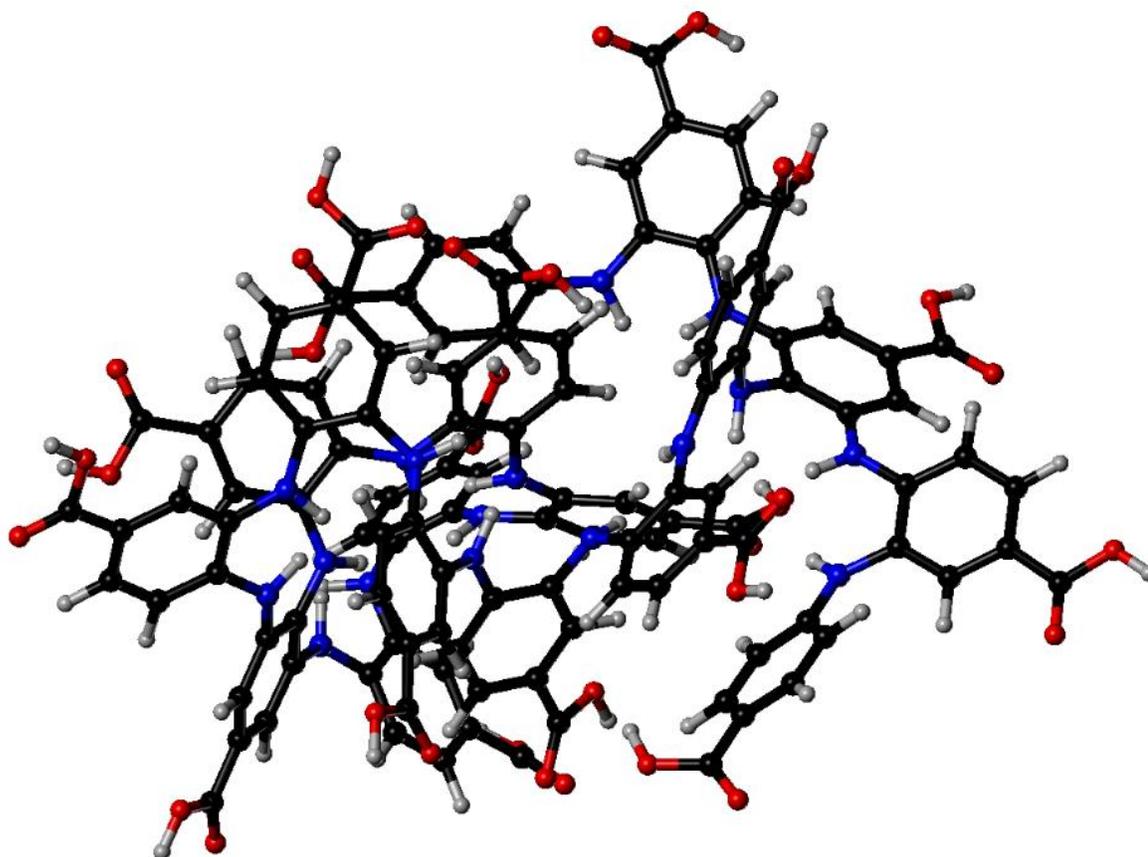
	4-aminobenzoic acid binding scores, kJ mol <sup>-1</sup>			
	Neutral	Anion	Cation	Zwitterion
Cocaine	-62.998	-64.894	-191.991	-123.775
Cocaine H+	-78.814	-220.869	-48.045	-203.020
Codeine	-91.341	-115.499	-147.214	-156.139
Codeine H+	-59.149	-138.478	-123.248	-127.997
Quinine	-78.709	-133.549	-141.248	-126.461
Quinine H+	-75.789	-195.384	-146.373	-252.517
(R)-MDMA	-41.325	-101.839	-79.977	-44.869
(R)-MDMA H+	-117.512	-266.140	---*	-299.148
(S)-MDMA	-39.037	-39.765	-146.984	-78.730
(S)-MDMA H+	-74.458	-384.091	-50.810	-256.207

\* No binding observed in screening

The results in Table 2.5 did suggest a lack of specificity but were considered too varied to provide meaningful answers. To further investigate the interactions of the polymer with the different drugs, a small oligomer was built from 17 4-aminobenzoic acid units, bound via secondary amine links ortho- to the amine of the monomer (Figure 2.5). The oligomer underwent five 100 ps molecular dynamics simulations followed by energy minimization with the MMFF94 force field and charges applied ( $\epsilon = 80$ ). After each minimization the dissociation constant of each of the 5 drugs (with and without protonated amine) for the oligomer was predicted using the Surflex docking program. The oligomer was then modified to have approximately half and then total deprotonation of the carboxylic acid groups, the cycles of MD, minimization and docking being repeated for each of these new structures. The mean value of the binding was taken from the three or more values giving the lowest deviation (Appendix 1).

Docking with half deprotonated oligomer gave zero binding for the majority of ligands on two occasions, only registering affinity for MDMA. These results were

omitted, and the process repeated. Fully deprotonated polymer only gave results for MDMA for almost every arrangement (approximately equal to those for the protonated and half protonated oligomer). The results for the protonated and half protonated oligomers are given in Table 2.6.



*Figure 2.5: 4-aminobenzoic acid based oligomer used to dock cocaine and several interferent drugs. The most suitable cavity was discovered automatically using the docking software and each of the drugs observed for their binding in this cavity.*

Table 2.6: Results of the oligomer-drug docking studies using the 4-aminobenzoic acid based structure shown in Figure 2.5.

	Ligand	$-\log(K_D)$	Standard deviation
<i>Protonated oligomer</i>	(R)-MDMA H+	3.19	0.14
	(S)-MDMA H+	2.94	0.12
	(S)-MDMA	2.92	0.48
	(R)-MDMA	2.65	0.13
	Quinine	2.59	0.09
	Quinine H+	2.45	0.23
	Cocaine	2.04	0.50
	Codeine	1.92	0.21
	Codeine H+	1.65	0.11
	Cocaine H+	1.38	0.17
<i>Half protonated oligomer</i>	(S)-MDMA H+	3.16	0.29
	(R)-MDMA H+	3.06	0.17
	(S)-MDMA	3.05	0.30
	Quinine	2.80	0.53
	(R)-MDMA	2.22	0.07
	Quinine H+	2.00	0.38
	Codeine H+	1.86	0.67
	Codeine	1.38	0.38
	Cocaine H+	0.85	0.33
	Cocaine	0	0

## 2.7. *Further applications of molecular mechanics modeling*

More novel applications of computational modeling related to imprinted polymers include the determination of oligomeric structure from mass spectrometry analysis. Linear molecularly imprinted polymers (LMIPs), small uncross-linked oligomers, were synthesized in the presence of a peptide associated with quorum sensing.<sup>227</sup> The advantage of LMIPs in terms of medicinal applications is their small size and simple structure, the sequence of monomer units potentially being determined by a combination of techniques (the lack of regular structure in standard MIP NPs makes them unlikely to pass regulatory standards for drug-like applications).

The LMIPs were found to interfere with bacterial quorum sensing and inhibit bacterial growth, with important implications for overcoming antibiotic resistance. To determine the sequence of the LMIP however a number of techniques were required. LC-MS analysis suggested the total mass of the oligomer was 1006 Da with evidence of fragments corresponding to copolymer trimers occurring as parts of the polymer, specifically acrylamide trimers, butylacrylamide trimers, and acrylic acid trimers and dimers being those fragments most strongly represented in the spectrum. In an attempt to provide some understanding of the complete structure, two different molecular modeling techniques were used based on the principle of imprinting, in that the fragments obtained are likely to exhibit strong interactions with the peptide in the appropriate positions relative to each other in the LIP.

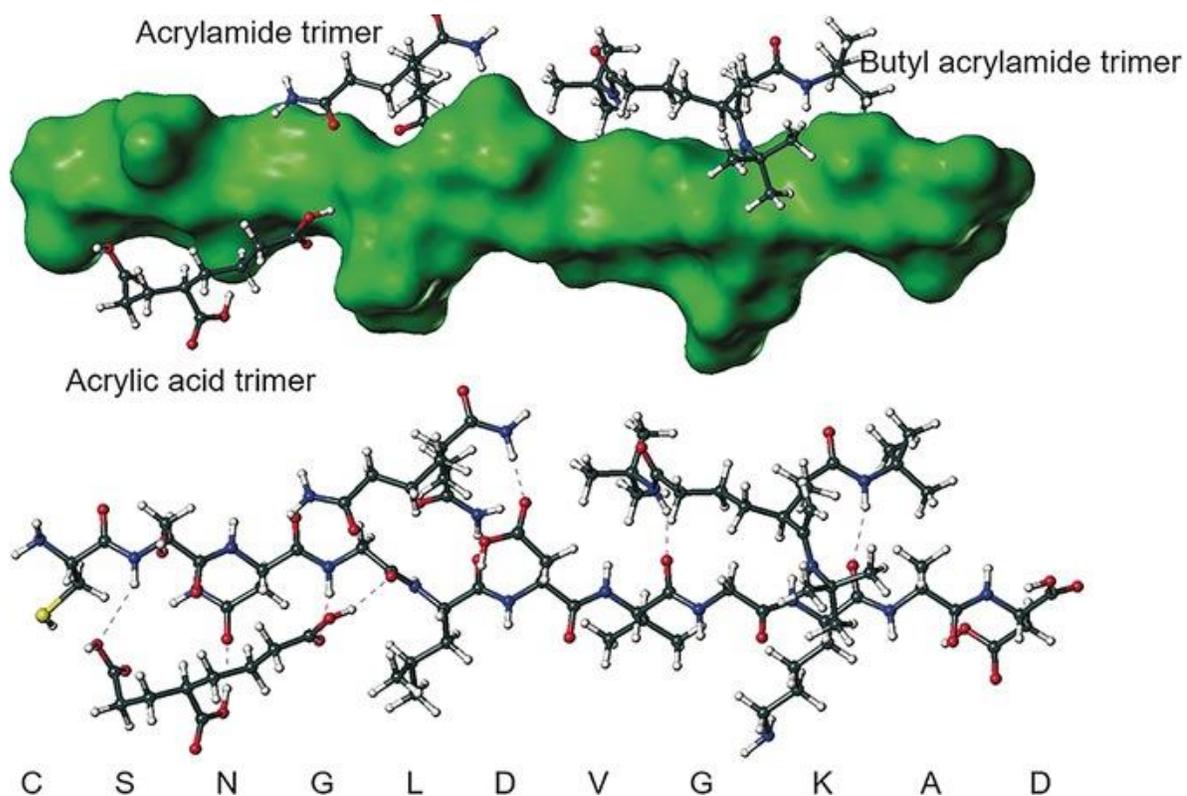
Molecular screening and docking programs provide the most efficient method of determining the strongest interactions between the peptide and the fragments. Surflex-Dock was used in these examination in combination with the Sybyl LeapFrog screening algorithm. The structures of the copolymer trimers were constructed based on their being the most energetically favorable in synthesis, and each of these was then analyzed via the two methods against the peptide. The docking and screening programs used work with the same principles of placing the fragments at (automatically) predetermined positions around the target. The strength of interaction between the fragment and the peptide is then measured to give a binding energy determined by an appropriate force field. Repeatedly moving the fragments to new points of interest and

further analysis of established locations around the peptide by small movements and changes to the fragment's conformation give the highest affinity positions for each of the molecules.

The peptide structure was built in both a charged and uncharged form for analysis and the four fragments described by LC-MS analysis were constructed in the arrangement most favored in polymerization. Molecules of acrylamide, butyl acrylamide and acrylic acid trimers were therefore designed with two subunits bound via the terminal alkene carbons of the monomers to the third, with the acrylic acid dimer bound through joining the terminal alkene of the monomers. All structures were energetically minimized using the Tripos force field with MMFF94 charges with an applied dielectric constant of 80 (water). Initial screening was performed with Leapfrog, taking the peptide as the analyte and observing the interactions with each of the fragments, generating the position of maximal interaction for each fragment around the peptide. Surflex-Dock was used in the same manner by setting the peptide as the area for docking, the 'cavity' encasing the peptide and the area of exploration set to extend outside of this assigned area. Relative positions of the fragments were obtained directly from the results of the applied methods with no rearrangements, substitutions or additions.

The acrylic acid trimer shows relatively high and equal affinity for either of two arrangements around the asparagine residue, forming either a four point interaction (Hydrogen bond donation to the C and N residue carbonyls, and accepting from the N amine and side chain amide) or five point interaction (hydrogen bond donation to the carbonyls of G and the N side chain, and acceptance from the N amide and amines of the G and N backbone) in which the two interactions with the G residue are relatively weaker. The strongest interaction found with the LeapFrog screening was that of a three point interaction between the acrylamide trimer the carbonyl and amine of the Leucine residue and the carboxylate of the aspartic acid residue. The butyl acrylamide trimer also showed three points of interaction, all hydrogen donation from the fragment amides to the backbone carbonyls of the V, G and K residues. The relative positions of these residues can be seen in Figure 2.6.

The total mass of these three fragments is  $810.981 \text{ g mol}^{-1}$ , and thus makes up the majority of the polymer. The polymer structure is therefore predicted to be X-AA-AA-AA-X-Ac-Ac-Ac-X-BA-BA-BA-X, where AA is acrylic acid, Ac acrylamide, BA butyl acrylamide, and X either a join between the adjacent subunits or another small section of the polymer, with the total mass of X being approximately 195 Da.



*Figure 2.6: The interactions observed between the linear molecularly imprinted polymer fragments and the quorum sensing signaling peptide (Motib, 2017). Initial predictions of the relative positions of the fragments were originally made with Leapfrog, screening the peptide with the different fragments, then refined with the Surflex docking software (reproduced with permission).*

## 2.8. *Conclusions*

Standard molecular mechanics modeling has been found very useful for MIP based research, providing a base for rational design of high affinity materials and insight into polymer structure and analyte binding. Further to these example of direct applications to specific systems include useful information that can be applied more generally, as in the general considerations required in imprinting small peptides.<sup>228</sup>

However the recurring theme through these examples is the insufficiency of monomer-template analysis alone. In the example of the electropolymerization studies it was necessary to construct a small oligomer to understand the interactions between the polymer and several different compounds; in this analysis it was observed that the monomer selected was unlikely to provide effective selectivity for the target cocaine, showing low affinity relative to the other compounds observed. Similarly in the studies of the peptide imprinted linear polymer, it was necessary to construct small oligomers in order to understand the structure of the ultimate LMIP, with monomer analysis being considered insufficient alone. Some further analysis of the problems with existing approaches to molecular modeling and some possible solutions are presented in the next chapter.

### 3. In silico synthesis of synthetic receptors

The MM and MD analysis discussed in the previous chapter are evidently useful in the design of functional polymeric materials and in more broad application. The limitations in polymer analysis however can be clearly observed to relate to the level of detail applied in standard computational analysis. This will be discussed first and followed with some examples of previous attempts from the literature to develop more realistic computational models of imprinted polymers. From here a description will be given of the development of a polymerization algorithm to overcome some of these issues of accuracy in MIP representation.

#### 3.1. *Problems with existing computational approaches*

It can be readily demonstrated that even simple modeling techniques offer significant advantages over alternative methods in MIP analysis; expenses aside, empirical analysis can be difficult to analyze and lack the precision required to observe the interactions of compounds of interest.<sup>229</sup> However, problems in the existing modeling protocol are easily demonstrated.

Yang *et al.* devised a two-part QM based design procedure in the development of magnetic MIP NPs for the removal of phenol from water.<sup>84</sup> After initial structure optimization of the template, monomers and complexes of different ratios by RHF/6-311G, natural bond orbital (NBO) charge transfer and Equation 1.1 were used to select an appropriate monomer and T-M ratio. NBO charge transfer models are used here to predict the strength of intermolecular hydrogen bonds in the complex, and importantly the difference in electron density associated with each atom on complexation. Analysis of this data and observation of the structures resulting from minimization clearly shows that in higher ratios of T-M (1:3, 1:4, etc.) the template shows relatively weaker interactions with the monomers due to M-M interactions being energetically favored in these systems. The  $\Delta E$  values calculated for the MAA complexes were  $-47.379 \text{ kJ mol}^{-1}$  and  $-102.060 \text{ kJ mol}^{-1}$  for the 1:1 and 1:4 systems respectively (a doubling in association

energy) despite the NBO charge transfer analysis showing a reduction in association to the template, the energetic stabilization originating in the monomer-monomer interactions.

This highlights the limitations of the standard form of Equation 1.1, where the potential energy of the individual components are taken as representative of the system without T, while interactions between other components obviously contribute in proportion to their number. Firstly, as the number of molecules present increases, in appropriate ratios, the more accurate the model becomes, and the interaction of the template with the other components becomes negligible relative to all the other interactions. The energy must therefore be observed for the system with and without T, not for the individual components. Secondly is screening of monomers; favorable interactions between template and monomer alone mean very little without consideration of the other intermolecular interactions that may occur.

Since the work of a few pioneer groups at the start of the new millennium the procedure generally applied has not advanced as required, centering on the relative strength of interactions between the template and different functional monomers. This approach, with the application of the formula  $\Delta E = E_C - (E_T + \sum E_M)$ , Equation 1.1, should not be applied for more than one monomer when the monomers are in close enough proximity to interact, i.e. almost all modeling in the design of MIPs. Further to this is the exclusion of other system components. As cross-linker molecules are usually omitted and the effects of solvents and physical parameters are accounted for very simplistically, problems of representations often arise. While correlations may be observed between more simple and more intensive approaches,<sup>230</sup> this is not always the case. Sobiech *et al.* studied the effect of modeling monomer template interactions alone compared with studies including the cross-linker, the results (Figure 3.1) demonstrating clearly that analysis of the M-T complex alone is insufficient for appropriate representation.<sup>74</sup>

Clearly the more closely the model approaches the real composition, the more accurate that model will become. It is for this reason that QM techniques can represent a huge increase in the precision of intramolecular structure and intermolecular binding while also representing a total loss in efficiency and an overall retardation of the

development of imprinted polymer research. The QM models used give accurate details of systems that are never produced: the interactions of various monomers with a compound of interest, in a vacuum without time.

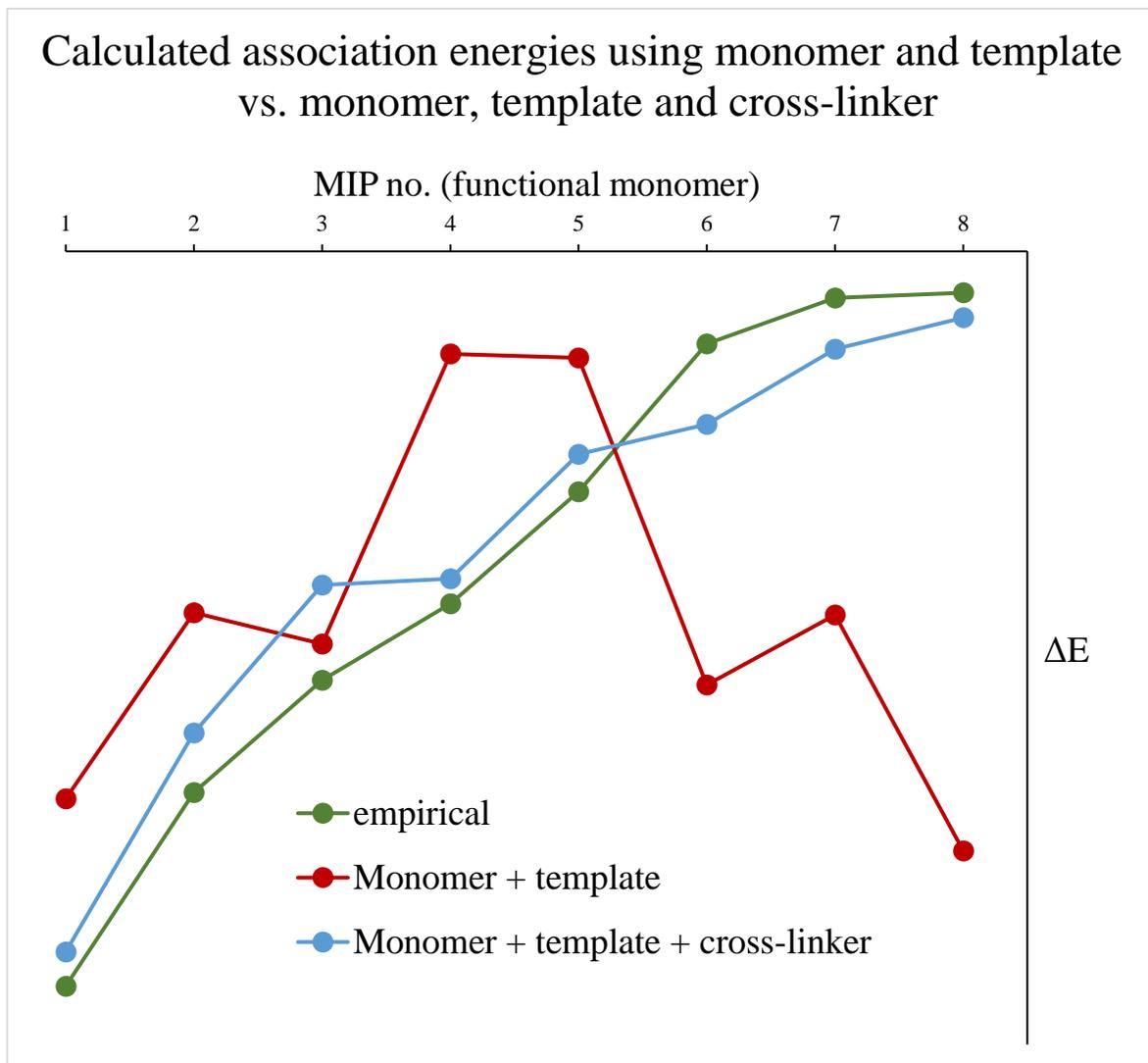


Figure 3.1: Comparison of monomer-template models and monomer-template-cross-linker interactions vs empirical binding (data from Sobiech et al. 2014, original graph). The data for the model including cross-linker molecules can be observed to correlate with the empirical values obtained, while the values for the monomer and template alone are less reliable.

Molecular dynamics (MD) exists as an alternative, permitting analysis of the whole pre-polymerization mixture and its evolution through time to a point of equilibrium, providing the most accurate model of the system currently available. Assumptions required in analysis of the monomer-template complex alone relating to its formation in solution and lack of response to other system components are not necessary in MD, and for this reason have been central in generating important theoretical MIP formation models and aiding in sophisticated design. Golker *et al.* have been consistent in producing quality work in this area, observing prepolymerization arrangements of thousands of molecules over many nanoseconds, with chemometric analysis and quantified hydrogen bond quality rankings, for demonstrable system reproduction and predictive capacity.<sup>120-122</sup>

This research with MD provides detailed analysis of the prepolymerization system, and therefore provides information on the likelihood of successful imprinting based on the M-T complex. It does not however provide a model of the polymer that is ultimately formed. Though the prepolymerization arrangement formed of components resulting from MD is accurate relative to the common alternatives, it cannot be assumed that binding affinities observed with system variation will be proportional to those observed empirically with polymers. For a more accurate model of imprinted polymers, it will be necessary to model imprinted polymers, and not monomer-template complexes or prepolymerization mixtures of reagents.

### 3.2. Molecular models of polymers

While similar work on non-imprinted sorbents predates this approach,<sup>231</sup> the first attempts to attempt to explicitly apply molecular modeling to the interaction between a MIP binding site and a specific target can likely be attributed to Rathbone and Ge in 2000.<sup>232</sup> Using AM1 semi-empirical calculations, a fluorescent monomer was modified to approximate the polymer equivalent structure (Figure 3.2) in an imprinted polymer specific to a set of antitubercular agents (N<sup>1</sup>-benzylidene pyridine-2-carboxamidrazones).

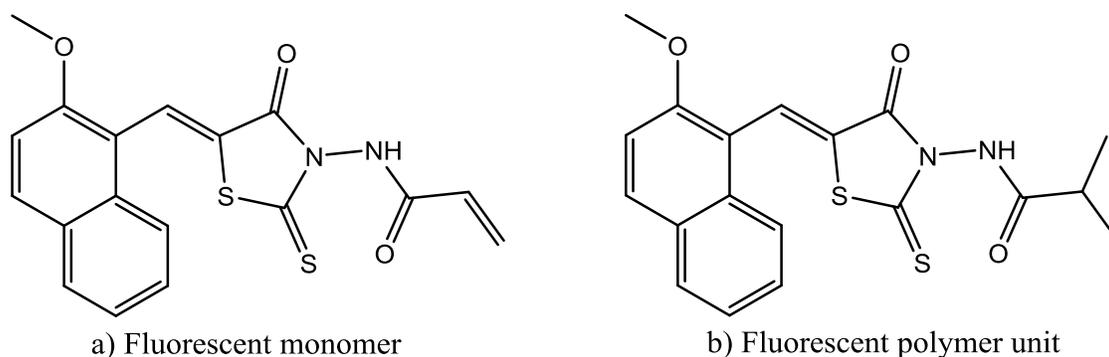
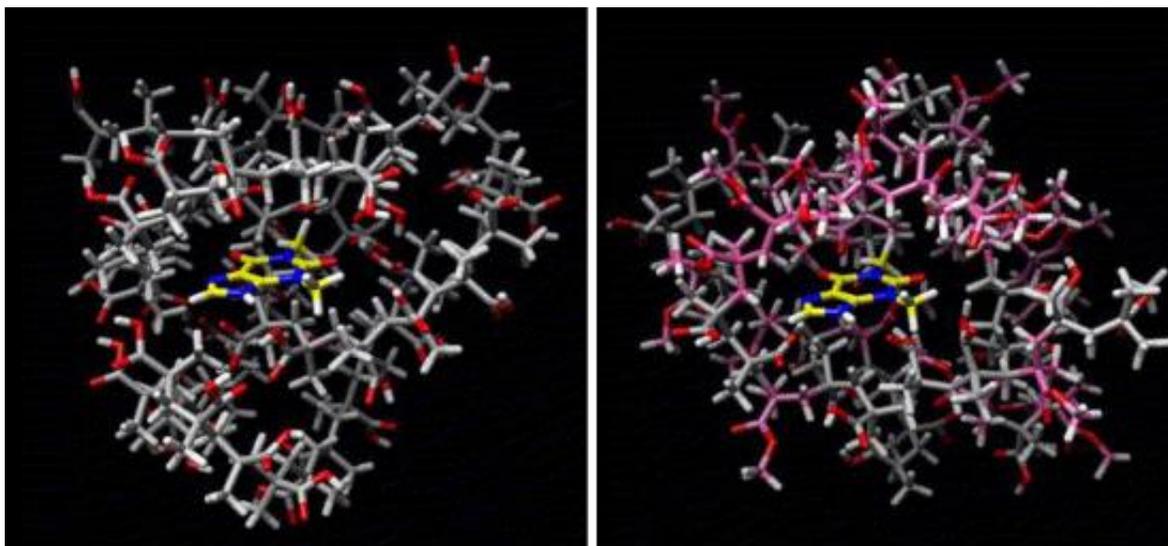


Figure 3.2: An early example of a simple approach to approximating polymer-template/target interactions by modifying the monomer to better represent the equivalent polymer unit.

Such studies are now relatively common with modified monomers, typically saturated alkenes, but the trend towards QM based techniques has stunted this area of research; the major advantage of MM and MD over QM is the ability to easily model large systems of molecules. In the past this has allowed researchers to build large oligomeric structures to aid in design and analysis, as in the work of Monti *et al.* who used these techniques to construct 50 unit oligomers composed of high affinity monomers to observe the interactions with their target molecule.<sup>233</sup>

The MIP designed by Monti *et al.* was designed for theophylline, a drug used to treat asthma and other respiratory diseases, using the General Amber Force Field (GAFF). In this work the group attempted to demonstrate the possible interactions the template theophylline had with the polymer, in addition to simply monomers, by

creating oligomer chains from 50 functional monomers. These oligomers each had a different composition or conformation, some of which were bound in the manner they were predicted to be found around the template in the monomer studies. The systems were then minimized with one molecule of theophylline, giving approximately 100 different conformations, the lowest energy forms of which are shown in Figure 3.3. These 50 unit polymers were found to be selective for theophylline over its analogues, with the homopolymer performing slightly better (2 kcal mol<sup>-1</sup> difference in total interaction energy and with greater selectivity) than the copolymer.

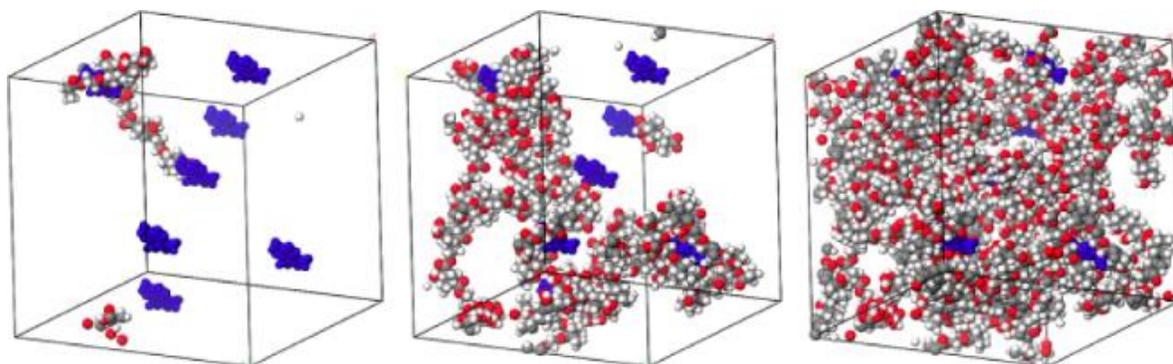


*Figure 3.3: Polymers constructed from 50 individual monomers and observed for interaction with theophylline. The figure on the left shows a polymer composed of MAA alone, and the right a copolymer of MAA and MMA (Monti et al., 2006, reproduced with permission).<sup>233</sup>*

More recent examples of modeling polymers to this level for MIP analysis are somewhat rare and vary considerably in their approach. Luo *et al* created a cubic mesh of poly(methacrylic acid) to model the interaction involved in a cholesterol imprinted polymer using MD and RDF, observing the effects of altering various factors in the model.<sup>162, 234</sup> Via a completely different approach, Huynh *et al.* simulated the polymerization process in the design of a 6-thioguanine specific synthetic receptor, using monomers which combined a large thiophene based segments for radical polymerization attached to a cytosyl functionality for binding to the target.<sup>143</sup> Molecular dynamics simulations were performed using eight template molecules with the equivalent number of monomers and cross-linkers, and accompanied by the addition of bonds between thiophene moieties when they came within 0.3 nm of each other. The analysis suggested that the complex and resultant polymer network were stable despite the large monomer units, and this was supported by fluorescence titration. The sensor produced was found to have a detection limit of 8  $\mu\text{M}$  for the target 6-thioguanine, and a sensitivity to this compound several times greater than that for all the observed target metabolites.

Among the most interesting articles relating to modeling of imprinted polymers published in recent years was that of Schauerl and Lewis in 2015, in which a polymer was grown around a template via an evolutionary process.<sup>129</sup> The researchers employed the program ZEBEDDE (zeolites by evolutionary *de novo* design) to randomly select monomers (functional and cross-linking, in a selected proportion) which are then bonded to the growing oligomer in a head-to-tail arrangement, with the growth being allowed only if the interactions between the template and the polymer are favorable.<sup>235</sup> Nicotine and theophylline were used as template molecules with MAA and EGDMA monomers in this simulation, and the polymer was allowed to grow until the density of the box reached  $0.65 \text{ g cm}^{-3}$ , a value likely similar to the surface of an MAA/EGDMA polymer. The resulting structures (shown in Figure 3.4) were then minimized, with or without prior dynamics simulation, and the adsorption of the targets were analyzed, giving good correlation with empirical data. This research demonstrates the closest

approximation to a polymerization process relevant to atomistic modeling of MIPs produced to that time.



*Figure 3.4: The polymerization simulation applied by Schauerl and Lewis, which incorporates a zeolite modeling tool to grow a polymer via an evolutionary mechanism (reproduced with permission).*

However, these examples of *in silico* polymer synthesis still exclude the appropriate level of detail required for a representative model. MD is required first to build an effective reproduction of the system observed; in readily including all chemical components and simulating the interaction of those components at a given temperature and pressure over time. MD is therefore required to accurately represent the pre-polymerization solution at a molecular level, and any attempt at decreasing the precision (representing molecules as single particles for example) or increasing the precision (e.g. QM methods) lead to a loss of accuracy in representing the intermolecular interactions by over-simplification.

### 3.3. *Quantum mechanical polymerization*

Technological development in computational hardware and software combined with the continued refinement of theoretical modeling methods has led to the combination of the classical and quantum approaches in quantum mechanics/molecular mechanics (QM/MM). Modeling programs, especially those which emphasize the quantum aspects of the model, increasingly contain algorithms which allow molecular simulations using each of these techniques. However, these studies include the use of each *sequentially*, while QM/MM uses quantum mechanics and molecular mechanics simultaneously, allowing a more accurate representation of the system. While there are accounts of quantum mechanics and molecular mechanics or ‘molecular dynamics/quantum mechanics’ in the literature, these do not combine the two (acceptance and sophistication of QM/MM modeling being a relatively recent development) and therefore do not produce results significantly different to those which are given by previously described methods.<sup>236, 237</sup> The importance of Quantum Mechanics/Molecular Mechanics is found in its capacity to simulate processes which quantum mechanics or molecular mechanics (or the use of both independently) cannot achieve alone.

Reported use of QM/MM in the design of MIPs is rare. Notable in this area was the Barkaline group, who in 2013 applied these techniques in their preparation of a tri-O-acetyladenosine (TOAA, an organically soluble ATP analogue) MIP.<sup>238</sup> In this work, the TOAA template and intermolecularly binding groups of the functional monomers were fitted with the 6-31G\*\* hybrid functional basis set while the remaining atoms were represented by an Amber force field, so as to better describe the hydrogen bonding which occurs in the complex than would be achieved using only molecular mechanics. The possible importance of this research, beyond its novel technique, was the allusion to modeling the polymerization process. As the researchers explain, the increase in efficiency found in using QM/MM becomes important in the second stage of the modeling process, in which cross-linkers are added to the system and the polymerization simulation is performed. Barkaline however does not elaborate on the

details regarding the polymerization process and has not produced any further material relating to this topic.

The QM/MM method involving the application of QM theory to directly reacting atoms and classical methods to the remainder of the system has been documented as far back as 1976, when Warshel and Levitt studies the enzymatic reactions of lysozyme.<sup>239</sup> The method was largely ignored however until the 1990s, but has since developed into a established tool in modeling biomolecular systems, in which DFT is usually used as the quantum method (or increasingly SCC-DFTB, Self-Consistent-Charge Density-Functional Tight-Binding, a DFT inspired semi-empirical method) in combination with either Amber, CHARMM or GROMOS force fields.<sup>240</sup> A similar technique to QM/MM has also emerged in the form of reactive force fields, which work on the same principles but emphasize bond order and demonstrate greater efficiency, though they tend to require relatively extensive tailoring to specific applications.<sup>241, 242</sup>

While this might appear to provide an answer however, this is unlikely to be suitable for *in silico* MIP synthesis. The reactions observed by QM techniques are constructed in such a way that possible transition states can be observed, comparisons of the energies then give information about the likelihood of a reaction occurring and the properties of the product. While this is useful to many areas of research (and beyond the capacities of MM), it is inappropriate for the required application. The interest with regard to MIP design is the interaction of some analyte (template, target or otherwise) with an imprinted polymer, and the synthesis of imprinted polymer binding sites. The former requirement may benefit from QM/MM, but only assuming a polymer has already been created. The latter point, regarding the synthesis of polymers, would be useful, but would require every reactive molecule to be analyzed by QM. This would also need to be done manually, which in addition to the fitting with QM functions would eliminate any potential benefits by massively reducing the size of the system that can be practically analyzed.

Related to the scale of the system observed is the size of the ultimate polymer. Conformational analysis can be an expensive process if appropriate methods are not employed, as the total number of bonding interactions increases with each additional reaction. This includes not just the potential energy associated with new bond but also the bond angles (1-3 interactions) and more importantly the torsional strains (1-4 interactions). Considering a polymer composed of saturated carbon-carbon bonds, if only the minima and maxima are acknowledged in the different conformations, which would be an unsuitable simplification, this still results in 6 geometries about the bond. The number of conformations to be processed will then increase rapidly with the size of the system, as a molecule will have  $6^n$  different 1-4 torsional conformations, where  $n$  is the number of rotatable bonds. MM/MD methods can be used in these situations, but with a given number of atoms, the probability of encountering limitations will increase with the number of bonds, i.e. effective simulation of a pre-polymerization system is no guarantee that the equivalent polymer can also be observed. QM cannot be applied on this scale, and molecular mechanics becomes the only suitable answer.

### 3.4. *Manual and automated polymerization*

Effective simulation of the prepolymerization system is the primary requirement for simulating the imprinted polymer network. In the design of polymeric adsorbents the specific arrangement of molecules in equilibrium is likely to be of less overall consequence, as a polymer of given single monomer will inherently adopt a narrow range of structures, on the microscopic scale, regardless of the conditions in which it is formed. With the inclusion of greater variability between individual reactive molecules (including two or more monomers for example) some equilibration experiments would be beneficial in predicting the structure of the product, particularly as this will likely vary with solvent etc. MIPs require the formation of a monomer-template complex, in which an equilibrium exists for each component between the complexed and fully solvated state, in addition to any interactions between like particles and additional system components (cross-linkers for example). For this reason functional monomers are selected to maximize the relative affinity between the monomer and the template. The equilibration of the observed system under conditions that most closely mimic the natural system is therefore paramount.

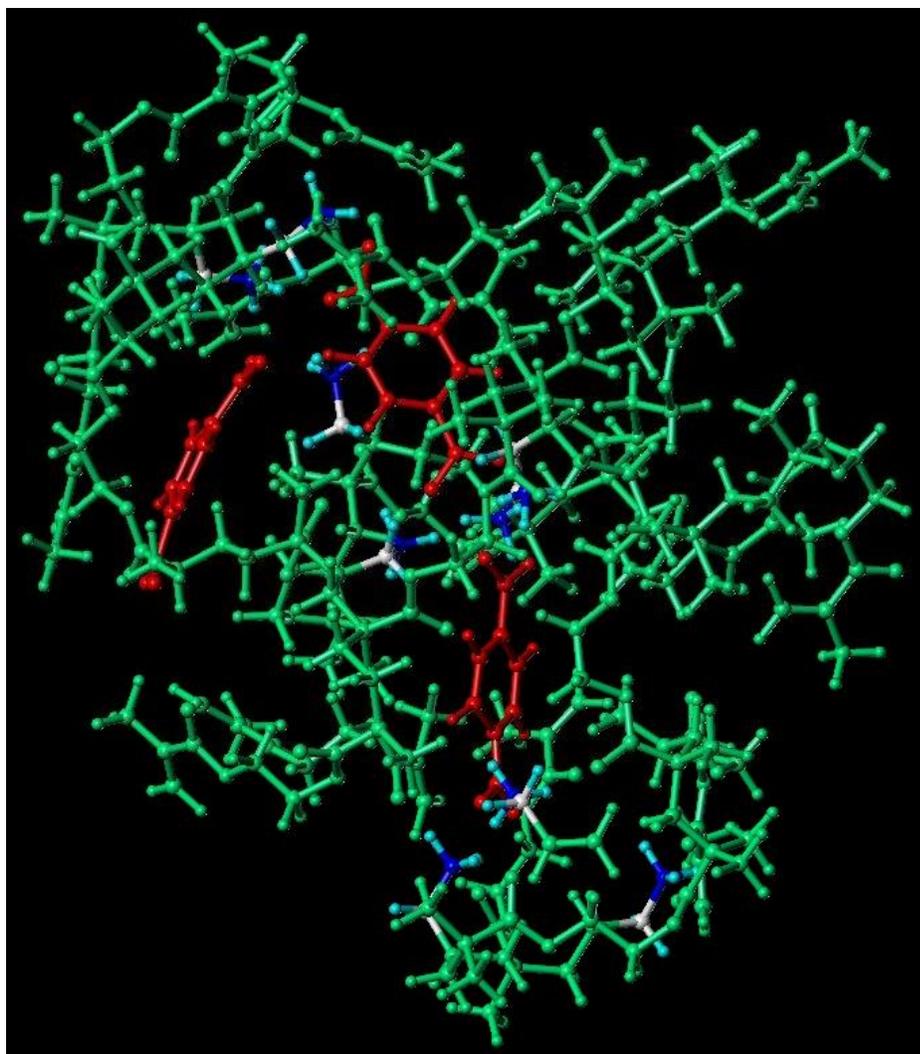
MD software are typically not capable of simulating chemical reactions, but approximations of the polymerization process have been performed by cross-linking reactive groups which lie in close proximity ( $< 6 \text{ \AA}$ ) and without steric resistance.<sup>243,244</sup> This is most likely what was actually performed by Barkaline *et al.* in their QM/MM studies. Attempts to reproduce this approach with equilibrated systems and manual cross-linking demonstrated the difficulties in developing this procedure as a practical modeling technique. Ignoring the time requirements to process even a small number of molecules, the manual step-wise approach requires that each individual stage of the reaction (bond breaking, bond formation, changes of hybridization, etc.) is represented in isolation, resulting in high energy structures unsuitable for representation by the force field.

Initial attempts at manual network formation were promising however. A simple composition was prepared which included charged allylamine monomers, benzene-1,4-dicarboxylate template molecules and cross-linkers to evaluate the plausibility of

simulated polymerization. This process was aided by (and later experiments would be near impossible without) an SPL algorithm previously written which allows the random mixing of selected molecules in quantities determined by the user.\* No equilibration was performed prior to manual cross-linking, which involved selection of possibly reactive atoms in an attempt to approximate a radical polymerization path. The small oligomer produced from this process showed an obvious binding site, which was maintained, with template bound, after 100 ps MD simulation at 300 K under low pressure, an environment likely to draw any weakly associated template away from the MIP (Figure 3.4). Removal of the template revealed that these benzene-1,4-dicarboxylate target molecules were supporting the polymer, molecular dynamics simulations and energy minimization after removal causing a loss of structure. This appears to be primarily a consequence of size, further cross-linking with a greater number of monomers likely resulting in a more rigid structure. Nevertheless, the core binding sites were apparently retained, as after reintroduction of six target molecules to the polymer and a simulation of 300 K for 500 ps, the benzene-1,4-dicarboxylates were found back within the polymer bound in a similar arrangement.

---

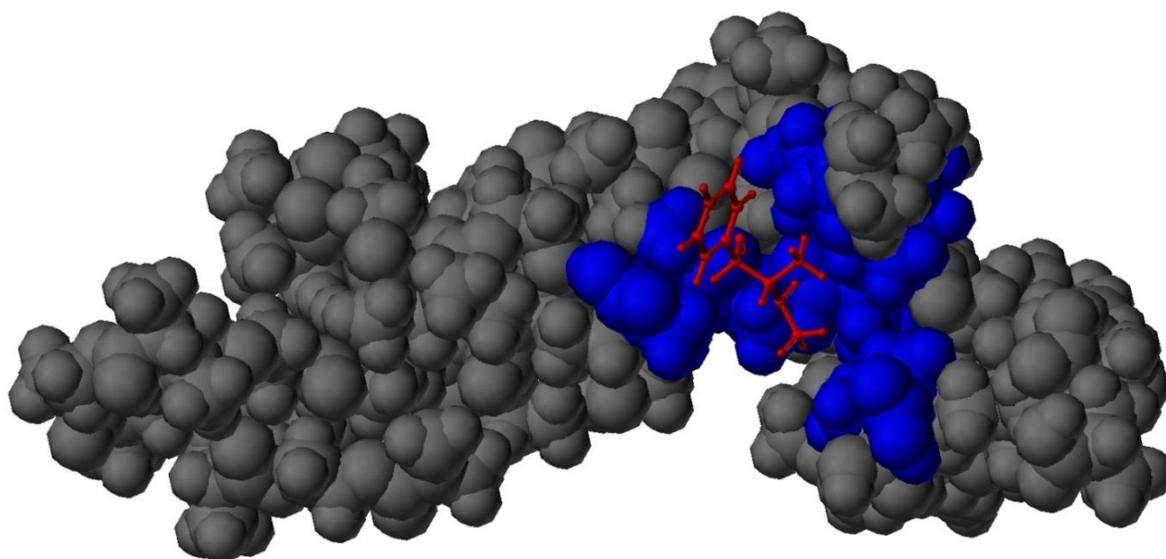
\*Credits for writing this script belong to Mirko Busato (Verona) and Dr Kal Karim.



*Figure 3.4: MIP produced manually using benzene-1,4-dicarboxylate as template (red) and allylamine functional monomers*

This composition however included highly charged monomers, in addition to oppositely charged template molecules, meaning a strong bias in favor of complexation and template rebinding. More realistic composition were selected with methacrylic acid and the hydroxyethyl methacrylate systems designed for ephedrine imprinted polymers, modeled according to their ratios applied in published work, equivalent to only one template molecule.<sup>52</sup> The systems underwent a dynamics experiment simulating 300 K for 100 ps before minimization, after which monomer alkene atoms closer than 6 Å were joined, with all polymerization reactions starting from the edge of

the droplet. On completion the system was reminimized, followed by 10 ps MD at 300 K, minimization, and a final observation for potential polymerization before initiating a final simulation of 300 K for 100 ps and a final minimization. During the section of the procedure after polymerization the movement of the template was observed to determine the validity of a proposed binding site and the system was 'washed' of monomers unbound to the primary polymer. The resulting structure of the methacrylic acid based polymer is shown in Figure 3.5 as an example.



*Figure 3.5: Manually produced methacrylic acid based polymer with ephedrine (red). Solvent chloroform is explicitly present but not displayed.*

While the targets was found to dwell predominantly around the binding sites of the polymers on dynamics simulation the affinity appeared relatively low, the ephedrine readily drifting away with the influence of the chloroform solvent, though binding in the imprint was found to be more favorable than in other regions. These studies suggested that progression onto writing an algorithm for application in the applied software with the described techniques was justified.

### 3.5. *Development of a polymerization algorithm*

The Sybyl software package has advantages beyond the previously described screening techniques (see Sections 2.4 and 2.5). While the program has a standard user interface allowing for typical experiments, additional operations can also be issued from the command line when given in Sybyl programming language (SPL). Commands entered in the Sybyl terminal can act as shortcuts to common actions (entering 'energy' for example to calculate the energy of the current system) and in following standard commands ('echo hello world'), including looping, if/else statements, etc. SPL scripts can also be written separately, retrieved and executed from the within the terminal. More important is the relative ease of manipulating the molecular structure by modifying atoms in type and position, their bonds and properties.

The polymerization algorithm was developed and published in 2016,<sup>245</sup> and the explanation of how it functions is explained schematically in Figure 3.6. The full script given in Appendix 2.1, contains approximately 600 lines of code, including many prompts for the user to control the many variables. On reading into the Sybyl terminal the user will be asked first to provide a filename, which is then used with suffixes to save and relevant points, and the option to run simulated annealing and equilibration dynamics. If selected the high and low temperatures of the simulated annealing, length of each period, gradient (temperature step size) of annealing and number of annealing cycles are all then given as required, along with the length of a final 300 K equilibration.

After the equilibration has been completed the polymerization simulation begins automatically from variables given prior to MD. This process begins from an automatically selected reactive molecule or by an initiator, as selected by the user. If applied via an initiator the simulation will proceed to follow the course of the radical reaction pathway, molecules remaining intact and untouched unless presented with an existing radical originating from the initiator. If an initiator is not selected, reactive molecules are selected sequentially and become radicals as if undergoing photolysis. This method is generally preferred, likely permitting the observed system to better represent a small section of a much larger prepolymerization mixture, where the course of reaction originates from outside.

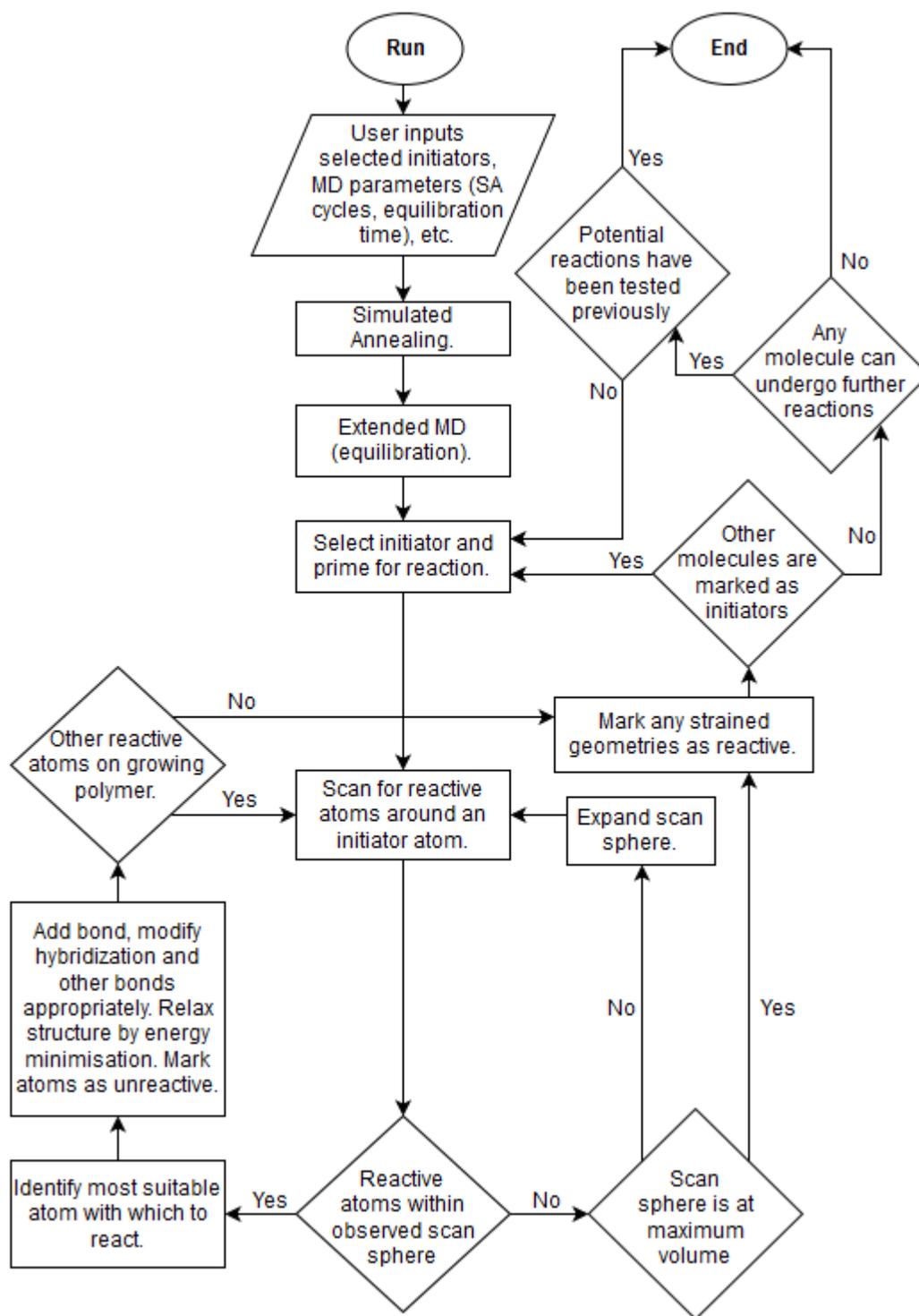


Figure 3.6: Explanation of the functioning of the polymerization algorithm. This scheme and the following images were originally published in *Macromolecular Rapid Communications* (Cowen, 2016).<sup>245</sup>

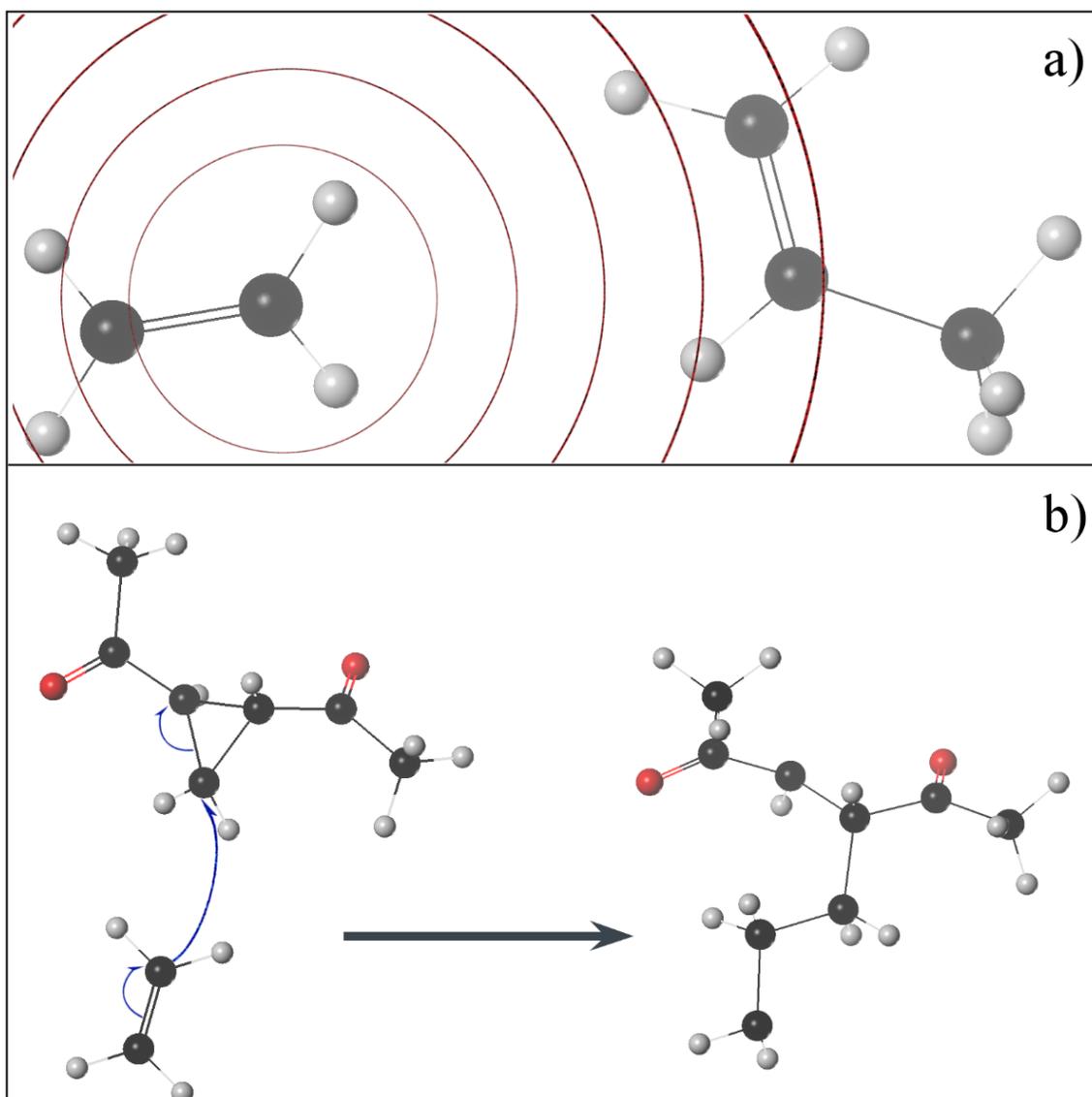


Figure 3.7: a) the reactive atom of the assigned molecule scans the surroundings in concentric spheres until an appropriately reactive atom is found with which to form a bond. Bond formation is favored with closer atoms, if multiple atoms are found within the same interval the more reactive is favored; b) reactions with sterically strained structures occur in the absence of suitable alkenes. Strained structures will not spontaneously break or react with each other.

Each reactive atom of the assigned molecule then scans the surroundings in concentric spheres until an appropriately reactive atom is found with which to form a bond. This new bond may form between separate molecules or between atoms of the same oligomers, with some restrictions. Highly strained structures are inhibited, as are illogical actions such as bonding to an already bound atom.

On successful bonding the atoms must be prevented from further reactions, the commands of the algorithm dominating any inhibitions of the Sybyl program (hypervalent atoms were common in development of the algorithm for example). In the current version, the user must give a maximum bonding distance, i.e. a largest sphere volume within which reaction can occur. Initial studies used a maximum distance of 6 Å, a value obtained from the literature,<sup>243, 244</sup> but the high energy of the resulting structures suggested a lower maximum distance was required for the systems observed. For the published work and analysis given in Chapter 3 the maximum bonding distance of 4.8 Å was used in all experiments. This value was determined by repeating the polymerization for various systems with different maximum bonding distances and calculating the final potential energy of the polymer system relative to the number of new bonds formed, therefore giving an indication of the most stable maximum bonding distance.

A more refined approach was applied to later experiments based on the intermolecular potential of the interacting components. As the algorithm was designed to simulate the radical polymerization of alkenes, two ethene molecules were used as a simple general representation of likely interactions of reactive groups. The two molecules were held parallel at a distance and the internal energy recorded as  $U_{\infty}$ . The molecules were then brought together with the energy associated with the intermolecular distance being recorded at regular intervals. The molecules were then rotated by 45 ° and then 90 ° in the plane or perpendicular to the plane of the interaction, giving a representation of the interaction from different approaches. The intermolecular energy approximately followed a Lennard-Jones potential, represented by Equation 3.1:

$$U_{1,2} = \frac{4\epsilon\sigma^{12}}{r_{1,2}^{12}} - \frac{4\epsilon\sigma^6}{r_{1,2}^6} \quad \text{Eq. 3.1}$$

As the intermolecular distance,  $r_{1,2}$ , is reduced, the interaction potential,  $U_{1,2}$ , decreases, representing an attractive interaction. The energy well reaches a maximum interaction energy of  $\varepsilon$ , after which  $U_{1,2}$  increases rapidly and the interaction decreases and becomes repulsive. As  $r_{1,2}$  decreases beyond  $\varepsilon$  the increase in  $U_{1,2}$  passes from negative to positive; the distance at which  $U_{1,2}$  is zero, i.e. at which  $U = U_{\infty}$ , is given by  $\sigma$ . Using a simple script (Appendix 2.3) the intermolecular potential for two ethene molecules was used to determine the different variables for Equation 3.1, with  $\varepsilon = 3.384$  kJ mol<sup>-1</sup> and  $\sigma = 3.633$  Å. From this a value of  $0.25\varepsilon$  was set as a minimum interaction required for reaction between atoms; this value was equal to 0.846 kJ mol<sup>-1</sup>, and occurred at a distance of 5.701 Å. This value was then applied as the maximum bonding distance, but future development of the polymerization algorithm will focus on making bonding probabilistic, dependent on  $r_{1,2}$ .

On scanning the volume surrounding a reactive atom, a bond will be formed with any reactive atom found within the observed sphere. On bond formation the atom hybridization is appropriately modified, any other bonds are changed accordingly, and the geometries are relaxed via a high termination gradient energy minimization of 2.062 kJ mol<sup>-1</sup> Å<sup>-1</sup> (0.5 kcal mol<sup>-1</sup> Å<sup>-1</sup>). The bonded atoms are then recorded as non-radical, but the reaction continues by any adjacent radical atom created by the previous reaction. If on scanning two atoms are approximately equally distant from the reactive atom (Figure 3.7a), bonding will occur with that assigned as more reactive in the original simulation preparation, the more reactive (least hindered) atom being named 'alpha' and the other 'beta'.

While reactions producing strained geometries are inhibited, in the absence of more suitable pathways these may occur. The system is regularly scanned however to identify these strained geometries. Typically these are three or four member ring, and are designed as still potentially reactive, though are considered less preferable than existing alkenes within range. Bonding can therefore occur between radical atoms and high energy groups in the absence of more suitable alternatives (Figure 3.7b) and the reaction propagated via the radical produced from the previously strained atoms. Strained groups are found by routine scanning, and labeled so as to be regarded as 'semi-reactive'. Bonds in strained groups will not break spontaneously, and reactions

will not occur between strained groups; only in the presence of a reactive group will the strained structure break.

### 3.6. Volume requirements

Molecular dynamics requires an appropriate volume, ‘box’, within which the simulation can occur. To establish appropriate volumes for polymerization reactions and other observations, a simple approximation was required which could be quickly applied. The result was Equation 3.2:

$$Volume\ T^{-1},\ nm^3 = \sum_i \frac{n_i M_{r_i} 10^{21}}{n_T N_A p_i} \quad \text{Eq. 3.2}$$

Where index  $i$  refers to the component (template, functional monomer, crosslinker or solvent), thus  $n_i$  is the number of moles of the component,  $M_{r_i}$  its relative mass and  $p_i$  its pure density, while  $n_T$  is the number of moles of template and  $N_A$  is the Avogadro constant. The volume is calculated per template molecule ( $T$ ) to facilitate scaling. Application of this formula will inevitably produce error due the volume occupied by each molecule being different in its pure form and in a mixture, but as a simple approximation this seems to be effective. References to this approach to volume calculation, or equations giving equivalent or improved approximations of the volume while retaining comparably low resource demands in calculation, could not be found in the literature (though they are likely available).

The previously mentioned Busato script can be used to produce a required mixture of molecules. The template, functional monomers, cross-linker, solvent, etc. are added to different molecular areas (individual systems) and the script is run, containing prompts for the number of each area to be included in the final single system. The molecules are added randomly, forming a sphere. An additional script was then required to rearrange the molecules to fit within the cubic box with dimensions determined from Equation 3.2. The resulting ‘auto-move’ script (Appendix 2.2) compresses all molecules into dimensions set by the user, which followed by minimization fill the box to give the appropriate conditions to begin the simulation.

### 3.7. Experiments with ephedrine imprinted polymers

The original Piletsky *et al.* experiments applying automated computational screening used (-)-ephedrine as a template and target in the design of an imprinted polymer.<sup>52</sup> All modeling was performed using the Sybyl 7.3 software package on an HP EliteDesk G1 Tower PC running CentOS Linux 7. Polymerization mixtures were prepared based on the original compositions, equal to five template molecules each, with the mixtures varying primarily by the functional monomer used. System A included methacrylic acid (MAA) as functional monomer, B used hydroxyethyl methacrylate (HEM), and C used 2-vinylpyridine (2-VP). The systems had 50 functional monomers, 104, 149 and 124 cross-linker EGDMA and 216, 309 and 256 solvent chloroform molecules for A, B and C respectively, following the composition in the empirical investigation. Applying Equation 3.2, Mixture A was found to have a volume of 68.506 nm<sup>3</sup>, B 99.077 nm<sup>3</sup>, and C 82.997 nm<sup>3</sup>, with each box being a cube.

5 cycles of stepwise annealing from 800 K to 300 K with 100 K intervals and 10,000 fs at each temperature were used before 2 ns equilibration periods at 300 K. All molecular dynamics were performed using an NTV canonical ensemble with the Tripos force field, MMFF94 charges and a non-bonding interaction cut-off distance of 8 Å. The inclusion of explicit solvent molecules allowed the dielectric constant to be set at one, and the initial velocities were selected from the Maxwell-Boltzmann distribution given for the relevant temperature. Polymerization was set to regard all reactive molecules as potential initiators, with scan spheres increasing in 10 pm increments to maximum bonding distance of 4.8 Å. Affinity analysis was performed on both the pre-polymerization and polymer systems by a combination of the direct measurement of the association energy of the system and by use of Surflex docking software. Inclusion of both sets of empirical data provides some element of consilience, the mesoscopic polymer properties (imprint density, extent of cavities within the polymer, etc.) which may positively or negatively affect each composition being cancelled to some extent by the different measurements.

The association energy,  $\Delta E$ , was calculated according to Equation 3.3, a modification of Equation 1.1:

$$\Delta E = E_{system} - (E_{template} + E_{system-template}) \quad \text{Eq. 3.3}$$

Each template molecule was removed for analysis of  $E_{template}$  and  $E_{system-template}$  and replaced sequentially for each system. The lowest value of  $\Delta E$  from each was then multiplied with the Surflex-dock determined dissociation constant  $K_D$  by Equation 3.4:

$$affinity\ score = \frac{\log(K_D)\Delta E}{100} \quad \text{Eq. 3.4}$$

The affinity score was then used to evaluate both the pre-polymerization mixture and the final polymer and compared with empirical data and LeapFrog based monomer rankings. The docking program was also used alone to predict the selectivity of the polymer and the pre-polymerization system. Here the selectivity value is given by Equation 3.5:

$$selectivity\ factor = \frac{\log(K_D)_{(-)}}{\log(K_D)_{(+)}} \quad \text{Eq. 3.5}$$

The numerator being the affinity of the (-)-ephedrine target and the denominator being that of the (+)-ephedrine enantiomer.

	A (MAA)	B (HEM)	C (2-VP)
Empirical MIP imprinting factor	2.87	2.67	1.00
Empirical MIP capacity factor	9.46	0.80	0.10
LeapFrog binding score, kcal mol <sup>-1</sup>	-14.62	-15.72	-1.82
Pre-polymerization affinity score	-4.093	-3.162	-3.382
Polymer affinity score	-4.530	-3.165	-2.316

*Table 3.1: Empirical affinity results and theoretical predictions with three MIP compositions. Empirical values and LeapFrog relative binding score were retrieved from Piletsky *et al.*, and pre-polymerization and polymer affinity scores were obtained using Equation 3.4. The abbreviations in parentheses are the functional monomers used in the syntheses: methacrylic acid (MAA), hydroxyethyl methacrylate (HEM), 2-vinylpyridine (2-VP).*

The theoretical results in Table 3.1 (the LeapFrog binding score and affinity scores) are all presented in terms of relative energetic stability with the more negative value representing stronger binding. The LeapFrog screening results from Piletsky *et al.* demonstrate the utility of the approach in eliminating the weakly interacting functional monomer (2-VP). This screening however incorrectly predicts that composition B would result in the strongest binding MIP. The pre-polymerization affinity score shows the interactions observed in the pre-polymerization equilibrated system, and the most effective mixture is correctly identified as that with the MAA monomers (A), but the other two compositions are incorrectly ordered in terms of relative affinity displayed. The score for C is of potential interest here in being relatively high and showing that the strength of the complex comes from interactions which are not present or reduced in the final binding cavity. The values obtained for each polymer model, produced with the polymerization algorithm, can be observed to follow the results obtained in the empirical investigation, with polymer C showing the weakest affinity and polymer A showing the greatest affinity for the (-)-ephedrine target.

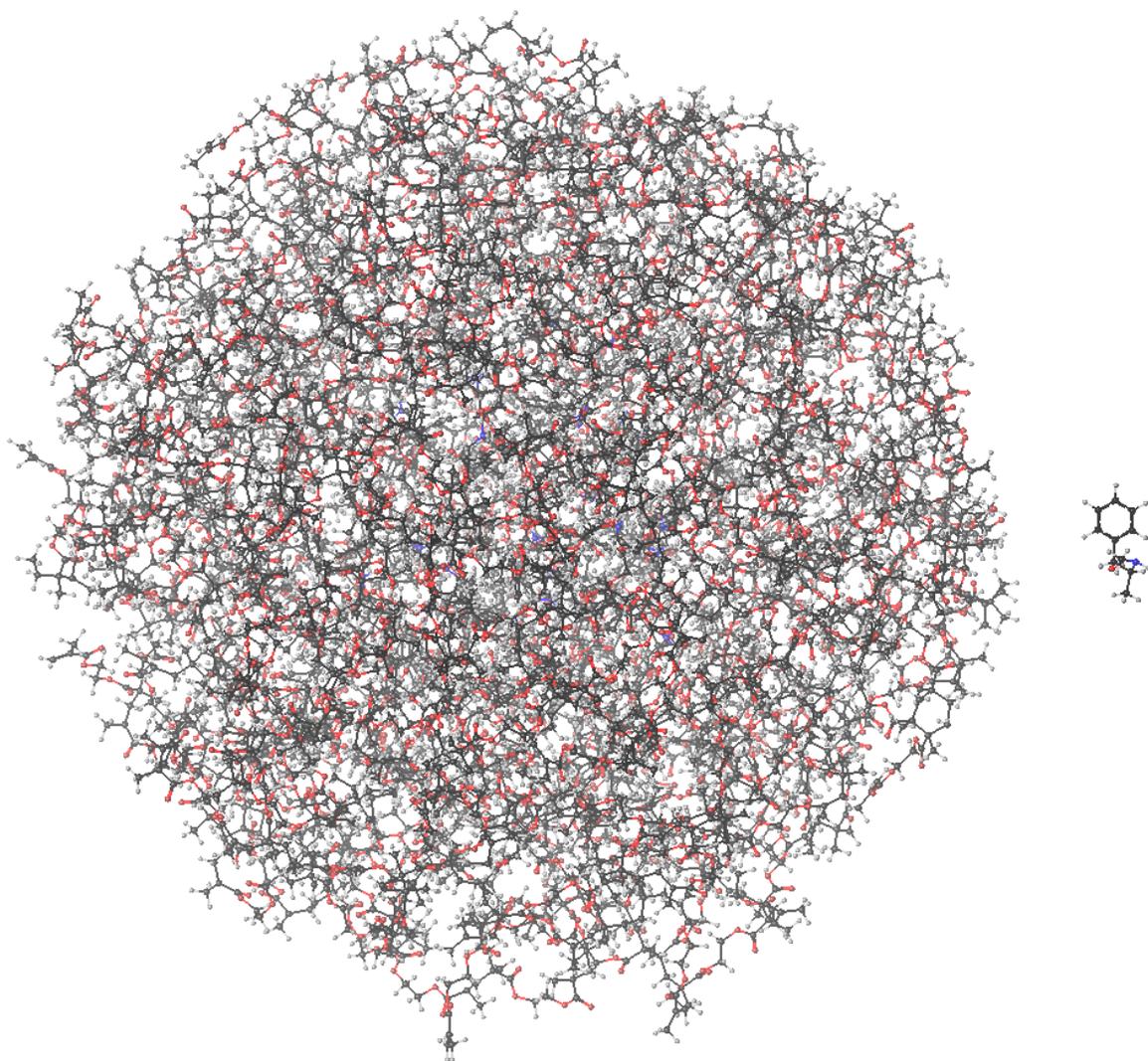
Binding site selectivity was observed with the Surflex docking software on both the pre-polymerization system and the resulting polymer with both the template (-)-ephedrine and its enantiomer (+)-ephedrine. The results are given in Table 3.2.

	A (MAA)	B (HEM)	C (2-VP)
Empirical selectivity $k^{\text{(-)}}/k^{\text{(+)}}$	1.34	1.42	1.00
Pre-polymerization selectivity factor	1.13	1.25	1.19
Polymer selectivity factor	1.54	1.57	1.04

*Table 3.2: Comparison of empirical selectivity, theoretical selectivity based on the pre-polymerization system and theoretical selectivity based on the model polymer.*

In the original article the empirical analysis gives HEM-based polymers (B) as giving the greatest selectivity, slightly above that of the MAA MIP (A). 2-VP (C) gives no measurable selectivity. In docking with the equilibrated pre-polymerization mixture, B is correctly identified as giving the greatest selectivity, but this is relatively underestimated and MIP C is predicted to demonstrate greater selectivity than A. The polymer models however perform comparatively well, allowing correct predictions of the order of selectivity between the three MIPs and giving a near equal binding for the two enantiomers in MIP C.

The algorithm therefore effectively mimics radical polymerization to the extent that it produces imprinted polymers molecular model with greater accuracy than alternative methods. Enantioselective analysis *in silico* is relatively rare and so the latter results should be of particular interest to those working in the design of MIPs.<sup>246</sup> This process is also not limited to only small compositions; a 104 kDa polymeric macromolecule (approximately that of a small MIP NP<sup>247</sup>) formed from 889 new bonds between EGDMA and MAA monomers is given in Figure 3.8, with an additional ephedrine molecule for comparison.



*Figure 3.8: a 104 kDa macromolecule synthesized in silico using the polymerization algorithm, with an ephedrine molecule for comparison.<sup>245</sup>*

### 3.8. *Conclusions*

The polymerization algorithm developed and described previously was shown to effectively produce molecularly imprinted polymers *in silico*. These polymers could then be analyzed via docking programs and other techniques to produce reliable representations of MIP characteristics. Comparison with template-monomer interactions, the most common method, and equilibrated pre-polymerization systems demonstrated that analysis of polymers gave more accurate predictions of the relative binding given in empirical investigation.

The polymerization algorithm developed gives an atomistic polymer structure that may then be analyzed. This is useful in binding studies and similar investigations but becomes computationally expensive with a larger number of atoms. For macroscopic analysis of imprinted polymers therefore it is necessary to find alternative methods of modeling.

## 4. Size and solubility of polymer nanoparticles

This chapter breaks from the previous discussion of molecular modeling to focus on empirical investigation and thermodynamics. Outlined below is an investigation into the relationship between polymer nanoparticle hydrodynamic diameter and the conditions of synthesis, specifically polymer solubility.\* As described below, control of polymer nanoparticle dimensions is of great importance for any researchers interested in the administration of MIP NP for direct biomedical applications, as described below.

The ultimate objective of the work presented below is to describe how the absolute hydrodynamic diameter of a polymeric particle may be determined from simple theoretical assumptions.

### 4.1. *Limits of atomistic modeling*

Molecular modeling is commonly employed in MIP design, but these approaches are highly limited by the scales appropriate to the models. For example, it is common amongst adherents to MM/MD methods to criticize QM methods for their inability to effectively reproduce solvent and thermodynamic properties, and for adherents to QM to respond with critiques of the overly simplistic functions used in MM to describe electronic effects. However, mesoscopic effects of polymerization are difficult to account for with molecular modeling, while having dramatic effects on the properties of the imprinted polymer. For example, Yang and Zhao performed detailed research into the development and optimization of an electrochemical sensor for chloramphenicol (an illegal antibiotic with application in livestock farming) in which they investigated the effect of using different ratios of template:monomer (T:M) and template:cross-linker (T:#).<sup>248</sup> In typical computational analysis the effect on binding affinity would be the subject of interest in these investigations, but here a more

---

\* This work would not be possible without the results of the preliminary experiments described in section 4.4, the experimental data for which was collected by Alistair Watson and Matthew Young, facilitated by Joanna Czulak and Antonio Guerreiro.

significant effect was found to result from the total number of imprinting sites formed (when the template made up a relatively small fraction of the polymerization mixture, i.e. high T:M and T:# ratios) and the thickness of the polymer on the electrode formed, affecting the conductivity and number of accessible sites, which resulted from higher relative template concentration. The differences in binding site affinity shown between these polymers was negligible in comparison to the variation in mesoscopic properties of the polymers.

Muzyka *et al.* advanced the design protocol for a vancomycin imprinted polymer for use in an immunoassay style analysis by applying statistical techniques to optimize the polymerization mixture.<sup>249</sup> With MODDE 9.0 software the group performed a design of experiment (DOE) procedure to determine the most successful conditions (monomer concentration, time required for irradiation and temperatures required and different points) to maximize the yield of nanoparticles. The group determined that in the solid-phase synthesis of these MIP nanoparticles the optimal conditions include a monomer content of approximately 1.8 – 3.25 % in the polymerization mixture with an irradiation period lasting approximately 2.5 minutes and that removal of low-affinity polymers and unreacted material should be performed at 10 °C. This work built on the researchers' previous chemometric studies into the appropriate concentration of radical polymerization initiator to be included in the reaction mixture, similarly using DOE with MODDE 6.0.<sup>250</sup> Many of these parameters are difficult to account for in molecular modeling.

The mesoscopic polymer properties are often overlooked in MIP modeling, and are often not observed in much detail in general polymer nanoparticle preparation. While the properties of nanoparticles will inevitably be dependent on numerous variables, one of the easiest to control is solubility.

## 4.2. *Introduction to solubility*

Solubility is often referred to in overly simplistic terms, a particular molecule often being described as 'soluble' or 'insoluble' in a particular medium. It is simultaneously commonly acknowledged that solubility exists as a spectrum dependent on the strength of interactions between particles. An approximation is often made in solution theory that the extent of chemical similarity between the two particles is proportional to the likelihood that the two will mix to form a single phase. This is obvious from the example of dissolving a liquid in that same liquid; water mixed into water gives water, and oils into oil.

The total potential energy of intermolecular interaction between components of a given material is described by its cohesive energy, given relative to the (negligible) intermolecular potential energy of the components in the gaseous state.<sup>251, 252</sup> The cohesive energy per volume of material is referred to as the cohesive energy density, and is proportional to the enthalpy of vaporization for many common applications.<sup>253</sup> Considering the above example of water added to water, a specific quantity of energy is required for intermolecular bond breaking to create a cavity that may accommodate an additional water molecule or number of water molecules. As the bonds forms between this additional 'solute' water and the 'solvent' water, an energy is released that is equal to the energy required for the original bond breaking, and the total energy change between the individual molecules is zero (though as an extensive property the total energy will vary with the number of molecules).

With the more practical example of a solute and solvent of different substances, the cohesive energy of each will be different and dependent on the chemistry of those materials, and the energy associated with solvation will equally depend on the bonding energy between the different particles. In these circumstances it is often practical to consider each component the solubility parameter of the material,  $\delta$ , which is equal to the square root of that material's cohesive energy. The solubility of a given solute in a particular solvent is then inversely proportional to the difference in solubility

parameter,  $\Delta\delta$ : 'like dissolves like', due to the similarities of the intermolecular interactions.<sup>251, 252, 254, 255</sup>

During the process of a chemical reaction the solubility of the solutes changes from that of the reactants to that of the product; in polymerization this change may be large, but it is the result of many small changes in solubility associated with forming the polymer. In the case of polymer nanoparticles, this basis has many consequences which have not been examined in detail previously, most importantly in the control of nanoparticle diameter.

### 4.3. *Nanoparticles in vivo*

The potential for applying nanoparticles in drug delivery and diagnostics, and the various ingenious methods, techniques and materials available, have been previously reviewed heavily in the literature.<sup>256-262</sup> Increased attention to the specifics of nanoparticle efficacy, toxicity and preparation for *in vivo* applications are more recent however, an obviously encouraging shift being observable from plausibility to practicality.<sup>263-265</sup> Amongst the more significant areas relate to cellular uptake, systemic circulation and excretion, each of which is dependent on the particle physical structure and chemical composition. The dimensions of nanoparticles intended for biological application are particularly important, particle diameter being inversely proportional to cellular uptake, but also to toxicity and nonspecific cellular absorption.<sup>266, 267</sup>

The optimal particle radius for cellular uptake has been observed as approximately 30-50 nm for various particles in different cells, smaller particles being required to cluster before absorption becomes energetically favorable, while particles larger than approximately 50 nm are hindered by the greater time required for membrane wrapping.<sup>268-271</sup> The rate of uptake however depends on a number of variables, notably the adhesive strength of the particle to the cell surface and the rigidity of the membrane, in addition to any possible receptor interaction and the mechanisms associated with ammonium induced endocytosis.<sup>272, 273</sup> Non-specific

cellular uptake may be increased 100-fold with peptide functionalization.<sup>274, 275</sup> However, the response of the cell to the nanoparticles will depend on both the cell and particle type, sometimes dramatically with small variations.<sup>276-278</sup>

Additionally, for certain applications maximizing the number of particles in the cell may not be desirable. Observations of hollow polymer particles of various diameter gives a maximal absorption number at 50 nm diameter, but a greater total mass of polymer absorbed for particles of 125 nm.<sup>279</sup> Further research has also observed an increase in reactive oxygen species (ROSs) generation, toxicity and cell death with particles below 100 nm, but reduced or zero significant adverse effects with particles over 100 nm.<sup>279, 280</sup> This is likely due to the greater surface area associated with having a larger number of smaller particles. Notable in terms of drug delivery however is that the reverse assumption of reduced practical efficacy in delivery with larger particles are generally unfounded, particle diameter typically correlating with higher drug release rates.<sup>280</sup>

However, while some success has been demonstrated applying polymer particles of approximately 250 nm,<sup>281, 282</sup> an upper limit of 200 nm has been suggested for particles applied *in vivo* to allow effective removal from the circulatory system and avoid accumulation in the spleen.<sup>260, 264, 283, 284</sup> Recent successes with polymer nanoparticles in suppressing tumor growth in animal models further support an ideal size of approximately 100nm.<sup>285, 286</sup> Doubtless the preferred particle diameter will vary with the application, but an effective basis for the control of dimensions is necessary to appropriately respond to such requirements.

The general trends between diameter and solubility in synthesis in the preparation of polymer nanoparticles has been noted previously, but in the absence of thermodynamic considerations this cannot account for several differences observed in practice.<sup>287-290</sup> Solubility parameters can be combined with Flory-Huggins theory (FHT) in the analysis of polymer nanoparticles, treating each as a phase with a critical point reached during polymerization, the prepolymerization mixture being analogous to a critical state.<sup>291</sup> Inclusion of FHT allows some predictions related to thermodynamic potentials, which in turn provides theoretical basis for the greater understanding of observed behavior of polymeric systems.

#### 4.4. Investigation of solubility parameters

Nanoparticles were prepared based on a protocol used in the synthesis of polymers tested *in vivo*.<sup>36, 44, 292</sup> Preliminary experiments were performed with both chemical and photopolymerization, various solvents, and nucleation adjacent to a solid surface, to produce a broad range of polymer nanoparticles which could be analyzed by DLS (dynamic light scattering). The results shown in Figure 4.1 demonstrate the clear correlation between the size of nanoparticles and the relative solubility (expressed in terms of the Flory parameter) of the polymer. Dependence of the diameter of nanoparticles on solvent viscosity was not observed, suggesting a primarily thermodynamic, and not kinetic basis for the differences in particle dimensions.<sup>293, 294</sup> Details of the experimental methodology are given in Appendix 3.3.

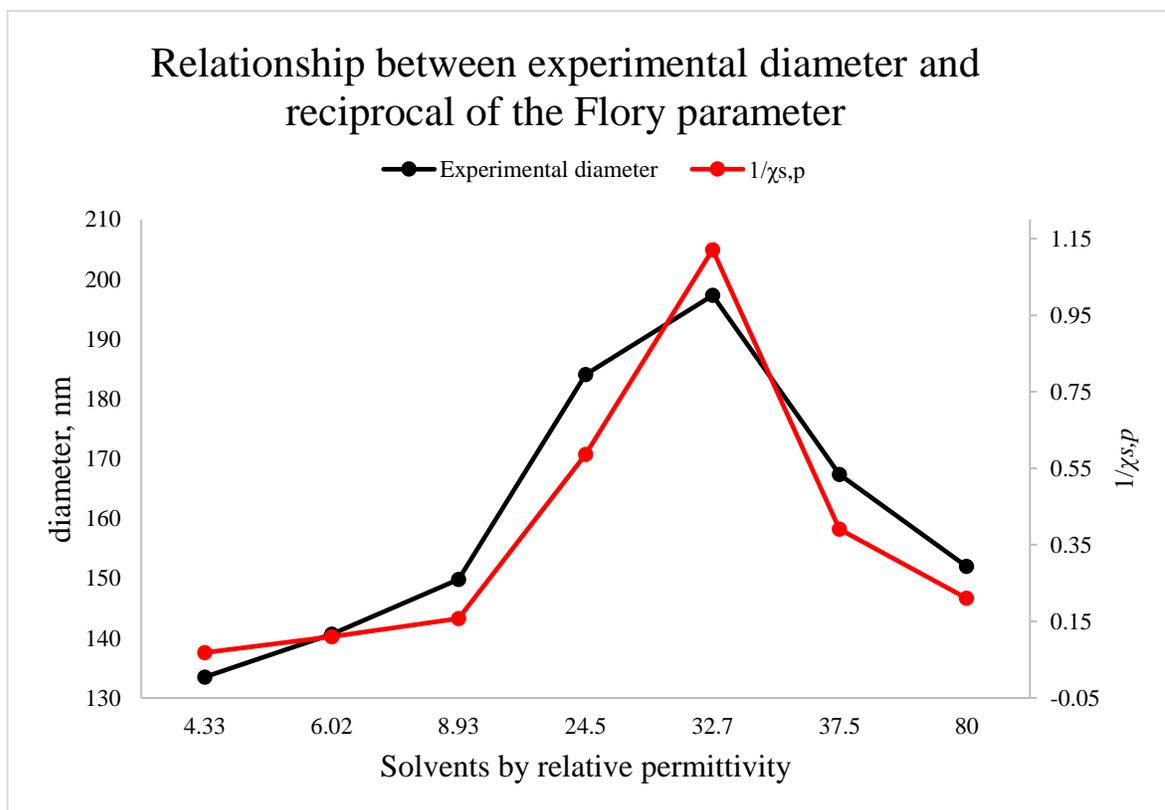


Figure 4.1: Preliminary experiments using a variety of particle synthesis techniques and reaction environments gave a correlation between particle diameter and solubility (expressed as the reciprocal of the Flory parameter  $1/\chi_{s,p}$ ).

The preliminary experiments suggested that the relative solubility may be approximated with the use of solubility parameter theory, in which a molecule may be described in terms of three intermolecular interaction components: dispersive forces,  $\delta_d$ , polarity,  $\delta_p$ , and hydrogen bonding,  $\delta_h$ . Combined together these values give the solubility parameter of the substance,  $\delta$ , representing the coordinates of the material's solubility in terms of  $\delta_d$ ,  $\delta_p$  and  $\delta_h$ . The three interaction components can be obtained directly from the literature for common materials or calculated based on group contribution, in which values for the molecule are approximated based on the number of methyl, methanediyl (methylene bridge), methylene, etc. groups. Here this approach is used to build a minimum statistical polymer unit (MSPU), the cross-linked equivalent of the repeat unit based on the initial monomer ratio, the number of each chemical group calculated from the number of that group in the initial monomer composition (accounting for the structure post-polymerization).<sup>291</sup> The polymers formed in this study were based on those of a standard protocol, giving a monomer ratio of Bis: AAC: tBAm: NiPAm of 1.000: 2.473: 20.000: 26.566. Group contribution tables vary with the method of treatment, and likely the two most accurate methods are those of Hoftyer-Van Krevelen and Hoy.<sup>295</sup> After some initial analysis, the Hoy method was found to be more appropriate for the materials studied.

Hoy's system begins with group contributions for  $F_t$ , the molar attraction function (the cohesive energy per volume),  $F_p$ , the polar component of the molar attraction function,  $V$ , the molar volume of the MSPU, and  $\Delta_T^{(P)}$ , a correction for non-ideality.<sup>295</sup> Each is calculated from the sum of the values given for each group  $i$  present in the MSPU:

$$F_t = \sum_i n_i F_{t,i} \quad \text{Eq. 4.1a}$$

$$F_p = \sum_i n_i F_{p,i} \quad \text{Eq. 4.1b}$$

$$V = \sum_i n_i V_i \quad \text{Eq. 4.1c}$$

$$\Delta_T^{(P)} = \sum_i n_i \Delta_{T,i}^{(P)} \quad \text{Eq. 4.1d}$$

There are also two auxiliary equations for polymers,  $\alpha^{(P)} = 777\Delta\tau^{(P)}/V$  and  $\bar{n} = 0.5/\Delta\tau^{(P)}$ , and the base value  $B = 277$ . The expressions for the solubility parameter components are then given in Equations 4.2a-d:<sup>296</sup>

$$\delta_t = \frac{F_t + B/\bar{n}}{V} \quad \text{Eq. 4.2a}$$

$$\delta_p = \delta_t \left( \frac{1}{\alpha^{(P)}} \frac{F_p}{F_t + B/\bar{n}} \right)^{\frac{1}{2}} \quad \text{Eq. 4.2b}$$

$$\delta_h = \delta_t \left( \frac{\alpha^{(P)} - 1}{\alpha^{(P)}} \right)^{\frac{1}{2}} \quad \text{Eq. 4.2c}$$

$$\delta_d = \left( \delta_t^2 - \delta_p^2 - \delta_h^2 \right)^{\frac{1}{2}} \quad \text{Eq. 4.2d}$$

These solubility parameter components can then be used to determine the relative solubility of the substance in another substance by the principle that ‘like dissolves like’ with Equation 4.3:

$$\Delta\delta = \left( (\delta_{d,p} - \delta_{d,s})^2 + (\delta_{p,p} - \delta_{p,s})^2 + (\delta_{h,p} - \delta_{h,s})^2 \right)^{\frac{1}{2}} \quad \text{Eq. 4.3}$$

Or often

$$R_{s,p} = \left( 4(\delta_{d,p} - \delta_{d,s})^2 + (\delta_{p,p} - \delta_{p,s})^2 + (\delta_{h,p} - \delta_{h,s})^2 \right)^{\frac{1}{2}} \quad \text{Eq. 4.4}$$

Where  $R_{s,p}$  is a modified form of  $\Delta\delta$  representing the ‘distance’ between the solubilities of two substances when plotted in three dimensions, with the constant 4 more accurately representing the dispersion forces.<sup>297</sup>

A greater value of  $\Delta\delta$  or  $R_{s,p}$  therefore indicates reduced affinity of the solute for the solvent, which in the example of precipitation polymerization is equivalent to smaller particles. A further modification includes the introduction of the Flory interaction parameter, which can be approximated using Equation 4.5:

$$\chi_{s,p} = \frac{1}{k_B T} \left( \varepsilon_{s,p} - \frac{1}{2} (\varepsilon_{p,p} + \varepsilon_{s,s}) \right) = \frac{V_s R_{s,p}^2}{RT} \quad \text{Eq. 4.5}$$

The Flory parameter is therefore a measure of the relative affinity of the polymer unit to the solvent, with smaller values indicating greater  $\varepsilon_{p,s}$  interactions, and so increased preference for miscibility.  $\chi_{s,p}$  can also be used in approximations of the system's entropy and Gibbs free energy, making it useful for further theoretical development. The possible relationship between diameter and solubility parameter is presented in Figure 4.1. The measurements associated with these preliminary experiments are varied, showing strong deviation within individual reaction conditions, but the mean trend justified further investigation.

The data in Table 4.1 shows calculated values of  $R_{s,p}$  ('distance in solubility'),  $\chi_{s,p}$  and measured diameters of polymer nanoparticles synthesized in solution of water and tetrahydrofuran mixed in different ratios. Values of  $V_s$ ,  $\delta_d$ ,  $\delta_p$ , and  $\delta_h$  for these mixtures were calculated from known values of pure solvents using the equation:  $\delta_{ij} = \sum \phi_i \delta_{ij}$ . From results obtained it can be concluded that only the Flory parameter yields successful prediction of fluctuation of diameter of polymer nanoparticles prepared in solution of different polarity.

*Table 4.1: Relationship between solubility parameter components,  $R_{s,p}$ ,  $\chi_{s,p}$ , and observed diameter of polymer nanoparticles. Only  $\chi_{s,p}$  is seen to effectively predict the initial reduction in polymer diameter with increasing THF content.*

Water vol%	$V_s, \text{cm}^3 \text{mol}^{-1}$	$R_{s,p}$	$\chi_{s,p}$	Diameter, nm
100	18.00	32.242	7.5526	$209 \pm 4$
95	21.19	30.470	7.941	$196 \pm 19$
90	24.37	28.700	8.102	$178 \pm 9$
85	27.56	26.936	8.071	$163 \pm 18$
80	30.74	25.175	7.864	$190 \pm 16$
75	33.93	23.421	7.512	$223 \pm 13$
70	37.11	21.675	7.037	$280 \pm 34$
65	40.30	19.936	6.465	$297 \pm 12$
60	43.48	18.213	5.821	335

#### 4.5. Gibbs energy of mixing calculation

In FHT the thermodynamics of mixing is constructed so as to account for the many configurations that a polymer chain may take. The polymeric particles formed here however can be better represented by spheres, interchangeable by rotation, and with each other. This requires an alternative model, but the general methodology and terminology of Flory will be followed.

The system is assumed to be a 3D lattice of cells, each cell being the volume of the MSPU, with solvents being represented as clusters of individual molecules amounting to the volume of one cell.<sup>298</sup> The total volume is given by  $n_0$ , total number of polymer particles by  $N$ , total number of MSPU units per particle (used in analogy to both the polymer volume and degree of polymerization) being  $x$ , and the total number of solvent clusters being given by  $n$ ; thus  $n_0 = n + xN$ .

If only one polymer particle (indistinguishable by rotation) is present, the number of positions it may occupy in the lattice is approximately equal to that of the particles central MSPU unit minus the radius of the particle multiplied by the area of the lattice, taken to be cubic for simplicity. The number of configurations that may be taken by one particle is then:

$$v = n_0 - \left(\frac{3x}{4\pi}\right)^{\frac{1}{3}} 6(\sqrt[3]{n_0})^2 \quad \text{Eq. 4.6}$$

The configurations available for a second particle are then equal to this volume minus the area of the previously added particle. The volume around the first particle that cannot be reached by the central MSPU of the second (i.e. the radius of a sphere multiplied by its area) is equal to three times the volume of the particle, giving a total of  $4x$ . The second particle thus has a total free volume of  $v - 4x(N - 1)$ , where  $N$  is currently 2. After adding all particles in this manner, the total number of configurations for the system is given by:

$$W_{particles} = \frac{v^N - (4x)^{N-1}(N-1)!}{N!} \quad \text{Eq. 4.7}$$

The solvent clusters can then be accounted for in the typical manner, the total remaining lattice sites being defined as solvent occupied, and the whole definition of  $W$  treated with Sterling's approximation ( $n! = (n/e)^n$ ). By the definition that  $\Delta S_m = (S_{particles} + S_{solvent}) - S_{mixture}$ , where both  $S_{particles}$  and  $S_{solvent}$  are taken to be zero, and the Boltzmann equation,  $S = k_B \ln W$ , the entropy of mixing can be given by Equation 4.8:

$$\Delta S_m = -k_B \left[ \ln \left( \left( \left( n_0 - \left( \frac{3x}{4\pi} \right)^{\frac{1}{3}} 6(\sqrt[3]{n_0})^2 \right) \frac{e}{N} \right)^N - (4x)^{N-1} \left( \frac{N-1}{N} \right)^N \right) + n \ln \left( \frac{n}{n_0} \right) \right]$$

Eq. 4.8

Computation with this equation is relatively difficult however and alternative entropy models are available. The Gibbs energy per segment may be more easily approximated using volume fractions, but must be adjusted for use of spherical particles from long chains.<sup>299, 300</sup>

An alternative model of the entropy of mixing can be applied instead, starting with simple entropy rules. The mixing of a number of particles can be described by  $\Delta S_m = n_x k_B \ln(V_2/V_1)$ , where  $n_x k_B$  describes the energy of some number of particles, with the logarithmic term giving the entropy as proportional to the difference between a starting volume ( $V_1$ ) and a final volume ( $V_2$ ). On mixing two different kinds of particles  $x$  and  $y$  the final volume is equal to  $V_x + V_y$ , and the total entropy is given as the sum of the two (Equation 4.9a):

$$\Delta S_m = \Delta S_x + \Delta S_y = n_x k_B \ln \left( \frac{V_x + V_y}{V_x} \right) + n_y k_B \ln \left( \frac{V_x + V_y}{V_y} \right) \quad \text{Eq. 4.9a}$$

In FHT the lattice cells are of equal volume (here, that of the MSPU) with solvents described as clusters of that same size. The volume terms can therefore be replaced with  $xN$  and  $n$  for the total number of polymer and solvent cells respectively. For application here therefore the entropy is given by Equation 4.9b:

$$\Delta S_m = \Delta S_{solvent} + \Delta S_{polymer} = n k_B \ln \left( \frac{n_0}{n} \right) + N k_B \ln \left( \frac{v}{xN} \right) \quad \text{Eq. 4.9b}$$

This includes the predefined volume that can be occupied by the polymer particles,  $v$ . The inverse of the logarithm is generally given as it represents the more manageable mole fraction in standard calculations, and the final form for the entropy can be presented as in Equation 4.10:

$$\Delta S_m = -k_B \left[ n \ln \left( \frac{n}{n_0} \right) + N \ln \left( \frac{xN}{v} \right) \right] \quad \text{Eq. 4.10}$$

This entropy equation appears overly simplistic compared with that of Equation 4.8, but the structure is generally the same, and the latter form is a minor variation on a common formula.

The enthalpy of mixing can be represented by a similarly simple equation, being given from the product of the volume fractions and a constant of interaction, expressed here as the Flory parameter.<sup>299, 300</sup> In the particle model only the surface area of the polymer can interact with the solvent, and so only this is included:

$$\Delta H_m = k_B T \phi_s A \chi_{s,p} \quad \text{Eq. 4.11}$$

Here  $\phi_s$  is the volume fraction of the solvent,  $(n_0 - Nx)/n_0$ , and the total area of all polymer material  $A$  is given by Equation 4.12:

$$A = N \pi^{\frac{1}{3}} (6x)^{\frac{2}{3}} \quad \text{Eq. 4.12}$$

The expression for  $\Delta H_m$  here is therefore equal to a constant of polymer unit-solvent interaction ( $\chi_{s,p}$ , here describing the interaction of the cell-cell interactions) scaled according to the rate of occurrence, accounting for the lack of interaction between solvent and polymer units within the polymeric sphere ( $\chi_{s,p}$  between identical substances is defined as zero, as defined by Equations 4.4 and 4.5). Combining these equations therefore gives a simple model for the Gibbs free energy of mixing:

$$\Delta G_m = k_B T \left[ \phi_s A \chi_{s,p} + n \ln \left( \frac{n}{n_0} \right) + N \ln \left( \frac{xN}{v} \right) \right] \quad \text{Eq. 4.13}$$

For these calculations it will be assumed that the conversion of monomer to polymer is 100 % and that the volume fraction occupied by the polymer is equal to that occupied by the reagents.

#### 4.6. Spinodal and binodal phase separation

The relationship described between  $\chi_{s,p}$  and particle diameter is also dependent on a minimal value for  $\chi_{s,p}$  for successful synthesis. Empirical investigation combined with the thermodynamic model outlined previously gives some insight into the basis of this minimal  $\chi_{s,p}$ , by approximating  $\Delta G_m$  values and thus predicting phase separation behavior. Phase separation from polymer solutions may occur by either nucleation and growth, if the system is metastable, or spinodal decomposition if the system is unstable.<sup>301</sup> In a system of two miscible components, the spinodal is typically defined by the condition below, in terms of the volume fraction  $\phi_p$ , where  $\phi_p = xN/n_0$ :<sup>302, 303</sup>

$$\left(\frac{\partial^2 \Delta G_m}{\partial \phi_p^2}\right)_{P,T} = \left(\frac{\partial^2 \Delta G_m}{\partial \left(\frac{xN}{n_0}\right)^2}\right)_{P,T} = 0 \quad \text{Eq. 4.14}$$

This inflection represents the point of spontaneous phase separation or spinodal decomposition, as the homogenous mixture reaches its limit of stability. The spinodal or spinodal curve can be observed relatively easily as it is the point beyond which a homogeneous phase will separate into two distinct phases, and in the case of polymer nanoparticle synthesis, precipitation occurs. Small droplets of polymer spontaneously separate from the solvent throughout the system, acting as nucleation sites, and the polymers are grown from the solution.<sup>304</sup> At higher concentrations spinodal decomposition produces membrane-like structures, the individual fibers of which rapidly expand to reduce the surface area of the new phase.<sup>300</sup>

The full spinodal is typically determined by measuring the inflection points at various temperatures, the combination of which gives the spinodal curve. Observation of different values of calculated  $\chi_{s,p}$  are however practically equivalent to observation of different temperatures, and are also commonly used.<sup>295, 305</sup> This may be understood by considering the internal energy of the system, where internal energy is equal to the sum of the potential energy and kinetic energy. When increasing the temperature through the UCST (upper critical solution temperature, the critical point between separated phases at lower temperatures and homogeneity at higher temperatures)

from a two phase to one phase system, the potential energy (the interatomic and intermolecular interactions resulting from electron density) will change very little, but the increase in kinetic (translational) energy reaches a value where it can overcome the more chemically favorable arrangement determined by the potential energy of the components. With greater kinetic energy therefore, the molecules have sufficient translational energy to overcome the restrictions of the potential energy and move freely between different molecules, resulting in a single phase. The Flory parameter  $\chi_{s,p}$  is related to the magnitude of the potential energy, where greater values are equivalent to a reduction in affinity between the two components, polymer and solvent. Increasing  $\chi_{s,p}$  is therefore analogous to increasing the potential energy, practically equivalent to reducing the kinetic energy, and at a sufficient value equating to an internal energy (or by derivation Gibbs energy) favoring separation.

Phase separation may occur for any value of  $\phi_p$  between the tangent points, above the low- $\phi_p$   $\Delta G_m$  extremum and the high- $\phi_p$   $\Delta G_m$  extremum (see below).<sup>295</sup> The difference in magnitude between the  $\Delta G_m$  local and global maxima is proportional to the degree of polymerization, and the relative position of each is moved to lower values of  $\phi_p$  with greater degree of polymerization. As  $\chi_{s,p}$  is reduced (or in UCST behavior, the temperature is increased) the tangent points become closer until they meet at the critical temperature ( $T_c$ ) or critical Flory parameter ( $\chi_c$ ), representing the critical point. Beyond the critical temperature/Flory parameter no separation occurs.

Under conditions other those of the critical point, the spinodal represents the boundary between that of phase separation and a region of homogeneous metastability. The metastable region is found beyond the point of equilibrium, the function referred to as the binodal, described by Equation 4.15:<sup>300</sup>

$$\left( \frac{\partial \Delta G_m \left( \frac{xN}{n_0} \right)'}{\partial \left( \frac{xN}{n_0} \right)} \right)_{P,T} = \left( \frac{\partial \Delta G_m \left( \frac{xN}{n_0} \right)''}{\partial \left( \frac{xN}{n_0} \right)} \right)_{P,T} \quad \text{Eq. 4.15}$$

Where prime and double prime refer to the two phases of the mixture, therefore representing the tangent line connecting point near the two extrema of the Free energy plot. The binodal thus representing the point of equilibrium between two phases as a typical phase diagram, while the spinodal shows the limit of stability. At  $\chi_c$ , the binodal

and spinodal are equivalent, and the point at which the second phase is favored is the same as that at which the existing system becomes unstable. Assuming the thermodynamic relationships to  $\chi_{s,p}$  are equivalent to those with temperature associated with the UCST, the relationship can be summarized as in Figure 4.2.

Phase separation from a homogeneous system in the metastable region beyond the binodal requires spontaneous phase separation of small droplets, followed by their growth by diffusion. The growth of the droplets proceeds readily by diffusion as the equilibrium point (Equation 4.15), i.e. the binodal has already been passed, but the point of stability (Equation 4.14), the spinodal, has not.<sup>303</sup> In the metastable state therefore, an energy barrier to phase separation must be overcome in order to induce precipitation from a homogeneous mixture, and in the event of such phase separation being induced, the resulting particles are formed extremely slowly and with high polydispersity.<sup>300</sup> In addition to this, random phase separation from a homogeneous mixture is associated with a quantity of work proportional to the volume of the new phase, with a minimum size requirement for energetically favorable development of the emerging phase.<sup>306</sup>

In a nucleated systems however, phase separation occurs via the equilibration of the growing polymers with the existing nucleant separate phase. As the early polymer develops the character of the MSPU, the affinity of the oligomers for the solvent decreases as it approaches  $\chi_{s,p}$ . In the presence of a suitable nucleant the equilibrium between the solvated state and complexed state shifts as the reaction proceeds, the equilibrium point between these being the binodal. Beyond the binodal in the metastable region, the nucleant is favored as a local fluctuation in  $\Delta G_m$  and polymerization proceeds from these sites. The specific system used in the experiments described was designed in such a way that the monomer would interact strongly with the nucleant, encouraging early binodal phase separation.

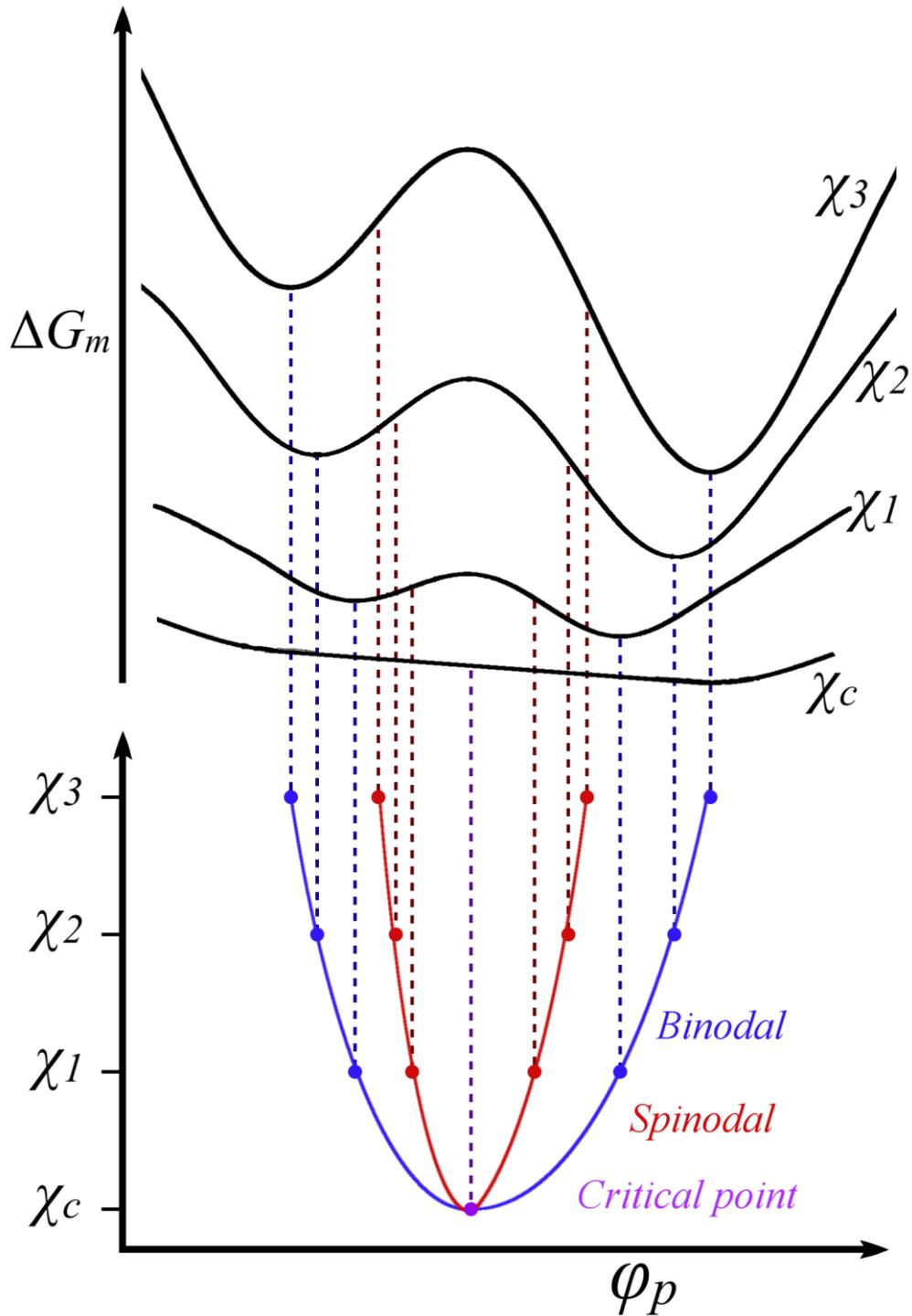


Figure 4.2: In idealized diagram for the theoretical basis of the phase boundaries. The binodal curve is built from the Gibbs free energy extrema at different temperatures or Flory parameter values. The spinodal occurs within the binodal, and is found from the inflection points. The critical point is the meeting point of the binodal and spinodal.

The relationship between the physical changes that occur at these boundaries and the free energy curve is explained by the response of the polymer at the described energy. Inside the spinodal the free energy is defined to occur where  $\partial^2\Delta G_m/\partial\phi_p^2 < 0$  (negative curvature), and the metastable binodal region where  $\partial^2\Delta G_m/\partial\phi_p^2 > 0$ . These can be visualized in Figures 4.2, 4.3 and 4.4. Inside the spinodal the system will spontaneously decompose due to small fluctuations, the coherence of the fluctuations lead to small decrease in Gibbs energy in separation due to the negativity of  $\partial^2\Delta G_m/\partial\phi_p^2$ .<sup>307</sup> The equivalent fluctuations between the spinodal and binodal however correspond to an increase in Gibbs energy in separation ( $\partial^2\Delta G_m/\partial\phi_p^2$  being positive), and therefore polymers in this region will not separate spontaneously.

Figure 4.2 represents an idealized system with  $x$  (degree of polymerization, polymer volume) close to one (monomers, or single MSPUs), effectively representing the mixing of approximately equally sized solvent and solute particles. With the large values of  $x$  found in the nanoparticles observed the position of the critical point will be shifted, the polymer fraction at which it occurs decreasing with increasing  $x$ . A more accurate representation is given in Figure 4.3.

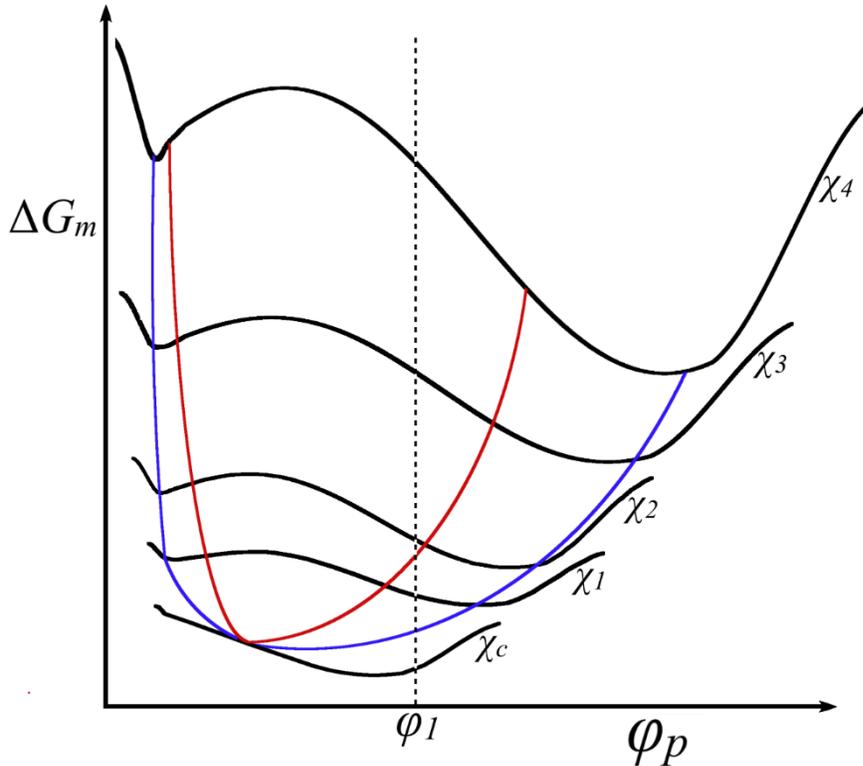


Figure 4.3: The relationship between free energy and the phase boundaries (spinodal in red, binodal in blue) in systems where individual polymers have a higher volume ( $x$ , degree of polymerization) than the solvent clusters. The dashed vertical line and  $\varphi_1$  are for illustrative purposes and explained in the text. This diagram too is exaggerated, in reality the phase boundaries stretch from near zero to near total polymer content.

The polymer concentrations ( $\varphi_p$  or  $xN/n_0$ ) observed were approximately in the range of the value  $\varphi_1$  in Figure 4.3, and can be used to restate the meaning of each term. At  $\chi_1$  the polymer is within the metastable region between the binodal and the spinodal, and will not spontaneously separate but may in the presence of a nucleant.  $\chi_2$ ,  $\chi_3$  and  $\chi_4$  are within the spinodal, and therefore will spontaneously separate, while  $\chi_c$  will remain in solution.

The binodal is found on the tangent between extrema (minima), so typically close to the minimum  $\partial\Delta G_m/\partial\varphi_p = 0$  as shown in Figure 4.2 and Figure 4.3. The spinodal is found at the inflection point of the  $\Delta G_m(\varphi_p)$  function, where the change in  $\partial\Delta G_m/\partial\varphi_p$

with  $\varphi_p$ ,  $\partial^2\Delta G_m/\partial\varphi_p^2 = 0$ . Before the inflection spontaneous separation will occur, and for a given value of  $\chi_{s,p}$ ,  $\partial\Delta G_m/\partial\varphi_p$  will decrease with  $\varphi_p$  beyond the inflection of this function. This point is illustrated in Figure 4.4.

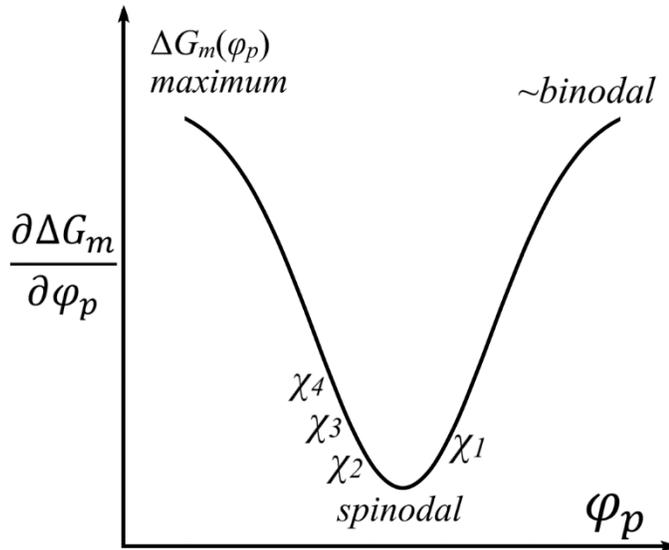


Figure 4.4: The positions of the various positions highlighted in Figure 4.3 at the point  $\varphi_1$ . The second derivative is negative from within the spinodal condition until the spinodal boundary at the minimum; from the spinodal the second derivative is positive, representing the metastable region between the spinodal and binodal.

At a given value of  $xN$  (total polymer content, equivalent to  $\varphi_p$ ) therefore, the rate of change of  $\Delta G_m$ ,  $\partial\Delta G_m/\partial\varphi_p$ , would be expected to be more negative for greater values of  $\chi_{s,p}$  close to but within the spinodal. Application of the equations for Gibbs energy described in Section 4.5 to polymer nanoparticles synthesized over a short range demonstrate that the relationship found in standard polymer theory can be applied on this scale.

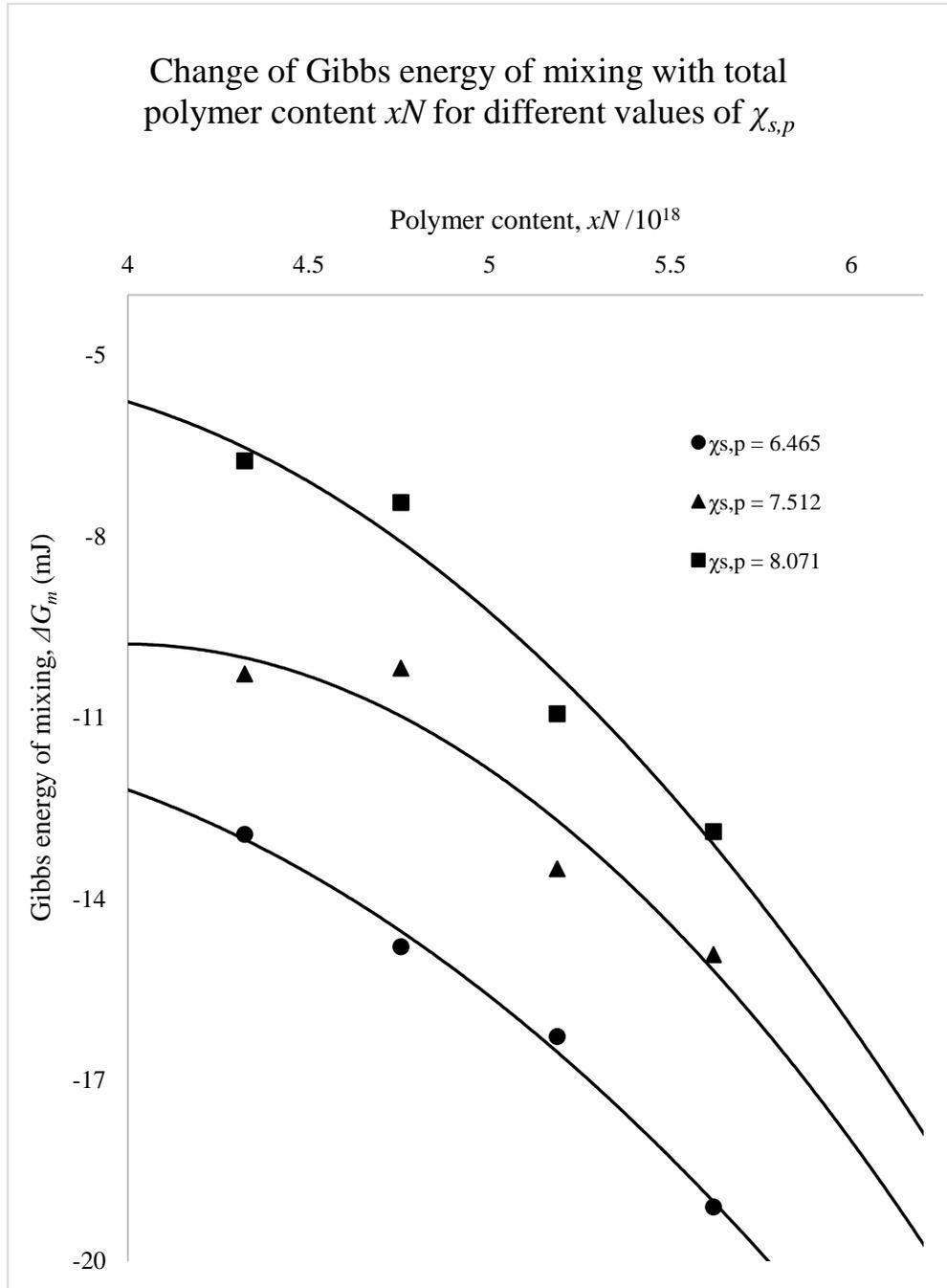


Figure 4.5: Gibbs energy of mixing over several polymer fractions with different Flory parameters;  $\partial\Delta G_m / \partial xN$  is observed to increase with greater  $\chi_{s,p}$  as the lower approaches the inflection. This graph and the following were originally published in *Polymer Chemistry* (Cowen, 2018).<sup>308</sup>

Figure 4.5 demonstrates the relationship between  $\Delta G_m$  and  $\chi_{s,p}$  over a short range of  $xN$ , the lower value of  $\chi_{s,p}$  giving a less negative  $\partial\Delta G_m/\partial xN$  as the inflection point spinodal is approached. The coordinate  $\chi_{s,p} = 5.821$ :  $xN = 3.026 \times 10^{18}$  is a suspected spinodal phase boundary based on success in polymerization at all greater values of  $\chi_{s,p}$  at the given  $xN$ , and greater  $xN$  at the given  $\chi_{s,p}$ , with consistent failure in polymerization with lower values; low concentrations of polydisperse polymer was found at the boundary. Due to the binodal being below the spinodal, polymers were formed in the presence of a nucleant at values of  $\chi_{s,p}$  below those where un-nucleated synthesis was unsuccessful.

The mechanisms required for system transition from single phase to two phases across the spinodal are different to those occurring during phase separation in the metastable region beyond the binodal, the mechanisms of the latter resembling those of crystallization.<sup>299</sup> This is an important addition as it suggests that polymers formed in the presence of a nucleant are fundamentally different from those without, and that regular arrangements representing the most energetically favorable interactions of monomers or oligomers around the seed compound may be found in the former. In the absence of detailed analysis of many different polymeric systems however this remains speculative. In the next section the practical value of finding the spinodal and binodal will be addressed.

#### 4.7. The predictive capacity of $\Delta\chi$

The spinodal and binodal on a  $\chi_{s,p}(xN)$  plot represent the minimal  $\chi_{s,p}$  required to induce phase separation at a given value of  $xN$  (total polymer content), depending on the absence or presence of high free energy fluctuations in the total system energy profile. The spinodal  $\chi_{s,p}$  (from here  $\chi_{spinodal}$ ) is that required for spinodal character phase separation (unnucleated polymerization), and the binodal  $\chi_{s,p}$  ( $\chi_{binodal}$ ) that required for binodal character phase separation (nucleated polymerization).

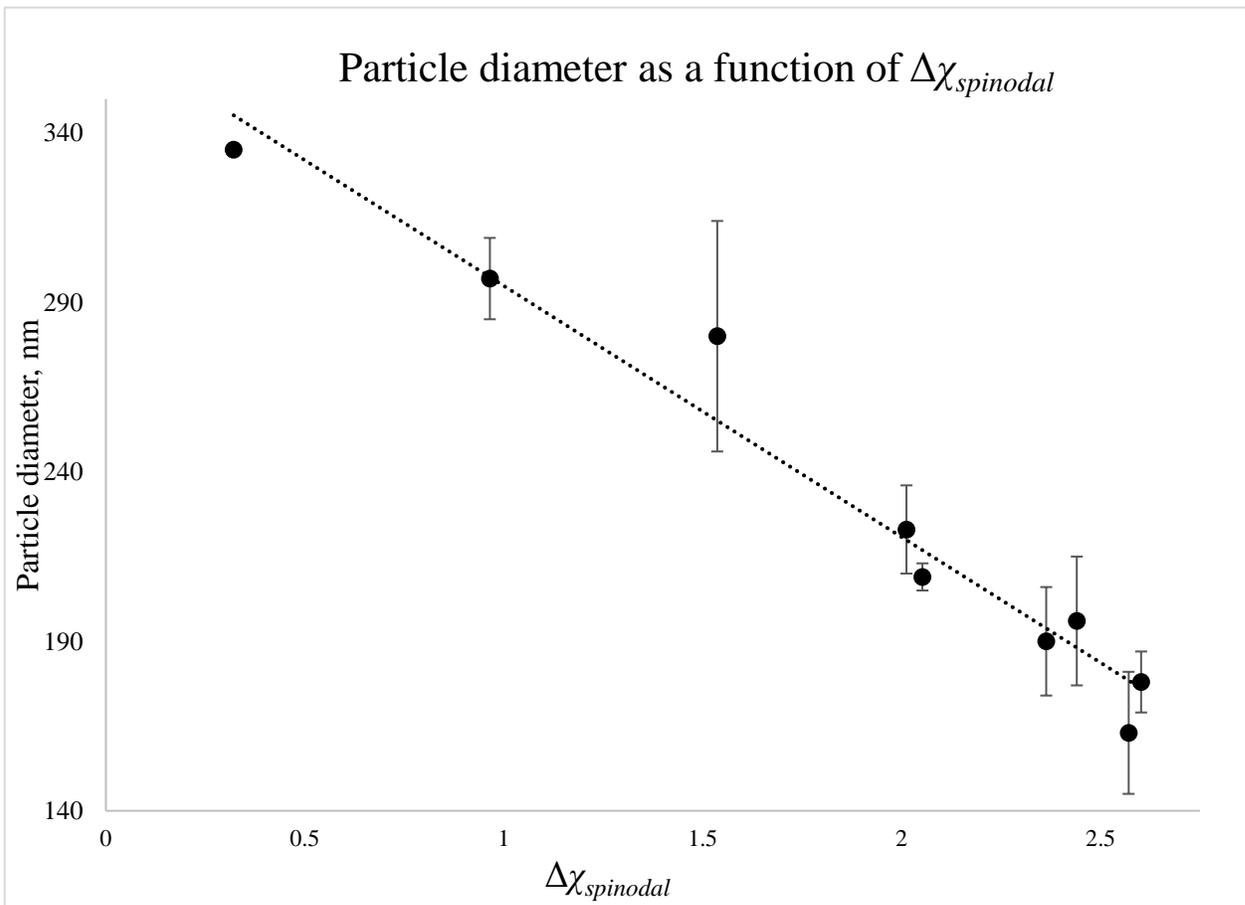


Figure 4.6: The relationship between particle diameter and  $\Delta\chi_{spinodal}$ , where  $\Delta\chi_{spinodal} = \chi_{s,p} - \chi_{spinodal}$ . At this  $xN$ ,  $D(nm) = -74\Delta\chi_{spinodal} + 367 nm$ .

As previously discussed,  $\chi_{spinodal}$  was determined to be approximately 5.8 at  $xN = 3.026 \times 10^{18}$ . Unsuccessful polymerization at  $xN = 4.323 \times 10^{18}$  (standard) with  $\chi_{s,p} = 5.131$  suggests that at this value of  $xN$  the value of  $\chi_{spinodal}$  is approximately 5.5. Taking this, a term  $\Delta\chi_{spinodal} = \chi_{s,p} - \chi_{spinodal}$  can be calculated, and a plot vs. nanoparticle diameter gives an approximately linear relationship  $D(nm) = -74\Delta\chi_{spinodal} + 367$  nm (Figure 4.6).

Repetition with alternative solvent validates this equation as a general trend at the given value of  $xN$ : water-methanol systems equivalent to  $\Delta\chi_{spinodal}$  1.965 (235 nm (predicted 222 nm)) and 1.721 (260 nm (predicted 240 nm)) demonstrating relatively good correlation despite problems often encountered with methanol in polymer synthesis. Successful spinodal-character polymerization also requires  $\Delta\chi_{spinodal}$  to be greater than zero, but the relationship given will be erased by the presence of nucleants, due the equivalent  $\Delta\chi_{binodal}$  then being dominant. This was demonstrated by successful polymerization in the presence of a streptomycin nucleant at a  $\Delta\chi_{spinodal}$  value of -1.083, far below the value where unnucleated attempts were successful. An overview of the model is given in Figure 4.7.

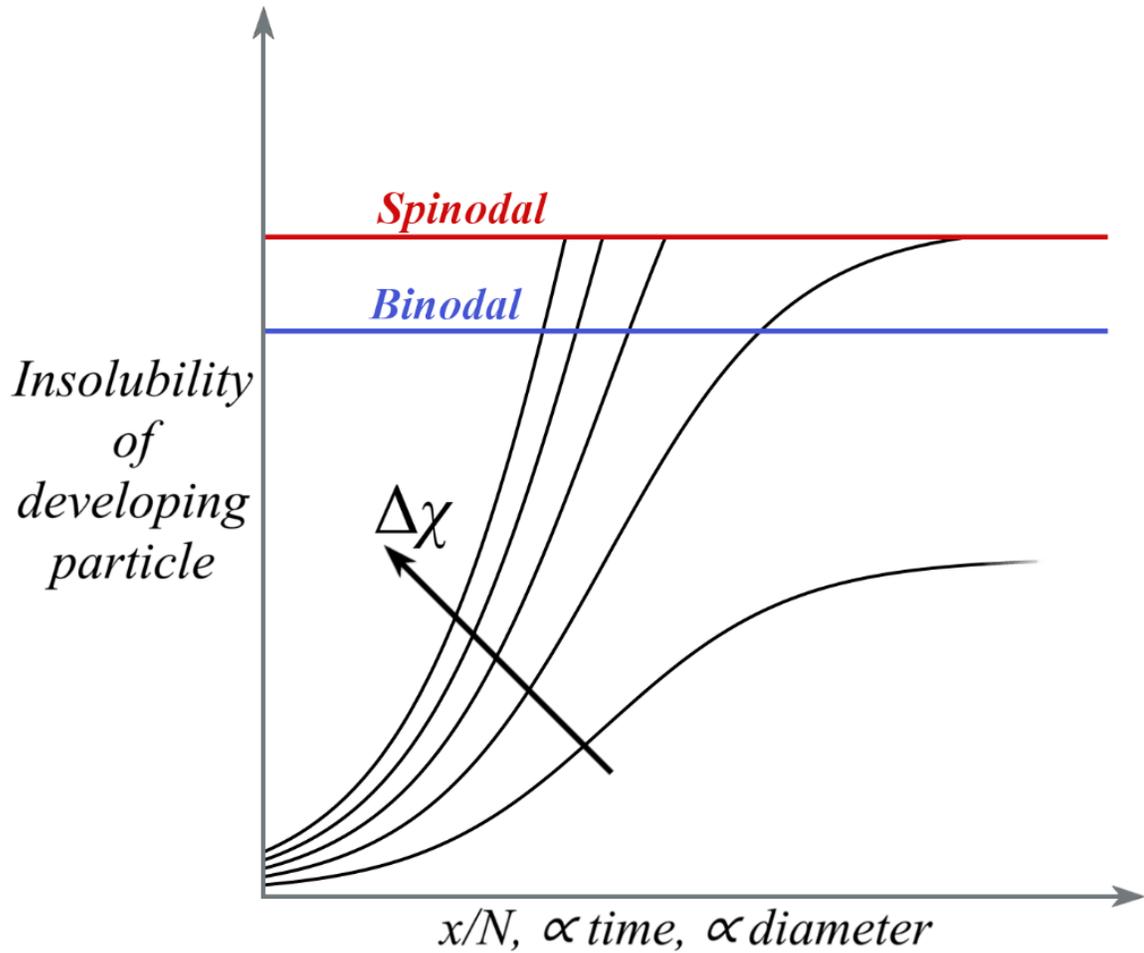


Figure 4.7: A possible explanation for the observations, showing insolubility of developing particles as a function of time,  $x/N$  and final particle diameter. The particle grows until it reaches  $\chi_{s,p}$ , the plateau of the sigmoid curve of insolubility. Separation occurs at the spinodal or from the binodal. The shape of the sigmoid curve, and thus the nanoparticle diameter, depends on  $\Delta\chi$  (the position of  $\chi_{s,p}$  relative to the spinodal or binodal).

The equation given above for  $\Delta\chi_{spinodal}$  however will not hold for values of  $xN$  other than that given ( $xN = 4.323 \times 10^{18}$ ). Simple analysis with the data collected previously with variation in  $xN$  for analysis of the Gibbs energy however gives a possible relationship. With the assumption that the trend in  $\chi_{spinodal}$  continues linearly as it does from  $xN = 3.026 \times 10^{18}$  to  $xN = 4.323 \times 10^{18}$ , through to  $xN = 5.620 \times 10^{18}$ , the plot of particle diameter as a function of  $(xN)^2/\Delta\chi_{spinodal}$  is given in Figure 4.8; the trend shown can be approximated to  $D(nm) = 173\ln[(xN)^2 10^{-36}/\Delta\chi_{spinodal}] - 193$  nm.

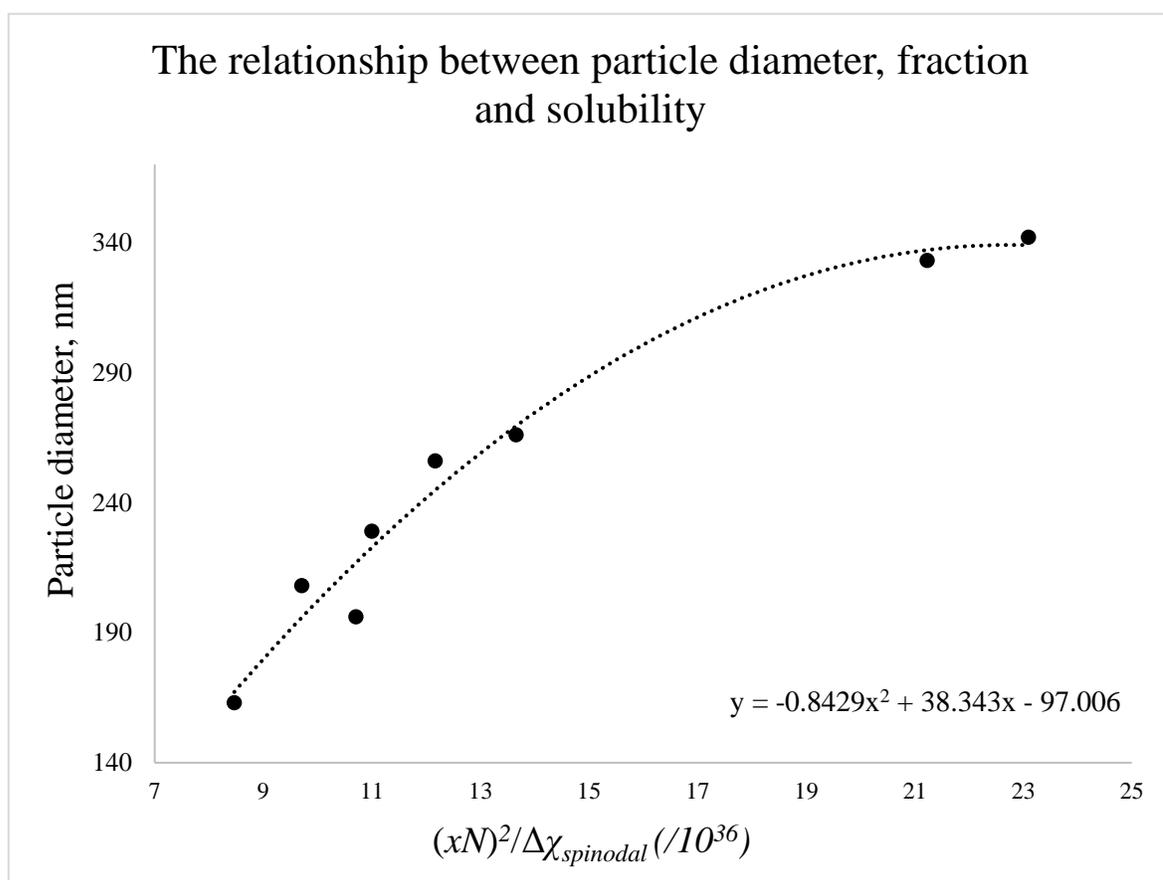


Figure 4.8: The relationship between particle diameter, fraction (expressed in values of  $xN$ ) and  $\Delta\chi_{spinodal}$ . The trend line shown is given by the equation in the Figure, but can be approximated to  $D(nm) = 173\ln[(xN)^2 10^{-36}/\Delta\chi_{spinodal}] - 193$  nm.

Further experiments are required however to fully determine the relationship beyond the relatively narrow range within which  $xN$  was varied. An additional point of further work centers on  $\Delta\chi_{binodal}$ . While inducing binodal-character phase separation of polymer nanoparticles is relatively simple by comparison (indeed often difficult to avoid entirely due the presence of contaminant) due to the lower value of  $\chi_{binodal}$  permitting a wider range of reaction conditions, generating relevant experimental data that helps with producing any consistent model of binodal-character phase separation will inherently be much more difficult than that associated with the spinodal. As Equation 4.15 shows, the position of the binodal is dependent on the relative affinity of the MSPU for the nucleant and the solvent. The relationship between polymer and solvent given by  $\chi_{s,p}$  then becomes irrelevant unless it is accompanied by an additional term expressing the interaction between polymer and nucleant, a relationship which can less easily be expressed with typical solubility parameters. Additionally, as the nucleant acts as a local fluctuation in the Gibbs energy in the profile of the system which the developing polymer may adsorb into, the solubility of the nucleant will also need to be included in any effective model of binodal-character phase separation.

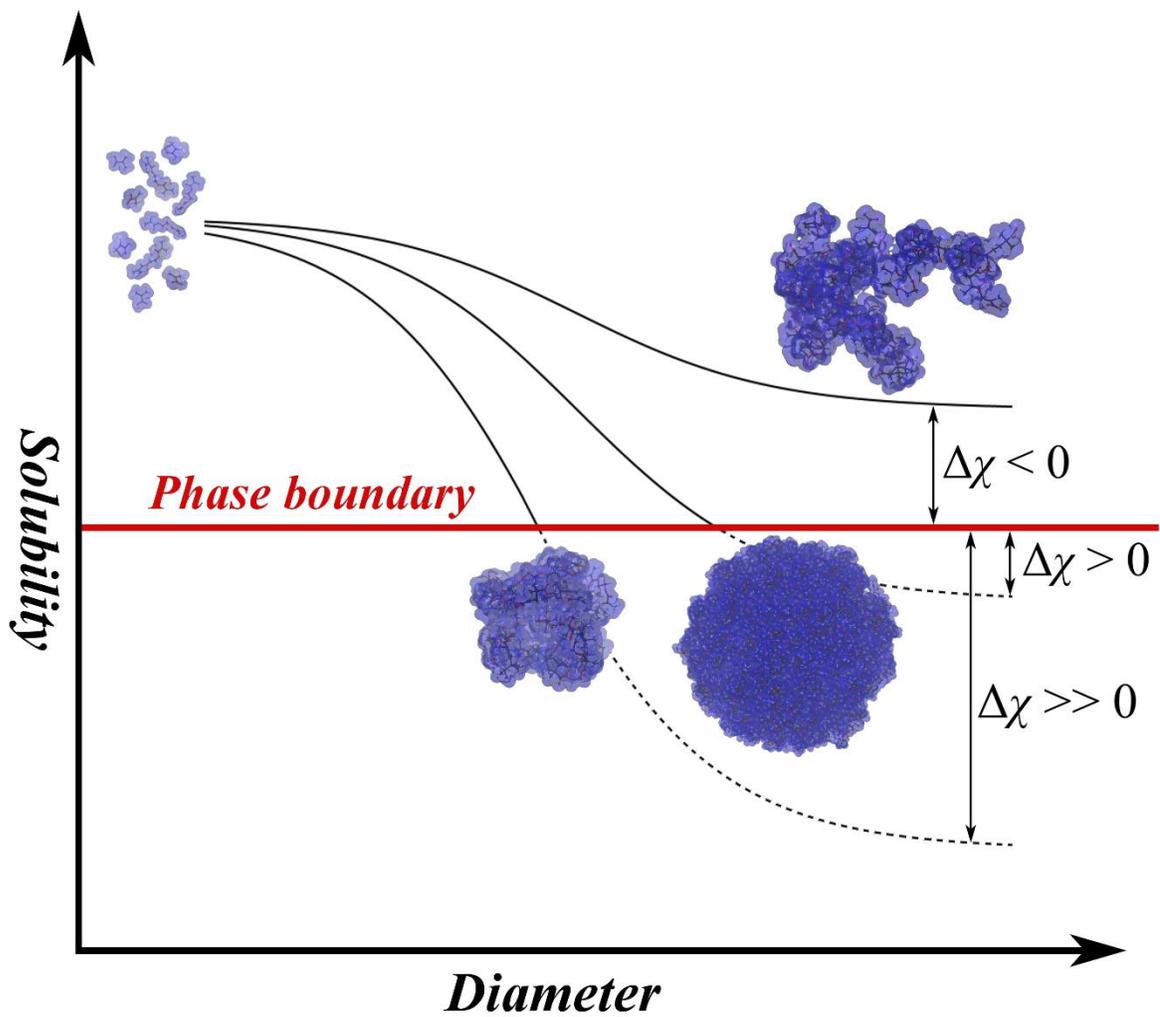


Figure 4.9: Simple diagram explaining the observed relationship between polymer MSPU solubility and nanoparticle diameter.

## 4.8. Conclusions

The relationship between polymer nanoparticle diameter and solubility has been established for phase separation at the spinodal in un-nucleated particles, being dependent on the value of  $\Delta\chi_{spinodal}$ . Increasing values of  $\Delta\chi_{spinodal}$  correlate with decreasing particle dimensions for  $\Delta\chi_{spinodal}$  greater than zero. Within some range this trend can also be extended to different polymer fractions, correlating with  $(xN)^2/\Delta\chi_{spinodal}$ . As control of nanoparticle dimensions is essential for applications *in vivo* this model should hopefully prove useful to researchers in the field.

Particle synthesis with  $\Delta\chi_{spinodal} < 0$  is associated with nucleated separation within the metastable region between the binodal and spinodal. The synthetic mechanisms of binodal character and spinodal character separation are therefore largely incomparable and further analysis of the equivalent  $\Delta\chi_{binodal}$  is required. This will be of particular relevance to research into imprinted polymers as due to the nucleating activity of the template. Predicting the size will be more difficult in these examples, being dependent on the solubility of both the template and polymer in the solvent in addition to the affinity of the developing oligomer for the template will need to be accounted for. However, the possibility of regularity of the nucleated particle core, i.e. the MIP binding site, should be of interest to those in the field. However, the dependence on solubility parameters inhibits the full potential of this area, leaving the base of the model on insufficient foundations. For expansion and application alternative more accurate methods of determining  $\chi_{s,p}$  are required.

## 5. Molecular mechanics of polymer solubility

This chapter focuses on applying the molecular modeling discussed in Chapter 2 and Chapter 3 to the study of polymer solubility and nanoparticle thermodynamics described in the previous chapter. The research is in its early stages and starts with an attempt to derive mesoscopic properties from interatomic interactions. The objective is to describe a method of describing the solubility of polymers with greater accuracy than that used in the previous chapter. This would allow more accurate prediction of polymer nanoparticle hydrodynamic diameters and possibly other properties. Applications of molecular mechanics and molecular dynamics, with the additional polymerization algorithm described in Chapter 3, is presented as a possible approach.

### 5.1. Modeling solubility

The major practical benefit of the model for the control of polymer nanoparticle size described in the previous chapter is the removal of the necessity of empirical investigation prior to synthesis (assuming prior knowledge of  $\Delta\chi_{spinodal}$ ). The solubility parameters however, from which the value of  $\chi_{s,p}$  and  $\Delta\chi_{spinodal}$  are determined, are obtained from old data-tables, and give approximations of the group contributions for each small section of the MSPU (Section 4.4).

QSPR (quantitative structure-property relationship, related to more common QSAR, quantitative structure-activity relationship) programs can be used in a similar manner to that described in Chapter 4 to determine relevant properties such as the partition coefficient ( $\log P$ ), but with much greater efficiency, and can include additional structural analysis to generate more accurate predictions of molecular solubility.<sup>309</sup> However, while the existing QSPR/QSPR programs are undoubtedly useful, particularly with the development of internet-based tools like VCCLAB's 'ALOGPS' and 'E-DRAGON',<sup>310</sup> there are still limitations in terms of accuracy due to the simplicity of the model and, more importantly for the analysis of polymers and the MSPU, often restrictions on the size of the structure that can be examined.

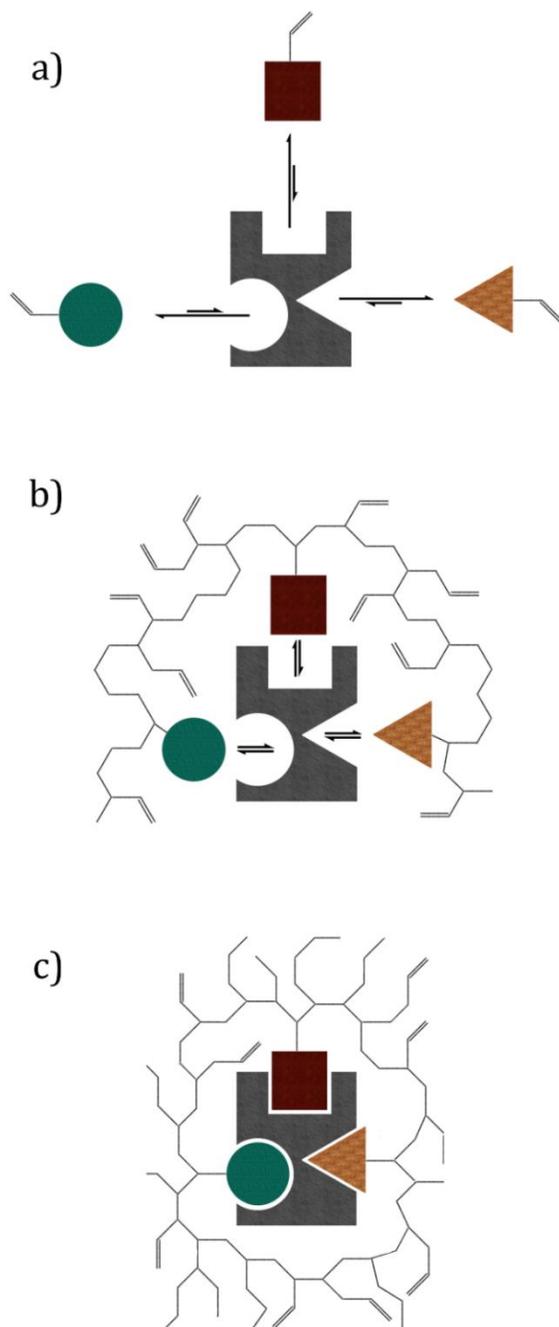
However, more accurate results may be possible with the application of molecular mechanics and the algorithm described in Chapter 3. MM, as described previously, can be applied for analysis at the level of interatomic interactions, giving good descriptions of the energies associated with these interactions with much greater accuracy than more simplistic techniques, while also capable of analyzing a large number of atoms and molecules efficiently, unlike QM techniques which would be unsuitable for solubility studies. With the application of the previously described polymerization algorithm, a structure representative of a polymer, the MSPU, can also be synthesized *in silico* efficiently and with good representation relative to manual production.

## 5.2. Models of imprinted polymer synthesis

The relevance of molecular modeling in the study of solubility to imprinted polymers arises from a synthetic mechanism of imprinted polymers that emerges from the empirical and theoretical data presented in Chapter 4. The model refers to a polymer-template affinity that develops through polymerization with the point of equilibrium being reached at the binodal. The binodal then represents a phase boundary, the minimum required  $\Delta\chi$  to achieve phase separation, and the point where the affinity of the oligomer for the template has developed to the point of equality with the affinity of the oligomer for the solvent (see Section 4.6).

This can be contrasted with the explanation given in the introduction (Section 1.1, shown in Figure 1.1). Here a complex is assumed to form spontaneously in solution before polymerization due to the strong affinity of the functional monomers for the template. However, if the monomers display sufficient affinity for the template to form a complex the binodal will rapidly be reached or has been passed before polymerization begins. This is not possible as the complexes would rapidly separate, aggregate and precipitate, resulting in no MIPs. The monomers must then be relatively low affinity, as would be expected from standard analysis of intermolecular interactions. If an

individual M-T interaction is less favored than each being solvated the probability of a complex with more than one monomer existing is very low (Figure 5.1).

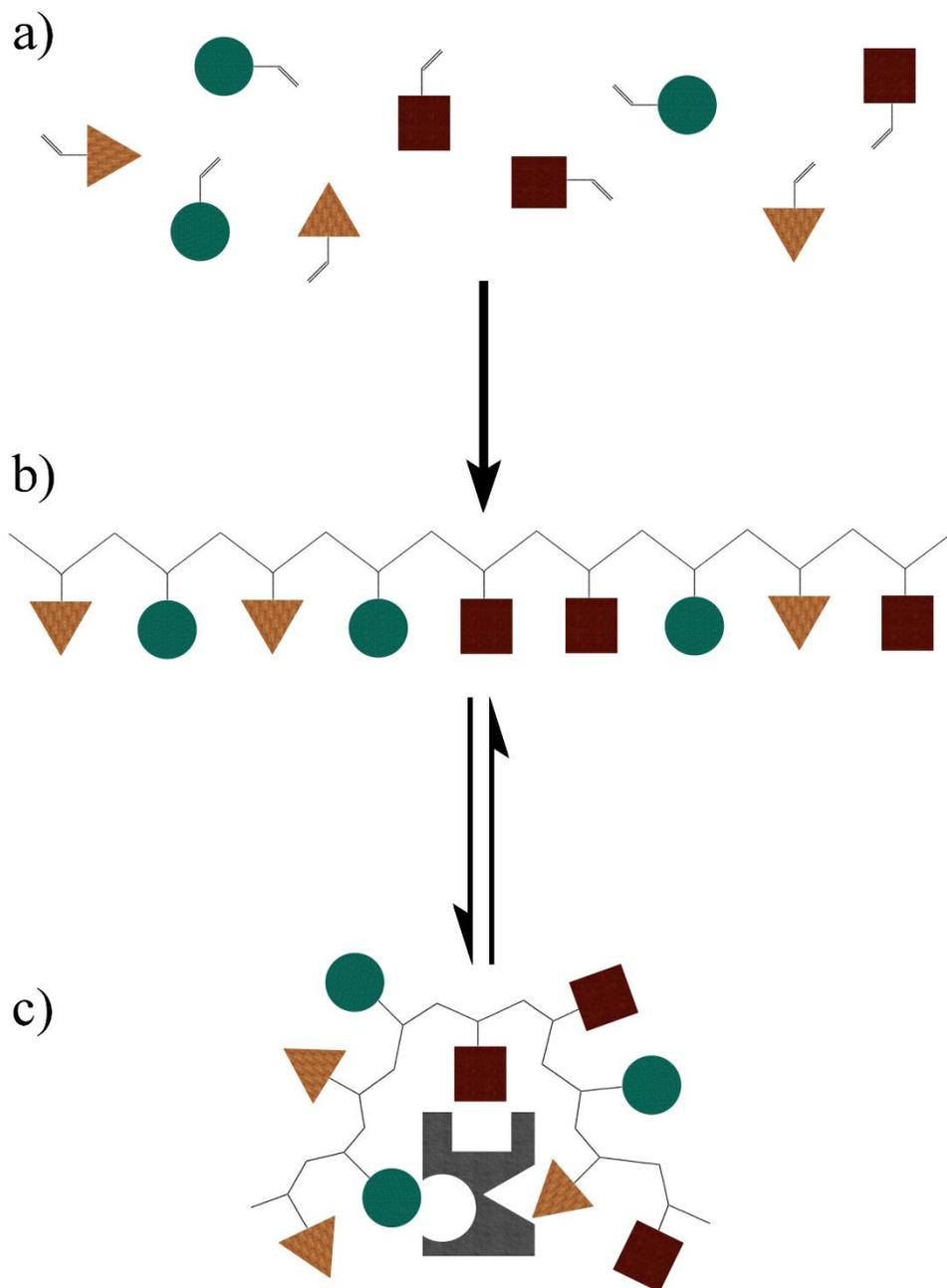


*Figure 5.1: States described by the current research; a) relatively low affinity monomers; b) the oligomer in equilibrium between template and solvent; c) the final imprinted polymer.*

The affinity of each monomer for the template will change very little on polymerization, and the contribution from adjacent polymer structure may be small (depending on the cross-linker and the template), but the affinity of the oligomers for the solvent is reduced relative to that of the monomer. The affinity of the template for what will become the binding site however will increase rapidly, being proportional to the sum of the functional groups plus any contribution from the polymer network, which may be large depending on the cross-linker.

As the polymer develops therefore the solubility decreases and the relative affinity for the template increases, reaching equilibrium between template and solvent at the binodal. As the polymer is already being synthesized and the binodal has not been reached, i.e. the polymer is more stable in solution, the binding site of the MIP must necessarily be synthesized predominantly in the absence of the template. This binding site formation must however occur during synthesis as the complete polymer will be non-imprinted if it reaches the spinodal phase boundary in the absence of template.

Binding site formation likely occurs close to the binodal therefore, when on encountering the template structural rearrangement is induced by the interactions due the increasing affinity of the polymer for the template relative to the affinity for the solvent. The polymer will also then act to tether itself to the template, any region of low affinity relative to solvation being held in proximity covalently. This hypothesis is given schematically in Figure 5.2. Interestingly this model is close to that proposed by Dickey in 1949, and Pauling's model of antibody production.



*Figure 5.2: Hypothetical mechanism of imprinted polymer synthesis (cross-linkers are omitted for schematic simplicity): a) Monomers are semi-randomly arranged in solution, and interactions between monomer and template are short lived; b) On polymerization the monomer units are arranged through the polymer largely without consideration for the template; c) As the polymer develops the affinity for the solvent is reduced and template binding occurs more frequently, with the random polymer being arranged into the most favorable arrangement for template binding.*

On reaching the binodal any further polymerization of the binding site will occur predominantly while bound to the template nucleant, the point of equilibrium being reached, but the polymer will then be in a metastable state and close to or within the process of separating from the solvent. Polymerization may continue from here, and it is possible that further reactions occur within the binding site, with the volume of the polymer that is phase-separated possibly undergoing a process similar to autoacceleration in the absence of terminating agents.<sup>311, 312</sup> More likely however the binding site and immediate surroundings will not undergo further reactions beyond the binodal as these would require chemical changes which might compromise the existing affinity. Further polymerization produces greater cross-linking which scaffolds the binding site cavity structure, in addition to further reducing the polymer solubility; both of these processes act to strengthen the affinity for the template, both absolutely and relatively by forcing the polymer deeper into the binodal condition.

Some empirical evidence is also available to verify this mechanism. MIP NPs are routinely prepared by a solid phase synthesis approach, in which a template is immobilized on a solid surface before polymerization is induced. A wash of cold solvent is then used to remove reagents and polymer with little/no affinity, before a second wash with hot solvent removes the high affinity MIPs. In the hot wash it can be observed that for every MIP NP synthesized many 'protoMIP' nanoparticles (pNPs) are also present. These particles are small, and often cannot be observed in routine analysis (e.g. DLS). While the pNPs show low affinity relative to the standard MIP NPs, the affinity is still high enough for them to be retained on the immobilized template through the cold wash.

There are two points of relevance here. Firstly, that the particles cannot be observed by DLS suggests that the particles do not behave as a separate phase in solvent, meaning that they have not reached the spinodal (spontaneous polymer phase separation). That the particles are retained on the template however suggests that they have passed the binodal, the pNPs necessarily favoring the template nucleant to the solvent to survive the cold wash. This provides further support to the conclusions drawn in Chapter 4.

The properties of pNPs also support the synthetic mechanism presented previously and in given in Figure 5.2. The template affinity, high relative to non-imprinted particles (which are removed in the cold wash) but low relative to standard MIP NPs, suggests that the proto-MIPs are the result of 'incomplete' synthesis, suspended between the binodal and spinodal. From this it can be concluded that MIP nanoparticles develop affinity in synthesis. If pNPs are a polymer beyond the binodal they are essentially an unsupported MIP binding site, without sufficient cross-linking and scaffolding to induce the full affinity. The significance of the pNP being an unsupported MIP binding site is that the binding site needs supporting. If the binding site is unstable relative to the whole MIP, with bonding groups presumably dissociating according to their individual equilibria regardless of cross-linking preventing full escape, the complex that precedes this cannot be stable in solution. The pre-polymerization mixture then will consist of predominantly randomly arranged monomers, which result in oligomer of random composition, which produce imprinted binding sites and ultimately imprinted polymers. For this procession to be followed there must be an intermediate state in which the random oligomer is organized into a favorable arrangement around the template, as in Figure 5.2.

By producing effective molecular models of polymer solvation it may be possible to test this hypothesis more thoroughly and develop more realistic theories of imprinted polymer synthesis. However, techniques designed for the analysis of polymers, or the MSPU, must first be produced.

### 5.3. Determining the Gibbs energy of solvation

The Gibbs energy of solvation  $\Delta G_{solv}$ , can be predicted using MM/MD and thermodynamic integration (Equation 5.1)<sup>313</sup>:

$$\Delta G_{solv} = \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda} \partial \lambda = \int_0^1 \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \partial \lambda \quad \text{Eq. 5.1}$$

Where  $\lambda$  is the coupling parameter, with a value from 1 (solute) to 0 (cavity), and  $U$  is the potential energy of the interaction. By repeatedly moving from  $\lambda = 1$  to 0 with intermittent MD simulations, an ensemble average can be recorded, the integral of which gives  $\Delta G_{solv}$ .<sup>314, 315</sup> The value and relevance of  $\Delta G_{solv}$  in this context is that it can be determined theoretically with molecular dynamics, and that it can be used to determine values of the Flory parameter of solvent polymer interaction,  $\chi_{s,p}$ , and potentially also the Gibbs energy of mixing,  $\Delta G_m$ .<sup>316</sup> If these can be established theoretically, the values of  $\chi_{spinodal}$  and  $\Delta \chi_{spinodal}$  could also be determined, permitting nanoparticle property prediction for any given starting mixture of monomers.

A short script was developed to automate the actions required for the calculation of  $\Delta G_{solv}$  by the above method (Appendix 2.4). Values for  $\lambda$  are first set to move from 1 to 0 in 0.05 intervals with picosecond MD simulations between each period. Before each MD simulation the values of all atomic charges in 'substructure 1' (the analyte) are multiplied by the  $\lambda$  value, and various commonly used atom types are redefined to possess an electronegativity and formal charge equal to their standard value multiplied by  $\lambda$ . This is possible if alternative atoms are used for the solvent molecules, in this case TIP3P water atoms in place of the standard oxygen and hydrogens. After each MD interval the energy of the system, analyte, and solvent with removed analyte are recorded and saved as a separate file for later analysis, and the cycle of  $\lambda$  from 1 to 0 is repeated 10 times. The result from the analysis of a single compound of interest by application of the script in Appendix 2.4 is therefore 600 individual files, each containing the recorded potential energy of the analyte, solvent or solution for each of the 20  $\lambda$  values. The difference in energy between the solution and the combined energy

of the solvent and analyte is then recorded, allowing calculation of the mean value for each  $\lambda$ , and therefore the change of potential energy with mean  $\lambda$ . This process was performed manually, but with automation could be an efficient method of determining the free energy of solvation.

Tripes force fields were applied with MMFF94 charges throughout the solubility studies. This combination was initially applied simply for consistency with previous experiments, and to facilitate application of the polymerization as described in Chapter 3. Observation of deviation in  $\Delta G_{solv}$  encourages analysis of alternative combinations, MMFF94 charges would be expected to produce more reliable results with MMFF94 force fields for example, or application of more broadly applicable charges (e.g. Gasteiger-Hückel) would at least give more broadly applicable to the methods applied. However, the algorithm written for analysis relies on modifying the electronegativity and charge of the atom, actions which are forbidden for all except Tripes atoms in the Sybyl software. Maintaining Tripes force fields and varying the charges was then attempted, but with all except MMFF94 the calculated energy of the aqueous analyte was extremely high, with ambient water density (as calculated automatically by the software, which were equal to those found using Equation 3.2) simulating high pressures and temperatures. This appears to be a result of using TIP3P water. The deviation from experimental (and theoretical) values of  $\Delta G_{solv}$  with the applied method was analyzed for any apparent correlation, which would then allow observation of a polymer system.

#### 5.4. Preliminary analysis of $\Delta G_{solv}$ calculation

Experimental values for the Gibbs free energy of solvation of various molecules was compared with values calculated using the molecular mechanics based approach described in the previous section. While alternative approaches using similar techniques but yielding more accurate results can be found readily in the literature, these methods are not compatible with the polymer analysis which was the ultimate purpose of this calibration.

In practice the script simulates the analyte having gradually reduced interactions with a solvent until all that effectively remains is a cavity occupying a volume equal to the analyte. The difference between the energy of the full system and that of the isolated analyte plus system without analyte (just solvent) then gives an interaction energy. Averaging over all discrete  $\lambda$  intervals obtained in repetition and plotted as a function of  $\lambda$  then allows estimation of  $\Delta G_{solv}$ . Initial experiments were performed using *N,N*-dimethyl-*p*-methoxybenzamide (DMMB), which has an experimental free energy of solvation of  $-11.01 \text{ kcal mol}^{-1}$  ( $-46.066 \text{ kJ mol}^{-1}$ ) with published theoretical calculations with various techniques all predicting higher values (to approximately  $-7 \text{ kcal mol}^{-1}$ ).<sup>317</sup> The method developed and described above gave a value of  $-15.741 \text{ kcal mol}^{-1}$  for the Gibbs free energy of solvation, a deviation which may occur from a number of difficult to avoid methodological variables (the force field-charge combination, use of TIP3P water, length of MD simulations, etc.) over simplicity in the written algorithm (some analyte intermolecular potential energy is retained at  $\lambda = 0$  for example) and/or inherent limitation of molecular mechanics.

To overcome the deviation between the model and empirical values of  $\Delta G_{solv}$  a number of small molecules were analyzed over 10 cycles of  $\lambda$  1-0 in 0.05 intervals with 5 ps equilibrations. The theoretical values obtained were plotted vs. experimental values for observation of any relationship which may justify continuation onto analysis of polymers. The results are given in Figure 5.3.

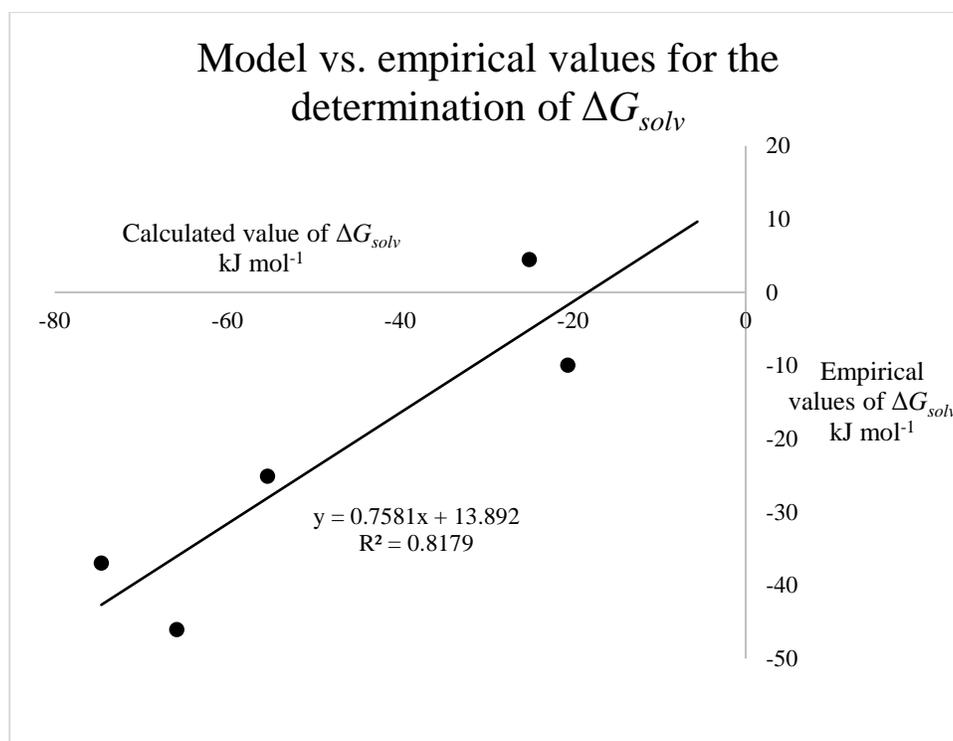


Figure 5.3: Values for the Gibbs energy of solvation as calculated with the script given in Appendix 2.4, and empirical values taken from the literature.<sup>317</sup>

Table 5.1: Calculated and empirical values of  $\Delta G_{solv}$  for various compounds.<sup>317</sup> The calculated values adjusted for residual analyte energy can be observed to be more accurate in the first three rows.

Compound	Calculated $\Delta G_{solv}$ , kJ mol <sup>-1</sup>	Calculated $\Delta G_{solv}$ plus residual energy, kJ mol <sup>-1</sup>	Empirical $\Delta G_{solv}$ , kJ mol <sup>-1</sup>
Diethyl propanedioate	-55.371	-26.623	-25.104
Glycerol triacetate	-74.605	-31.468	-36.986
Benzyl bromide	-20.588	-11.643	-9.958
<i>m</i> -bis(trifluoromethyl) benzene	-25.080	-16.034	+4.477
<i>N,N</i> -dimethyl- <i>p</i> -methoxybenzamide	-65.862	-19.854	-46.066

All calculated values of  $\Delta G_{solv}$  are below their empirically determined values, though an approximately linear relationship can be observed between the two. More appropriate values were also given when the mean residual analyte potential energy was added to the calculated free energy of solvation, as shown in Table 5.1. This method appeared to work some of the compounds analyzed (top three rows) but not for others. For the analysis of the polymer therefore, both would be used to predict the Gibbs energy of solvation.

### 5.5. Calculation of $\Delta G_{solv}$ for polymers

An analogue of the polymers used in Chapter 4 equivalent to two MSPU was initially synthesized *in silico*. From the ratios applied in the empirical studies, 2 Bis, 5 AAc, 40 TBAm and 53 NiPAm were initially added randomly to form a sphere. This sphere was then solvated with 1000 TIP3P water molecules and the script given in Appendix 2.2 was applied to fit the molecules into a cube of dimensions given by Equation 3.2. A brief examination was performed to ensure there were no irregularities before the algorithm given in Appendix 2.4 was retrieved to initiate MD equilibration polymerization. Simulated annealing was performed on the prepolymerization mixture of solvated monomers in five cycles of 800 K to 300 K with 100 K steps for 1 ps at each temperature. A 3 ns equilibration then followed at 300 K before the polymerization was initiated, with a maximum bonding distance set at 5.7 Å, as was determined appropriate in Chapter 3. This process was then repeated with five 100 ps and two 250 ps cycles of MD re-equilibrations followed by polymerization. After this the polymerization was repeated with maximum bonding distances of 30 Å and 100 Å, with a final manual bond forming of (ring) oligomers to form one single polymer macromolecule. The MSPU polymer then underwent a brief *in vacuo* MD simulation and minimization, before being solvated in 10,000 TIP3P water molecules. The box dimensions were then given by the total volume of the of the reagent monomer and water molecules (320.662 nm<sup>3</sup>) and an

equilibration simulation of 10 ps at 300 K was performed before the  $\Delta G_{solv}$  script (Appendix 2.4) was recalled. The value of  $\lambda$  was reduced from 1 to 0 in 0.05 increments over ten cycles with 1000 fs equilibration periods at each  $\lambda$ .

A problem that was not encountered with the previous studies was an automatic increase in the size of the simulation volume during the course of the molecular dynamics equilibration. MD had been performed with the NTV canonical ensemble for all previous studies, but possibly due to the large size of the system analyzed, the software was found to automatically adjust to NTP or NEP, before the software became unresponsive. Varying the parameters was not found to resolve this issue, and no way to inhibit this modification was found. From the data obtain a value of  $\Delta G_{solv}$  equal to 799.336 kJ mol<sup>-1</sup> was given, accounting for residual energy, 2502.868 kJ mol<sup>-1</sup>, or by the formula given in Figure 5.3, 619.869 kJ mol<sup>-1</sup>.

## 5.6. Conclusions

Attempts were made to determine  $\Delta G_{solv}$  for the polymer MSPU in order to improve calculations of  $\Delta\chi$ , thus giving more accurate predictions of nanoparticle size from the synthetic conditions. A semi-automated process was designed and engineered for the calculations, employing the algorithm in Appendix 2.4. Initial studies with small molecules suggested that the results would be inaccurate but may correlate with experimental values. Applying molecular dynamics and the polymerization algorithm described in Chapter 3, a small representation of the polymers produced for Chapter 4 was synthesized *in silico*. Attempts to determine  $\Delta G_{solv}$  using this polymer representation have so far been unsuccessful however, and inconsistency in the results for the small molecule analysis suggests an alternative method of calculating  $\Delta G_{solv}$  should be sought.

Additionally, methods superior to  $\Delta G_{solv}$  calculation are likely available, and could be pursued in the attempts to improve polymer nanoparticle property prediction. The solubility parameter model has demonstrated sufficiency, and the potential benefit of more accurate values from MD analysis may not necessarily justify the loss of efficiency in practice. One possible future area could be the development of a technique to automatically calculate solubility parameters, using the standard values, from a given ratio of monomers or polymer MSPU. Given sufficient prior experimental data, thermodynamics could also be calculated and ultimate nanoparticle properties predicted from a given starting monomer composition. For now however this work will be suspended, and any immediate calculations that may be necessary will be performed using data tables with ink and paper.

## 6. Conclusion

Karl Popper believed that the work of scientists, and thus that the process of scientific discovery, 'consists in putting forward and testing theories'.<sup>318</sup> In the language of Thomas Kuhn's work these correspond roughly to revolutionary, paradigm shifting science and 'normal science' respectively.<sup>319</sup> While both of these authors have declined in popularity, including accusations that both their philosophies are fundamentally impractical and irrational,<sup>320</sup> some strict descriptions of science may be helpful in contextualizing the work presented in this thesis. Popper describes four necessary tests of a scientific theory: internal consistency of conclusions, logical consistency of the theory, comparison with other theories to determine whether the theory constitutes an advancement, and empirical testing.<sup>318</sup>

The research described in Chapter 4 appears to conform to the first, second and fourth test. The third requirement that the work presents an advance in the context of existing knowledge. The application of models typically applied for polymer blends and macroscopic polymer synthesis to the preparation of polymer nanoparticles constitutes a natural development of existing theory, including the addition of the emphasis on the value of the  $\Delta\chi$  term. It is inherently 'normal science' as it is based on FHT and pre-existing phase boundary definitions. The emphasis on the binodal character of MIP synthesis and the mechanism described in Section 5.2 may yet contribute to something more significant in the area of molecular imprinting, but further analysis will be required before this presumption is properly applicable.

One possible method of gather further empirical data for the testing of these hypotheses would be experimentation with polymerization in helium nanodroplets. These droplets are composed of a controllable number of between a thousand and a million individual atoms, and readily incorporate both volatile and non-volatile materials.<sup>321</sup> As helium is chemically inert and can form quantum vortices (a hole in the superfluid around which the helium atoms circulate), the droplet can be manipulated with sufficient precision to form one-dimensional particles with even the most reactive substances.<sup>322</sup> Typically however, particles which enter the droplet cluster together,

and reactions forming covalent bonds between neutral species at appropriate rates have been observed.<sup>323</sup> These helium droplets therefore have potential as polymerization reactor nano-vessels, enabling the formation of polymer particles with just a few monomers, and permitting detailed analysis of MIP binding site formation.

As the majority of research into imprinted polymers remains focused on technological development however, work like that described in Chapter 2 and 3 will likely be of more immediate interest to researchers in the field. More useful concepts here than those of Popper and Kuhn could be those imagined by Francis Bacon: *lucifera*, bringers of light, and *fructifera*, bringers of fruit, analogues to Freeman Dyson's 'concept-driven science' and 'tool-driven science'.<sup>324</sup> To describe the research presented in Chapter 4 as *lucifera* may be somewhat grandiose, but the applications of molecular mechanics for MIP design and analysis described in Chapter 2, and to some limited extent the *in silico* polymerization tool described in Chapter 3, certainly appear to be effective in bringing fruit.

The polymerization algorithm as a technology was applied for the beginning of a more detailed study into developing the model relating polymer solubility to nanoparticle size and observations of the mechanisms involved in this phenomenon. The methods applied were based on existing theory and require much more study, but regardless lay foundations for further development of the model described in Chapter 4 and the hypothetical synthetic mechanism proposed in Section 5.2. The latter of these, that the imprinted polymer is formed largely in the absence of the template, could have significant implications for future MIP design and production. The evidence collected is interesting in that it supports the original proposal by Dickey that MIP synthesis is analogous to Pauling's model of antibody formation, individual solvated polymers being formed in the absence of template but induced into binding sites when the template is presented.

How this relates the prediction also given by binodal-character phase separation of a regular structure to MIP binding sites resulting from the separation mechanism has yet to be determined. The Pauling-style mechanism of synthesis and the regular binding site are not incompatible however as a polymer with sufficient flexibility is likely to adopt an arrangement of highest interaction with the template nucleant. Additionally,

both conclusions can be correct and still not invalidate the typical monomer selection procedure described in Chapter 1 and demonstrated in Chapter 2, as the success rate and affinity will still depend on the strength of monomer-template interactions. However, if subsequent evidence and theoretical analysis can verify the mechanism, further measures could be applied to further control the properties of imprinted polymers by application of these models.

Individual examples of research and extended research programs specifically studying models of MIPs are relatively rare currently, the focus generally being focused on the development of new imprinted polymer based technologies and the adaptation of existing technologies for commercial production. This is demonstrably not due to an abundance of existing MIP theory however, and the future advance of imprinted polymer based biotechnology and nanotechnology will inevitably be proportional to the foundational science.

## Appendix 1. Docking of drugs with 4-aminobenzoic acid oligomer

		(S)-MD+	(S)-MD	(R)-MD+	(R)-MD	Quin+	Quin	Cod+	Cod	Coca+	Coca
<b>Protonated</b>		3.97	4	3.7	3.02	2.63	3.12	2.27	2.37	1.94	2.61
		3.77	3.46	3.33	2.8	2.52	3.07	1.99	2.16	1.78	1.82
		3.03	2.74	3.18	2.59	2.19	2.69	1.74	1.8	1.5	1.69
		2.99	2.56	3.06	2.57	2.03	2.54	1.68	1.8	1.46	0
		2.81	1.39	2.82	2.15	0.32	2.53	1.53	1.25	1.19	0
SD	All	0.519115	0.99	0.32775	0.322692	0.936413	0.286094	0.291153	0.42665	0.292541	1.171508
	Exc. High	0.424264	0.858152	0.215465	0.27232	0.9847	0.252504	0.191572	0.375533	0.241437	1.014639
	Exc. 2 High	0.117189	0.733008	0.183303	0.248462	1.036549	0.089629	0.108167	0.317543	0.168622	0.975722
	Exc. Low	0.503455	0.665432	0.277894	0.210792	0.279926	0.284546	0.269196	0.281824	0.229492	1.098029
	Exc. 2 Low	0.495177	0.632139	0.267644	0.215019	0.228983	0.23516	0.265141	0.288271	0.222711	0.497896
	Exc. High & Low	0.439242	0.476235	0.135277	0.12741	0.249867	0.273191	0.164418	0.207846	0.174356	1.015332
mean		2.943333	2.92	3.19	2.653333	2.446667	2.586667	1.65	1.92	1.383333	2.04
<hr/>											
		(S)-MD+	(S)-MD	(R)-MD+	(R)-MD	Quin+	Quin	Cod+	Cod	Coca+	Coca
<b>half deprotonated</b>		4.12	3.4	4.18	3.5	3.92	3.25	3.11	3.28	2.83	2.6
		3.83	2.88	3.18	2.29	2.54	2.94	2.64	1.81	1.19	0.96
		3.49	2.87	3.14	2.23	2.01	2.22	1.53	1.25	0.83	0
		3.03	1.39	2.87	2.15	1.78	0.32	1.42	1.09	0.54	0
		2.96	0	2.8	0	1.68	0	0.03	0	0	0
All	All	0.501228	1.396879	0.553968	1.265279	0.919772	1.499527	1.199721	1.198345	1.071807	1.134337
	Exc. High	0.409257	1.38064	0.190504	1.113145	0.384046	1.433736	1.069533	0.757424	0.502726	0.48
	Exc. 2 High	0.287924	1.435235	0.179536	1.26503	0.169214	1.200056	0.836082	0.680221	0.421228	0
	Exc. Low	0.468713	0.866122	0.575058	0.640904	0.959318	1.314493	0.831966	0.997309	1.02347	1.22654
	Exc. 2 Low	0.315331	0.30315	0.589237	0.716543	0.986019	0.528425	0.811316	1.048443	1.066083	1.314737
	Exc. High & Low	0.401497	0.85738	0.168622	0.070238	0.389744	1.353563	0.674858	0.378065	0.325628	0.554256
mean		3.16	3.05	3.063333	2.223333	2.0025	2.803333	1.863333	1.383333	0.853333	0
<hr/>											
		(S)-MD+	(S)-MD	(R)-MD+	(R)-MD	Quin+	Quin	Cod+	Cod	Coca+	Coca
<b>total deprotonation</b>		3.46	3.18	3.24	3.08	2.2	1.79	2.55	2.25	2.19	2.14
		2.56		2.87							
		1.32									
		3.91	3.38	3.96	3.32	2.4	2.51	2.74	3.02	1.72	2.2
		1.02									
		3.4	3.27	3.78	3	2.95	1.34	2.79	1.44	1.46	2.87
		3.15		3.31	2.81		2.52		1.15		
	3.27										

## Appendix 2. Algorithms

Below are a series of algorithms used frequently in the computational work. All are written in Sybyl programming language (SPL) and were designed for use with Sybyl 7.3.

### 2.1. Polymerization algorithm

for just\_to\_tidy in 1

```
setvar SName %prompt(filename x "filename for MD files" "x")
setvar sim_ann %prompt(YES_NO YES "perform dynamics and/or simulated annealing?")
if %streql( $sim_ann YES)
setvar SAs %prompt(int 10000 "fs at each temperature" "10000")
setvar starting_temp %prompt(int 1000 "maximum temperature, K" "1000")
setvar end_temp %prompt(int 300 "minimum temperature, K" "300")
setvar step %prompt(int 100 "step size/gradient" "100")
setvar cycles %prompt(int 10 "number of annealing cycles" "100")
setvar SAend %prompt(int 10000 "fs of final 300 K" "10,000")
setvar record %prompt(int 1000 "fs between recording" "1000")
endif

setvar max_bonding_distance %prompt(int 460 "maximum bonding distance (pm)" "the maximum interatomic distance at
which a bond can form between reactive species")
setvar max_distance %math( $max_bonding_distance / 100 )
setvar range_increments %prompt(int 20 "incremental interatomic binding distance (pm)" "The reactive atom atom scans
the surroundings for other reactive atoms. The range of this scan begins at 300 pm (or less) and will increase to the
maximum bonding distance by the selected incremental binding distance.")
setvar range_increment %math( $range_increments / 100 )
setvar grad %prompt(anything 0.5 "minimisation gradient" "minimisation gradient")

setvar reg_min %prompt(YES_NO NO "include regular minimisation" "not including these regular energy minimisation
may my give faster synthesis and possibly more accurate structure")

setvar final_min %prompt(YES_NO NO "end polymerisation with low gradient minimisation" "energy minimisation of
max 100,000 iterations" )

setvar cycle_save %prompt(YES_NO NO "save the system after each cycle" "saved as [name]_cycle_[cycle number]" )

setvar all %prompt(YES_NO YES "all reactive molecules are initiators" "yes for all reactive molecules as potential
initiators, no to select an initiator molecule" )

if %streql( $sim_ann YES)
for SAX in %range(1 $cycles )

for SAy in %range(0 1000)
if %gt( %math( $starting_temp - %math( $step * $SAy )) $end_temp )
dynamics m1 setup $SName YES DONE interval_length $SAs data_write $record temperature %math( $starting_temp
- %math( $step * $SAy )) coupling_factors 100 0 DONE FINISHED interactive
endif
endifor

endifor

dynamics m1 setup $SName YES DONE interval_length $SAend data_write $record temperature 300 coupling_factors
100 0 DONE FINISHED interactive
```

```

if %streql( $sim_ann YES)
  mol2 out m1 %cat( $SAname _ prepoly )
endif
endif

if %streql( $all YES )
  setvar ks *
else
  setvar select_k %prompt(int 1 "select initiator molecule" "type the initiator molecule's number")
  setvar kt %subst( %cat( # $select_k ) )
  modify substructure name $kt initiator
  setvar ks initiator
endif
setvar starting_bonds %count(%bonds(*))
setvar starting_time %time()
setvar starting_mols %count(%subst(*))
setvar number_bonds 1

for cycling in %range( 1 1000 )
  if %not(%eq( %count(%bonds(*)) $number_bonds )
setvar number_bonds %count(%bonds(*))
echo xxxxxxxx cycle: $cycling time: %time() xxxxxxxx
  if %streql( $cycle_save YES )
    if %not(%eq( $cycling 1))
      mol2 out m1 %cat( $SAname _ cycle_ $cycling )
    endif
  endif
endif

for k in %subst( $ks )
  if %or( %atoms((alpha)-({#1} )
))) %atoms((beta)-({$k})))
TAILOR SET MAXIMIN2 RMS_GRADIENT $grad ^

if %streql( $reg_min YES )
tailor set maximin2 MAXIMUM_ITERATIONS 2000 ^
maximin2 m1 DONE interactive
endif
tailor set maximin2 MAXIMUM_ITERATIONS 50 ^

for name in %range(1 1000)
  if %or(%atoms(m1(ALPHA-({$k} ) )))
  modify atom name %choose(%atoms(m1(ALPHA-({$k} ) ))) %cat( origin_ $name )
  endif
endfor
for name in %range(1 1000)
  if %or(%atoms(m1(BETA-({$k} ) )))
  modify atom name %choose(%atoms(m1(BETA-({$k} ) ))) %cat( originB_ $name )
  endif
endfor
for set in %range(1 1000)
  if %or(%atoms(m1(%cat( origin_ $set )-({$k} ) )))
  setvar %cat( origin_ $set ) %atoms(m1(%cat( origin_ $set )-({$k} ) )))
  endif
endfor
for set in %range(1 1000)
  if %or(%atoms(m1(%cat( originB_ $set )-({$k} ) )))
  setvar %cat( originB_ $set ) %atoms(m1(%cat( originB_ $set )-({$k} ) )))
  endif
endfor
for atom in %range(1 1000)
  if %or( %atoms(%cat( origin_ $atom ) %atoms(%cat( originB_ $atom ) )
  for origin_atom in %cat( origin_ $atom ) %cat( originB_ $atom )

```

```

if %or( %atoms( $origin_atom ) )
  setvar var_close_1
  setvar var_close_2
  setvar close_a
for close in %atoms(<<c.2>>+(REACTIVE))
if %lteq(%distance( %atoms( $origin_atom ) $close ) $max_distance )
if %not(%streql( %atoms( $origin_atom ) $close ))
  setvar var_close_1 $close
  setvar var_close_2 $var_close_1 $close_a
  setvar close_a $var_close_2
endif
endif
endfor
for distance in %range(3 $max_distance $range_increment )
for close_atom in $close_a
if %lteq(%distance( %atoms( $origin_atom ) $close_atom ) $distance )

if %streql( %atom_info( $close_atom ) ALPHA )
if %eq(%count(%atom_info( $origin_atom neighbors)) 3)
if %eq(%count(%atom_info( $close_atom neighbors)) 3)
  setvar between %cat( %atoms( $origin_atom ) : %atoms( $close_atom ) )
if %not(%eq(%count(%atoms( $between )) 4))
  modify atom name $origin_atom new_bond
  modify atom name $close_atom new_bondx
for g in %atom_info( %atoms(new_bond) neighbors)
  modify bond type %bond_info( %cat( $g = %atoms(new_bond) ) id ) 1
endfor
for g in %atom_info( %atoms(new_bondx) neighbors)
  modify bond type %bond_info( %cat( $g = %atoms(new_bondx) ) id ) 1
endfor
add bond new_bond new_bondx 1
maximin2 m1 MARK_CHARGES_VALID DONE interactive
modify atom type %atoms(new_bond) c.3
modify atom type %atoms(new_bondx) c.3
modify atom name %atoms(new_bond) done_x
modify atom name %atoms(new_bondx) done_y
maximin2 m1 MARK_CHARGES_VALID DONE interactive

for x in %atoms(m1(alpha-((*)- ({{ $k }}))) )
if %streql(%atom_info( $x type) C.3)
  modify atom name $x sp3
endif
endfor
for x in %atoms(m1(beta-((*)- ({{ $k }}))) )
if %streql(%atom_info( $x type) C.3)
  modify atom name $x sp3
endif
endfor
for n in %range(1 1000)
for x in %atoms(m1(alpha-((*)- ({{ $k }}))) )
if %or( $x )
  if %not(%atoms(m1(%cat( origin_ $n )-((*)- ({{ $k }}))) )
    modify atom name $x %cat( origin_ $n )
  endif
else
  return
endif
endfor
endfor
for n in %range(1 1000)
for x in %atoms(m1(beta-((*)- ({{ $k }}))) )
if %or( $x )

```

```

    if %not(%atoms(m1(%cat( originB_ $n )-((*)- ({{ $k }}) )))
      modify atom name $x %cat( originB_ $n )
    endif
  else
    return
  endif
endfor
endfor
endif
endif
endif
endif
for number in %range(1 1000)
  if %or(%atoms(%cat( origin_ $number )))
    if %streql( %atom_info( $close_atom name ) %CAT( ORIGIN_ $number ) )
    if %not(%streql( %atom_info( $origin_atom name ) %CAT( ORIGIN_ $number ) ))
    if %not(%streql( %atoms( $close_atom ) %max(%atom_info( $origin_atom neighbors)) ))
    if %not(%streql( %atoms( $close_atom ) %min(%atom_info( $origin_atom neighbors)) ))
    if %not(%streql( %atoms( $close_atom ) %stats(%atom_info( $origin_atom neighbors) median) ))
    if %eq(%count(%atom_info( $origin_atom neighbors)) 3)
    if %eq(%count(%atom_info( $close_atom neighbors)) 3)
    setvar between %cat( %atoms( $origin_atom ) : %atoms( $close_atom ) )
    if %not(%eq(%count(%atoms( $between ) ) 4))
      modify atom name $origin_atom new_bond
      modify atom name $close_atom new_bondx
      for g in %atom_info( %atoms(new_bond) neighbors)
        modify bond type %bond_info( %cat( $g = %atoms(new_bond) ) id ) 1
      endfor
      for g in %atom_info( %atoms(new_bondx) neighbors)
        modify bond type %bond_info( %cat( $g = %atoms(new_bondx) ) id ) 1
      endfor
      add bond new_bond new_bondx 1
      maximin2 m1 MARK_CHARGES_VALID DONE interactive
      modify atom type new_bond c.3
      modify atom type new_bondx c.3
      modify atom name %atoms(new_bond) done_x
      modify atom name %atoms(new_bondx) done_y
      maximin2 m1 MARK_CHARGES_VALID DONE interactive
      for x in %atoms(m1(alpha-((*)- ({{ $k }}) )))
        if %streql(%atom_info( $x type) C.3)
          modify atom name $x done
        endif
      endfor
      for x in %atoms(m1(beta-((*)- ({{ $k }}) )))
        if %streql(%atom_info( $x type) C.3)
          modify atom name $x done
        endif
      endfor
      for n in %range(1 1000)
        for x in %atoms(m1(alpha-((*)- ({{ $k }}) )))
          if %or( $x )
            if %not(%atoms(m1(%cat( origin_ $n )-((*)- ({{ $k }}) )))
              modify atom name $x %cat( origin_ $n )
            endif
          else
            return
          endif
        endfor
      endfor
      for n in %range(1 1000)
        for x in %atoms(m1(beta-((*)- ({{ $k }}) )))
          if %or( $x )

```



```

    if %not(%atoms(m1(%cat( originB_ $n )-((*)- ({{ $k }} ) )))
      modify atom name $x %cat( originB_ $n )
    endif
  else
    return
  endif
endfor
endif
endif
endif
endif
for number in %range(1 1000)
  if %or(%atoms(%cat( originB_ $number )))
    if %streql( %atom_info( $close_atom name ) %CAT(ORIGINB_ $number ) )
    if %not(%streql( %atom_info( $origin_atom name ) %CAT(ORIGINB_ $number ) ) )
    if %not(%streql( %atoms( $close_atom ) %max(%atom_info( $origin_atom neighbors)) ) )
    if %not(%streql( %atoms( $close_atom ) %min(%atom_info( $origin_atom neighbors)) ) )
    if %not(%streql( %atoms( $close_atom ) %stats(%atom_info( $origin_atom neighbors) median) ) )
    if %eq(%count(%atom_info( $origin_atom neighbors)) 3)
    if %eq(%count(%atom_info( $close_atom neighbors)) 3)
    setvar between %cat( %atoms( $origin_atom ) : %atoms( $close_atom ) )
    if %not(%eq(%count(%atoms( $between ) ) 4))
      modify atom name $origin_atom new_bond
      modify atom name $close_atom new_bondx
      for g in %atom_info( %atoms(new_bond) neighbors)
        modify bond type %bond_info( %cat( $g = %atoms(new_bond) ) id ) 1
      endfor
      for g in %atom_info( %atoms(new_bondx) neighbors)
        modify bond type %bond_info( %cat( $g = %atoms(new_bondx) ) id ) 1
      endfor
      add bond new_bond new_bondx 1
      maximin2 m1 MARK_CHARGES_VALID DONE interactive
      modify atom type new_bond c.3
      modify atom type new_bondx c.3
      modify atom name %atoms(new_bond) done_x
      modify atom name %atoms(new_bondx) done_y
      maximin2 m1 MARK_CHARGES_VALID DONE interactive
      for x in %atoms(m1(alpha-((*)- ({{ $k }} ) )))
        if %streql(%atom_info( $x type) C.3)
          modify atom name $x sp3
        endif
      endfor
      for x in %atoms(m1(beta-((*)- ({{ $k }} ) )))
        if %streql(%atom_info( $x type) C.3)
          modify atom name $x sp3
        endif
      endfor
      for n in %range(1 1000)
        for x in %atoms(m1(alpha-((*)- ({{ $k }} ) )))
          if %or( $x )
            if %not(%atoms(m1(%cat( origin_ $n )-((*)- ({{ $k }} ) )))
              modify atom name $x %cat( origin_ $n )
            endif
          else
            return
          endif
        endfor
      endfor
      for n in %range(1 1000)
        for x in %atoms(m1(beta-((*)- ({{ $k }} ) )))
          if %or( $x )

```



```

        modify atom name $cbs reactive
      endif
    endfor
    if %streql( %atom_info( $cbs name) ALPHA )
      modify atom type $cbs c.2
      modify atom name $cbs alpha
    endif
  endif
  endif
  for cbs_n in %atom_info( $cbs neighbors)
    if %not(%streql( %atom_info( $cbs_n type) O.2 ) %streql( %atom_info( $cbs_n type) H ) %or( %streql(
%atom_info( $cbs_n name) DONE ) %streql( %atom_info( $cbs_n name) NEW_BOND ) %streql( %atom_info( $cbs_n
name) NEW_BOND ) )) )
      if %streql( %count(%atom_info( $cbs_n neighbors )) 4 )
        modify atom name $cbs_n done
      else
        if %streql( %count(%atom_info( $cbs_n neighbors )) 3 )
          if %streql( %atom_info( $cbs_n type ) C.3 )
            modify atom name $cbs_n alpha
            for j in %atom_info( $cbs_n neighbors)
              if %streql( %atom_info( $j type) C.2 )
                modify atom name $cbs_n reactive
              endif
            endfor
            if %streql( %atom_info( $cbs_n name) ALPHA )
              modify atom type $cbs_n c.2
              modify atom name $cbs_n alpha
            endif
          endif
        endif
        for cbs_nn in %atom_info( $cbs_n neighbors)
          if %not(%or( %streql( %atom_info( $cbs_nn type) O.2 ) %streql( %atom_info( $cbs_nn type) H ) %streql(
%atom_info( $cbs_nn name) DONE ) %streql( %atom_info( $cbs_nn name) NEW_BOND ) %streql( %atom_info(
$cbs_nn name) NEW_BOND ) )) )
            if %streql( %count(%atom_info( $cbs_nn neighbors )) 4 )
              modify atom name $cbs_nn done
            else
              if %streql( %count(%atom_info( $cbs_nn neighbors )) 3 )
                if %streql( %atom_info( $cbs_nn type ) C.3 )
                  modify atom name $cbs_nn alpha
                  for j in %atom_info( $cbs_nn neighbors)
                    if %streql( %atom_info( $j type) C.2 )
                      modify atom name $cbs_nn reactive
                    endif
                  endfor
                  if %streql( %atom_info( $cbs_nn name) ALPHA )
                    modify atom type $cbs_nn c.2
                    modify atom name $cbs_nn alpha
                  endif
                endif
              endif
            endif
            for cbs_nnn in %atom_info( $cbs_nn neighbors)
              if %not(%or( %streql( %atom_info( $cbs_nnn type) O.2 ) %streql( %atom_info( $cbs_nnn type) H ) %streql(
%atom_info( $cbs_nnn name) DONE ) %streql( %atom_info( $cbs_nnn name) NEW_BOND ) %streql( %atom_info(
$cbs_nnn name) NEW_BOND ) )) )
                if %streql( %count(%atom_info( $cbs_nnn neighbors )) 4 )
                  modify atom name $cbs_nnn done
                endif
              endif
            endfor
          endif
        endfor
      endif
    endfor
  endfor

```





```

if %or(%atoms(%cat( origin_ $no )))
  for a in %atoms(%cat( origin_ $no ))
    modify atom name $a ALPHA
  endfor
endif
endif
endif
for no in %range(1 1000)
  if %or(%atoms(%cat( originB_ $no )))
    for a in %atoms(%cat( originB_ $no ))
      modify atom name $a BETA
    endfor
  endif
endif
endif
endif
else
^
endif
endif
if %streql( $final_min YES )
  tailor set maximin2 MAXIMUM_ITERATIONS 100000 ^
  TAILOR SET MAXIMIN2 RMS_GRADIENT 0.01 ^
  maximin2 m1 DONE interactive
endif
tailor set maximin2 MAXIMUM_ITERATIONS 1000 ^
maximin2 m1 MARK_CHARGES_VALID DONE interactive
mol2 out m1 %cat( $SAname _ poly )
echo polymerisation started on $starting_time and finished on %time() . There are a total of %count(%bonds(*)) bonds;
%math( %count(%bonds(*)) - $starting_bonds ) new bonds have been formed. There were $starting_mols molecules
before the polymerisation, there are/is now %count(%subst(*)) individual molecule(s) in total.

endif

```

## 2.2. System preparation

This permits the molecules to be moved and compressed into a cube of selected dimensions

```
setvar max_x %prompt("anything" "0" "maximum x coordinate")
setvar min_x %prompt("anything" "0" "minimum x coordinate")
setvar max_y %prompt("anything" "0" "maximum y coordinate")
setvar min_y %prompt("anything" "0" "minimum y coordinate")
setvar max_z %prompt("anything" "0" "maximum z coordinate")
setvar min_z %prompt("anything" "0" "minimum z coordinate")

for atom in %atoms(*-(lrf+llf+ulf+urf+urr+ulr+lrr+llr))
if %gt( %atom_info( $atom x) $max_x )
for repeat in %range(1 50)
if %gt( %atom_info( $atom x) $max_x )
for k in %subst( %atom_info( $atom substructure))
for sub_atoms in %atoms({{ $k }})
modify atom coordinates $sub_atoms %math( %atom_info( $sub_atoms x) - 1 ) %atom_info( $sub_atoms y )
%atom_info( $sub_atoms z )
endfor
endfor
endif
endif
endif
if %lt( %atom_info( $atom x) $min_x )
for repeat in %range(1 50)
if %lt( %atom_info( $atom x) $min_x )
for k in %subst( %atom_info( $atom substructure))
for sub_atoms in %atoms({{ $k }})
modify atom coordinates $sub_atoms %math( %atom_info( $sub_atoms x) + 1 ) %atom_info( $sub_atoms y )
%atom_info( $sub_atoms z )
endfor
endfor
endif
endif
endif
if %gt( %atom_info( $atom y) $max_y )
for repeat in %range(1 50)
if %gt( %atom_info( $atom y) $max_y )
for k in %subst( %atom_info( $atom substructure))
for sub_atoms in %atoms({{ $k }})
modify atom coordinates $sub_atoms %atom_info( $sub_atoms x) %math(%atom_info( $sub_atoms y) - 1 )
%atom_info( $sub_atoms z )
endfor
endfor
endif
endif
endif
if %lt( %atom_info( $atom y) $min_y )
for repeat in %range(1 50)
if %lt( %atom_info( $atom y) $min_y )
for k in %subst( %atom_info( $atom substructure))
for sub_atoms in %atoms({{ $k }})
modify atom coordinates $sub_atoms %atom_info( $sub_atoms x) %math( %atom_info( $sub_atoms y) + 1 )
%atom_info( $sub_atoms z )
endfor
endfor
endif
endif
endif
```

```

endfor
endif
if %gt( %atom_info( $atom z) $max_z )
for repeat in %range(1 50)
if %gt( %atom_info( $atom z) $max_z )
for k in %subst( %atom_info( $atom substructure))
for sub_atoms in %atoms({$k})
modify atom coordinates $sub_atoms %atom_info( $sub_atoms x) %atom_info( $sub_atoms y) %math(%atom_info(
$sub_atoms z) - 1 )
endfor
endfor
endif
endif
endif
if %lt( %atom_info( $atom z) $min_z )
for repeat in %range(1 50)
if %lt( %atom_info( $atom z) $min_z )
for k in %subst( %atom_info( $atom substructure))
for sub_atoms in %atoms({$k})
modify atom coordinates $sub_atoms %atom_info( $sub_atoms x) %atom_info( $sub_atoms y) %math(%atom_info(
$sub_atoms z) + 1 )
endfor
endfor
endif
endif
endif
endif
endfor

```

### 2.3. Intermolecular potential

```
#intermolecular potential
for loop in 1

setvar select_k %prompt(anything 1 "moving molecule number(s)" "use % range(1 [no. of molecules]) to move many")
#setvar x_coord %prompt(anything 0 "Change in x coordinate" "e.g. atom at (5,0,0) is moved to (0,0,0) with -5")
#setvar y_coord %prompt(anything 0 "Change in y coordinate" "e.g. atom at (0,5,0) is moved to (0,0,0) with -5")
#setvar z_coord %prompt(anything 0 "Change in z coordinate" "e.g. atom at (0,0,5) is moved to (0,0,0) with -5")

setvar change %prompt(anything -0.001 "movement, angstroms, of mol")
setvar measurements %prompt(int 100000 "no. of measurements made")
setvar subst1 %prompt(anything 1 "atom ids from group one")
setvar subst2 %prompt(anything 1 "atom ids from group two")
setvar U_in %prompt(YES_NO YES "input value for internal energy at infinite distance")
setvar r_min
setvar U_min

if %streql( $U_in NO )
energy m1 done
setvar U_infinity $energy_total
setvar E1 $energy_total
endif

if %streql( $U_in YES )
setvar U_in2 %prompt(anything 0 "enter value")
setvar U_infinity $U_in2
setvar E1 $U_in2
energy m1 done
endif

setvar x_coord $change
setvar y_coord 0
setvar z_coord 0

for range in %range(1 $measurements )

for mult in $select_k

setvar kt %substs( %cat( # $mult ) )

modify substructure name $kt move
setvar ks move

for k in %substs( $ks )
for x in %atoms(({$k}))
modify atom coordinates $x %math( %atom_info( $x x ) + $x_coord ) %math( %atom_info( $x y ) + $y_coord ) %math(
%atom_info( $x z ) + $z_coord )
endfor
endfor
modify substructure name $kt %cat( n %irand() )
endfor

#setvar atoms_1 %prompt(anything 1 "atoms of molecule 1")
#setvar atoms_2 %prompt(anything 1 "atoms of molecule 2")
setvar atoms_1 $subst1
```

```

setvar atoms_2 $subst2
for x in %range(1 16)
for x1 in $atoms_1
for x2 in $atoms_2
setvar %cat( x_ $x ) %distance( $x1 $x2 )
endfor
endfor
endfor

energy m1 done
setvar E2 $energy_total
echo E1: $E1 kcal/mol
echo E2: $E2 kcal/mol
echo r %stats( $x_1 $x_2 $x_3 $x_4 $x_5 $x_6 $x_7 $x_8 $x_9 $x_10 $x_11 $x_12 $x_13 $x_14 $x_15 $x_16
mean) A
setvar r %stats( $x_1 $x_2 $x_3 $x_4 $x_5 $x_6 $x_7 $x_8 $x_9 $x_10 $x_11 $x_12 $x_13 $x_14 $x_15 $x_16
mean)

if %lteq( $E2 $E1 )
setvar E1 $energy_total
endif
if %gt( $E2 $E1 )

setvar U_min $E1

if %not(%or( $r_min ))
setvar r_min %math( $r - $change )
endif

echo U_min is $U_min
echo r_min is $r_min
endif

if %or( $U_min )
if %gt( $E2 $U_infinity )
GOTO EarlyExit
endif
endif

endfor

EarlyExit:
setvar r_sigma %stats( $x_1 $x_2 $x_3 $x_4 $x_5 $x_6 $x_7 $x_8 $x_9 $x_10 $x_11 $x_12 $x_13 $x_14 $x_15 $x_16
mean)
setvar U_sigma $energy_total
setvar U_sigma_prev
echo At $r_min A is epsilon, the minimum energy: $U_min kcal/mol ###
echo At $r_sigma is sigma ( $U_sigma kcal/mol), where U ~ U infinity ( $U_infinity kcal/mol)
echo Before sigma is $E1 kcal/mol at r %math( $r_sigma - $change ), also ~ U infinity ( $U_infinity kcal/mol)
endfor

```

## 2.4. Analysis of $\Delta G_{solv}$

```
setvar name %prompt(filename x "name")
setvar reps %prompt(int 5 "repeats")
setvar solvent %prompt(YES_NO YES "solvent is present? 'NO' for vacuum")
setvar MD %prompt(int 10000 "MD equilibration in fs")
if %streql( $solvent YES)
setvar xaxis %prompt(anything 35 "x axis side length")
setvar yaxis %prompt(anything 35 "y axis side length")
setvar zaxis %prompt(anything 35 "z axis side length")
endif

for repeats in %range(1 $reps )

  for lambda in %range(1 0 0.05)
    for x in %atoms({#1})
      modify atom charge $x %math( %atom_info( $x charge) * $lambda )
    endfor
    parameter modify atom_def C.3 c 4 th c.3 6 1.52 white 0 %math(2.55 * $lambda )
no no 0 19.74 12.01078 1 c3 | |
    parameter modify atom_def n.am n 3 tg n.am 7 1.45 blue 0 %math(3.04 * $lambda )
yes no 0 16.44 14.00672 28 nam | |
    parameter modify atom_def c.2 c 3 tg c.2 6 1.53 white 0 %math(2.55 * $lambda )
no no 0 20.13 12.01078 2 c2 | |
    parameter modify atom_def o.2 o 3 tg o.2 8 1.36 red 0 %math(3.44 * $lambda ) no
yes 2 13.26 15.99943 9 o2 | |
    parameter modify atom_def n.pl3 n 3 tg n.pl3 7 1.5 blue 0 %math(3.04 * $lambda
) yes no 0 18.11 14.00672 19 np13 | |
    parameter modify atom_def n.2 n 3 tg n.2 7 1.48 blue 0 %math(3.04 * $lambda )
yes yes 1 17.43 14.00672 6 n2 | |
    parameter modify atom_def n.3 n 4 th n.3 7 1.45 blue 0 %math(3.04 * $lambda )
yes yes 1 16.44 14.00672 5 n3 | |
    parameter modify atom_def c.cat c 3 tg c.2 6 1.53 white %math(1 * $lambda )
%math(2.55 * $lambda ) no no 0 20.13 12.01078 33 c2 | |
    parameter modify atom_def o.co2 o 3 tg o.2 8 1.36 red %math(-0.5 * $lambda )
%math(3.44 * $lambda ) no yes 2 13.26 15.99943 32 o2 | |
    parameter modify atom_def n.4 n 4 th n.3 7 1.45 blue %math(1 * $lambda )
%math(3.04 * $lambda ) yes no 0 16.44 14.00672 31 n3+ | |
    parameter modify atom_def o.3 o 4 th o.3 8 1.36 red 0 %math(3.44 * $lambda )
yes yes 2 13.26 15.99943 8 o3 | |
    parameter modify atom_def h h 1 11 h 1 1.08 cyan 0 %math(2.2 * $lambda ) no no
0 81.97 1.007947 13 h |1 |
    parameter modify atom_def f f 4 th f 9 1.3 green 0 %math(3.98 / $lambda ) yes
yes 3 11.41 18.998403 16 f | |
    parameter modify atom_def br br 4 th br 35 1.8 green 0 %math(2.96 / $lambda )
no no 3 20.13 79.904100 14 br | |

    dynamics m1 setup %cat( $name _ $lambda ) done interval_length $MD done
finished interactive

mol2 out m1 %cat( $name _temp)

    PHOTO ON %cat( $name _ lambda_ $lambda _ $repeats _ system )
energy m1 done
    PHOTO OFF %cat( $name _ lambda_ $lambda _ $repeats _ system )

if %streql( $solvent YES)

remove substructure M1((*)-({#1}))
```

```

    PHOTO ON %cat( $name _ lambda_ $lambda _ $repeats _ analyte )
    energy m1 mark_charges_valid done
    PHOTO OFF %cat( $name _ lambda_ $lambda _ $repeats _ analyte )
zap m1

mol2 in m1 %cat( $name _temp)

remove substructure M1(((#1)))

    PHOTO ON %cat( $name _ lambda_ $lambda _ $repeats _ solvent )
    energy m1 mark_charges_valid done
    PHOTO OFF %cat( $name _ lambda_ $lambda _ $repeats _ solvent )

zap m1

mol2 in m1 %cat( $name _temp)

#energy m1 electrostatics calculate_electrostatics get_new_charges MMFF94
boundary_conditions apply_pbc $xaxis $yaxis $zaxis done

#charge m1 undisplay
endif

    for x in %atoms((#1))
        modify atom charge $x %math( %atom_info( $x charge) / $lambda )
    endfor

    parameter modify atom_def C.3 c 4 th c.3 6 1.52 white 0 %math(2.55 / $lambda )
no no 0 19.74 12.01078 1 c3 | |
    parameter modify atom_def n.am n 3 tg n.am 7 1.45 blue 0 %math(3.04 / $lambda )
yes no 0 16.44 14.00672 28 nam | |
    parameter modify atom_def c.2 c 3 tg c.2 6 1.53 white 0 %math(2.55 / $lambda )
no no 0 20.13 12.01078 2 c2 | |
    parameter modify atom_def o.2 o 3 tg o.2 8 1.36 red 0 %math(3.44 / $lambda ) no
yes 2 13.26 15.99943 9 o2 | |
    parameter modify atom_def n.pl3 n 3 tg n.pl3 7 1.5 blue 0 %math(3.04 / $lambda
) yes no 0 18.11 14.00672 19 npl3 | |
    parameter modify atom_def n.2 n 3 tg n.2 7 1.48 blue 0 %math(3.04 / $lambda )
yes yes 1 17.43 14.00672 6 n2 | |
    parameter modify atom_def n.3 n 4 th n.3 7 1.45 blue 0 %math(3.04 / $lambda )
yes yes 1 16.44 14.00672 5 n3 | |
    parameter modify atom_def c.cat c 3 tg c.2 6 1.53 white %math(1 / $lambda )
%math(2.55 / $lambda ) no no 0 20.13 12.01078 33 c2 | |
    parameter modify atom_def o.co2 o 3 tg o.2 8 1.36 red %math(-0.5 / $lambda )
%math(3.44 / $lambda ) no yes 2 13.26 15.99943 32 o2 | |
    parameter modify atom_def n.4 n 4 th n.3 7 1.45 blue %math(1 / $lambda )
%math(3.04 / $lambda ) yes no 0 16.44 14.00672 31 n3+ | |
    parameter modify atom_def o.3 o 4 th o.3 8 1.36 red 0 %math(3.44 / $lambda )
yes yes 2 13.26 15.99943 8 o3 | |
    parameter modify atom_def h h 1 11 h 1 1.08 cyan 0 %math(2.2 / $lambda ) no no
0 81.97 1.007947 13 h | |
    parameter modify atom_def f f 4 th f 9 1.3 green 0 %math(3.98 / $lambda ) yes
yes 3 11.41 18.998403 16 f | |
    parameter modify atom_def br br 4 th br 35 1.8 green 0 %math(2.96 / $lambda )
no no 3 20.13 79.904100 14 br | |

    if %lt( $lambda 0.05)
        parameter modify atom_def C.3 c 4 th c.3 6 1.52 white 0 2.55 no no 0 19.74
12.01078 1 c3 | |
        parameter modify atom_def n.am n 3 tg n.am 7 1.45 blue 0 3.04 yes no 0 16.44
14.00672 28 nam | |

```

```

parameter modify atom_def c.2 c 3 tg c.2 6 1.53 white 0 2.55 no no 0 20.13
12.01078 2 c2 | |
parameter modify atom_def o.2 o 3 tg o.2 8 1.36 red 0 3.44 no yes 2 13.26
15.99943 9 o2 | |
parameter modify atom_def n.pl3 n 3 tg n.pl3 7 1.5 blue 0 3.04 yes no 0 18.11
14.00672 19 npl3 | |
parameter modify atom_def n.2 n 3 tg n.2 7 1.48 blue 0 3.04 yes yes 1 17.43
14.00672 6 n2 | |
parameter modify atom_def n.3 n 4 th n.3 7 1.45 blue 0 3.04 yes yes 1 16.44
14.00672 5 n3 | |
parameter modify atom_def c.cat c 3 tg c.2 6 1.53 white 1 2.55 no no 0 20.13
12.01078 33 c2 | |
parameter modify atom_def o.co2 o 3 tg o.2 8 1.36 red -0.5 3.44 no yes 2 13.26
15.99943 32 o2 | |
parameter modify atom_def n.4 n 4 th n.3 7 1.45 blue 1 3.04 yes no 0 16.44
14.00672 31 n3+ | |
parameter modify atom_def o.3 o 4 th o.3 8 1.36 red 0 3.44 yes yes 2 13.26
15.99943 8 o3 | |
parameter modify atom_def h h 1 l1 h 1 1.08 cyan 0 2.2 no no 0 81.97 1.007947
13 h | |
parameter modify atom_def f f 4 th f 9 1.3 green 0 3.98 yes yes 3 11.41
18.998403 16 f | |
parameter modify atom_def br br 4 th br 35 1.8 green 0 2.96 no no 3 20.13
79.904100 14 br | |
charge m1 compute MMFF94
charge m1 undisplay
endif
endfor
endif
endfor

```

## Appendix 3. Experimental section for Chapter 4

### 3.1. Materials and methods

The reaction conditions used here are adapted from that which has been widely used in the preparation of imprinted polymer nanoparticles.<sup>44</sup> *N*-isopropylacrylamide (NiPAm), *N,N'*-methylenebis(acrylamide) (Bis), ammonium persulfate (APS), *N,N,N',N'*-tetramethylethylenediamine (TEMED), glutaraldehyde, acrylamide and (3-aminopropyl)trimethoxysilane (APTMS) were obtained from Sigma-Aldrich, *N-tert*-butylacrylamide (TBAm) and acrylic acid (AAc) from Acros, *N,N*-diethyldithiocarbamic acid benzyl ester (iniferter) from TCI Europe and streptomycin sulfate from Amresco. Glass beads (Spherglass A 2429) was obtained from Potters Industries. Phosphate buffer saline (PBS, Gibco) 1X pH 7.4 was prepared according to the manufacturers instruction. Particle diameters were analyzed by DLS using a Zetasizer Nano S (Malvern). The collected DLS data are given in the Appendices.

### 3.2. Primary experiments

NiPAm (19.5 mg, 0.193 mmol), TBAm (16.5 mg, 0.143 mmol), Bis (1 mg, 6.486  $\mu$ mol) and AAc (1.1  $\mu$ l, 16.02  $\mu$ mol) were dissolved with or without the nucleant streptomycin sulfate (2.4 mg, 3.245  $\mu$ mol). Solvents were varied by percentage volume in 5 % increments, giving solvents 100 %-50 % water with the remainder the relevant solvent. Nitrogen was passed through each solution for 25 minutes before initiation with an aqueous solution of APS (15 mg, 65.73  $\mu$ mol) and TEMED (15  $\mu$ l, 0.100 mmol). Each mixture was then topped to 50 ml total volume to give the appropriate solvent ratio and briefly degassed again with nitrogen before being left to incubate for approximately 22 hours.<sup>311</sup> The reactions were quenched by passing oxygen gas through each flask for 25 minutes. Variations on this composition involved scaling all components by a common value.

Particle diameters were analyzed by DLS, the appropriate density, viscosity and refractive index for each solvent mixture was calibrated using previously published empirical data.<sup>325-328</sup> For DLS analysis all data for each conditions was combined and ordered by polydispersity index (PdI); measurements with count rate values (kcps) below 1000 being rejected. The majority of measurements used in the final analysis displayed a polydispersity index below 0.06 (difficulties obtaining low-PdI, high-count measurements with water and water-methanol mixtures required a greater tolerance).

### 3.3. Preliminary experiments

A range of conditions were initially observed with consistent monomer mixture in order to evaluate early hypotheses of particle diameter dependence. Diethyl ether, ethyl acetate, dichloromethane, ethanol, methanol, acetonitrile and water were used alone in each case as solvent.

Glass beads (60 g) were boiled in aqueous sodium hydroxide solution ( $1 \text{ mol dm}^{-3}$ , 48 ml) for 15 minutes, then washed with deionized water ( $8 \times 200 \text{ ml}$ ). PBS (300 ml) was added to the beads before being again washed with deionized water ( $3 \times 200 \text{ ml}$ ) and acetone ( $2 \times 200 \text{ ml}$ ) before being left at  $80 \text{ }^\circ\text{C}$  for 3 hours. The beads were then incubate in a 2 vol% APTMS toluene solution (24 ml) for 24 hours then washed with acetone ( $8 \times 200 \text{ ml}$ ) and methanol ( $1 \times 200 \text{ ml}$ ). The prepared solid phase was then incubated with glutaraldehyde (7 vol%, PBS  $0.01 \text{ mol dm}^{-3}$  pH 7.2 solution) for 2 hours and washed with deionized water ( $8 \times 200 \text{ ml}$ ) under vacuum before being submerged in a  $3.5 \text{ mmol dm}^{-3}$  vancomycin PBS ( $0.01 \text{ mol dm}^{-3}$ ) solution. Residual vancomycin and PBS was washed with deionized water ( $8 \times 200 \text{ ml}$ ).

Acrylamide (68 mg, 0.9567 mmol) and Bis (3 mg, 19.46  $\mu\text{mol}$ ) were dissolved in diethyl ether, ethyl acetate, methanol, acetonitrile or water, and combined with the prepared beads. Separate acrylamide-Bis solutions were prepared for polymerization without the beads, with and without the solvated nucleant vancomycin. The mixtures were purged with nitrogen before initiation by one of two methods: i) TEMED (30  $\mu\text{l}$ , 0.1999 mmol) was added a solution of APS (30 mg, 0.1315 mmol) in 0.5 ml of water, which was then added to the monomer mixture, or ii) iniferter (0.75 g, 3.133 mmol) was added to the mixture, followed by photopolymerization by exposure to UV light for 12 minutes.

Measurements were recorded by DLS as in the primary experiments but lower standards were set in the data analysis (PdI below 0.4 and count rate above 600) due to the limited solubility of the components. The measurements recorded for each condition were combined according to solvent and are shown in Figure 4.1.

Full DLS data can be provided on request.

## References

1. G. L. Flynn, S. H. Yalkowsky and T. J. Roseman, *J. Pharm. Sci.*, 1974, **63**, 479-510.
2. H. S. Andersson and I. A. Nicholls, *Molecularly imprinted polymers, volume 23: man-made mimics of antibodies and their application in analytical chemistry*, University of Mainz, Mainz, Germany, 2000.
3. M. V. Polyakov, *Zhurnal fizicheskoi khimii*, 1931, **2**, 799-805.
4. R. G. Haldeman and P. H. Emmett, *The Journal of Physical Chemistry*, 1955, **59**, 1039-1043.
5. F. H. Dickey, *Proceedings of the National Academy of Sciences*, 1949, **35**, 227-229.
6. L. Pauling, *J Am. Chem. Soc.*, 1940, **62**, 2643-2657.
7. G. Wulff and A. Sarhan, *Angewandte Chemie-International Edition*, 1972, **11**, 341-&.
8. Takagishi, T and I. M. Klotz, *Biopolymers*, 1972, **11**, 483-&.
9. F. G. Tamayo, E. Turiel and A. Martín-Esteban, *Journal of Chromatography A*, 2007, **1152**, 32-40.
10. L. I. Andersson and K. Mosbach, *Journal of Chromatography*, 1990, **516**, 313-322.
11. D. J. O'Shannessy, L. I. Andersson and K. Mosbach, *Journal of molecular recognition : JMR*, 1989, **2**, 1-5.
12. B. Sellergren, B. Ekberg and K. Mosbach, *Journal of Chromatography*, 1985, **347**, 1-10.
13. O. Norrlof, M. O. Mansson and K. Mosbach, *Journal of Chromatography*, 1987, **396**, 374-377.
14. B. Sellergren, M. Lepisto and K. Mosbach, *J Am. Chem. Soc.*, 1988, **110**, 5853-5860.
15. J. Matsui, Y. Miyoshi, O. Doblhoffdier and T. Takeuchi, *Anal. Chem.*, 1995, **67**, 4404-4408.
16. S. A. Piletsky, S. Alcock and A. P. F. Turner, *Trends Biotechnol.*, 2001, **19**, 9-12.
17. N. Pérez, M. J. Whitcombe and E. N. Vulfson, *J. Appl. Polym. Sci.*, 2000, **77**, 1851-1859.
18. L. Ye, P. A. G. Cormack and K. Mosbach, *Analytical Communications*, 1999, **36**, 35-38.
19. M. T. Muldoon and L. H. Stanker, *Chem. Ind.*, 1996, 204-207.
20. D. K. Robinson and K. Mosbach, *Journal of the Chemical Society, Chemical Communications*, 1989, 969-970.
21. G. Vlatakis, L. I. Andersson, R. Müller and K. Mosbach, *Nature*, 1993, **361**, 645.
22. L. I. Andersson, R. Muller, G. Vlatakis and K. Mosbach, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 4788-4792.
23. I. Chianella, A. Guerreiro, E. Moczko, J. S. Caygill, E. V. Piletska, I. Sansalvador, M. J. Whitcombe and S. A. Piletsky, *Anal. Chem.*, 2013, **85**, 8462-8468.
24. Y. Yonamine, Y. Hoshino and K. J. Shea, *Biomacromolecules*, 2012, **13**, 2952-2957.
25. T. S. Bedwell and M. J. Whitcombe, *Anal. Bioanal. Chem.*, 2016, **408**, 1735-1751.
26. R. V. Shutov, A. Guerreiro, E. Moczko, I. P. de Vargas-Sansalvador, I. Chianella, M. J. Whitcombe and S. A. Piletsky, *Small*, 2014, **10**, 1086-1089.

27. K. Haupt, *Nat. Mater.*, 2010, **9**, 612-614.
28. Y. Hoshino and K. J. Shea, *J. Mater. Chem.*, 2011, **21**, 3517-3521.
29. M. J. Whitcombe, I. Chianella, L. Larcombe, S. A. Piletsky, J. Noble, R. Porter and A. Horgan, *Chem. Soc. Rev.*, 2011, **40**, 1547-1571.
30. R. Schirhagl, P. A. Lieberzeit and F. L. Dickert, *Adv. Mater.*, 2010, **22**, 2078-+.
31. L. D. Bolisay, J. N. Culver and P. Kofinas, *Biomaterials*, 2006, **27**, 4165-4168.
32. A. Ramanaviciene and A. Ramanavicius, *Biosens. Bioelectron.*, 2004, **20**, 1076-1082.
33. D. F. Tai, C. Y. Lin, T. Z. Wu and L. K. Chen, *Anal. Chem.*, 2005, **77**, 5140-5143.
34. C. H. Lu, Y. Zhang, S. F. Tang, Z. B. Fang, H. H. Yang, X. Chen and G. N. Chen, *Biosens. Bioelectron.*, 2012, **31**, 439-444.
35. T. Kusunoki and T. Kobayashi, *J. Appl. Polym. Sci.*, 2010, **117**, 565-571.
36. Y. Hoshino, H. Koide, T. Urakami, H. Kanazawa, T. Kodama, N. Oku and K. J. Shea, *J. Am. Chem. Soc.*, 2010, **132**, 6644-+.
37. A. Kasturiratne, A. R. Wickremasinghe, N. de Silva, N. K. Gunawardena, A. Pathmeswaran, R. Premaratna, L. Savioli, D. G. Laloo and H. J. de Silva, *PLOS Medicine*, 2008, **5**, e218.
38. F. Yanez, L. Martikainen, M. E. M. Braga, C. Alvarez-Lorenzo, A. Concheiro, C. M. M. Duarte, M. H. Gil and H. C. de Sousa, *Acta Biomater.*, 2011, **7**, 1019-1030.
39. H. Hashemi-Moghaddam, S. Kazemi-Bagsangani, M. Jamili and S. Zavareh, *Int. J. Pharm.*, 2016, **497**, 228-238.
40. F. Puoci, G. Cirillo, M. Curcio, O. I. Parisi, F. Iemma and N. Picci, *Expert Opin. Drug Deliv.*, 2011, **8**, 1379-1393.
41. S. S. Wang, D. Y. Yin, W. J. Wang, X. J. Shen, J. J. Zhu, H. Y. Chen and Z. Liu, *Sci Rep*, 2016, **6**, 11.
42. B. Sellergren and C. J. Allender, *Adv. Drug Deliv. Rev.*, 2005, **57**, 1733-1741.
43. S. A. Piletsky, E. V. Piletskaya, A. V. Elgersma, K. Yano, I. Karube, Y. P. Parhometz and A. V. El'skaya, *Biosens. Bioelectron.*, 1995, **10**, 959-964.
44. F. Canfarotta, A. Poma, A. Guerreiro and S. Piletsky, *Nat. Protoc.*, 2016, **11**, 443-455.
45. J. Hou, H. Y. Li, L. Wang, P. Zhang, T. Y. Zhou, H. Ding and L. Ding, *Talanta*, 2016, **146**, 34-40.
46. G. Wulff, *Microchimica Acta*, 2013, **180**, 1359-1370.
47. T. A. Sergeeva, S. A. Piletsky, A. A. Brovko, E. A. Slinchenko, L. M. Sergeeva and A. V. El'skaya, *Anal. Chim. Acta*, 1999, **392**, 105-111.
48. B. Castro, G. Ramirez, M. F. Rubio, M. J. Whitcombe, E. N. Vulfson, R. Vazquez-Duhalt and E. Barzana, *Applying molecular modeling and computational chemistry*, 1998.
49. B. Castro, M. J. Whitcombe, E. N. Vulfson, R. Vazquez-Duhalt and E. Bárzana, *Anal. Chim. Acta*, 2001, **435**, 83-90.
50. B. Chen, R.M. Day, S. Subrahmanyam, S.A. Piletsky, O.V. Piletska, A.P.F. Turner, US pat., 808 641 5B2, 2000.
51. T. Cowen, K. Karim and S. Piletsky, *Anal. Chim. Acta*, 2016, **936**, 62-74.
52. S. A. Piletsky, K. Karim, E. V. Piletska, C. J. Day, K. W. Freebairn, C. Legge and A. P. F. Turner, *Analyst*, 2001, **126**, 1826-1830.
53. D. Batra and K. J. Shea, *Curr. Opin. Chem. Biol.*, 2003, **7**, 434-442.

54. S. Subrahmanyam, K. Karim and S. A. Piletsky, in *Designing Receptors for the Next Generation of Biosensors*, eds. A. S. Piletsky and J. M. Whitcombe, Springer Berlin Heidelberg, Berlin, Heidelberg, Editon edn., 2013, pp. 131-165.
55. E. V. Piletska, B. H. Abd, A. S. Krakowiak, A. Parmar, D. L. Pink, K. S. Wall, L. Wharton, E. Moczko, M. J. Whitcombe, K. Karim and S. A. Piletsky, *Analyst*, 2015, **140**, 3113-3120.
56. I. Bakas, N. Ben Oujji, G. Istamboulie, S. Piletsky, E. Piletska, E. Ait-Addi, I. Ait-Ichou, T. Noguier and R. Rouillon, *Talanta*, 2014, **125**, 313-318.
57. I. Bakas, A. Hayat, S. Piletsky, E. Piletska, M. M. Chehimi, T. Noguier and R. Rouillon, *Talanta*, 2014, **130**, 294-298.
58. K. Karim, L. Giannoudi, E. Piletska, I. Chianella, O. Y. F. Henry, P. Laitenberger, S. A. Piletsky and T. Cowen, *J. Chin. Adv. Mater. Soc.*, 2015, **3**, 149-160.
59. K. Zhang, W. Zou, H. Zhao, P. Dramou, C. Pham-Huy, J. He and H. He, *RSC Adv.*, 2015, **5**, 61161-61169.
60. M. J. Abdin, Z. Altintas and I. E. Tothill, *Biosens. Bioelectron.*, 2015, **67**, 177-183.
61. S. P. Wren, S. A. Piletsky, K. Karim, P. Gascoine, R. Lacey, T. Sun and K. T. V. Grattan, *J. Lightwave Technol.*, 2015, **33**, 2572-2579.
62. C. L. Wang, X. L. Hu, P. Guan, L. W. Qian, D. F. Wu and J. Li, *Sep. Sci. Technol.*, 2015, **50**, 1768-1775.
63. C. L. Wang, X. L. Hu, P. Guan, L. W. Qian, D. F. Wu and J. Li, *Int. J. Polym. Anal. Charact.*, 2014, **19**, 70-82.
64. C. L. Wang, X. L. Hu, P. Guan, D. F. Wu, L. W. Qian, J. Li and R. Y. Song, *J. Pharm. Biomed. Anal.*, 2015, **102**, 137-143.
65. S. Hou, Y. F. Wang, N. Liu and J. Liu, *Adsorpt. Sci. Technol.*, 2014, **32**, 833-843.
66. A. Wong, M. V. Foguel, S. Khan, F. M. de Oliveira, C. R. T. Tarley and M. Sotomayor, *Electrochim. Acta*, 2015, **182**, 122-130.
67. Z. Altintas, B. France, J. O. Ortiz and I. E. Tothill, *Sens. Actuator B-Chem.*, 2016, **224**, 726-737.
68. J. Mistry, A. Guerreiro, E. Moczko, E. Piletska, K. Karim and S. Piletsky, *Molecular Imprinting*, 2015, **2**, 83-92.
69. M. Bitar, E. Bou-Maroun, A. Lerbret, N. Ouaini and P. Cayot, *J. Sep. Sci.*, 2015, **38**, 3607-3614.
70. Y. W. Wang, T. Zhao, P. Dai, N. Jiang and F. Li, *Chemphyschem*, 2016, **17**, 893-901.
71. M. Perez, R. Concu, M. Ornelas, M. Cordeiro, M. Azenha and A. F. Silva, *Spectroc. Acta Pt. A-Molec. Biomolec. Spectr.*, 2016, **153**, 661-668.
72. R. J. Uzuriaga-Sanchez, S. Khan, A. Wong, G. Picasso, M. I. Pividori and M. D. T. Sotomayor, *Food Chem.*, 2016, **190**, 460-467.
73. M. J. U. Toro, L. D. Marestoni and M. D. T. Sotomayor, *Sens. Actuator B-Chem.*, 2015, **208**, 299-306.
74. M. Sobiech, T. Zolek, P. Lulinski and D. Maciejewska, *Analyst*, 2014, **139**, 1779-1788.
75. F. Ahmadi and E. Karamian, *Iran. J. Pharm. Res.*, 2014, **13**, 417-430.
76. L. D. Marestoni, A. Wong, G. T. Feliciano, M. R. R. Marchi, C. R. T. Tarley and M. Sotomayor, *J. Braz. Chem. Soc.*, 2016, **27**, 109-118.
77. H. Li, H. L. He, J. J. Huang, C. Z. Wang, X. L. Gu, Y. K. Gao, H. J. Zhang, S. H. Du, L. N. Chen and C. S. Yuan, *Biomed. Chromatogr.*, 2016, **30**, 117-125.

78. A. G. Ayankojo, A. Tretjakoy, J. Reut, R. Boroznjak, A. Opik, J. Rappich, A. Furchner, K. Hinrichs and V. Syritski, *Anal. Chem.*, 2016, **88**, 1476-1484.
79. P. Lulinski, M. Dana and D. Maciejewska, *Talanta*, 2014, **119**, 623-631.
80. H. Zayas, C. I. Holdsworth, M. C. Bowyer and A. McCluskey, *Org. Biomol. Chem.*, 2014, **12**, 6994-7003.
81. H. M. Bi, J. P. Hu, Y. Liu, X. Q. Chai, L. X. Tong and C. H. Dong, *Asian J. Chem.*, 2014, **26**, 161-163.
82. H. L. He, X. L. Gu, L. Y. Shi, J. L. Hong, H. J. Zhang, Y. K. Gao, S. H. Du and L. N. Chen, *Anal. Bioanal. Chem.*, 2015, **407**, 509-519.
83. A. Nezhadali, Z. Rouki and M. Nezhadali, *Talanta*, 2015, **144**, 456-465.
84. W. Yang, L. Liu, X. Ni, W. Zhou, W. Huang, H. Liu and W. Xu, *J. Sep. Sci.*, 2016, **39**, 503-517.
85. X. P. Luo, C. Z. Li, Y. Q. Duan, H. H. Zhang, D. Zhang, C. Zhang, G. B. Sun and X. B. Sun, *J. Appl. Polym. Sci.*, 2016, **133**, 11.
86. K. A. M. Attia, N. M. El-Abasawi and A. H. Abdel-Azim, *Mater. Sci. Eng. C-Mater. Biol. Appl.*, 2016, **61**, 773-781.
87. M. Appell, M. A. Jackson, L. J. C. Wang, C. H. Ho and A. Mueller, *J. Sep. Sci.*, 2014, **37**, 281-286.
88. B. B. Prasad and A. Kumar, *J. Mat. Chem. B*, 2015, **3**, 5864-5876.
89. E. Roy, S. Patra, R. Madhuri and P. K. Sharma, *Talanta*, 2014, **120**, 198-207.
90. H. L. He, H. Li, Y. K. Gao, D. Chen, L. Y. Shi, J. Peng, S. H. Du and L. N. Chen, *Anal. Lett.*, 2015, **48**, 47-60.
91. A. Nezhadali and R. Shadmehri, *Sens. Actuator B-Chem.*, 2014, **202**, 240-251.
92. F. Ahmadi, E. Yawari and M. Nikbakht, *Journal of Chromatography A*, 2014, **1338**, 9-16.
93. A. Nezhadali and M. Mojarrab, *Sens. Actuator B-Chem.*, 2014, **190**, 829-837.
94. X. B. Yu, J. B. Liu and X. Yuan, *Chin. J. Struct. Chem.*, 2015, **34**, 488-496.
95. K. K. Tadi, R. V. Motghare and V. Ganesh, *Electroanalysis*, 2014, **26**, 2328-2336.
96. P. P. Qi, X. Y. Wang, X. Q. Wang, H. Zhang, H. Xu, K. Z. Jiang and Q. Wang, *J. Sep. Sci.*, 2014, **37**, 2955-2965.
97. L. A. de Barros, L. A. Pereira, R. Custodio and S. Rath, *J. Braz. Chem. Soc.*, 2014, **25**, 619-628.
98. M. S. Khan, S. Pal and R. J. Krupadam, *Journal of Molecular Recognition*, 2015, **28**, 427-437.
99. W. Zhang, N. Tan, X. H. Jia, G. P. Wang, W. Long, X. L. Li, S. Liao and D. Hou, *Mater. Sci. Eng. C-Mater. Biol. Appl.*, 2015, **53**, 166-174.
100. M. Torkashvand, M. B. Gholivand and F. Taherkhani, *Mater. Sci. Eng. C-Mater. Biol. Appl.*, 2015, **55**, 209-217.
101. Z. Q. Dai, J. B. Liu, S. S. Tang, Y. Wang, Y. M. Wang and R. F. Jin, *Journal of Molecular Modeling*, 2015, **21**, 9.
102. A. Nezhadali, S. Senobari and M. Mojarrab, *Talanta*, 2016, **146**, 525-532.
103. X. Qiu, X.-Y. Xu, Y. Liang, Y. Hua and H. Guo, *Journal of chromatography. A*, 2016, **1429**, 79-85.
104. N. Maouche, N. Ktari, I. Bakas, N. Fourati, C. Zerrouki, M. Seydou, F. Maurel and M. M. Chehimi, *Journal of Molecular Recognition*, 2015, **28**, 667-678.
105. L. S. Fernandes, P. Homem-de-Mello, E. C. de Lima and K. M. Honorio, *Eur. Polym. J.*, 2015, **71**, 364-371.

106. M. C. Fonseca, C. S. Nascimento and K. B. Borges, *Chemical Physics Letters*, 2016, **645**, 174-179.
107. B. B. Prasad, R. Singh and A. Kumar, *Carbon*, 2016, **102**, 86-96.
108. K. M. Muzyka, *J. Nano Electron. Phys.*, 2015, **7**, 5.
109. Y. Wang, J. B. Liu, S. S. Tang, R. F. Jin and H. B. Chang, *Chem. J. Chin. Univ.-Chin.*, 2015, **36**, 945-954.
110. Y. Wang, J. B. Liu, S. S. Tang, R. F. Jin and H. B. Chang, *Acta Polym. Sin.*, 2015, 641-649.
111. Y. J. Huang and Q. J. Zhu, *J. Chem.*, 2015, 9.
112. K. K. Tadi, R. V. Motghare and V. Ganesh, *RSC Adv.*, 2015, **5**, 99115-99124.
113. W. M. Yang, P. F. Ma, T. Fan, Z. P. Zhou, H. Liu and W. Z. Xu, *J. Appl. Polym. Sci.*, 2015, **132**, 10.
114. A. Nezhadali and M. Mojarrab, *Journal of Electroanalytical Chemistry*, 2015, **744**, 85-94.
115. B. Liu, L. L. Ou, F. Y. Zhang, Z. J. Zhang, H. Y. Li, M. Y. Zhu and S. Wang, *J. Sep. Sci.*, 2014, **37**, 3579-3586.
116. S. J. Pace, E. Nguyen, M. P. Baria and E. R. E. Mojica, *Microchimica Acta*, 2015, **182**, 69-76.
117. P. Lulinski, M. Sobiech, T. Zolek and D. Maciejewska, *Talanta*, 2014, **129**, 155-164.
118. M. Sobiech, T. Zolek, P. Lulinski and D. Maciejewska, *Talanta*, 2016, **146**, 556-567.
119. L. L. Vallo, R. T. Laxamana, F. U. Paredes, I. H. J. Arellano and S. D. Arco, *Mater. Lett.*, 2015, **159**, 317-320.
120. K. Golker and I. A. Nicholls, *Eur. Polym. J.*, 2016, **75**, 423-430.
121. K. Golker, B. C. G. Karlsson, J. G. Wiklander, A. M. Rosengren and I. A. Nicholls, *Eur. Polym. J.*, 2015, **66**, 558-568.
122. K. Golker, B. C. G. Karlsson, A. M. Rosengren and I. A. Nicholls, *Int. J. Mol. Sci.*, 2014, **15**, 20572-20584.
123. Y. Kong, N. W. Wang, X. N. Ni, Q. Y. Yu, H. Liu, W. H. Huang and W. Z. Xu, *J. Appl. Polym. Sci.*, 2016, **133**, 11.
124. C. X. Qiu, W. M. Yang, Z. P. Zhou, Y. S. Yan and W. Z. Xu, *J. Appl. Polym. Sci.*, 2015, **132**, 15.
125. Y. C. Wang, N. W. Wang, X. N. Ni, Q. Q. Jiang, W. M. Yang, W. H. Huang and W. Z. Xu, *RSC Adv.*, 2015, **5**, 73424-73433.
126. C. X. Qiu, Y. H. Xing, W. M. Yang, Z. P. Zhou, Y. C. Wang, H. Liu and W. Z. Xu, *Appl. Surf. Sci.*, 2015, **345**, 405-417.
127. J. Li, X. L. Hu, P. Guan, D. M. Song, L. W. Qian, C. B. Du, R. Y. Song and C. L. Wang, *J. Sep. Sci.*, 2015, **38**, 3279-3287.
128. V. Yilmaz, Z. Arslan, O. Hazer and H. Yilmaz, *Microchem J.*, 2014, **114**, 65-72.
129. M. Schauerl and D. W. Lewis, *J. Phys. Chem. B*, 2015, **119**, 563-571.
130. S. Shoravi, G. D. Olsson, B. C. G. Karlsson and I. A. Nicholls, *Int. J. Mol. Sci.*, 2014, **15**, 10622-10634.
131. D. Cleland, G. D. Olsson, B. C. G. Karlsson, I. A. Nicholls and A. McCluskey, *Org. Biomol. Chem.*, 2014, **12**, 844-853.
132. B. D. B. Tiu, R. B. Pernites, S. B. Tin and R. C. Advincula, *Colloid Surf. A-Physicochem. Eng. Asp.*, 2016, **495**, 149-158.

133. H. Kempe, A. P. Pujolras and M. Kempe, *Pharm. Res.*, 2015, **32**, 375-388.
134. L. Chen, Y. K. Lee, Y. Manmana, K. S. Tay, V. S. Lee and N. A. Rahman, *e-Polymers*, 2015, **15**, 141-150.
135. R. V. Motghare, K. K. Tadi, P. Dhawale, S. Deotare, A. K. Kawadkar, R. Chillawar and S. Khan, *Electroanalysis*, 2015, **27**, 825-832.
136. N. Karimian, M. B. Gholivand and F. Taherkhani, *Journal of Electroanalytical Chemistry*, 2015, **740**, 45-52.
137. N. A. El Gohary, A. Madbouly, R. M. El Nashar and B. Mizaikoff, *Biosens. Bioelectron.*, 2015, **65**, 108-114.
138. E. M. Saad, A. Madbouly, N. Ayoub and R. M. El Nashar, *Anal. Chim. Acta*, 2015, **877**, 80-89.
139. N. Martins, E. P. Carreiro, A. Locati, J. P. P. Ramalho, M. J. Cabrita, A. J. Burke and R. Garcia, *Journal of Chromatography A*, 2015, **1409**, 1-10.
140. J. B. Liu, Y. Wang, T. T. Su, B. Li, S. S. Tang and R. F. Jin, *J. Sep. Sci.*, 2015, **38**, 4105-4110.
141. L. Li, L. Chen, H. Zhang, Y. Yang, X. Liu and Y. Chen, *Materials science & engineering. C, Materials for biological applications*, 2016, **61**, 158-168.
142. A. Wojnarowicz, P. S. Sharma, M. Sosnowska, W. Lisowski, T. P. Huynh, M. Pszona, P. Borowicz, F. D'Souza and W. Kutner, *J. Mat. Chem. B*, 2016, **4**, 1156-1165.
143. T. P. Huynh, A. Wojnarowicz, M. Sosnowska, S. Srebnik, T. Benincori, F. Sanniccolo, F. D'Souza and W. Kutner, *Biosens. Bioelectron.*, 2015, **70**, 153-160.
144. N. Ktari, N. Fourati, C. Zerrouki, M. Ruan, M. Seydou, F. Barbault, F. Nal, N. Yaakoubi, M. M. Chehimi and R. Kalfat, *RSC Adv.*, 2015, **5**, 88666-88674.
145. A. S. Ogunlaja, M. J. Coombes, N. Torto and Z. R. Tshentu, *React. Funct. Polym.*, 2014, **81**, 61-76.
146. A. S. Ogunlaja, C. du Sautoy, N. Tort and Z. R. Tshentu, *Talanta*, 2014, **126**, 61-72.
147. Y. X. Zhang, X. Qu, J. P. Yu, L. C. Xu, Z. Q. Zhang, H. Hong and C. S. Liu, *J. Mat. Chem. B*, 2014, **2**, 1390-1399.
148. L. M. Feng, L. Tan, H. Li, Z. G. Xu, G. X. Shen and Y. W. Tang, *Biosens. Bioelectron.*, 2015, **69**, 265-271.
149. T. Alizadeh, A. Bagherzadeh and A. N. Shamkhali, *Materials science & engineering. C, Materials for biological applications*, 2016, **63**, 247-255.
150. T. Alizadeh and A. N. Shamkhali, *J. Chromatogr. B*, 2016, **1009**, 96-106.
151. L. Jun-Bo, S. Yang, T. Shan-Shan and J. Rui-Fa, *J. Sep. Sci.*, 2015, **38**, 1065-1071.
152. R. Bujak, R. Gadzala-Kopciuch, A. Nowaczyk, J. Raczak-Gutknecht, M. Kordalewska, W. Struck-Lewicka, M. J. Markuszewski and B. Buszewski, *Talanta*, 2016, **146**, 401-409.
153. N. L. A. N. D. O. V. S. V. F. F. A. V. Cao, *Sorption and Chromatography Processes*, 2015, **15**, 421-428.
154. T. P. Huynh, B. K. C. Chandra, M. Sosnowska, J. W. Sobczak, V. N. Nesterov, F. D'Souza and W. Kutner, *Biosens. Bioelectron.*, 2015, **64**, 657-663.
155. X. Li, Y. F. He, F. Zhao, W. Y. Zhang and Z. L. Ye, *RSC Adv.*, 2015, **5**, 56534-56540.
156. A. Pardo, L. Mespouille, B. Blankert, P. Trouillas, M. Surin, P. Dubois and P. Duez, *Journal of Chromatography A*, 2014, **1364**, 128-139.

157. B. D. B. Tiu, R. J. Krupadam and R. C. Advincula, *Sens. Actuator B-Chem.*, 2016, **228**, 693-701.
158. F. S. Mehamod, K. KuBulat, N. F. Yusof and N. A. Othman, *Int. J. Technol.*, 2015, **6**, 546-554.
159. S. Pardeshi, R. Dhodapkar and A. Kumar, *Compos. Interfaces*, 2014, **21**, 13-30.
160. B. Gao, X. P. He, Y. Jiang, J. T. Wei, H. Suo and C. Zhao, *J. Sep. Sci.*, 2014, **37**, 3753-3759.
161. P. S. Sharma, A. Wojnarowicz, M. Sosnowska, T. Benincori, K. Noworyta, F. D'Souza and W. Kutner, *Biosens. Bioelectron.*, 2016, **77**, 565-572.
162. D. H. Luo, Z. J. Zhao, L. Zhang, Q. Wang and J. Wang, *Mol. Simul.*, 2014, **40**, 431-438.
163. I. A. Nicholls, H. S. Andersson, C. Charlton, H. Henschel, B. C. G. Karlsson, J. G. Karlsson, J. O'Mahony, A. M. Rosengren, K. J. Rosengren and S. Wikman, *Biosens. Bioelectron.*, 2009, **25**, 543-552.
164. Y. B. Malykhanov and M. V. Gorshunov, *Journal of Applied Spectroscopy*, 2013, **80**, 631-636.
165. F. Calvo, E. Pahl, M. Wormit and P. Schwerdtfeger, *Angewandte Chemie-International Edition*, 2013, **52**, 7583-7585.
166. J. J. P. Stewart, *Journal of Computer-Aided Molecular Design*, 1990, **4**, 1-103.
167. M. J. S. Dewar and W. Thiel, *J Am. Chem. Soc.*, 1977, **99**, 4907-4917.
168. R. C. Bingham, M. J. S. Dewar and D. H. Lo, *J Am. Chem. Soc.*, 1975, **97**, 1285-1293.
169. J. J. P. Stewart, *Journal of Molecular Modeling*, 2007, **13**, 1173-1213.
170. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J Am. Chem. Soc.*, 1985, **107**, 3902-3909.
171. M. J. S. Dewar and D. M. Storch, *J Am. Chem. Soc.*, 1985, **107**, 3898-3902.
172. J. J. P. Stewart, *Journal of Computational Chemistry*, 1989, **10**, 209-220.
173. J. J. P. Stewart, *Journal of Computational Chemistry*, 1989, **10**, 221-264.
174. P. Repasky Matthew, J. Chandrasekhar and L. Jorgensen William, *Journal of Computational Chemistry*, 2002, **23**, 1601-1622.
175. A. D. Becke, *Journal of Chemical Physics*, 1993, **98**, 1372-1377.
176. K. Kim and K. D. Jordan, *The Journal of Physical Chemistry*, 1994, **98**, 10089-10094.
177. C. T. Lee, W. T. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785-789.
178. B. D. Alexander and T. J. Dines, *Journal of Physical Chemistry A*, 2004, **108**, 146-156.
179. L. Hung and E. A. Carter, *Chemical Physics Letters*, 2009, **475**, 163-170.
180. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J Am. Chem. Soc.*, 1996, **118**, 2309-2309.
181. J. Wang, R. Guo, J. Chen, Q. Zhang and X. Liang, *Anal. Chim. Acta*, 2005, **540**, 307-315.
182. R. Jacob, M. Tate, Y. Banti, C. Rix and D. E. Mainwaring, *Journal of Physical Chemistry A*, 2008, **112**, 322-331.
183. J. Aihara, *Journal of Physical Chemistry A*, 1999, **103**, 7487-7495.
184. Z. Adali-Kaya, B. T. S. Bui, A. Falcimaigne-Cordin and K. Haupt, *Angewandte Chemie-International Edition*, 2015, **54**, 5192-5195.

185. J. W. Li, Y. L. Wang, S. Yan, X. J. Li and S. Y. Pan, *Food Chem.*, 2016, **192**, 260-267.
186. B. Rezaei, M. K. Boroujeni and A. A. Ensafi, *Biosens. Bioelectron.*, 2015, **66**, 490-496.
187. G. R. Deviprasad and F. D'Souza, *Chem. Commun.*, 2000, 1915-1916.
188. Y. Diñeiro, M. I. Menéndez, M. C. Blanco-López, M. J. Lobo-Castañón, A. J. Miranda-Ordieres and P. Tuñón-Blanco, *Anal. Chem.*, 2005, **77**, 6741-6746.
189. Y. Diñeiro, M. I. Menéndez, M. C. Blanco-López, M. J. Lobo-Castañón, A. J. Miranda-Ordieres and P. Tuñón-Blanco, *Biosens. Bioelectron.*, 2006, **22**, 364-371.
190. Y. Yoshimi, A. Narimatsu, K. Nakayama, S. Sekine, K. Hattori and K. Sakai, *Journal of Artificial Organs*, 2009, **12**, 264-270.
191. K. Hattori, M. Hiwatari, C. Iiyama, Y. Yoshimi, F. Kohori, K. Sakai and S. A. Piletsky, *Journal of Membrane Science*, 2004, **233**, 169-173.
192. W. J. Wizeman and P. Kofinas, *Biomaterials*, 2001, **22**, 1485-1491.
193. A. Mirmohseni, R. Pourata and M. Shojaei, *Ieee Sensors Journal*, 2014, **14**.
194. H. Seong, H. B. Lee and K. Park, *Journal of Biomaterials Science-Polymer Edition*, 2002, **13**, 637-649.
195. B. Okutucu and S. Onal, *Talanta*, 2011, **87**, 74-79.
196. Z. Shariatinia, M. F. Erben, C. O. Della Vedova, M. Abdous and S. Azodi, *Structural Chemistry*, 2011, **22**, 1347-1352.
197. K. K. Tadi and R. V. Motghare, *Journal of Molecular Modeling*, 2013, **19**, 3385-3396.
198. T. A. Halgren, *Journal of Computational Chemistry*, 1996, **17**, 490-519.
199. T. Ogawa, K. Hoshina, J. Haginaka, C. Honda, T. T. Moto and T. Uchida, *J. Pharm. Sci.*, 2005, **94**, 353-362.
200. T. Sagawa, K. Togo, C. Miyahara, H. Ihara and K. Ohkubo, *Anal. Chim. Acta*, 2004, **504**, 37-41.
201. M. C. Norell, H. S. Andersson and I. A. Nicholls, *Journal of Molecular Recognition*, 1998, **11**, 98-102.
202. S. Azodi-Deilami, M. Abdouss and M. Javanbakht, *Appl. Biochem. Biotechnol.*, 2011, **164**, 133-147.
203. K. Rostamizadeh, M. Vahedpour and S. Bozorgi, *Int. J. Pharm.*, 2012, **424**, 67-75.
204. K. Karim, T. Cowen, A. Guerreiro, E. Piletska, J. M. Whitcombe and A. S. Piletsky, *Global Journal of Biotechnology and Biomaterial Science*, 2017, **3**, 1-7.
205. N. L. Allinger, *J Am. Chem. Soc.*, 1977, **99**, 8127-8134.
206. N. L. Allinger, Y. H. Yuh and J. H. Lii, *J Am. Chem. Soc.*, 1989, **111**, 8551-8566.
207. N. L. Allinger, L. Yan and K. Chen, *Journal of Computational Chemistry*, 2004, **15**, 1321-1330.
208. N. L. Allinger, K. Chen and J. H. Lii, *Journal of Computational Chemistry*, 1996, **17**, 642-668.
209. K. Gundertofte, T. Liljefors, P. o. Norrby and I. Pettersson, *Journal of Computational Chemistry*, 1996, **17**, 429-449.
210. P. K. Weiner and P. A. Kollman, *Journal of Computational Chemistry*, 1981, **2**, 287-303.
211. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman, *Computer Physics Communications*, 1995, **91**, 1-41.

212. P. Hobza, M. Kabeláč, J. Šponer, P. Mejzlík and J. Vondrášek, *Journal of Computational Chemistry*, 1997, **18**, 1136-1150.
213. R. Salomon-Ferrer, D. A. Case and R. C. Walker, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2013, **3**, 198-210.
214. M. Clark, R. D. Cramer and N. Vanopdenbosch, *Journal of Computational Chemistry*, 1989, **10**, 982-1012.
215. K. Gundertofte, J. Palm, I. Pettersson and A. Stamvik, *Journal of Computational Chemistry*, 1991, **12**, 200-208.
216. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of Computational Chemistry*, 2004, **25**, 1157-1174.
217. E. V. Piletska, R. Burns, L. A. Terry and S. A. Piletsky, *Journal of Agricultural and Food Chemistry*, 2012, **60**, 95-99.
218. E. V. Piletska, K. Karim, M. Cutler and S. A. Piletsky, *J. Sep. Sci.*, 2013, **36**, 400-406.
219. I. Bakas, N. Ben Oujji, E. Moczko, G. Istamboulie, S. Piletsky, E. Piletska, I. Ait-Ichou, E. Ait-Addi, T. Noguier and R. Rouillon, *Anal. Chim. Acta*, 2012, **734**, 99-105.
220. I. Chianella, S. A. Piletsky, I. E. Tothill, B. Chen and A. P. F. Turner, *Biosens. Bioelectron.*, 2003, **18**, 119-127.
221. S. D. Azevedo, D. Lakshmi, I. Chianella, M. J. Whitcombe, K. Karim, P. K. Ivanova-Mitseva, S. Subrahmanyam and S. A. Piletsky, *Industrial & Engineering Chemistry Research*, 2013, **52**, 13917-13923.
222. I. Sánchez-Barragán, K. Karim, J. M. Costa-Fernández, S. A. Piletsky and A. Sanz-Medal, *Sensors and Actuators*, 2007, **123**, 798-804.
223. S. Subrahmanyam, S. A. Piletsky, E. V. Piletska, B. Chen, K. Karim and A. P. F. Turner, *Biosens. Bioelectron.*, 2001, **16**, 631-637.
224. D. Lakshmi, A. Bossi, M. J. Whitcombe, I. Chianella, S. A. Fowler, S. Subrahmanyam, E. V. Piletska and S. A. Piletsky, *Anal. Chem.*, 2009, **81**, 3576-3584.
225. E. Piletska, D. Cowieson, C. Legge, A. Guerreiro, K. Karim and S. Piletsky, *Anal. Methods*, 2013, **5**, 6954-6959.
226. A. Florea, T. Cowen, S. Piletsky and K. De Wael, *Talanta*, 2018, **186**, 362-367.
227. M. Anfal, G. Antonio, A. B. Firas, P. Elena, M. Irfan, S. Sulman, K. Anagha, K. Oscar, H. Luisa, C. Todd, P. Sergey, A. P. W. and Y. Hasan, *Angewandte Chemie International Edition*, 2017, **56**, 16555-16558.
228. J. Settapani, K. Karim, A. Chauvin, S. M. Ibnou-Ali, F. Paille-Barrere, E. Mirkes, A. Gorban, L. Larcombe, M. J. Whitcombe, T. Cowen and S. A. Piletsky, *J. Chin. Adv. Mater. Soc.*, 2018, 1-10.
229. K. Farrington, E. Magner and F. Regan, *Anal. Chim. Acta*, 2006, **566**, 60-68.
230. M. S. Khan, P. S. Wate and R. J. Krupadam, *Journal of Molecular Modeling*, 2012, **18**, 1969-1981.
231. B. J. Brune, J. A. Koehler, P. J. Smith and G. F. Payne, *Langmuir*, 1999, **15**, 3987-3992.
232. D. L. Rathbone and Y. Ge, *Anal. Chim. Acta*, 2001, **435**, 129-136.
233. S. Monti, C. Cappelli, S. Bronco, P. Giusti and G. Ciardelli, *Biosens. Bioelectron.*, 2006, **22**, 153-163.
234. Z. J. Zhao, Q. Wang, L. Zhang and T. Wu, *J. Phys. Chem. B*, 2008, **112**, 7515-7521.

235. D. W. Lewis, D. J. Willock, C. R. A. Catlow, J. M. Thomas and G. J. Hutchings, *Nature*, 1996, **382**, 604-606.
236. K. Farrington and F. Regan, *Biosens. Bioelectron.*, 2007, **22**, 1138-1146.
237. C. K. Dong, X. Li, Z. C. Guo and J. Y. Qi, *Anal. Chim. Acta*, 2009, **647**, 117-124.
238. V. V. Barkaline, Y. V. Douhaya and A. Tsakalof, *Journal of Molecular Modeling*, 2013, **19**, 359-369.
239. A. Warshel and M. Levitt, *Journal of Molecular Biology*, 1976, **103**, 227-249.
240. H. M. Senn and W. Thiel, *Angewandte Chemie-International Edition*, 2009, **48**, 1198-1229.
241. A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, *Journal of Physical Chemistry A*, 2001, **105**, 9396-9409.
242. J. Behler, *Chemical Modelling: Application and Theory*, RSC Publishing, Cambridge, 2010.
243. I. Yarovsky and E. Evans, *Polymer*, 2002, **43**, 963-969.
244. A. Bandyopadhyay, P. K. Valavala, T. C. Clancy, K. E. Wise and G. M. Odegard, *Polymer*, 2011, **52**, 2445-2452.
245. T. Cowen, M. Busato, K. Karim and S. A. Piletsky, *Macromolecular Rapid Communications*, 2016, **37**, 2011-2016.
246. J. Terracina Jacob, T. Sharfstein Susan and M. Bergkvist, *Journal of Molecular Recognition*, 2017, **31**, e2612.
247. P. Cakir, A. Cutivet, M. Resmini, B. T. S. Bui and K. Haupt, *Adv. Mater.*, 2013, **25**, 1048-1051.
248. G. M. Yang and F. Q. Zhao, *Biosens. Bioelectron.*, 2015, **64**, 416-422.
249. K. Muzyka, K. Karim, A. Guerreiro, A. Poma and S. Piletsky, *Nanoscale Research Letters*, 2014, **9**, 154.
250. I. Mijangos, F. Navarro-Villoslada, A. Guerreiro, E. Piletska, I. Chianella, K. Karim, A. Turner and S. Piletsky, *Biosens. Bioelectron.*, 2006, **22**, 381-387.
251. A. F. M. Barton, *Chemical Reviews*, 1975, **75**, 731-753.
252. D. J. Greenhalgh, A. C. Williams, P. Timmins and P. York, *J. Pharm. Sci.*, 1999, **88**, 1182-1190.
253. R. F. Fedors, *Polymer Engineering & Science*, 1974, **14**, 147-154.
254. B. C. Hancock, P. York and R. C. Rowe, *Int. J. Pharm.*, 1997, **148**, 1-21.
255. G. DiPaola-Baranyi and J. E. Guillet, *Macromolecules*, 1978, **11**, 228-235.
256. Y. Zhu and L. Liao, *Journal of Nanoscience and Nanotechnology*, 2015, **15**, 4753-4773.
257. H. Meng, M. Liong, T. Xia, Z. Li, Z. Ji, J. I. Zink and A. E. Nel, *ACS Nano*, 2010, **4**, 4539-4550.
258. S. K. Sahoo and V. Labhasetwar, *Molecular Pharmaceutics*, 2005, **2**, 373-383.
259. S. A. A. Rizvi and A. M. Saleh, *Saudi Pharmaceutical Journal*, 2018, **26**, 64-70.
260. E. Pérez-Herrero and A. Fernández-Medarde, *European Journal of Pharmaceutics and Biopharmaceutics*, 2015, **93**, 52-79.
261. K. Ulbrich, K. Holá, V. Šubr, A. Bakandritsos, J. Tuček and R. Zbořil, *Chemical Reviews*, 2016, **116**, 5338-5431.
262. C. Wu, B. Bull, C. Szymanski, K. Christensen and J. McNeill, *ACS Nano*, 2008, **2**, 2415-2423.
263. A. A. Shemetov, I. Nabiev and A. Sukhanova, *ACS Nano*, 2012, **6**, 4585-4602.

264. X. Yu, I. Trase, M. Ren, K. Duval, X. Guo and Z. Chen, *Journal of Nanomaterials*, 2016, **2016**, 15.
265. J. J. Shi, P. W. Kantoff, R. Wooster and O. C. Farokhzad, *Nat. Rev. Cancer*, 2017, **17**, 20-37.
266. J. M. Rosenholm, A. Meinander, E. Peuhu, R. Niemi, J. E. Eriksson, C. Sahlgren and M. Lindén, *ACS Nano*, 2009, **3**, 197-206.
267. M. P. Desai, V. Labhasetwar, E. Walter, R. J. Levy and G. L. Amidon, *Pharm. Res.*, 1997, **14**, 1568-1573.
268. H. Gao, W. Shi and L. B. Freund, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 9469-9474.
269. B. D. Chithrani and W. C. W. Chan, *Nano Letters*, 2007, **7**, 1542-1550.
270. B. D. Chithrani, A. A. Ghazani and W. C. W. Chan, *Nano Letters*, 2006, **6**, 662-668.
271. F. Osaki, T. Kanamori, S. Sando, T. Sera and Y. Aoyama, *J Am. Chem. Soc.*, 2004, **126**, 6520-6521.
272. J. Agudo-Canalejo and R. Lipowsky, *ACS Nano*, 2015, **9**, 3704-3720.
273. T. Nakai, T. Kanamori, S. Sando and Y. Aoyama, *J Am. Chem. Soc.*, 2003, **125**, 8465-8475.
274. L. Josephson, C.-H. Tung, A. Moore and R. Weissleder, *Bioconjugate Chemistry*, 1999, **10**, 186-191.
275. H. W. Child, P. A. del Pino, J. M. De La Fuente, A. S. Hursthouse, D. Stirling, M. Mullen, G. M. McPhee, C. Nixon, V. Jayawarna and C. C. Berry, *ACS Nano*, 2011, **5**, 7910-7919.
276. T. Xia, M. Kovochich, J. Brant, M. Hotze, J. Sempf, T. Oberley, C. Sioutas, J. I. Yeh, M. R. Wiesner and A. E. Nel, *Nano Letters*, 2006, **6**, 1794-1807.
277. M. Mahmoudi, J. Meng, X. Xue, X. J. Liang, M. Rahman, C. Pfeiffer, R. Hartmann, P. R. Gil, B. Pelaz, W. J. Parak, P. del Pino, S. Carregal-Romero, A. G. Kanaras and S. Tamil Selvan, *Biotechnology Advances*, 2014, **32**, 679-692.
278. A. Schädlich, H. Caysa, T. Mueller, F. Tenambergen, C. Rose, A. Göpferich, J. Kuntsche and K. Mäder, *ACS Nano*, 2011, **5**, 8710-8720.
279. W.-K. Oh, S. Kim, M. Choi, C. Kim, Y. S. Jeong, B.-R. Cho, J.-S. Hahn and J. Jang, *ACS Nano*, 2010, **4**, 5301-5313.
280. J. M. Caster, S. K. Yu, A. N. Patel, N. J. Newman, Z. J. Lee, S. B. Warner, K. T. Wagner, K. C. Roche, X. Tian, Y. Z. Min and A. Z. Wang, *Nanomed.-Nanotechnol. Biol. Med.*, 2017, **13**, 1673-1683.
281. N. Jaulin, M. Appel, C. Passirani, G. Barratt and D. Labarre, *Journal of Drug Targeting*, 2000, **8**, 165-172.
282. R. Alyautdin, A. Reichel, R. Löbenberg, P. Ramge, J. Kreuter and D. Begley, *Interaction of Poly(butylcyanoacrylate) Nanoparticles with the Blood-Brain Barrier in vivo and in vitro*, 2010.
283. A. Prokop and J. M. Davidson, *J. Pharm. Sci.*, 2008, **97**, 3518-3590.
284. J. P. M. Almeida, A. L. Chen, A. Foster and R. Drezek, *Nanomedicine*, 2011, **6**, 815+.
285. X. Chen, L. Zhao, Y. Kang, Z. He, F. Xiong, X. Ling and J. Wu, *Frontiers in Pharmacology*, 2018, **9**.
286. S. Ashton, Y. H. Song, J. Nolan, E. Cadogan, J. Murray, R. Odedra, J. Foster, P. A. Hall, S. Low, P. Taylor, R. Ellston, U. M. Polanska, J. Wilson, C. Howes, A. Smith,

- R. J. A. Goodwin, J. G. Swales, N. Strittmatter, Z. Takáts, A. Nilsson, P. Andren, D. Trueman, M. Walker, C. L. Reimer, G. Troiano, D. Parsons, D. De Witt, M. Ashford, J. Hrkach, S. Zale, P. J. Jewsbury and S. T. Barry, *Science Translational Medicine*, 2016, **8**, 325ra317-325ra317.
287. S. Pardeshi, R. Dhodapkar and A. Kumar, *Food Chem.*, 2014, **146**, 385-393.
288. P. Hazot, T. Delair, C. Pichot, J.-P. Chapel and A. Elaissari, *Comptes Rendus Chimie*, 2003, **6**, 1417-1424.
289. B. Elmas, M. Tuncel, S. Şenel, S. Patir and A. Tuncel, *Journal of Colloid and Interface Science*, 2007, **313**, 174-183.
290. L. Yang, Z. Fan, T. Wang, W. Cai, M. Yang, P. Jiang, M. Zhang and X. Dong, *Anal. Lett.*, 2011, **44**, 2617-2632.
291. A. L. Medina-Castillo, J. F. Fernandez-Sanchez, A. Segura-Carretero and A. Fernandez-Gutierrez, *Macromolecules*, 2010, **43**, 5804-5813.
292. Y. Hoshino, T. Urakami, T. Kodama, H. Koide, N. Oku, Y. Okahata and K. J. Shea, *Small*, 2009, **5**, 1562-1568.
293. A. M. North and G. A. Reed, *Transactions of the Faraday Society*, 1961, **57**, 859-870.
294. R. J. Frank-Finney and M. Gupta, *Langmuir*, 2016, **32**, 11014-11020.
295. D. W. Van Krevelen and K. Te Nijenhuis, in *Properties of Polymers (Fourth Edition)*, Elsevier, Amsterdam, Editon edn., 2009, pp. 189-227.
296. E. A. Grulke, in *Polymer Handbook*, eds. J. Brandrup, E. H. Immergut and E. A. Grulke, John Wiley & Sons, inc., New York, USA, Editon edn., 1999, vol. Four, p. VII.
297. C. M. Hansen, *Hansen Solubility Parameters: A User's Handbook*, second edn., CRC Press, Boca Raton, FL, 2007.
298. P. J. Flory, *The Journal of Chemical Physics*, 1942, **10**, 51-61.
299. E. Manias and L. A. Utracki, in *Polymer Blends Handbook*, eds. L. A. Utracki and C. A. Wilkie, Springer Netherlands, Dordrecht, Editon edn., 2014, pp. 171-289.
300. F. S. Bates, *Science*, 1991, **251**, 898-905.
301. E. P. Favvas and A. C. Mitropoulos, *J. Eng. Sci. Tech. Rev.*, 2008, **1**, 25-27.
302. B. Grosdidier, S. M. Osman and A. Ben Abdellah, *Physical Review B*, 2008, **78**, 024205.
303. N. P. Young and N. P. Balsara, in *Encyclopedia of Polymeric Nanomaterials*, eds. S. Kobayashi and K. Müllen, Springer Berlin, Heidelberg, Editon edn., 2015, pp. 777-782.
304. N. P. Young and N. P. Balsara, in *Encyclopedia of Polymeric Nanomaterials*, eds. S. Kobayashi and K. Müllen, Springer Berlin, Heidelberg, Editon edn., 2015, pp. 747-755.
305. T. Ougizawa and T. Inoue, *Polym J*, 1986, **18**, 521-527.
306. D. Kashchiev, *Nucleation : Basic Theory with Applications*, Elsevier Science, Oxford, 2000.
307. D. A. Porter and K. E. Easterling, *Phase transformations in metals and alloys*, Second edn., Springer-Science+Business Media, B.V., 1992.
308. T. Cowen, K. Karim and S. A. Piletsky, *Polymer Chemistry*, 2018, **9**, 4566 - 4573.
309. D. Butina and J. M. R. Gola, *Journal of Chemical Information and Computer Sciences*, 2003, **43**, 837-841.

310. I. V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk and V. V. Prokopenko, *Journal of Computer-Aided Molecular Design*, 2005, **19**, 453-463.
311. J. S. Downey, R. S. Frank, W.-H. Li and H. D. H. Stöver, *Macromolecules*, 1999, **32**, 2838-2844.
312. G. A. O'Neil, M. B. Wisnudel and J. M. Torkelson, *Macromolecules*, 1996, **29**, 7477-7490.
313. E. J. Smith, T. Bryk and A. D. J. Haymet, *The Journal of Chemical Physics*, 2005, **123**, 034706.
314. S. A. Martins, S. F. Sousa, M. J. Ramos and P. A. Fernandes, *J. Chem. Theory Comput.*, 2014, **10**, 3570-3577.
315. D. Petrov, X. Daura and B. Zagrovic, *Biophysical Journal*, 2016, **110**, 1499-1509.
316. R. L. C. Akkermans, *The Journal of Physical Chemistry B*, 2017, **121**, 1675-1683.
317. A. Nicholls, D. L. Mobley, J. P. Guthrie, J. D. Chodera, C. I. Bayly, M. D. Cooper and V. S. Pande, *J. Med. Chem.*, 2008, **51**, 769-779.
318. K. Popper, 1st English edn., Routledge, London, 1934.
319. T. S. Kuhn, *The structure of scientific revolutions*, The University of Chicago Press, Ltd., London, 1962.
320. D. Stove, *Scientific irrationalism*, Taylor & Francis, Oxford, 2001.
321. A. Boatwright, C. Feng, D. Spence, E. Latimer, C. Binns, A. M. Ellis and S. Yang, *Faraday Discussions*, 2013, **162**, 113-124.
322. E. Latimer, D. Spence, C. Feng, A. Boatwright, A. M. Ellis and S. F. Yang, *Nano Letters*, 2014, **14**, 2902-2906.
323. S. A. Krasnokutski and F. Huisken, *Journal of Physical Chemistry A*, 2010, **114**, 13045-13049.
324. P. Atkins, *Galileo's Finger*, Oxford University Press, Oxford, 2003.
325. J. N. Nayak, M. I. Aralaguppi, B. V. Kumar Naidu and T. M. Aminabhavi, *Journal of Chemical & Engineering Data*, 2004, **49**, 468-474.
326. P. S. Albright and L. J. Gosting, *J Am. Chem. Soc.*, 1946, **68**, 1061-1063.
327. C. Carr and J. A. Riddick, *Industrial & Engineering Chemistry*, 1951, **43**, 692-696.
328. J. V. Herráez and R. Belda, *Journal of Solution Chemistry*, 2006, **35**, 1315-1328.