Computational optimization and characterization of molecularly imprinted polymers

Jacob Jordan Terracina
University at Albany, State University of New York, jake.terracina@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Materials Science and Engineering Commons, Nanoscience and Nanotechnology Commons, and the Other Chemistry Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Computational Optimization and Characterization of

Molecularly Imprinted Polymers

By

Jacob J. Terracina

A Dissertation

Submitted to the University at Albany, State University of New York

In Partial Fulfillment of Requirements for the Degree of

Doctor of Philosophy

College of Nanoscale Science and Engineering

2017
PREVIOUS PUBLICATIONS

Studies discussed in Chapter 2, Chapter 3, and Chapter 4 have been published in the following citations:


Published material in this dissertation is included with permission from publishers. The wording has been revised and updated to better suit the progression of this thesis document. These previously published materials are included because they were part of a programmatic line research that comprises this dissertation and their inclusion provides a coherent and sequenced scientific investigation. I, the author of this dissertation, was the primary researcher for these works.

License for use of the Journal of Molecular Modeling article has been granted by Springer.
License for use of the Journal of Molecular Recognition article has been granted by John Wiley and Sons. Terms and Conditions for the use of these articles are provided in the appendix.
ABSTRACT

Molecularly imprinted polymers (MIPs) are a class of materials containing sites capable of selectively binding to the imprinted target molecule. Computational chemistry techniques were used to study the effect of different fabrication parameters (the monomer-to-target ratios, pre-polymerization solvent, temperature, and pH) on the formation of the MIP binding sites. Imprinted binding sites were built in silico for the purposes of better characterizing the receptor–ligand interactions. Chiefly, the sites were characterized with respect to their selectivities and the heterogeneity between sites.

First, a series of two-step molecular mechanics (MM) and quantum mechanics (QM) computational optimizations of monomer–target systems was used to determine optimal monomer-to-target ratios for the MIPs. Imidazole- and xanthine-derived target molecules were studied. The investigation included both small-scale models (one-target) and larger scale models (five-targets). The optimal ratios differed between the small and larger scales. For the larger models containing multiple targets, binding-site surface area analysis was used to evaluate the heterogeneity of the sites. The more fully surrounded sites had greater binding energies. Molecular docking was then used to measure the selectivities of the QM-optimized binding sites by comparing the binding energies of the imprinted target to that of a structural analogue. Selectivity was also shown to improve as binding sites become more fully encased by the monomers. For internal sites, docking consistently showed selectivity favoring the molecules that had been imprinted via QM geometry optimizations. The computationally imprinted sites were shown to exhibit size-, shape-, and polarity-based selectivity. This represented a novel approach to investigate the selectivity and heterogeneity of imprinted polymer binding sites, by applying the rapid orientation screening of MM docking to the highly accurate QM-optimized geometries.
Next, we sought to computationally construct and investigate binding sites for their enantioselectivity. Again, a two-step MM → QM optimization scheme was used to “computationally imprint” chiral molecules. Using docking techniques, the imprinted binding sites were shown to exhibit an enantioselective preference for the imprinted molecule over its enantiomer. Docking of structurally similar chiral molecules showed that the sites computationally imprinted with R- or S-tBOC-tyrosine were able to differentiate between R- and S-forms of other tyrosine derivatives. The cross-enantioselectivity did not hold for chiral molecules that did not share the tyrosine H-bonding functional group orientations. Further analysis of the individual monomer–target interactions within the binding site led us to conclude that H-bonding functional groups that are located immediately next to the target’s chiral center, and therefore spatially fixed relative to the chiral center, will have a stronger contribution to the enantioselectivity of the site than those groups separated from the chiral center by two or more rotatable bonds. These models were the first computationally imprinted binding sites to exhibit this enantioselective preference for the imprinted target molecules.

Finally, molecular dynamics (MD) was used to quantify H-bonding interactions between target molecules, monomers, and solvents representative of the pre-polymerization matrix. It was found that both target dimerization and solvent interference decrease the number of monomer–target H-bonds present. Systems were optimized via simulated annealing to create binding sites that were then subjected to molecular docking analysis. Docking showed that the presence of solvent had a detrimental effect on the sensitivity and selectivity of the sites, and that solvents with more H-bonding capabilities were more disruptive to the binding properties of the site. Dynamic simulations also showed that increasing the temperature of the solution can significantly decrease the number of H-bonds formed between the targets and monomers. It is believed that the monomer
– target complexes formed within the pre-polymerization matrix are translated into the selective binding cavities formed during polymerization. Elucidating the nature of these interactions *in silico* improves our understanding of MIPs, ultimately allowing for more optimized sensing materials.
ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Susan Sharfstein and Dr. Magnus Bergkvist for all their guidance, scientific and otherwise, throughout my graduate education. I was fortunate to have advisors that not only tolerated but actively encouraged the direction of the project to be guided by my curiosity, rather than strictly by a pre-determined set of objectives. I would also like to thank my committee members, Dr. Jodi O’Donnell, Dr. Alan Chen, and Dr. Nate Cady, each of whom went out of his or her way to help me become the researcher that I am today.

The path from college student to PhD candidate is a bumpy one, and I could not have made it were it not for the friendships formed along the way. Thanks to Mike Rizzolo, Tyler Michalak, and Martin Tomov (now all PhDs themselves) for playing dual roles as roommates and mentors. A huge thanks to Jake Scoggin, Avery Green, and Bobby Schramm for making commiseration about research troubles an enjoyable pastime.

I must take this opportunity to acknowledge my girlfriend, Aurora Cooper, for supporting me through all the ups and downs of a doctoral candidacy. At this point, I believe she could present my research as competently as I could. For every misspelling and grammatical error that sneak into this thesis, a thousand others were caught and corrected by her.

I’d like to extend an enormous thank you to my family. I am forever grateful to my parents for their constant (and frequently embarrassing) insistence that I could be whatever I wanted to be. The friendly (and often not-at-all friendly) competition with my brother helped drive us both to achieve great things academically. Now there will be TWO Dr. Terracinas, and I’m certain my mom’s friends are already tired of hearing about it.

Finally, I would like to thank my puppy, Nana, who is sleeping on my feet as I type this.
# Table of Contents

Abstract ........................................................................................................................................... iii  
Acknowledgments ....................................................................................................................... vi  
Table of Contents ......................................................................................................................... vii  
List of Figures ............................................................................................................................... xii  
List of Tables ................................................................................................................................ xv  
List of Abbreviations .................................................................................................................... xvii  

**Chapter 1  Introduction to Molecularly Imprinted Polymers** ......................................................... 1  
1.1 Molecular Imprinting ................................................................................................................ 1  
  1.1.1 Historical Perspective ......................................................................................................... 1  
  1.1.2 MIPs in a Nutshell ............................................................................................................ 3  
1.2 Computational Modeling ........................................................................................................ 5  
  1.2.1 Quantum Mechanics (QM) Methods ................................................................................... 6  
    1.2.1.1 Ab initio ....................................................................................................................... 6  
    1.2.1.2 Semi-Empirical ........................................................................................................... 7  
    1.2.1.3 Basis Sets .................................................................................................................. 8  
  1.2.2 Molecular Mechanics (MM) and Molecular Dynamics (MD) Methods ......................... 9  
    1.2.2.1 Force Fields ............................................................................................................... 9  
    1.2.2.2 Pre-MD Equilibration ............................................................................................... 11  
    1.2.2.3 Simulated Annealing .............................................................................................. 12  
    1.2.2.4 Docking .................................................................................................................... 13  
1.3 *In Silico* Characterization and Analysis of MIPs ............................................................... 15  
  1.3.1 QM Methods for MIPs ...................................................................................................... 17
Chapter 4  Docking Selectivity of QM-Optimized Binding Sites ..........................59

4.1 Introduction to MIP Selectivity .................................................................59

4.2 Methods ....................................................................................................62

4.2.1 Size, Shape, and Polarity-Based Selectivity Modeling .......................62

4.2.2 Enantioselectivity Modeling .................................................................63

4.2.2.1 Chemical Components .................................................................63

4.2.2.2 Enantioselectivity .................................................................64

4.2.3 Cross-Enantioselectivity Modeling ....................................................66

4.2.4 H-bond Neutralization ...............................................................66

4.3 Results and Discussion .................................................................69

4.3.1 Size, Shape, and Polarity-Based Selectivity .....................................69

4.3.2 Enantioselectivity .................................................................73

4.3.3 Cross-Enantioselectivity ...............................................................77

4.3.4 H-bond Neutralization ...............................................................80

4.4 Conclusions .................................................................86

4.5 References .................................................................88

Chapter 5  QM and MD Evaluations of Protonation States, Solvents, and Temperature Effects on Molecular Imprinting .............................................................91

5.1 Introduction to Protonation States, Solvent Effects, and Temperature Effects ..........91

5.1.1 Monomer – Target Protonation States ............................................91

5.1.2 Solvent and Temperature ...............................................................92
6.3.3 Development of Nanoscale Interdigitated Electrode Transducer .................133
6.3.4 Preparation of Surface-Polymerized MIPs ........................................133
6.3.5 Sensor Analysis ..............................................................................133

6.4 References ...............................................................................................134

Post-script – A Few Words on Bulk MIP Fabrication and Analysis ....................136

Appendix .......................................................................................................139
### List of Figures

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Figure 1.1</td>
<td>Diagram of MIP Formation</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Figure 1.2</td>
<td>Flow chart of minimization scheme</td>
<td>13</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Figure 2.1</td>
<td>Theoretical behavior of the number of binding sites formed (N) as the concentration of target is increased</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Figure 2.2</td>
<td>Chromatographic selectivity vs. mole percent template for the optimization of a methacrylic acid-based nicotine-imprinted polymer</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Figure 2.3</td>
<td>Diagram of binding site formation with excess of monomers</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Figure 2.4</td>
<td>Diagram of binding site formation below critical concentration of monomers</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Figure 2.5</td>
<td>Molecular structures of the targets and monomer species investigated</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Figure 2.6</td>
<td>Small-scale and larger-scale modeled systems</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Figure 2.7</td>
<td>Small-scale BE vs. monomer-to-target ratio</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Figure 2.8</td>
<td>Larger-scale BE per target vs. monomer-to-target ratio</td>
<td>36</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Figure 3.1</td>
<td>Homogenous and heterogeneous binding site diagram</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Figure 3.2</td>
<td>Target concentration dependence of the separation factor</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Figure 3.3</td>
<td>Scatchard plot of a Langmuir isotherm</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Figure 3.4</td>
<td>Scatchard plot for ethyl adenine-9-acetate imprinted polymer with the bi-Langmuir fit lines</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Figure 3.5</td>
<td>Proposed explanation for bi-modal fit of early MIP binding</td>
<td>46</td>
</tr>
</tbody>
</table>
Figure 3.6 A log-log plot of the Freundlich isotherm ..................................................47

Figure 3.7 Affinity Distribution of (S)-cyclohexyl(phenyl)amine imprinted MIP based on Freundlich isotherm .................................................................48

Figure 3.8 Binding-site surface-area factor diagram ......................................................52

Chapter 4

Figure 4.1 Diagram of binding site selectivity .............................................................60

Figure 4.2 Molecular structures of the monomer species and the S-forms of the targets and analogues investigated .................................................................64

Figure 4.3 Monomer – target interactions .................................................................67

Figure 4.4 Two views of the MAA – target H-bonding arrangements after MM \( \rightarrow \) QM optimization of 30 MAAs and S-CBZ-Tyr .........................................................68

Figure 4.5 Highest scoring docking conformation of R-BOC-Tyr overlaid with the MM \( \rightarrow \) QM-optimized geometry .................................................................75

Figure 4.6 S-BOC-Tyr conformations after MM \( \rightarrow \) QM optimization and docking ............75

Figure 4.7 R-CBZ-Tyr conformations after MM \( \rightarrow \) QM optimization and docking ............76

Figure 4.8 S-CBZ-Tyr conformations after MM \( \rightarrow \) QM optimization and docking ............76

Figure 4.9 R-CBZ-Tyr and S-CBZ-Tyr docking in binding site optimized with R-BOC-Tyr ......78

Figure 4.10 R-BOC-Tyr and S-BOC-Tyr after docking into R-BOC-Tyr imprinted 30 MAA binding site. The MAAs of the MAA=COOH interaction, the MAA=NH+O1 interaction, the MAA-O2 interaction, and the MAA=PhenolOH interaction are shown as well ..................84

Figure 4.11 BOC-Tyr molecules modified with additional rotatable bonds .........................85

Figure 4.12 The imprinted and enantiomer BOC-Tyr molecules with an additional -CH2- between the chiral center and the carboxylic acid group .................................................................85
Chapter 5

Figure 5.1 Structures of the protonation states of histamine and methacrylic acid ..................92

Figure 5.2 Molecular structures of the targets, monomer species, and solvents .......................98

Figure 5.3 ATB generated charge group distribution for BOC-Tyr ........................................100

Figure 5.4 Starting geometry of system containing 1000 acetonitrile molecules, 100 methacrylic acid molecules, and 20 caffeine molecules .................................................................101

Figure 5.5 3x3 matrix of 20:100 BOC-Tyr:MAA and zoomed-in image of central unit ..........103

Figure 5.6 Relative impedance signal of MIP and NIP in response to addition of histamine under various pH conditions .................................................................106

Figure 5.7 Speciation graphs for histamine and methacrylic acid ........................................107

Figure 5.8 Caffeine with atom labels used in the H-bond counts ...........................................111

Figure 5.9 Histamine with atom labels used in the H-bond counts .......................................111

Figure 5.10 BOC-Tyrosine with atom labels used in the H-bond counts ...............................112

Figure 5.11 Methacrylic acid with atom labels used in the H-bond counts ...........................112

Figure 5.12 π-π stacking of CAF molecules after SA-optimization .................................115

Figure 5.13 Box-and-whisker plot of the individual MAA – target H-bond counts for the 20 Target + 100 MAA systems after SA optimization ..........................................................116

Chapter 6

Figure 6.1 Determination of the optimal percent functional monomer ...............................130

Figure 6.2 Field lines over nano-IDEs ..............................................................................133
List of Tables

Chapter 1

Table 1.1 AutoDock Vina scoring function terms and weights ........................................15

Chapter 3

Table 3.1 FBS, BE, and BE / FBS values of each of the five binding sites within each larger scale imprinted system .................................................................55

Chapter 4

Table 4.1 FBS of each binding site and α relative to other species .................................72
Table 4.2 α values of 20:1 MAA–Target imprinted sites .............................................73
Table 4.3 BE and αE values of R/S-BOC-Tyr docked into BOC-Tyr imprinted sites ..............74
Table 4.4 BE and αE values of R/S-CBZ-Tyr docked into CBZ-Tyr imprinted sites .............74
Table 4.5 BE and αC-E of BOC-Tyr and analogues docked into BOC-Tyr imprinted sites .........79
Table 4.6 Monomer – target bond energies ................................................................80
Table 4.7 BEs of BOC-Tyr docked into BOC-Tyr binding sites after neutralizing individual H-bonding MAAs .................................................................81
Table 4.8 Effect of each MAA - target interaction on BE and αE ........................................82
Table 4.9 The BE and αE values of the modified BOC-Tyr molecules, calculated via docking into the imprinted site .................................................................86

Chapter 5

Table 5.1 Protonation states of methacrylic acid histamine found throughout the range of pH values .........................................................................................................92
Table 5.2 Dielectric constant and H-bonding capabilities of porogens, and the surface areas, pore volumes, average pore sizes, and enantioseparation factors of MAA/EGDMA-based MIPs ……94

Table 5.3 Classifications and properties of the solvents .................................................................97

Table 5.4 QM calculated from BEs of 4:1 MAA:HA systems at five potential protonation state combinations ….................................................................105

Table 5.5 Number of H-bonds between various components after SA optimization step ………109

Table 5.6 Details of MAA – CAF H-bonds in 1:100 CAF:MAA system after optimization ……..113

Table 5.7 Details of MAA – HA H-bonds in 1:100 HA:MAA system after optimization ………113

Table 5.8 Details of MAA – BOC-Tyr H-bonds in 1:100 BOC-Tyr:MAA system after optimization ….................................................................114

Table 5.9 Average, range, and standard deviation of the number of MAA – target H-bonds in SA-optimized 20 target systems .................................................................117

Table 5.10 BE and α values for SA-optimized R-BOC-Tyr, CAF, and HA binding sites ………119

Table 5.11 Average number of MAA – target H-bonds during final 750 ps of MD …………..120
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-chloro-1-phenylethanol</td>
<td>ClPhEthOH</td>
</tr>
<tr>
<td>Acetyl-tyrosine</td>
<td>Acetyl-Tyr</td>
</tr>
<tr>
<td>Affinity Distribution</td>
<td>AD</td>
</tr>
<tr>
<td>Automated Topology Builder</td>
<td>ATB</td>
</tr>
<tr>
<td>Binding Energy</td>
<td>BE</td>
</tr>
<tr>
<td>Binding Site Surface Area Factor</td>
<td>F_{BS}</td>
</tr>
<tr>
<td>Caffeine</td>
<td>CAF</td>
</tr>
<tr>
<td>Canonical ensemble</td>
<td>NVT</td>
</tr>
<tr>
<td>Carbobenzoxy-serine</td>
<td>CBZ-Ser</td>
</tr>
<tr>
<td>Carbobenzoxy-tyrosine</td>
<td>CBZ-Tyr</td>
</tr>
<tr>
<td>Cross-enantioselectivity</td>
<td>α_{C-E}</td>
</tr>
<tr>
<td>D-Histidine</td>
<td>D-H</td>
</tr>
<tr>
<td>Enantioselectivity</td>
<td>α_{E}</td>
</tr>
<tr>
<td>Energy minimization</td>
<td>EM</td>
</tr>
<tr>
<td>Ethylene glycol dimethacrylate</td>
<td>EGDMA</td>
</tr>
<tr>
<td>Gaussian-type orbital</td>
<td>GTO</td>
</tr>
<tr>
<td>Hartree-Fock</td>
<td>HF</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>-------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Histamine</td>
<td>HA</td>
</tr>
<tr>
<td>Isothermal-isobaric ensemble</td>
<td>NPT</td>
</tr>
<tr>
<td>Linear combination of atomic orbitals</td>
<td>LCAO</td>
</tr>
<tr>
<td>L-Histidine</td>
<td>L-H</td>
</tr>
<tr>
<td>Methacrylic Acid</td>
<td>MAA</td>
</tr>
<tr>
<td>Molecularly imprinted polymer</td>
<td>MIP</td>
</tr>
<tr>
<td>Molecular Dynamics</td>
<td>MD</td>
</tr>
<tr>
<td>Molecular Mechanics</td>
<td>MM</td>
</tr>
<tr>
<td>Neglect of diatomic differential overlap</td>
<td>NDDO</td>
</tr>
<tr>
<td>Non-imprinted polymer</td>
<td>NIP</td>
</tr>
<tr>
<td>Quantum Mechanics</td>
<td>QM</td>
</tr>
<tr>
<td>Selectivity</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Self-consistent field</td>
<td>SCF</td>
</tr>
<tr>
<td>Simulated Anneal</td>
<td>SA</td>
</tr>
<tr>
<td>Slater-type Orbital</td>
<td>STO</td>
</tr>
<tr>
<td>tert-butoxycarbonyl-tyrosine</td>
<td>BOC-Tyr</td>
</tr>
<tr>
<td>Theobromine</td>
<td>TB</td>
</tr>
<tr>
<td>Theophylline</td>
<td>THO</td>
</tr>
</tbody>
</table>

CHAPTER 1

Introduction to Molecular Imprinting and Computational Modeling

1.1 Molecular Imprinting

1.1.1 Historical Perspective

Molecularly imprinted polymers (MIPs) are synthetic materials that have the ability to selectively bind to specific target analytes of interest. In 1973, Wulff et al. were able to produce what they called “enzyme-analogue built polymers” [1]. 2:1 covalent conjugates of p-vinylbenzeneboronic acid (the functional monomer) and 4-nitrophenyl-α-D-mannopyranoside (the target) were synthesized and copolymerized with methyl methacrylate and ethylene dimethacrylate (a crosslinker). After polymerization, the boronic acid ester was cleaved, removing the 4-nitrophenyl-α-D-mannopyranoside from the matrix. The polymer then exhibited a selectivity toward the target. It was hypothesized that the conformation of the two boronic acid groups were “frozen” in the polymer, creating a “memory” or “imprint” of the target structure derived from the position and orientation of these groups. The target was therefore able to covalently bind to the site when re-introduced.

In 1981, Mosbach and Arshady demonstrated that covalent linkages between the functional monomer and the target were not necessarily required for imprinting non-covalent monomer–target interactions were sufficient for the “memory” of the target to be retained after polymerization [2, 3]. This revelation significantly simplified the fabrication process, and non-covalent imprinting became the primary focus of academic study. The advantages and disadvantages of the two techniques are highlighted below (adapted from Komiyama et al. [3]):
• Covalent Imprinting
  o Advantages
    ▪ Monomer – target conjugates are stable and stoichiometric, making the imprinting process and the structure of the binding sites relatively simple to predict and understand
    ▪ A wide range of polymerization conditions can be employed (high temperature, high and low pH, wide range of solvents), as the conjugates are formed by relatively stable covalent linkages
  o Disadvantages
    ▪ Monomer – target conjugation is non-trivial and less economical than non-covalent methods
    ▪ The number of reversible covalent linkages available is limited
    ▪ The imprinting effect is significantly diminished when cleaving the covalent linkages, which requires severe conditions
    ▪ Binding and release are slow
• Non-covalent Imprinting
  o Advantages
    ▪ Synthesis of covalent conjugates not required
    ▪ Target can be removed from the polymer matrix more easily
    ▪ Binding and release are fast
  o Disadvantages
    ▪ The monomer – target complexation process is not strictly stoichiometric, which makes the formation of binding sites less predictable and uniform
Polymerization conditions must be carefully chosen to facilitate monomer–target complexation in the mixtures.

Monomers must exist in excess in order to displace the equilibrium for the formation of the complexes, which provides non-specific binding sites.

The work presented in this thesis is focused on non-covalent imprinting, and henceforth all mentions of imprinting are in reference to non-covalent techniques.

1.1.2 MIPs in a Nutshell

The target analytes are typically small, organic molecules, such as drugs, toxins, or amino acids, with molecular weights on the order of 100-300 g/mol, though imprinting of macromolecules has become increasingly common in recent years [4, 5]. The target and monomer precursors are mixed with a porogen solvent and incubated to allow monomer–target complexes to form. Inducing polymerization locks these complexes into place, where extensive washing to remove the target leaves behind a binding site with size-, shape-, and polarity-based complementarity toward that particular target molecule. Figure 1.1 below diagrams the formation of a MIP binding site. The formation of selective binding sites makes MIPs interesting alternatives to antibodies and other bio-based systems for use in separations [6], catalysis [7], and sensing applications [8]. In addition to being relatively inexpensive to produce, the highly crosslinked polymeric structure enables MIPs to function in a wide range of temperatures, pH values, and solvents not suitable for antibody-based binding elements [9, 10].

Compared to monoclonal antibodies, MIP sites tend to exhibit weaker binding affinities and are typically less selective. While a cluster of monoclonal antibodies will share a monovalent affinity and selectivity for the target molecule, MIPs will inherently contain a range of binding
sites varying from strong and selective to weak and indiscriminate [11]. Optimization of the MIP formulation can improve the selectivity and binding affinity, bringing them closer to the values presented by their biological counterparts. MIPs have historically been optimized experimentally. However, the large number of variables (choice of monomer, choice of crosslinker, choice of porogen, stoichiometric ratios between each component, reaction time and temperature, etc.) complicates the effort to optimize MIPs for a particular target, as each one may require a unique optimization scheme [11–13].

Despite their apparent simplicity, there is still much to learn about the underlying mechanisms that ultimately define the sensitivity and selectivity of a MIP. Today, it is generally accepted that the recognition capabilities of imprinted binding sites can be correlated to the nature of the monomer – target complexes that are formed in the pre-polymerization matrix [13–17]. Unfortunately, it has proven exceedingly difficult to characterize these interactions experimentally. It has been said that the greatest problem facing MIP technology is the lack of understanding of exactly how and when the imprinted receptor sites are produced [18]. Toward this end, computational approaches have gained attention for the purpose of rationally optimizing and characterizing MIPs [19–22]. Greater understanding of the interactions between the target and the monomers that will become the binding site is necessary for MIPs to achieve their maximum potential as extracting and sensing agents, as well as to recognize the theoretical limitations of the technology. To explore these limitations, it is necessary to replicate the imprinting process and analyze the properties contributing to the behavior of selective binding sites [23].

A detailed history and analysis of MIPs more specific to the work presented here can be found in the following chapters. **Chapter 2** focuses on the monomer-to-target ratios; **Chapter 3**
delves into binding site heterogeneity; Chapter 4 discusses binding site selectivity, and Chapter 5 concentrates on the solvent, temperature, and protonation states of MIP fabrication.

Figure 1.1: Diagram of MIP formation: (A) Targets and monomers are dissolved into a solvent; (B) Monomer – target complexes form via H-bonding; (C) Polymerization locks the monomers into the monomer – target complexation orientation; (D) Chemical rinsing removes the target from the binding site; (E) The target can then selectively and reversibly bind with the imprinted site.

1.2 Computational Modeling

Broadly speaking, computational chemistry is the use of computer simulations to investigate chemical properties. These properties include the molecular structure, interaction energies, and charge densities, as well as the reactivities of both static and dynamic interactions. Simplified or idealized molecular systems are defined in mathematical terms to facilitate the
calculation and prediction of chemical behavior. Computational expense (the amount of time, memory, and disk space required to complete a given task) scales dramatically with the number of atoms within the simulated system. Systems consisting of a few small molecules can be modeled with a high degree of accuracy using quantum chemical methods to solve the Schrödinger equation associated with the molecular Hamiltonian, without including any known experimental values into the parameterization. To simulate larger systems, the mathematical models must be simplified in order to attain reasonable computational expenses. The methods relevant to the work presented in the following chapters are further described below.

### 1.2.1 Quantum Mechanics (QM) Methods

Electronic structure methods, commonly known as quantum mechanics (QM) methods, are those based on quantum chemistry, for which the electronic Schrödinger equation is solved. QM methods explicitly describe the behavior of electrons with mathematical representations, making it possible to derive chemical properties that are dependent on the distribution of electrons [24]. The goal of QM methods is generally to find the minimum energy state of the system. Due to the computational expense of QM, the final optimized geometry is highly dependent on the starting geometry.

#### 1.2.1.1 Ab Initio

*Ab initio* methods are those derived purely from theoretical principles; all approximations used to solve the Schrödinger equation are defined on first principles. One of the more common techniques for approximating wave functions is known as the Hartree-Fock (HF) method, also known as the self-consistent field (SCF) method. HF is a method for determining the energy of a
many-body quantum system while in a stationary state [24]. HF was developed to solve the electronic Schrödinger equation that results from the time-independent Schrödinger equation after invoking the Born-Oppenheimer approximation (the assumption that the atomic motion of the nuclei and electrons can be separated, allowing for the wavefunction to be broken into separate electronic and nuclear components) [25]. Modern HF calculations approximate wave functions by a linear combination of atomic orbital (LCAO). Slater-type orbitals (STOs) are atomic orbital functions commonly used in the LCAO method. However, Gaussian-type orbitals (GTOs) have become more common. Gaussian function properties guarantee that the product of two GTOs centered on two different atoms results in a finite sum of Gaussians centered on a point along the axis between them, allowing for a reduction in computational expense by several orders of magnitude over STO methods [26]. Even with GTOs, \textit{ab initio} methods are very computationally expensive, and are therefore typically only used for very small systems (~100 atoms).

QM methods computationally solve the Roothaan-Hall equation. The Roothaan-Hall equation (eq. 1.1) is a representation of the HF equation.

\[ FC = SC\epsilon \]  

\text{eq. 1.1}

Here \( F \) is the Fock matrix, \( C \) is a matrix of electron interaction coefficients, \( S \) is the overlap matrix of the basis functions (see “Basis Sets” section below), and \( \epsilon \) is the matrix of orbital energies [27]. Detailed derivations of each component in eq. 1.1 are beyond the scope of this work, but can be found in any computational chemistry textbook.

1.2.1.2 Semi-Empirical

Semi-empirical methods are derived from the same quantum chemistry formalism as \textit{ab initio} methods (typically using HF) but parameterization includes approximations based on
empirical data to simplify the computations. The primary approximation included in the semi-empirical methods used for this work is the neglect of diatomic differential overlap (NDDO) [28]. NDDO theory neglects the differential overlap between atomic orbitals on different atoms [29]. This replaces the overlap matrix in eq. 1.1 with an identity matrix, simplifying the Roothaan-Hall equation:

\[ FC = C \epsilon \]  

\text{eq. 1.2}

The HF calculation is typically further simplified by considering only the valence electrons. Core electrons are subsumed into the nuclear core [30]. These simplifications reduce the computational expense without sacrificing much of the accuracy associated with QM methods, allowing for the simulation of larger systems (~1000 atoms).

1.2.1.3 Basis Sets

The set of atomic functions combined to define the molecular orbitals and their corresponding interactions with other atoms is known as a basis set [24]. There are hundreds of different basis sets, with varying degrees of computational expense and accuracy. Minimal basis sets are those containing only the number of functions required to accommodate all the filled orbitals in each atom. Doubling the number of functions used for each orbital creates what is known as a double zeta basis set. Double zeta basis sets generally improve upon results obtained with minimal basis sets. An alternative to the double zeta approach is to only double the functions used for valence electrons. This is known as a split-valence double zeta basis set. Triple and quadruple zeta basis sets are also commonly used. Basis sets also vary on whether polarization and/or diffuse functions are included.
1.2.2 Molecular Mechanics (MM) and Molecular Dynamics (MD) Methods

Unlike QM methods, Molecular Mechanics (MM) methods do not involve any quantum chemistry calculations. Instead, the atoms are modeled using classical Newtonian principles; particles are treated as point-charges and bonds are modeled as harmonic oscillators. Newtonian approximations allow for the modeling of very large systems (100,000+ atoms). However, the reductions in computational expense are typically paired with reductions in accuracy, particularly when investigating chemical properties that depend on the electronic distribution in a molecule [31].

MM methods can be used to simulate interactions over a fixed body of time, modeling the dynamic evolution of the system. These simulations are described as molecular dynamics (MD, or MM/MD). By simulating molecules in motion, the dependence on the starting orientation inherent to QM methods is eliminated. Typical timeframes for MD simulations are on the order of picoseconds to nanoseconds, though longer simulations are possible [32].

1.2.2.1 Force Fields

MM/MD simulations rely on functional forms and parameter sets to calculate the potential energy of the system, known collectively as the force field. The functional form includes bonded and non-bonded terms for describing atomic interactions, and can be described:

\[ E_{total} = E_{bonded} + E_{nonbonded} \] \hspace{1cm} \text{eq. 1.3}

where

\[ E_{bonded} = E_{bonds} + E_{angles} + E_{torsions} \] \hspace{1cm} \text{eq. 1.4}
\[ E_{\text{nonbonded}} = E_{\text{vander Waals}} + E_{\text{electrostatic}} \quad \text{eq. 1.5} \]

Common functional forms for MM force field bonded and nonbonded terms are shown in eq. 1.6 and 1.7, respectively.

\[ E_{\text{bonded}} = \sum_{\text{bonds}} \frac{k_{b,i}}{2} (l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_{\theta,i}}{2} (\theta_i - \theta_{i,0})^2 \]

\[ + \sum_{\text{torsions}} \frac{V_{n,i}}{2} (1 + \cos(n\tau - \Phi)) \quad \text{eq. 1.6} \]

Here, \( l_{i,0}, \theta_{i,0}, \) and \( \Phi \) represent the equilibrium reference length, bond angle, and torsion angle respectively. \( k_b \) and \( k_{\theta} \) are empirical stiffness constants related to the stretching and bending of each individual bond or pair of bonds. \( V_n \) is another empirical constant related to the amplitude of the energy range as a bond rotates, with a periodicity of \( n \).

Non-bonded interactions are the basis of the monomer – target complexation that defines MIP functionality, as the monomer – target interactions are comprised of non-covalent intermolecular forces. Electrostatic interactions are typically computed using Coulomb’s law, while the van der Waals interactions are computed using the Lennard-Jones potential, as shown in eq. 1.7 below.

\[ E_{\text{nonbonded}} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left( 4\varepsilon_{ij} \left[ \frac{1}{r_{ij}} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right) \quad \text{eq. 1.7} \]

Here, \( r \) is a measure of the distance between particles, \( \varepsilon \) is the “well depth”, a measure of how strongly two particles attract each other, \( \sigma \) is the distance at which the intermolecular potential between the two particles is zero, and \( q \) is the charge of the particle.

The parameterization, the assignment of values to the constants used in the force field equation, is generally derived from empirical values or high level ab-initio QM calculations. These
parameters include atomic masses, bond lengths and angles, van der Waals radii, partial charges, effective spring constants, and the Lennard-Jones parameters described above. The parameterization values, as well as the weighting of each term’s contribution to the potential energy, varies between different force fields, which are generally optimized for particular fields of chemistry [31]. When using MM to investigate any chemical behavior, one must select a force field that has been designed for and proven capable of accurately representing phenomena similar to those to be studied. For example, a force field parameterized to model protein – ligand interactions would likely be capable of simulating other organic molecule interactions (such as those between a monomer and an organic target). However, if the goal of the study was to model the thermal properties of silicon dioxide, it would likely fail because it has not been parameterized with the properties for appropriate the materials and conditions [33]. Common force fields include AMBER [34], GAFF [35], GROMOS [36], CHARMM [37], and OPLS [38], among others.

1.2.2.2 Pre-MD equilibration

In order to prepare a system for MD, it is important to first equilibrate it to the conditions desired for the MD simulation. Energy minimizations (EM) allow for relaxation of the molecules within the system, and are typically run until the maximum force found within the system is below a specified target threshold, determined by the user. This reduces the likelihood of unstable conformations and steric interference. The system can then be brought to the temperature and pressure desired for the MD run with a two-phase ensemble equilibration. Using a canonical ensemble (NVT), the number of atoms (N), the volume (V) and the temperature (T) are conserved. Energy is exchanged with a thermostat until the temperature converges at a desired value. For many applications, the target temperature may be ambient room temperature (~290 K) or
physiological temperature (~310 K). Next, an isothermal-isobaric ensemble (NPT) can be employed, whereby the number of atoms (N), pressure (P), and temperature (T) are conserved. The thermostat is combined with a barostat to converge at the appropriate pressure.

1.2.2.3 Simulated Annealing

The EM step described above will typically find only a local minimum potential energy configuration, which is highly dependent on the starting orientations. Consequently, MM-based EM is useful for equilibrating systems prior to running MD simulations, but should not be considered as a reliable means of obtaining representations of the global minimum. However, the influx of thermal energy into the system from the NVT and NPT ensembles allows the system the opportunity to overcome the local energy barriers as the molecules reconfigure in the MD step. The system can then be gradually cooled, increasing the likelihood of settling in a configuration that is more representative of the global minimum [39]. This process is known as simulated annealing (SA), and is considered one of the most reliable means to find the optimal solution to problems which have a large number of possible solutions [40]. Figure 1.2 below shows a flowchart of the MD production steps used in this work.
1.2.2.4 Docking

Molecular docking is a MM-based technique for determining the optimal orientation of one molecule, the ligand, when bonding to another molecule or group of molecules, the receptor. Rapid orientation screening searches for a binding mode that optimizes the bond energy of the
ligand – receptor pair. Docking algorithms are capable of generating a large number of possible structures and must include a means to score each structure to identify those of most interest [41, 42]. The central problem that docking algorithms seek to solve is how to manage the many degrees of freedom of the molecules. Each molecule has six degrees of translational freedom (up, down, left, right, forward, backward), six degrees of rotational freedom (clockwise and counter-clockwise relative with respect to each axis), and any degrees of conformational freedom afforded by each rotatable single-bond within the ligand. The ideal docking method would allow both the ligand and receptor to fully explore all possible degrees of conformational freedom. However, the exponential nature with which the computational expense scales with the number of degrees of freedom limits the ability to conduct such a complete search. Consequently, most modern docking algorithms treat the receptor molecule(s) as a rigid body [41, 43].

AutoDock Vina (a molecular docking program) was used to calculate interaction energies in the following chapters [44]. Vina’s interaction functions, $f_{t_i,t_j}$ (eq. 1.9), are defined relative to the surface distance, $d_{ij}$:

$$d_{ij} = r_{ij} - R_{t_i} - R_{t_j} \quad \text{eq. 1.8}$$

where $R_t$ is the van der Waals radius of atom type $t$.

$$f_{t_i,t_j}(r_{ij}) = h_{t_i,t_j}(d_{ij}) \quad \text{eq. 1.9}$$

Here, $h_{t_i,t_j}$ is a weighted sum of steric interactions (which are identical for all atom pairs), hydrophobic interactions, and hydrogen bonding. The steric interactions are described by two attractive Gaussian functions (eqs. 1.10 and 1.11) and one repulsive function (eq. 1.12).
The hydrophobic and hydrogen bonding terms are defined in eq. 1.13 and eq. 1.14, respectively.

\[
\text{gauss}_1(d) = e^{-\left(\frac{d}{0.5\text{Å}}\right)^2} \quad \text{eq. 1.10}
\]

\[
\text{gauss}_2(d) = e^{-\left[\left(\frac{d-3\text{Å}}{2\text{Å}}\right)^2\right]} \quad \text{eq. 1.11}
\]

\[
\text{repulsion}(d) = \begin{cases} d^2, & \text{if } d < 0 \\ 0, & \text{if } d \geq 0 \end{cases} \quad \text{eq. 1.12}
\]

The hydrophobic and hydrogen bonding terms are defined in eq. 1.13 and eq. 1.14, respectively.

\[
\text{hydrophobic}(d) = \begin{cases} 1, & d < 0.5\text{Å} \\ 1.5 - d, & 0.5 \leq d \leq 1.5\text{Å} \\ 0, & d > 1.5\text{Å} \end{cases} \quad \text{eq. 1.13}
\]

\[
\text{hydrogen bonding}(d) = \begin{cases} 1, & d < -0.7\text{Å} \\ -\frac{d}{0.7}, & -0.7\text{Å} \leq d \leq 0\text{Å} \\ 0, & d > 0\text{Å} \end{cases} \quad \text{Eq. 1.14}
\]

Table 1.1 below shows the empirically determined weight given to each term in the scoring function. These weights can be adjusted to better align the model with experimental results.

<table>
<thead>
<tr>
<th>Term</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{gauss}_1</td>
<td>-0.0356</td>
</tr>
<tr>
<td>\text{gauss}_2</td>
<td>-0.00516</td>
</tr>
<tr>
<td>\text{repulsion}</td>
<td>0.840</td>
</tr>
<tr>
<td>\text{hydrophobic}</td>
<td>-0.0351</td>
</tr>
<tr>
<td>\text{hydrogen bonding}</td>
<td>-0.587</td>
</tr>
</tbody>
</table>

Table 1.1: AutoDock Vina scoring function terms and default weights, adapted from Trott and Olson [44]

1.3 \textit{In Silico} Characterization and Analysis of MIPs

Historically, MIPs have been optimized by exhaustive empirical experiments. While functional, this approach is time consuming and expensive. As mentioned above, the molecular imprinting process is one that is surprisingly simple to outline; monomers surround a target molecule in solution, polymerization locks the components in place, and the target molecule is
removed, leaving behind a binding site with a size-, shape-, and polarity-based selective affinity for that particular molecule. However, MIP characterization and optimization has proven to be deceptively complex. The number of experimental variables available for optimization, which includes the choice of monomer(s) and crosslinker(s), their respective molar ratios with the target, the porogen solvent, and the polymerization conditions represents a significant challenge that makes formulation of a MIP toward a “new” analyte target extremely time intensive. The amorphous nature of MIPs, along with the inherent binding site heterogeneity, obscures the underlying mechanisms that define the MIP’s selective properties. To date, analysis of individual binding sites has proven to be a challenge beyond the reach of current experimental techniques and instrumental limitations. The number of binding sites can only be estimated, adding a high degree of error to calculations of binding affinity and heterogeneity.

To combat these challenges, both combinatorial and computational approaches have been explored [19, 45, 46]. In the early 2000s, researchers began to utilize computational techniques in an effort to elucidate the hidden origins of MIP binding properties [47]. In silico characterization and rational design of MIPs has become much more common in the last decade, as the rapid improvements in computing power and modeling software have made the simulation of increasingly complex molecular systems more feasible. Two thorough reviews of computational methods for rational design and characterization of MIPs have been published in the last two years [20, 21], so this section will not attempt to serve as a comprehensive summary of previous contributions to the field. Instead, it will highlight the major strategies for in silico MIP studies, with a focus on the seminal publications for various computational MIP screening techniques. Further review of the computational studies most relevant to the work presented in the following chapters can be found the chapter introductions, where applicable.
As molecular modeling can save significant experimental time if carefully applied, computational analyses of MIPs are key to facilitating their development and practical implementation. Computational modeling of MIPs using QM, MM, and MD methods offers an attractive approach to gain insights into MIP characteristics and accelerate MIP development through *in silico* optimization [19].

### 1.3.1 QM Methods for MIPs

Electronic structure methods have become a fairly common toolset in the optimization of MIP systems [21]. The high degree of accuracy afforded by QM techniques allows for the non-covalent interaction energies between monomer – target systems to be quantified. The computational cost significantly restricts the size of the system, which typically consists of only one target molecule and 1-20 monomers. Crosslinkers and solvents are usually excluded from QM systems due to the computational limitations, though solvents can be implicitly incorporated via a polarizable continuum model (PCM) [48–51]. Despite the limited system size, QM analysis of monomer – target interactions can offer insight into two critical MIP design parameters: the choice of the monomer and the monomer-to-target ratio.

In 2001, Castro et al. incorporated semi-empirical QM techniques into the rational design of a polymer imprinted with dibenzothiophene sulphone (DBTS) [47]. Using MOPAC [28], they determined the interaction energies between the target and two monomers, methacrylic acid (MAA) and 5-methacryloxy-1,3-bis(4-ethenylphenyl)-benzene dicarboxamide (MBEPB). Interaction energies, $BE$, were calculated:

\[
BE = [\Delta H_f(target + monomers) - \Delta H_f(target) - \Delta H_f(monomers)]
\]
where $\Delta H_f$ is the total heat of formation determined via the QM energy minimization. Systems containing one DBTS and one MBEPB were found to produce a stronger $BE$ than those with one DBTS and two MAAs (the ratios were chosen because DBTS contains two functional groups and it was believed that one MBEPB complexes with both groups, while MAAs could only interact with one group each). The computational $BE$ results lead the researchers to produce MIPs using the stronger monomer. Experimental HPLC binding data correlated with the computational results, marking the first seemingly successful attempt to computationally optimize a MIP design parameter.

A few years later, Holdsworth et al. used a similar technique to determine the optimal monomer (between MAA and 4-vinylpyridine (4VP)) for a cocaine-imprinted polymer, as well as the optimal monomer-to-target ratio [52]. The $BE$s were calculated for QM-optimized systems containing one cocaine molecule and 1-14 monomers (MAA or 4VP). NMR analysis validated both the computationally proposed optimal monomer and the optimal monomer-to-target ratio, further bolstering arguments for the viability of computational MIP optimization.

### 1.3.2 MM and MD Methods for MIPs

While computational methods built on Newtonian principles cannot fully replicate the quantum nature of molecular interactions, the simplification of the algorithms allows for the simulation of molecular systems that would otherwise be computationally cost-prohibitive for electronic structure methods. The development and refinement of a variety of force fields enables the study of biomolecules and small organic molecules with reasonable accuracy, as the force fields have been tuned to provide accurate reproductions of empirical data.
In 2001, the same year QM methods for MIP optimization were first being explored, Piletsky et al. presented a MD-based technique for rapid monomer screening [53, 54]. Using a library of 20 monomers, interaction energies were determined between the monomers and a target molecule (ephedrine [53] or creatinine [54]) after a simulated anneal. Molinelli et al. introduced explicit solvent molecules (water or chloroform) to MD simulations of one target molecule (2,4-dichlorophenoxyacetic acid) and one monomer molecule (4VP) in 2005 [55]. The following year, Monti et al. explored the potential for molecular docking into MD-optimized binding sites [56]. In 2007, Wei et al. took advantage of the low computational costs of MD simulations by including an excess of monomers around a target molecule, 17β-estradiol [57]. Reducing the computational expense allowed for the inclusion of monomer dimerization into the simulation, bringing the simplified computational system one step closer to the complex, real-world, pre-polymerization matrix.

In more recent years, attempts have been made to model the full pre-polymerization matrix, including explicit solvents, multiple targets, monomers, and even crosslinkers [19, 58–61].

1.4 Scope of Thesis

The trial-and-error approach to MIP design demonstrates a clear need for an improved understanding of the effects of MIP design parameters. Even with experimental screening as a means to optimize MIP fabrication protocols, the difficulties in characterizing the complex, heterogeneous, amorphous structures significantly obfuscate the nature and origins of selective binding. Recently, improvements in computational modeling techniques have allowed for a glimpse into the chemical properties behind MIP selectivity [20, 21, 62].

The aim of our study focused on the use of modern computational chemistry techniques
for the characterization of systems representative of the MIP pre-polymerization matrix. We hypothesized that a computational approach, combining small-scale QM studies and larger system MM/MD simulation studies could significantly accelerate the identification of specific MIP formulations for polar analytes. More specifically, we investigated the monomer – target interactions of the pre-polymerization matrix that ultimately define the imprinted polymer binding site. Using energy minimization techniques, we optimized these interactions to create binding sites in silico. Because individual MIP binding sites cannot be probed experimentally, the in silico binding sites afforded us the opportunity to investigate the nature of imprinting and selective binding in ways that are not possible with experimental methods. These “computationally imprinted” binding sites were optimized with respect to their monomer-to-target ratios, and then characterized in silico to study the selectivity of individual binding sites. These were the first in silico models of full MIP binding sites, and were shown to exhibit selectivity, enantioselectivity, and cross-enantioselectivity comparable to experimental results. Starting from these small models containing only one target molecule, and building from there to models that contain five targets, we were able qualitatively elucidate properties governing MIP heterogeneity. Finally building to more complex models containing 20 target molecules, we were able to simulate solvent and temperature effects that have been demonstrated in the literature.

With the broad goal of discovering and evaluating methods to optimize and characterize MIPs in silico, we found success with both QM and MM/MD techniques by computationally imprinting binding sites and then focusing on the interactions between the monomers and targets. For the first time, we were able to evaluate the selectivity and heterogeneity of imprinted sites by modeling the sites as a collection of monomers optimized around a target molecule. With some overlap, the monomer – target interaction investigations are presented with the following chapters:
In **Chapter 2**, MM and QM energy minimization techniques were used to investigate the influence of monomer-to-target ratios on the binding energies of target molecules. Ratios were investigated for small, one-target systems and larger, five-target systems. **Chapter 3** builds on this work, using the optimized five-target systems as an opportunity to investigate the heterogeneity of binding sites within MIPs. The methods and results presented in these two chapters were published in 2016 [63].

**Chapter 4** focuses on the selectivity of QM-imprinted binding sites, i.e. the imprinted target’s binding energy with the binding site relative to the binding energy of a structural analogue into the same site. Selectivity based on size, shape, and polarity was qualified, as was enantioselectivity. These methods and results were published in 2016 and 2017 [63, 64].

The influence of solvents on the monomer – target interactions was analyzed with MD simulations, presented in **Chapter 5**. The MD studies include an investigation of temperature effects. These studies build on the elements of the previous chapters, including selectivity and heterogeneity. These methods and results were submitted for publication and were under review at the time this was written.

**Chapter 5** also includes a computational study on the effect of the target and monomer protonation states on the binding energy.

### 1.5 References


27. Roothaan CCJ (1951) New Developments in Molecular Orbital Theory. Rev. Mod. Phys. 23:


2.1 Introduction to MIP Stoichiometry

For non-covalent molecular imprinting, the optimal monomer-to-target ratios have traditionally been determined empirically through experimental means [1, 2]. It was originally though that MIPs could be optimized by saturating the pre-polymerization system with target molecules. This theory was founded on the idea that the maximum number of monomer – target complexes would form, and any excess targets would not influence the system. A graphical representation of this theoretical outcome is shown in Figure 2.1. It was also expected that the optimal ratio would be directly related to the number of functional groups in the target capable of H-bonding with the surrounding monomer.

Figure 2.1: Theoretical behavior of the number of binding sites formed (N) as the concentration of target is increased [3]
However, optimizing monomer-to-target ratios for MIPs soon proved to be significantly more challenging. Experimental data repeatedly showed the affinity and selectivity the MIPs to increase with target concentration up to a point, and then drop significantly, losing all selective properties. Monomer-to-target ratios of 6:1 to 2:1 may produce strong binding sites, but beyond a threshold, increasing the target concentration had deleterious effects on the final product. Figure 2.2 shows such a case.

![Figure 2.2: Chromatographic selectivity vs. mole percent template (relative to monomer concentration) for the optimization of a methacrylic acid-based nicotine-imprinted polymer [4]](image)

Due to the difficulties inherent in characterizing binding sites after polymerization, understanding of monomer-to-target ratio optimization is limited and largely derived from complex empirical results, which rely on assumptions about the number of sites and the heterogeneity of the sites. One possible explanation for the target ratio threshold may be that an increase in concentration will result in an increase in low-affinity sites as a result of the limited monomers available for the formation of high-affinity sites. At low target concentrations, each
target can fully complex with multiple monomers. Figure 2.3 shows a fully surrounded target. At higher target concentrations, there will be fewer monomers for each target (Figure 2.4: Scheme A). The likelihood of target dimerization also increases at higher concentrations. During polymerization, dimerized targets may form large, overlapping binding sites with little or no selectivity for the individual imprinted molecules (Figure 2.4: Scheme B) [5]. Under these assumptions, the number of selective binding sites formed can be expected to increase until there are no longer enough monomers to fully surround and complex with each target. As the target concentration increases, the density of selective binding sites will also increase. At a critical concentration of targets, the monomer-to-target ratio is balanced such that the binding sites are densely populated throughout the imprinted polymer without overlapping. At higher concentrations, the likelihood of dimerization and the formation of low-affinity sites reduces the overall selectivity and affinity of the polymer.

Figure 2.3: At and below the critical concentration of targets, there are enough monomers to fully complex with and surround each target.
Figure 2.4: Beyond the critical concentration of targets, the number of selective binding decreases. This is likely due to the creation more low-affinity sites (Scheme A) and the dimerization of targets, producing overlapping binding sites.

In the late 1990s, attempts were made to predict optimal monomer-to-target ratios based on thermodynamic models, with limited success [1]. *In silico* ratio analysis, on the other hand, can offer insight into monomer – target complexation that cannot be attained through experimental means. Relative to MM and MD methods, QM methods provide a more accurate model of molecular interactions but with higher computational expense. For practical purposes, this expense typically limits the scale of the models to a few hundred atoms. The classical mechanics calculation approaches used for MM and MD methods are significantly less computationally expensive than
QM calculations, allowing for the simulation of much larger systems. However, these classical force field methods oversimplify or ignore many of the molecular interactions that arise from QM principles. Because H-bonding, the driving force of MIP interactions, is partially quantum mechanical in nature and can therefore not be fully represented with the point-charge approximations of MM and MD methods, we chose to use QM energy minimization to obtain our geometries. This is not to say that MM and MD methods have no value in computational MIP investigations; there have been many reported successes with these techniques, particularly in virtual monomer library analysis [6–8], as well as with our docking method discussed in later chapters. Improved understanding of the properties that contribute to the optimal monomer-to-target ratio will reduce time-intensive experimental MIP screening. Here, we present a method to optimize the monomer-to-target ratios using QM energy minimizations of one-target and five-target systems. Though one-target QM ratio optimizations have been explored, the five-target systems, to the best of our knowledge, represent the first time a QM-level ratio optimization accounted for the presence of multiple target molecules.

2.2 Methods

2.2.1 Components

The targets were divided into two categories. The first includes imidazole-derived structures: histamine (HA), L-histidine (L-H), and D-histidine (D-H). The second includes xanthine-derived structures: theophylline (THO), caffeine (CAF), and theobromine (TB). These targets were selected for their biological relevance as drugs (THO, CAF, TB), biomarkers (HA), and amino acids (L-H, D-H). Analyzing two groups of compounds allowed for a comparison of the relative binding energies of molecules with varying structural similarity to the imprinted
molecule. Methacrylic acid (MAA) was selected as the monomer for these studies based on a strong literature precedent and previous observations of strong interactions with the selected target molecules [9–19]. The structures of each component can be seen in Figure 2.5.

![Molecular structures of the targets and monomer species investigated, created with ACD/ChemSketch][20]

2.2.2 Stoichiometric Ratio Optimization

Ratio optimization was computed at both small (one-target) and larger (five-target) scales. Avogadro, a cross platform molecular editor (avogadro.cc), was used to build the pre-polymerization systems, which include the target(s) and monomers [21, 22]. The porogen was included via implicit solvation, using the Conductor-like Screening Model (COSMO) [23]. Chloroform was chosen as the implicit solvent for all QM optimizations. Chloroform is capable of solvating the targets and monomers, does not form strong H-bonds with either, and does not interfere with complexation and the formation of binding sites. Consequently, chloroform has been shown to produce more effective MIPs than many other pre-polymerization solvents [24, 25]. First,
the small systems consisting of one target and 2-6 monomers were built in Avogadro. The starting orientations of the molecules were chosen such that the carboxylic acid functional groups of the monomers were in approximate co-location to the polar groups of the target. The hydrophobic moieties of the monomers, which form the backbone of the MIP, were directed away from the target. The systems were then geometrically optimized using the PM6-DH2 [26] basis set with MOPAC [27], which has been called the “gold standard” of semi-empirical QM calculations [28]. This method was parameterized to correct for dispersion and H-bonding within the PM6 Hamiltonian, making it suitable for the noncovalent interactions that drive MIP binding [26].

For the larger systems, five targets and the proportional numbers of monomers (10, 15, 20, 25, or 30) were built into the system. The starting positions of targets and monomers were similar to those described for the one-target systems. Within one geometry file, five separate one-target systems were built containing the relevant number of monomers (2-6). The distances between the one-target systems were then decreased until the monomers of one system were able to interact with those of the adjacent system. The MOZYME method was utilized to manage the size of the model. With MOZYME, a localized molecular orbital (LMO) method replaces the SCF procedure to reduce computational expense. To further manage the processing load, the minimizations were made in two steps. First, coarse geometric optimizations were computed with an exit gradient of 20 kcal/mol/Å and a cutoff distance of 6 Å. This cutoff determines the range beyond which neglect of diatomic differential overlap (NDDO) interactions and polarization functions are ignored; beyond the cutoff, only point charge interactions are considered. Second, using these coarse output geometries as new starting points, fine optimizations were computed with an exit gradient of 10 kcal/mol/Å. The cutoffs were removed for the fine optimizations.

All binding energies, $BE$, are defined:
BE = \left[ \Delta H_f(templates + monomers) - \Delta H_f(templates) - \Delta H_f(monomers) \right] / N \quad \text{eq. 2.1}

where $\Delta H_f$ is the total heat of formation and $N$ is the number of targets in the system. For the small system analysis, $N=1$. For the larger systems, $N=5$. Examples of both the small and larger model matrices are shown in Fig. 2.6, which was made with AutoDock Tools [29, 30]. This method provides the average binding energy of targets in the five-target system, rather than specific binding energies for each site. Efforts to investigate variations between these individual binding sites are described in Chapter 3.

Figure 2.6: Small scale (left) and larger scale (right) modeled systems, displayed with AutoDock Tools [30]. The targets (in this case CAF) are bound to the MAA molecules, represented here by a continuous molecular surface.

2.3 Results and Discussion

2.3.1 Small Scale Stoichiometric Optimization

All systems consisting of one target show the strongest binding energy when monomers are present at a monomer-to-target ratio of 4:1. The xanthine-derived targets (THO, CAF, TB)
have greater binding energies than the imidazole derived targets (HA, L-H, D-H) as shown in Figure 2.7.

The binding energy peak at a monomer-to-target ratio of 4:1 can be attributed to the ability of four monomers to surround the target site without interfering with each other. Larger ratios will see crowding of monomers around the target, with this steric hindrance leading to weaker binding interaction. This can be observed when orienting the functional groups of the monomers toward the corresponding target groups; at higher ratios, the monomers will repel each other and weaken the interactions with the target molecule.

While small scale QM geometry optimizations can correlate to experimentally optimized monomer-to-target ratios, the results of the larger system help identify the limitations of these minimal system models. The monomer formation around the lone target represents an idealized binding site; all monomers are forming H-bonds with the target functional groups, and the target is encased. Realistically, the pre-polymerization system will be a complex amalgamation of less-than-ideal interactions, including: dimerization of monomers, dimerization of targets resulting in overlapping binding sites, solvent interference, crosslinker – target H-bonding, and other factors that contribute to the heterogeneity of binding sites. That is to say that while the models offer useful approximations of optimal experimental ratios, these optimized binding energy values will likely only qualitatively correlate with experimental results.
2.3.2 Larger scale stoichiometric optimization

When scaled up five-fold, the strongest binding energy shifts to a 5:1 monomer-to-target ratio, though 4:1 is the second strongest in most cases, as seen in Figure 2.8. This shift is likely due to the opportunity of the monomers to interact with each other in “bulk” and have less crowding around the target. These results suggest that though a 4:1 monomer-to-target ratio may form the most energetically favorable single binding site, in practice, a higher proportion of monomers will be required to account for monomer dimerization and reduce the likelihood of target dimerization by separating the binding pockets. These results correspond well with several independent experiments, which found the optimized MAA-to-HA ratio to be between 4:1 and 6:1 [15, 16]. Interestingly, the trend of xanthine targets showing stronger binding energies relative to the imidazole targets that was seen in the small scale models is no longer apparent in the larger scale models. Considering the larger scale models with a 5:1 monomer-to-target ratio showed the strongest average binding energy, they were selected for further heterogeneity and selectivity studies.
2.4 Conclusions

In this study, we demonstrated a QM approach for larger scale MIP modeling involving five target molecules and up to 30 monomer species, and compared the results with a smaller scale (one-target) system. While a one-target system model offers some insight in predicting an optimal monomer-to-target stoichiometry, the number of factors ignored by these models limits their practical value. We found that although the ratios were similar, larger models using the same QM basis set showed deviations from the small scale optimal ratios. One-target systems may be representative of an ideal interaction with the surrounding monomers, but larger models allow for the monomer–monomer interactions that could be expected in solution, making for a more accurate representation of the complexities of real world monomer–target H-bonding. It should be noted that even if a one-target system appears to be optimized when only two or three monomers are present, a higher ratio will be required to practically separate the targets and prevent target dimerization. The methods and results presented in this chapter have been published in the Journal of Molecular Modeling [31]. Target dimerization is revisited in Chapter 5.
2.5 References


CHAPTER 3

In Silico Evaluation of MIP Binding Site Heterogeneity

3.1 Introduction to MIP Heterogeneity and Binding Isotherms

Rather than uniform, homogeneous binding sites, MIPs will typically express a wide range of affinities and selectivities between binding sites (Figure 3.1). The binding site heterogeneity inherent to imprinted polymers has complicated experimental MIP analysis and limited their utility since the introduction of the science. Unlike sensitive binding elements with homogeneous binding sites (such as enzymes or monoclonal antibodies), the binding properties of MIPs are highly dependent on the particular concentration of analytes during the analysis [1]. This concentration dependence reduces the ability to extrapolate data points to predict binding properties at different concentration ranges, and can be seen in the chromatographic binding data presented by Wulff et al. in one of the seminal MIP publications in 1977 [2]. Their data (Figure 3.2) shows the chromatographic separation factor ($\alpha$) for an L-phenylalanine anilide (L-PAA) imprinted polymer. The $\alpha$ value is defined here as:

$$\alpha = \frac{[\text{bound } L_{PAA}]/[\text{free } L_{PAA}]}{[\text{bound } D_{PAA}]/[\text{free } D_{PAA}]}$$

eq. 3.1
Figure 3.1: Homogenous (A) and heterogeneous (B) binding sites. Black arrows represent functional groups of the polymer.

Figure 3.2: Target concentration dependence of the separation factor ($\alpha$) as measured for a MIP imprinted with L-PAA, measured in batch rebinding studies [2].

The graph indicates a relatively stable separation factor at concentrations greater than 0.5 mM. At lower concentrations, however, the separation factor becomes highly concentration dependent.
As early as 1995, the heterogeneity of imprinted binding sites was identified as the cause of peak broadening and asymmetry in chromatographic studies with imprinted polymers serving as the stationary phase [3]. Not long after, binding site heterogeneity was found to be a major contributor to the low selectivity of MIPs when used as enzyme analogues for catalysis [4]. However, nearly a decade would pass before binding models began to be applied in attempt to understand the nature of MIP binding site heterogeneity [1].

Once the heterogeneous nature of MIPs had been realized, the next step was to determine the shape of the distribution [5]. The shapes are most commonly represented with affinity distribution graphs (AD), which plot the number of sites (N) vs. the log of the association constant (log K). Because log K is proportional to the binding energy, \( \Delta G \), the AD plots are also known as site-energy distributions [1].

Binding isotherms, defined here as the concentration of analyte bound to a polymer, \( B \), versus the concentration of free analyte in solution, \( F \), at a constant temperature, are useful for investigating heterogeneity. The bound and free concentrations can be measured using batch rebinding studies, in which a known mass of polymer is equilibrated with a range of analyte concentrations. The concentration of bound analyte is calculated as the difference between the starting and final free concentrations. In order to derive binding parameters such as \( N \) and \( K \) from the isotherm, specific binding models must be applied. Each model specifies the mathematical relationship between bound and free analytes in solution, and makes certain assumptions about the AD. Due to the time-intensive nature of MIP experimentation, it is not uncommon for MIP publications to present binding data without first constructing a thorough isotherm. Instead, binding in small concentrations ranges, or even single concentrations will be presented. In order for MIP binding data to be transferred to other applications or conditions, the isotherm must be
properly measured [6]. Many models have been applied to MIP binding, including Langmuir, bi-
Langmuir, Freundlich, and Langmuir-Freundlich isotherms [5, 7]. The selection of the binding
model is based on its correlation to the trend of the experimental isotherm data. Binding modes
are grouped into two classes, discrete and continuous binding distribution models, and are further
described below.

3.1.1 Discrete binding models

Discrete binding models are those that consider a finite number of binding site “classes” or
“types”. This is the most simplified means of interpreting experimental data. They are especially
attractive due to the ease with which \( N \) and \( K \) can be derived from experimental binding data.

3.1.1.1 Langmuir and Bi-Langmuir Isotherms

Irving Langmuir won the Nobel Prize in 1932 for his contributions to the field of surface
chemistry. The Langmuir isotherm, first presented in 1916, makes several assumptions about
binding: (1) The surface containing the absorption sites is perfectly flat and homogeneous; (2) The
binding gas/particle/molecule adsorbs into an immobile state; (3) All sites are equivalent; (4) Sites
can hold at most one molecule; (5) There are no interactions between molecules in adjacent sites
[8]. The Langmuir isotherm is derived:

\[
A + S \leftrightarrow AS
\]

where \( A \) is the analyte, \( S \) is the surface, and \( AS \) represents an analyte molecule bound to the surface.

The equilibrium constant \( K_{ads} \) can then be defined:

\[
K_{ads} = \frac{[AS]}{[A][S]}
\]

\[ eq. 3.2 \]

\[ eq. 3.3 \]
where \([A]\) is the concentration of the analyte, and \([AS]\) and \([S]\) are two-dimensional terms analogous to concentration, expressed in terms of number of molecules per unit area. The Langmuir isotherm defines adsorption in relation to surface coverage, \(\theta\), defined as the fraction of binding sites that are filled with analyte molecules. Therefore,

\[
\frac{[AS]}{[S]} = \frac{\theta}{\theta - 1} \quad \text{eq. 3.4}
\]

The concentration of analyte \([A]\) can also be expressed as \(F\), the free analyte term defined earlier.

\[
K_{ads} = \frac{\theta}{F(\theta - 1)} \quad \text{eq. 3.5}
\]

Rearranging, we obtain the common form of the Langmuir adsorption isotherm:

\[
\theta = \frac{K_{ads}F}{1 + K_{ads}F} \quad \text{eq. 3.6}
\]

The surface coverage, \(\theta\), can also be defined as:

\[
\theta = \frac{B}{N} \quad \text{eq. 3.7}
\]

Therefore, with a little reconfiguring:

\[
\frac{B}{F} = -K_{ads}B + N \quad \text{eq. 3.8}
\]

When the binding data is presented with a Scatchard Plot \((B/F \text{ vs. } B)\), the number of binding sites and the association constant \(K_{ads}\) and are shown as the x-intercept and the negative of the slope, respectively (Figure 3.3).
Figure 3.3: The Scatchard plot of a Langmuir isotherm produces a straight line with a slope of $-K_{ads}$ and an x-intercept of $N$.

The linearity of the Scatchard plot is dependent on the Langmuir assumption that binding sites are homogeneous. While the MIP binding data may resemble a linear trend over small ranges of concentration, the heterogeneity of binding sites creates non-linear trends when graphed on a Scatchard plot [9]. Early explanations for the curvature were that MIPs contained two distinct binding classes, with two unique sets of binding parameters ($K_1$, $N_1$, and $K_2$, $N_2$). Fitting two separated straight lines to the Scatchard curve could yield the parameters (Figure 3.4). This two-line fit is known as the bi-Langmuir isotherm.
The bi-Langmuir fit is mathematically described:

\[ B = \frac{N_1K_1F}{1+K_1F} + \frac{N_2K_2F}{1+K_2F} \]  

\text{eq. 3.9}

The bi-Langmuir model stems from the idea of two discrete sets of homogeneous high-affinity sites and homogeneous low-affinity sites perhaps as result of the target having two functional groups capable of H-bonding with the monomers and binding sites forming with either one functional group H-bonding (low-affinity) or both functional groups H-bonding (high-affinity) (Figure 3.5). However, this was proven to be inaccurate [10].
Figure 3.5: Proposed explanation for seeming bi-modal fit of early MIP binding Scatchard plots. Many analytes were believed to be capable of forming two H-bonds with the functional monomers. If both bonds were present during the formation of the binding site, a high-affinity site would be created (A). One H-bond present during polymerization would result in a low-affinity binding site (B).

Measurements of MIP ADs have revealed broad, heterogeneous distributions that cannot be fit with Langmuir or bi-Langmuir models over the full range of binding [10]. The discrepancy between the experimental MIP system and mathematical model leads to inconsistencies in the calculation of binding parameters [1].

3.1.2 Continuous binding models

In response to the limitations and insufficiencies of the discrete binding models, continuous distribution models have been more recently applied to the characterization of MIP binding site heterogeneity.

3.1.2.1 Freundlich Isotherm

Described by Herbert Freundlich in 1909, the Freundlich isotherm is the oldest mathematical representation of binding behavior. It is an empirically derived relation between the
concentration of a solute bound to a surface ($B$) and the concentration in solution ($F$). The Freundlich equation is defined:

$$B = aF^m$$  \hspace{1cm} \text{eq. 3.10}

where $a$ and $m$ are empirically determined fitting parameters. The factor $a$ is a measure of both the capacity ($N$) and the average binding affinity ($K_0$), though the individual contributions of $N$ and $K_0$ to $a$ cannot be determined without first making further assumptions [11]. The exponential factor $m$ is commonly known as the heterogeneity index, which can range from zero (completely heterogeneous; each binding site is unique) to one (homogeneous) [1]. The Freundlich equation can be rewritten:

$$\log B = m \log F + \log a$$  \hspace{1cm} \text{eq. 3.11}

Freundlich models are commonly applied by plotting the experimental binding data in terms of $\log(B)$ vs. $\log(F)$. Using the log-log plot, systems that fit the Freundlich model will form straight lines with a slope of $m$ and a y-intercept of $\log(a)$ (Figure 3.6). Modeling the isotherm as a linear function necessitates fewer experimental data points to accurately define the trend than a curved function.

![Figure 3.6: A log-log plot of the Freundlich isotherm produces a straight line with a slope of $m$ and a y-intercept of $\log(a)$.](image)

Figure 3.6: A log-log plot of the Freundlich isotherm produces a straight line with a slope of $m$ and a y-intercept of $\log(a)$. 
To extract the AD from the Freundlich plot, assumptions must be made regarding the contributions to the $a$ term, and several methods have arisen to correlate assumptions with experimental data [5, 12]. One of the more popular equations for the AD using the Freundlich model parameters is described:

$$N(K_i) = a \frac{\sin(\pi m)}{\pi} K_i^{-m}$$

eq. 3.12

The shape of an AD derived from the Freundlich isotherm is generally an exponentially decaying function with respect to $\log(K)$. Figure 3.7 shows the Freundlich fit AD of a MIP imprinted with (S)-cyclohexyl(phenyl)amine using the above equation [13].

![Figure 3.7: Affinity Distribution of (S)-cyclohexyl(phenyl)amine imprinted MIP based on Freundlich isotherm [13]](image)

One of the major limitations of the Freundlich isotherm is that, because it is incapable of modeling saturation behavior, it can only accurately represent a portion of the binding isotherm. Therefore, deviations become increasingly problematic at higher concentrations.
3.1.2.2 Langmuir-Freundlich Isotherm

The Langmuir-Freundlich model is a hybrid approach that can accommodate both the saturation and subsaturation regions [5]. The general formula for the Langmuir-Freundlich isotherm is described:

$$B = \frac{N_T K_0^m F^m}{1 + K_0^m F^m}$$  \hspace{1cm} \text{eq. 3.13}

where $N_T$ is the total number of binding sites and $K_0$ is the overall average affinity constant. The heterogeneity index $m$ is identical to the Freundlich isotherm parameter. Deriving the AD function then becomes:

$$N(K) = 2.3 N_T m K_0^m K^{-m} \times \frac{(1+2K_0^m K^{-m}+K_0^{2m} K^{-2m}+4K_0^m K^{-m} K^{-2m} m^2-K_0^{2m} K^{-2m} m^2)}{(1+K_0^m K^{-m})^4}$$  \hspace{1cm} \text{eq. 3.14}

This hybrid model can accurately fit many MIP binding isotherms over a wide range of concentrations. However, like the Freundlich isotherm, many assumptions about the system must be made in order to produce the AD. Many distributions could theoretically be derived from the same binding data.

3.1.3 \textit{In Silico} Heterogeneity

The difficulties in accurately producing values for $N$ and $K$ without making assumptions about the MIP system has led to many competing theories on the nature of heterogeneous binding. It is difficult to draw conclusions about the specifics of heterogeneity from binding data that stems from an unknown distribution of binding sites. Below we outline \textit{in silico} methods for characterizing binding site heterogeneity.

Because of the high computational expense, QM models of MIPs to date have been limited to one-target systems, determining the theoretical binding energy of a single target with the monomer species. These one-target QM models have been applied to selecting a specific monomer
suitable for the target [14–17], approximating optimal monomer-to-target ratio [18], and predicting which solvent porogen will result in a stronger imprint [19]. However, these systems do not capture many of the molecular interactions that would be seen in larger models and real MIPs, including monomer – monomer and target – target interactions, which contribute to the heterogeneity. Because a one-target system forms only one binding site, there is limited opportunity to extract information on binding site heterogeneity. Interestingly, MD investigation of MIP heterogeneity are less explored. Karlsson et al. utilized MD methods to evaluate heterogeneity, but they did so by comparing separate one-target simulations rather than by simulating a system containing multiple targets [20].

Using methods designed to lessen the computational expense of QM energy minimizations, we hypothesized that larger systems containing multiple target molecules and more monomer units could be optimized. This would allow for the complexities of the pre-polymerization system to be more fully represented. Lessening the computational expense would also create an opportunity for binding-site heterogeneity analyses while maintaining QM accuracy. From there, a MM docking technique could be implemented to quantify the selectivity of these binding sites. Together, these approaches would help predict the effectiveness of a proposed MIP formulation in silico, reducing the amount of experimental screening required to optimize a particular MIP fabrication protocol.

Using the MOZYME approach, where the self-consistent field (SCF) procedure is replaced with a localized molecular orbital (LMO) method, the QM computational expense can be significantly reduced [21]. Applying the MOZYME approach allows for MIP systems consisting of five targets and up to 30 monomers to be geometrically optimized in less than an hour on a commercial laptop computer. Further, a MM force field-based molecular docking technique was used to evaluate the interaction energies of the binding sites formed in these QM-optimized MIPs.
Molecular docking is a computational strategy for determining binding modes and affinities between receptors and ligands, playing an important role in modern rational drug screening [22]. Using the MIP binding site as the receptor and comparing the “docking” of the target into each of the five binding sites allowed us to quantify the heterogeneity of these sites. With docking, the low computational expense allows for rapid orientation screening and scoring of binding modes, making it a potential technique for MIP heterogeneity analysis [23].

The variations in binding energy between different binding sites within the same QM-optimized system were used to evaluate the heterogeneity. This heterogeneity was also quantified using correlations between the binding site surface area and the corresponding binding energies. In this chapter, we outline an approach for in silico analysis of MIP heterogeneity using QM-optimized systems containing five targets.

3.2 Methods

3.2.1 Binding Site Surface Area Calculations

The five-target systems evaluated in this section are the same as those described in the previous chapter. The number of monomers H-bonding with a target in the pre-polymerization system will directly determine the size and shape of the binding site. Targets that bind with greater surface area contact to the site are likely to have greater binding energies and selectivities. Here we define the Binding Site Surface Area Factor, $F_{BS}$, as the ratio of the surface area of the binding site, $S_{ABS}$, to that of the target molecule, $S_{At}$. For a fully enclosed target molecule, the surface areas of the binding site and the target will be approximately equal; therefore $F_{BS} \approx 1$. For an incomplete binding site at the interface of the MIP and the solvent, $F_{BS} \approx 0$. Those interfacial sites with lower $F_{BS}$ values form weak and non-selective bonds. This concept is shown graphically in
Figure 3.8. The solvent-excluded surface area (SES) was used to define all SA values, with a probe-radius of 1.4Å. The equation derived to calculate $F_{BS}$ is shown below:

$$
F_{BS} = \frac{SA_{M-T} - SA_{M+T} + 1}{2SA_T} = \frac{SA_{BS}}{SA_T}
$$

where $SA_{M-T}$ and $SA_{M+T}$ are the SES values of the larger scale system without and with the target bound to the binding site of interest, respectively.

It should be noted that the contact area between the target and the monomers is not the only factor determining binding energies; the presence and position of the monomer functional groups within this shared space ultimately determines the strength of the bond. It is therefore possible for a binding site with a lower $F_{BS}$ value to outperform one with a higher $F_{BS}$ value if the former contains more complementary interactions. $F_{BS}$ values serve as a simple means of representing the structural aspect of the binding site. Though the interpretations of the $F_{BS}$ values are not straightforward, they can be easily calculated and correlate reasonably well with binding energy and selectivity.

![Figure 3.8](image.png)

Figure 3.8: $F_{BS}$ for internal, intermediate, and interfacial binding sites (left to right, respectively). Dotted lines represent H-bonds.

### 3.3 Results and Discussion

The larger models used here, while still simplifying the pre-polymerization system, allow for more of the expected interactions that are ignored at the small scale, including: monomer
dimerization, target dimerization, and variations between the monomer-to-target ratios per binding site throughout the system. This last effect allows for the existence of the surface binding sites where the targets are not fully surrounded by monomers. As surface sites will be the most accessible to the target, their contributions to the overall MIP behavior should not be ignored. The difference in optimal monomer-to-target ratios between the small and larger model systems may be due to the increased spatial freedom of the monomers. The resulting 3D models show that the number of monomers bound to each target is not consistent; targets in the center of the system will typically have 4-6 monomers in their immediate vicinity, while targets near the periphery of the system may only be interacting with 1-3 monomers. This variation leads to a heterogeneity of binding sites, with each site having a different surface area and binding energy. Binding site heterogeneity has long been observed in experimental MIP analysis [13, 24]. MM investigation of MIP heterogeneity has garnered some attention [25]. However, QM studies of this heterogeneity have, to the best of our knowledge, been less explored.

Table 3.1 shows the $F_{BS}$ values for the five binding sites within each matrix, as well as the binding energies calculated via docking. $F_{BS}$ values correspond to the extent monomers are surrounding the target, and binding site heterogeneity can be attributed to variations in the number of monomers H-bonding to each target in the pre-polymerization matrix. For each simulated MIP, the binding energy correlates strongly with the surface area of the binding site, with the highest energies corresponding to sites with the highest $F_{BS}$ values (in bold) in all but one case, THO site 5. This exception is a case in which the site with the largest $F_{BS}$ value, THO site 2, allowed for fewer and/or weaker H-bonds than site 5, which had the second largest $F_{BS}$. Higher $F_{BS}$ values will typically allow for more H-bond formation. However, it is possible for a site with lower a $F_{BS}$ value to be more functionalized for H-bonding. Likewise, a higher $F_{BS}$ site could have fewer
energetic interactions if the monomers’ hydrophobic groups are more exposed in the site. Normalizing the binding energies to the $F_{BS}$ values (seen in the $BE / F_{BS}$ column in Table 1) illustrates this, where a smaller site may have a greater binding energy per unit surface area. The value of a site’s $BE / F_{BS}$ reflects how strongly the target is interacting with each of the surrounding monomers. In the case of the THO MIP, the $F_{BS}$ value of site 2 was 15% higher than that of site 5, but site 5 had a nearly 60% higher $BE / F_{BS}$ (i.e. higher binding energy per surface area), giving site 5 the highest overall binding energy. Therefore, for a strong and selective site, the adjacent monomers must at least partially surround the target and must also be oriented in such a way that they can form H-bonds with the imprinted molecule.
Table 3.2: $F_{BS}$, $BE$, and $BE / F_{BS}$ values of each of the five binding sites within each larger scale imprinted system. Sites with the highest $F_{BS}$ values and their corresponding $BE$s are in bold.

<table>
<thead>
<tr>
<th>Imprint</th>
<th>Site</th>
<th>$F_{BS}$ (kcal/mol)</th>
<th>$BE$ (kcal/mol)</th>
<th>$BE / F_{BS}$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAF</td>
<td>1</td>
<td>0.539</td>
<td>-2.6</td>
<td>-4.8</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.570</td>
<td>-3.2</td>
<td>-5.6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.559</td>
<td>-2.7</td>
<td>-4.8</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td><strong>0.874</strong></td>
<td><strong>-4.6</strong></td>
<td><strong>-5.3</strong></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.618</td>
<td>-2.2</td>
<td>-3.6</td>
</tr>
<tr>
<td>THO</td>
<td>1</td>
<td>0.530</td>
<td>-3.4</td>
<td>-6.4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><strong>0.649</strong></td>
<td><strong>-2.7</strong></td>
<td><strong>-4.2</strong></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.461</td>
<td>-3.6</td>
<td>-7.8</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.500</td>
<td>-3.1</td>
<td>-6.2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td><strong>0.563</strong></td>
<td><strong>-3.8</strong></td>
<td><strong>-6.7</strong></td>
</tr>
<tr>
<td>TB</td>
<td>1</td>
<td>0.545</td>
<td>-3.5</td>
<td>-6.4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.671</td>
<td>-4.0</td>
<td>-6.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.630</td>
<td>-2.8</td>
<td>-4.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td><strong>0.779</strong></td>
<td><strong>-4.8</strong></td>
<td><strong>-6.2</strong></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.543</td>
<td>-3.5</td>
<td>-6.4</td>
</tr>
<tr>
<td>HA</td>
<td>1</td>
<td><strong>0.802</strong></td>
<td><strong>-3.8</strong></td>
<td><strong>-4.7</strong></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.602</td>
<td>-3.0</td>
<td>-5.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.497</td>
<td>-3.2</td>
<td>-6.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.635</td>
<td>-3.5</td>
<td>-5.5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.560</td>
<td>-3.3</td>
<td>-5.9</td>
</tr>
<tr>
<td>L-H</td>
<td>1</td>
<td>0.583</td>
<td>-3.5</td>
<td>-6.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.569</td>
<td>-3.5</td>
<td>-6.2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><strong>0.801</strong></td>
<td><strong>-4.3</strong></td>
<td><strong>-5.4</strong></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.756</td>
<td>-4.0</td>
<td>-5.3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.675</td>
<td>-4.0</td>
<td>-5.9</td>
</tr>
<tr>
<td>D-H</td>
<td>1</td>
<td>0.464</td>
<td>-3.4</td>
<td>-7.3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><strong>0.726</strong></td>
<td><strong>-4.4</strong></td>
<td><strong>-6.1</strong></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.594</td>
<td>-3.4</td>
<td>-5.7</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.659</td>
<td>-3.9</td>
<td>-5.9</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.573</td>
<td>-2.5</td>
<td>-4.4</td>
</tr>
</tbody>
</table>

It is common to mathematically describe MIPs with two discrete binding affinities, a strong site affinity and a weak site affinity, and fit the data accordingly. The most common of these methods, the bi-Langmuir isotherm, relies on the assumption that $KN(\text{strong}) \gg KN(\text{weak})$, where
\( K \) and \( N \) are the binding affinity and number of binding sites, respectively [26]. Experimentally however, the bi-modal binding model does not fit well [1, 24, 27]. The modeling data presented in Table 3.1 capture these experimental results, indicating no bi-modal discretization of binding sites, but rather a continuous range. Thus, continuous affinity distribution models, such as the Freundlich and Langmuir-Freundlich isotherms, seem more representative models for MIP binding behavior [5, 27]. The continuous distribution can be attributed to orientation variability of the monomer functional groups. If the number of adjacent monomers were the only important factor, one would still expect a discrete number of possible binding affinities (one for each possible monomer-to-target ratio until the target is completely surrounded). However, because the many possible positions and orientations of the monomers and their H-bonding groups, binding energies can instead be expected to show continuous variance between weak, non-selective sites and idealized, optimized sites.

### 3.4 Conclusions

In this study, we demonstrated a QM approach for larger scale MIP modeling involving five target molecules and 25 monomer species. These models also allowed for an analysis of MIP binding site heterogeneity, using a comparison of binding site surface areas to rebinding energies calculated with a MM docking approach. We found that the interaction area of the target with the binding site, quantified here by the \( F_{BS} \) value, correlates with both the \( BE \) and \( \alpha \) values of the site. As expected, sites with an \( F_{BS} \) of 1 were shown to be the strongest and most selective. However, the \( F_{BS} \) value is not the only indicator determining the properties of a binding site; the possible orientations of monomer functional groups in relation to the target within the binding site can significantly affect the site’s selective binding capabilities. The re-binding models used here show
a continuous range of binding energies, and this heterogeneity supports the use of continuous distribution models rather than bi-modal models for experimental MIP binding analysis. The methods and results presented in this chapter were published in the Journal of Molecular Modeling [28].

3.5 References


CHAPTER 4

Docking Selectivity of QM-Optimized Binding Sites

4.1 Introduction to MIP Selectivity

The selectivity factor, $\alpha$, is defined as the ratio of binding constants between the target molecule and an analogue, typically a molecule similar to the target in size and shape. Figure 4.1 below diagrams the formation of an imprinted binding site with a selective preference for the target over an analogue. In our previous work, we used quantum mechanics (QM) and molecular mechanics (MM) modeling techniques to evaluate the ability of simulated MIP binding sites to select between molecules of different sizes, shapes, and polarities [1]. However, the differences in polarity, hydrophobicity, and ionization states between the targets and the analogues complicate the analysis of what exactly is causing the imprinted selectivity [2]. It has therefore been proposed that enantioselectivity, the ability to select an imprinted chiral molecule over its enantiomer, is the best probe to evaluate the imprinting effect [3]. In this case, the only difference between the target and the analogue is the topological orientation of the functional groups in three-dimensional space. The enantioselectivity factor, $\alpha_E$, the ratio of binding constants between an imprinted chiral molecule and its enantiomer, provides a clear metric when analyzing and optimizing the MIP system. The enantioselective properties of non-covalent MIPs have long been the subject of experimental investigation [4–10], and more recent studies have computationally probed orientations of monomers complexing with chiral targets [11]. However, to the best of our knowledge, this work represents a novel in silico method for molecular imprinting and analysis of
enantioselective binding sites by investigating the factors contributing to the selective nature of the fully formed binding pocket.

Figure 4.1: Diagram of the imprinting of a target molecule, forming a selective binding site. Though the analogue molecule may be able to bind with the receptor, $BE_{\text{Target}} > BE_{\text{Analogue}}$ (hence the sizes of the arrows).

A MM force field-based molecular docking technique was used to evaluate the interaction energies of the binding sites formed in the QM-optimized MIPs as discussed in Chapters 2 and 3. Molecular docking is a computational strategy for determining binding modes and affinities
between receptors and ligands, playing an important role in modern rational drug screening [12]. Using the MIP binding site as the receptor and comparing the “docking” of the target with structurally similar molecules allowed us to quantify these selectivities. QM-based selectivity models have proven useful [13], but the QM selectivity optimization’s dependence on the user-selected starting orientation drastically limits the effectiveness of the technique [14]. In contrast with the initial QM optimization of the target and monomers, in a selectivity analysis, the monomers are locked in place and can therefore no longer be rationally oriented to form H-bonds with the analogue molecule. Due to the limited range of movement within QM geometry optimizations, locking the monomer geometries places the burden of finding a suitable starting orientation for the analogue on the user. With docking, the lower computational expense allows for rapid orientation screening and scoring of binding modes, making it a promising technique for MIP selectivity analysis [15].

Finally, taking advantage of quantum mechanics (QM) and molecular mechanics (MM) methods, we developed a means to probe the enantioselectivity of MIP binding sites in silico. The chiral target molecules were selected for their relevance in pre-existing experimental enantioselectivity studies [4, 16]. Surrounding chiral targets with monomers and then energetically minimizing the system using a two-step approach, whereby a MM-optimized geometry was used as the starting orientation for a QM-optimization (MM \(\rightarrow\) QM), created “computationally imprinted” binding sites. A MM docking technique was then used to compare the binding energies of the imprinted target with its enantiomer. Molecular docking has recently gained attention in MIP research [1, 11, 15, 17, 18].

The influence of individual H-bonding functional groups on the selectivity was determined by individually neutralizing the monomers that formed H-bonds with the target or analogue during
docking. Further, cross-enantioselectivity, the ability for a binding site imprinted with a chiral molecule to select between the chiral forms of molecules structurally similar to the imprinted target, was also studied using a similar docking technique. In this work, we outline an approach for computational imprinting and characterization of enantioselective binding sites. The positioning and chemical make-up of H-bonding functional groups within both the binding sites and targets were studied in order to better understand the properties contributing to the chirality-based binding preferences.

4.2 Methods

4.2.1 Size, Shape, and Polarity-Based Selectivity Modeling

The larger scale geometrically optimized models described in Chapter 2 were investigated for selectivity (See Section 2.2.1 for Components). Using the optimized structure files from the models with the monomer-to-target ratios corresponding to the strongest binding energies, the target molecules were removed, leaving behind the imprinted binding sites. Substituting structural analogues into the binding sites and scoring the binding modes with AutoDock Vina (a molecular docking program) [19] gives a direct comparison between the binding energies. AutoDock uses the AMBER force field, which has a proven record with proteins, nucleic acids, and other organics [20]. Grid boxes were centered in the binding site, with dimensions of 11x11x11Å, and an exhaustiveness of 8 was used to allow for sufficient sampling of docking orientations. Here, the selectivity factor, $\alpha$, is defined as the ratio between the binding energies of the target and the analogue. Within the larger scale models, selectivities were computed for each of the system’s five binding sites. The selectivity was compared to the $F_{BS}$ values of each site, showing how binding site heterogeneity influences the overall selectivity of the MIP.
In order to investigate the selectivity of a fully encased, idealized binding site ($F_{BS}\approx1$), models consisting of one target and 20 monomers were optimized with the PM6-DH2 basis set. Saturating the model with an excess of monomers ensured the optimized geometry would include one fully internal binding site. The docking method described above was then used to compute the selectivity of these sites.

### 4.2.2 Enantioselectivity Modeling

**4.2.2.1 Chemical Components**

Three of the selected targets were tyrosine derivatives, chosen for the location of the chiral center, as well as their history of experimental study in MIP literature: R/S-tert-butoxycarbonyl-tyrosine (BOC-Tyr), R/S-carbobenzoxy-tyrosine (CBZ-Tyr), and R/S-acetyl-tyrosine (Acetyl-Tyr) [4, 16]. The other two targets, R/S-2-chloro-1-phenylethanol (ClPhEthOH) and R/S-carbobenzoxoxy-serine (CBZ-Ser), were chosen to represent molecules that did not exhibit experimental cross-enantioselectivity in the BOC-Tyr imprinted polymer studies of Meador and Spivak, 2014 [16]. Methacrylic acid (MAA) was chosen as the monomer for the studies based on the strong literature precedent. The structure of each component used is shown in Figure 4.2, with the chiral centers circled in red.
4.2.2 Enantioselectivity

Avogadro, a cross platform molecular editor (Avogadro.cc) [22, 23] was used to generate the binding site, where the target molecule was placed into a box containing 30 MAA molecules. The quantity of MAAs was selected because it allowed for the target molecules to be surrounded by two layers of monomers. The first layer consists of monomers interacting with the target via H-bonding or hydrophobic forces. The second layer of monomers interacts with the first layer in ways that can influence the monomer – target interactions of the first layer, creating a more realistic representation of the pre-polymerization matrix. Further, the optimizations of systems containing 30 monomers and one target carried a significant but not unreasonable computational expense, with run times of 4-8 hours. The binding site geometry was energetically optimized in a two-step MM \( \rightarrow \) QM scheme. First, a MM steepest descent optimization with the MMFF94 force field consisting of 500 steps and a convergence of 10E-7 was used to obtain rough starting orientations for the QM step. The MM optimized geometries were then further refined using a semi-empirical QM method with MOPAC [24] and the PM7 basis set. An exit gradient of 1.0 kcal/mol/Å was
incorporated in the QM optimization. This multiscale “coarse \( \rightarrow \) fine” technique minimizes the disadvantages of using either MM or QM separately, while emphasizing their advantages. MM optimizations are much less computationally expensive, and are therefore orders of magnitude faster than QM methods. However, the accuracy of the MM results is limited by the simplified mathematical models that allow for this reduced computational expense. While more accurate, the computational expense of QM models introduces the problem of users selecting a suitable starting orientation for interacting molecules. Feeding the MM optimized geometry into the QM optimization reduces the impact of the user-defined starting orientation and drastically lowers the computing time. After optimizing the system, the imprinted target molecule was deleted from the QM optimized geometry, leaving behind an “imprinted” binding site. AutoDock Vina, a molecular docking program [19], was used to probe the enantioselectivity of the computationally imprinted binding sites. AutoDock uses the AMBER force field, which has been proven to accurately simulate the interactions of proteins, nucleic acids, and other organic molecules [20]. Docking of the imprinted molecule and its opposite chiral form allowed for a measure of the enantioselectivity of the binding site. Grid boxes were centered in the binding site and sized to fit within the MAA matrix. In order to fit the oblong binding sites, grid box dimensions of 16x12x12Å were used. An exhaustiveness of 8 allowed for sufficient sampling of docking orientations [25]. In order to emphasize the role of H-bonds in the binding site, the weight of H-bonds in the scoring algorithm was increased two-fold, from -0.6 to -1.2 [18, 19]. The enantioselectivity factor, \( \alpha_E \), is defined here as the ratio of the binding energy (\( BE \)) of the target to that of its enantiomer obtained from docking. An \( \alpha_E \) value of >1 represents a preference for the imprinted molecule over the opposite chirality. Systems imprinted with both enantiomers of BOC-Tyr (R-BOC-Tyr and S-BOC-Tyr), as well as with both enantiomers of CBZ (R-CBZ-Tyr and S-CBZ-Tyr) were investigated.
4.2.3 Cross-Enantioselectivity Modeling

Cross-enantioselectivity is defined here as the ability of a site imprinted with a chiral target to distinguish between enantiomers of other structurally similar chiral molecules. The ability of R-BOC-Tyr and S-BOC-Tyr imprinted polymers to select for the R- and S-forms of other tyrosine derivatives has been shown experimentally [16]. Using the same R- and S-BOC optimized binding sites and docking procedure described above, both forms of CBZ-Tyr, Acetyl-Tyr, ClPhEthOH, and CBZ-Ser were docked into each BOC-Tyr imprinted site. The cross-enantioselectivity factor, $\alpha_{C-E}$, is defined as the ratio of the $BE$ of the enantiomer matching the chirality of the imprinted BOC-Tyr to that of the opposite enantiomer into the same site: i.e. the $BE$ of R-CBZ-Tyr docked into the R-BOC-Tyr imprinted site, divided by the $BE$ of S-CBZ-Tyr docked into the R-BOC-Tyr imprinted site. A $\alpha_{C-E}$ value of >1 represents a preference for the enantiomer with the same R/S handedness as the originally imprinted molecule.

4.2.4 H-bond Neutralization

In order to investigate the H-bond functionalization of the binding site, the individual bond energies between the functional (carboxylic acid) groups of the monomers and the H-bonding groups of the targets were computed. Four MAA molecules were found to form H-bonds with such functional groups of R/S-BOC-Tyr during the geometry optimizations. Henceforth, the term “functional group” will refer to the four components of the target participating in H-bonding interactions with the surrounding monomers, unless otherwise stated. These four MAA – target arrangements are shown in Figure 4.3. The descriptions, as well the abbreviations for the arrangements that will be used henceforth, are as follows: (a) MAA=COOH, a MAA forms two H-bonds with the carboxylic acid group; (b) MAA=NH+O1, a MAA forms two H-bonds with the
single-bonded oxygen and the secondary amine’s hydrogen in the urethane group; (c) MAA-O2, a MAA forms one H-bond with the double-bonded oxygen in the urethane group; (d) MAA=PhenolOH, a MAA forms two H-bonds with the alcohol of the phenol group. Monomer – monomer dimerization energies were also calculated: (e) MAA-MAA, a MAA forms one H-bond with a second MAA; (f) MAA=MAA, a MAA forms two H-bonds with a second MAA. It should be noted that while there may be many other energetically favorable MAA – target arrangements than those described here, these four interactions repeatedly formed during the MM → QM optimizations of the targets in boxes of 30 randomly oriented monomers.

![Figure 4.3: Monomer – target interactions: (a) MAA=COOH; (b) MAA=NH+O1; (c) MAA-O2; (d) MAA=PhenolOH. Monomer dimerizations: (e) MAA-MAA; (f) MAA=MAA](image)

First, Avogadro was used to build each of the four monomer – target arrangements that were found in the QM optimizations of the systems containing 30 MAA molecules and R/S-BOC-Tyr, as well as the two monomer – monomer dimerization arrangements. The carboxylic acid group of the monomer was oriented toward the various functional groups of the target. A two-step MM → QM geometry optimization, similar to the one described in the Enantioselectivity section
above, was then used to compute the BEs of each functional group pairing. The QM steps were computed in vacuum with an exit gradient of 0.01 kcal/mol/Å, which is appropriate for high-precision work with small systems [26]. All BEs were calculated:

\[
BE = \left[ \Delta H_f(target + monomer) - \Delta H_f(target) - \Delta H_f(monomer) \right] \quad \text{eq. 4.1}
\]

where \( \Delta H_f \) is the total heat of formation. In the case of the monomer dimerizations, the BE was calculated as the difference in \( \Delta H_f \) of the bonded and unbonded monomers. The four MAA – target interactions were calculated for BOC-Tyr, as well as CBZ-Tyr and Acetyl-Tyr. CBZ-Tyr and Acetyl-Tyr contain identical orientations of the functional groups that participated in BOC-Tyr’s H-bonding, and would therefore be expected to have similar BE values. Furthermore, the optimizations of 30 MAA systems imprinted with R/S-CBZ-Tyr showed the same four MAA – target H-bonding arrangements (Figure 4.4).

Figure 4.4: Two views of the MAA – target H-bonding arrangements after MM \( \rightarrow \) QM optimization of 30 MAAs and S-CBZ-Tyr. All but the H-bonding MAAs are hidden to simplify the image. The benzyl group of CBZ-Tyr is also hidden for the same reason. On the left view, the MAA=NH+O1 interaction (bottom, highlighted in blue) and the MAA-O2 interaction (top, highlighted blue) can be seen. On the right view, the MAA=COOH (leftmost MAA of the view on the right, highlighted in blue) and the MAA=PhenolOH (rightmost MAA highlighted in blue) can be seen. The yellow dashed lines indicate H-bonds. These same four H-bonding arrangements were found after optimized R- and S-BOC-Tyr, as well as R- and S-CBZ-Tyr with 30 MAAs.
To further study the effect of the H-bonds between the binding site and the docking molecules, H-bonding MAAs were individually neutralized prior to docking analysis. The carboxylic acid groups of the MAAs H-bonding with the target after the MM $\rightarrow$ QM optimization were converted to methyl groups, which eliminates the ability to form H-bonds without significantly altering the steric contributions to the site. The changes in $BE$ and $\alpha_E$ values were calculated by docking the imprinted molecule and its enantiomer into the binding sites after neutralizing H-bonding MAAs one at a time. The overall effect of the H-bonding MAAs was calculated by docking the enantiomers into binding sites after neutralizing all four of the contributing monomers.

4.3 Results and Discussion

4.3.1 Size, Shape, and Polarity-Based Selectivity

Table 4.1 summarizes the binding site selectivity data gathered through molecular docking. The $\alpha$ values represent the binding site’s preference for the imprinted molecule relative to the analogue species, and is calculated:

$$\alpha = \frac{BE_{Target}}{BE_{Analogue}}$$  

Selectivities greater than 1 indicate a positive preference for the imprinted molecule, whereas a selectivity less than one indicates a preference for the analogue. As predicted, binding sites with greater $F_{BS}$ values have greater preference for their imprinted targets relative to the weaker, less embedded binding sites. The data also show consistent similarities with the selectivities for or against molecules of the same group (imidazole or xanthine). This indicates shape-based selection. The poor selectivity of CAF imprinted binding sites is likely size related. As CAF is the largest of the molecules imprinted, these larger binding sites allow invading
analogues the freedom to position themselves into more energetically favorable binding orientations. THO and TB contain the same polar nitrogen and oxygen groups in the same positions as CAF, but having one fewer methyl group makes for tighter, and therefore more exclusive, binding sites. Enantioselectivity, the preference of D-H over L-H in D-H imprinted sites and vice-versa, was not observed in these larger docking models. The overall weak selectivities of the binding sites formed in the five-target optimizations can be attributed to the open geometry and the frequent overlapping with other binding sites, minimizing the intended ability for the sites to exclude invasive analogues based on their sizes and shapes.

To further study the selective properties of an idealized binding pocket, our focus shifted to fully enclosed sites. As described in the methods section, these 20:1 MAA-to-target models did not allow for target hybridization or interfacial binding sites to form, creating one site with \( F_{BS} \approx 1 \). (Note: Actual \( F_{BS} \) values could not be calculated because the complete internalization of the site allows it to “hide” from the SES probe algorithm.) The results of these studies are shown in Table 4.2. As expected, these \( F_{BS} \approx 1 \) sites are much more selective than the more open and incomplete binding sites of the 25:5 MAA-to-target models. Interestingly, while the CAF imprinted site cannot select against the other xanthine-derived molecules (for reasons described above), it can now selectively exclude the imidazole group molecules to some extent. Because CAF is the largest of the molecules, this exclusion cannot be purely size-based. This indicates a shape-based preference for the xanthine group. The selectivities of L-H and D-H imprinted sites over HA are also telling. HA could not be excluded from these sites by size or shape. The selection of the sites against HA is therefore attributed to polarity-based preferences; HA lacks the carboxylic acid group found in L-H and D-H that allows them to form another H-bond with the cavity. All imidazole group molecules selectively screen against xanthine group molecules based on size, with \( \alpha \)-values of at
least 2.5. The smallest imprint, HA, completely excludes CAF; the larger molecule could not find any energetically favorable positions within the HA-imprinted site. It should be noted that *all* binding sites showed a positive-preference toward the imprinted molecule relative to the analogues, with the one exception being the L-H imprinted site’s slight preference for D-H. These data points represent successful modeling of molecularly imprinted selectivity using a novel QM-optimized MM docking method. It should be noted, however, that binding sites so fully encapsulated by the monomers would not be accessible for target extraction and rebinding. This study serves to demonstrate the ability of docking analysis to show size-, shape-, and polarity-based complementarity, quickly and with little computational expense.
Table 4.1: $F_{BS}$ of each binding site and $\alpha$ relative to other species. Sites with the highest $F_{BS}$ values are shown in bold.

<table>
<thead>
<tr>
<th>Imprint</th>
<th>Site</th>
<th>$F_{BS}$</th>
<th>CAF</th>
<th>THO</th>
<th>TB</th>
<th>HA</th>
<th>L-H</th>
<th>D-H</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAF</td>
<td>1</td>
<td>0.539</td>
<td>0.867</td>
<td>0.788</td>
<td>0.897</td>
<td>0.788</td>
<td>0.765</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.570</td>
<td>1.000</td>
<td>0.941</td>
<td>0.970</td>
<td>0.941</td>
<td>0.842</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.559</td>
<td>0.931</td>
<td>0.964</td>
<td>0.692</td>
<td>0.771</td>
<td>0.771</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.874</td>
<td>0.885</td>
<td>0.902</td>
<td>1.179</td>
<td>0.902</td>
<td>0.920</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.618</td>
<td>0.846</td>
<td>0.710</td>
<td>0.667</td>
<td>0.786</td>
<td>0.688</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.530</td>
<td>1.097</td>
<td>1.030</td>
<td>1.214</td>
<td>0.944</td>
<td>1.063</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.649</td>
<td>1.038</td>
<td>1.038</td>
<td>1.000</td>
<td>0.844</td>
<td>1.038</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.461</td>
<td>1.161</td>
<td>1.000</td>
<td>1.333</td>
<td>1.000</td>
<td>1.200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.500</td>
<td>1.033</td>
<td>1.348</td>
<td>1.069</td>
<td>0.861</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.563</td>
<td>1.056</td>
<td>0.905</td>
<td>1.267</td>
<td>1.118</td>
<td>1.188</td>
<td></td>
</tr>
<tr>
<td>THO</td>
<td>1</td>
<td>0.545</td>
<td>1.094</td>
<td>1.000</td>
<td>1.207</td>
<td>1.129</td>
<td>1.094</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.671</td>
<td>1.081</td>
<td>1.111</td>
<td>1.333</td>
<td>0.976</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.630</td>
<td>1.273</td>
<td>0.933</td>
<td>1.120</td>
<td>0.848</td>
<td>0.848</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.779</td>
<td>1.171</td>
<td>1.000</td>
<td>1.231</td>
<td>0.980</td>
<td>1.043</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.543</td>
<td>1.029</td>
<td>0.833</td>
<td>0.897</td>
<td>0.897</td>
<td>0.875</td>
<td></td>
</tr>
<tr>
<td>TB</td>
<td>1</td>
<td>0.802</td>
<td>1.652</td>
<td>1.027</td>
<td>1.267</td>
<td>1.000</td>
<td>0.927</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.602</td>
<td>0.882</td>
<td>0.833</td>
<td>0.789</td>
<td>0.833</td>
<td>0.732</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.497</td>
<td>1.067</td>
<td>0.914</td>
<td>0.941</td>
<td>0.889</td>
<td>0.821</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.635</td>
<td>0.921</td>
<td>0.897</td>
<td>0.875</td>
<td>0.972</td>
<td>0.946</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.560</td>
<td>0.917</td>
<td>0.846</td>
<td>0.825</td>
<td>0.868</td>
<td>0.868</td>
<td></td>
</tr>
<tr>
<td>HA</td>
<td>1</td>
<td>0.583</td>
<td>1.029</td>
<td>0.875</td>
<td>1.000</td>
<td>1.129</td>
<td>1.029</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.569</td>
<td>1.061</td>
<td>0.921</td>
<td>0.897</td>
<td>1.207</td>
<td>1.094</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.801</td>
<td>1.024</td>
<td>0.935</td>
<td>0.811</td>
<td>1.194</td>
<td>0.956</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.756</td>
<td>1.053</td>
<td>0.976</td>
<td>0.976</td>
<td>1.250</td>
<td>0.976</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.675</td>
<td>1.053</td>
<td>0.952</td>
<td>0.976</td>
<td>1.212</td>
<td>0.976</td>
<td></td>
</tr>
<tr>
<td>L-H</td>
<td>1</td>
<td>0.464</td>
<td>1.063</td>
<td>0.919</td>
<td>0.872</td>
<td>1.172</td>
<td>1.097</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.726</td>
<td>1.048</td>
<td>0.936</td>
<td>0.863</td>
<td>1.222</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.594</td>
<td>1.030</td>
<td>0.850</td>
<td>0.971</td>
<td>1.097</td>
<td>1.030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.659</td>
<td>1.026</td>
<td>0.951</td>
<td>0.975</td>
<td>1.147</td>
<td>0.951</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.573</td>
<td>1.000</td>
<td>0.833</td>
<td>0.926</td>
<td>1.136</td>
<td>1.042</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2: $\alpha$ values of 20:1 MAA–Target imprinted sites

<table>
<thead>
<tr>
<th>Imprint</th>
<th>CAF</th>
<th>THO</th>
<th>TB</th>
<th>HA</th>
<th>L-H</th>
<th>D-H</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAF</td>
<td>0.967</td>
<td>0.937</td>
<td>1.372</td>
<td>1.180</td>
<td>1.073</td>
<td></td>
</tr>
<tr>
<td>THO</td>
<td>2.875</td>
<td>1.394</td>
<td>1.045</td>
<td>1.070</td>
<td>1.179</td>
<td></td>
</tr>
<tr>
<td>TB</td>
<td>3.933</td>
<td>1.093</td>
<td>1.341</td>
<td>1.113</td>
<td>1.229</td>
<td></td>
</tr>
<tr>
<td>HA</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>1.400</td>
<td>1.355</td>
<td></td>
</tr>
<tr>
<td>L-H</td>
<td>12.500</td>
<td>2.778</td>
<td>2.500</td>
<td>1.190</td>
<td>0.980</td>
<td></td>
</tr>
<tr>
<td>D-H</td>
<td>17.000</td>
<td>3.000</td>
<td>2.684</td>
<td>1.214</td>
<td>1.041</td>
<td></td>
</tr>
</tbody>
</table>

4.3.2 Enantioselectivity

The $BE$ and $\alpha_E$ values of R- and S-BOC-Tyr docking into R- and S-BOC-Tyr imprinted sites are shown in Table 4.3, with R- and S-CBZ-Tyr docking values into R- and S-CBZ-Tyr imprinted sites in Table 4.4. It can be seen that all of the binding sites exhibit enantioselectivity; all $\alpha_E$ values are greater than 1, indicating a preference for the imprinted molecule over its enantiomer. R-BOC-Tyr bonded to the R-imprinted site 9.7% more strongly than its enantiomer, and S-BOC-Tyr bonded to the S-imprinted site 10.9% more strongly than R-BOC-Tyr. The R-CBZ-Tyr imprinted site bonded to the R-CBZ-Tyr 4.7% more strongly than the S-form, while the S-CBZ-Tyr imprinted site bonded to S-CBZ-Tyr 5.7% more strongly than the R-form. These values were calculated as the differences in the $BE$s between the target and its enantiomer, divided by the $BE$ of the enantiomer and multiplied by 100. When docking the imprinted enantiomer, the highest scoring conformation consistently aligned well with the target geometry from the MM $\rightarrow$ QM optimization. Made with AutoDock Tools [27], Figure 4.5 shows an overlay of the docking geometry of R-BOC-Tyr (blue) and the MM $\rightarrow$ QM-optimized geometry (pink). The overlay of the MM $\rightarrow$ QM optimized conformation and the docking conformation for S-BOC-Tyr, as well as those for R- and S-CBZ-Tyr, are shown in Figures 4.6, 4.7, and 4.8, respectively. The differences in the $BE$ and $\alpha_E$ values between R- and S-imprinted sites indicate a variance between the two sites that extends beyond chirality. Though both sites included the same four arrangements of H-
bonding MAAs, the exact positioning of these MAAs, as well as the other non-bonding monomers making up the remaining surface of the site, differs between sites. This means that the R- and S-imprinted sites are not perfect mirror images of each other. There is a degree of randomness to the starting orientation of the system that will produce differing final geometries. One would expect that if numerous optimizations were to be computed with large numbers of starting orientations, the average $BE$ and $\alpha_E$ values between the R- and S-imprinted sites would converge. However, given the computational costs, we were satisfied with the results of one minimization for each target system serving as a proof-of-principle enantioselectivity model.

Table 4.3: $BE$ and $\alpha_E$ values of R/S-BOC-Tyr docked into BOC-Tyr imprinted sites

<table>
<thead>
<tr>
<th>Ligand</th>
<th>R-BOC-Tyr</th>
<th>S-BOC-Tyr</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-BOC-Tyr</td>
<td>-10.2</td>
<td>-9.2</td>
</tr>
<tr>
<td>S-BOC-Tyr</td>
<td>-9.3</td>
<td>-10.2</td>
</tr>
<tr>
<td>$\alpha_E$</td>
<td>1.10</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Table 4.4: $BE$ and $\alpha_E$ values of R/S-CBZ-Tyr docked into CBZ-Tyr imprinted sites

<table>
<thead>
<tr>
<th>Ligand</th>
<th>R-CBZ-Tyr</th>
<th>S-CBZ-Tyr</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-CBZ-Tyr</td>
<td>-9.0</td>
<td>-8.4</td>
</tr>
<tr>
<td>S-CBZ-Tyr</td>
<td>-8.6</td>
<td>-8.9</td>
</tr>
<tr>
<td>$\alpha_E$</td>
<td>1.05</td>
<td>1.06</td>
</tr>
</tbody>
</table>
Figure 4.5: Highest scoring docking conformation of R-BOC-Tyr (blue) overlaid with the MM $\rightarrow$ QM-optimized geometry (pink), as displayed with AutoDock Tools. The figure on the right is a 90-degree rotated view of the binding site. Chiral centers are indicated with arrows matching the color of the molecule to which they are pointing.

Figure 4.6: S-BOC-Tyr conformations after MM $\rightarrow$ QM optimization (pink) and docking (green). MAAs not shown
Figure 4.7: R-CBZ-Tyr conformations after MM $\rightarrow$ QM optimization (pink) and docking (blue). MAAs not shown

Figure 4.8: S-CBZ-Tyr conformations after MM $\rightarrow$ QM optimization (pink) and docking (blue). MAAs not shown
4.3.3 Cross-Enantioselectivity

The results of the cross-enantioselectivity docking studies are summarized in Table 4.5. These data show that the R- and S-BOC-Tyr imprinted sites exhibit a cross-enantioselective preference for the R- and S-forms of the structurally similar tyrosine derivatives, CBZ-Tyr and Acetyl-Tyr. CBZ-Tyr and Acetyl-Tyr differ from BOC-Tyr only in the replacement of one non-polar non-bonding group. In the place of BOC-Tyr’s tert-butyl group, CBZ-Tyr has a benzyl group and Acetyl-Tyr has only a methyl group. Interestingly, the BOC-Tyr imprinted sites show greater selection between R- and S-CBZ-Tyr than the CBZ-Tyr imprinted sites. This is likely due to the size differences in the molecules. The solvent-excluded surface area (SES) of BOC-Tyr is approximately 9.6% smaller than that of CBZ-Tyr (SES values were calculated with a probe-radius of 1.4 Å). Therefore, the BOC-Tyr imprinted site will be proportionally smaller than the CBZ-Tyr imprinted site, giving the opposite enantiomer of CBZ-Tyr less “freedom” to find energetically favorable orientations. Figure 4.9 shows the highest scoring docking orientations of R-CBZ-Tyr (top, green) and S-CBZ-Tyr (bottom, blue) into the R-BOC-Tyr imprinted site, overlaid with the MM → QM-optimized orientation of R-BOC-Tyr (pink). R-CBZ-Tyr is well-aligned with the orientation of the imprinted molecule. In order to for its functional groups to bond with the respective monomers, S-CBZ-Tyr must rotate in different directions than its enantiomer, forcing the backbone out of line with the original imprint, thus increasing steric hindrance with the site and weakening the $BE$. 
Figure 4.9: R-CBZ-Tyr (top, green) and S-CBZ-Tyr (bottom, blue) docking in binding site optimized with R-BOC-Tyr (pink). Figures on the right side are side-views (rotated 90 degrees) of those on the left side. Chiral centers are indicated with arrows matching the color of the molecule to which they are pointing.
Table 4.5: $BE$ and $\alpha_{C,E}$ values of BOC-Tyr and analogues docked into BOC-Tyr imprinted sites

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Imprinted molecule (BEs in kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-BOC-Tyr</td>
</tr>
<tr>
<td>R-BOC-Tyr</td>
<td>-10.2</td>
</tr>
<tr>
<td>S-BOC-Tyr</td>
<td>-9.3</td>
</tr>
<tr>
<td>$\alpha_{E}$</td>
<td>1.10</td>
</tr>
<tr>
<td>R-CBZ-Tyr</td>
<td>-8.9</td>
</tr>
<tr>
<td>S-CBZ-Tyr</td>
<td>-7.2</td>
</tr>
<tr>
<td>$\alpha_{C,E}$</td>
<td>1.24</td>
</tr>
<tr>
<td>R-Acetyl-Tyr</td>
<td>-9.9</td>
</tr>
<tr>
<td>S-Acetyl-Tyr</td>
<td>-9.0</td>
</tr>
<tr>
<td>$\alpha_{C,E}$</td>
<td>1.10</td>
</tr>
<tr>
<td>R-ClPhEthOH</td>
<td>-6.6</td>
</tr>
<tr>
<td>S-ClPhEthOH</td>
<td>-6.5</td>
</tr>
<tr>
<td>$\alpha_{C,E}$</td>
<td>1.01</td>
</tr>
<tr>
<td>R-CBZ-Ser</td>
<td>-9.0</td>
</tr>
<tr>
<td>S-CBZ-Ser</td>
<td>-9.5</td>
</tr>
<tr>
<td>$\alpha_{C,E}$</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Little to no cross-enantioselectivity is observed for the non-tyrosine molecules, ClPhEthOH and CBZ-Ser. Though chiral, the make-up and positioning of functional groups differs significantly from BOC-Tyr and the other tyrosine derivatives. The tyrosine-specific cross-enantioselectivity seen for BOC-Tyr imprinted sites correlates well the experimental data gathered by Meador and Spivak [16]. The calculated $BE$s correlate with their results, which showed that HPLC retention times of CBZ-Ser were similar to those of the tyrosine derivatives, but without the differentiation between chiralities. The HPLC data also showed that ClPhEthOH had the lowest retention times of all the analogues tested with the R/S-BOC-Tyr MIP columns; similarly, we found ClPhEthOH to have the lowest $BE$ of the molecules used in our docking experiments.
4.3.4 H-bond Neutralization

The QM calculated $BE$s of the selected monomer – monomer and monomer – target interactions are shown in Table 4.6. As expected, the tyrosine derivatives share similar $BE$ trends for the MAA – target interactions. The MAA=COOH interactions have the greatest $BE$s, all of which are stronger than the MAA=MAA dimerization. All other monomer – target bonds, MAA=NH+O1, MAA-O2, and MAA=PhenolOH, are weaker than the MAA=MAA double H-bond but stronger than the MAA-MAA single H-bond. The strength of the monomer – target bonds relative to the monomer – monomer bonds is important, as it determines the energetic favorability of the H-bond formations that will ultimately define the make-up of the binding site.

<table>
<thead>
<tr>
<th>System</th>
<th>BOC-Tyr BE (kcal/mol)</th>
<th>CBZ-Tyr BE (kcal/mol)</th>
<th>Acetyl-Tyr BE (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) MAA=COOH</td>
<td>-19.7330</td>
<td>-18.6889</td>
<td>-18.7057</td>
</tr>
<tr>
<td>(b) MAA=NH+O1</td>
<td>-15.5696</td>
<td>-17.2195</td>
<td>-13.7548</td>
</tr>
<tr>
<td>(c) MAA-O2</td>
<td>-14.4219</td>
<td>-13.0722</td>
<td>-12.1547</td>
</tr>
<tr>
<td>(d) MAA=PhenolOH</td>
<td>-11.7682</td>
<td>-13.1854</td>
<td>-11.7057</td>
</tr>
<tr>
<td>(e) MAA-MAA</td>
<td>-9.46381</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(f) MAA=MAA</td>
<td>-17.6764</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6: Monomer - target bond energies

Table 4.7 shows the $BE$s of R- and S-BOC-Tyr docked into R- and S-BOC-Tyr imprinted sites after replacing individual H-bonding MAAs with a methyl group. An analysis of this data is shown in Table 4.8. The average bond strength was calculated as the average of the differences in docking $BE$s between the normal sites (with all H-bonding interactions) and neutralized binding sites (no H-bonding with that particular functional group), and is presented as a percentage of the total contributed to the normal $BE$. Interestingly, the strength of the interactions within the binding site qualitatively differs from those derived from the single monomer – target calculations. Within
the optimized binding site, the MAA=NH+O1 interaction contributes most to the total $BE$, with MAA-O2 a distant second. MAA=COOH, the strongest interaction in the single monomer – target studies, is only the third strongest contribution to the $BE$ of the site as a whole. This is likely because the single monomer – target models represent an idealized, isolated interaction. The interaction formed as part of the 30 MAA optimization better represents the much more complex reality of monomer – target complexing, where imperfect orientations and misalignments occur in response to adjacent monomers and other outside forces. The QM-optimized single monomer – target interactions can be considered the most energetically favorable orientations for the functional groups to interact. When optimized with the additional 29 monomers, the MAA participating in the H-bonding interaction can be pushed or pulled away from its optimized orientation by steric and electrostatic forces from neighboring monomers.

Table 4.7: $BE$s of BOC-Tyr docked into BOC-Tyr binding sites after neutralizing individual H-bonding MAAs. The interaction labeled at the top of each column represents the interaction neutralized within binding site prior to docking.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Normal</th>
<th>MAA=COOH</th>
<th>MAA=NH+O1</th>
<th>MAA-O2</th>
<th>MAA=PhenolOH</th>
<th>All H-bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-BOC-Tyr</td>
<td>R</td>
<td>-10.2</td>
<td>-9.2</td>
<td>-8.7</td>
<td>-9.8</td>
<td>-9.3</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>-9.2</td>
<td>-8.7</td>
<td>-7.8</td>
<td>-8.8</td>
<td>-8.2</td>
</tr>
<tr>
<td>S-BOC-Tyr</td>
<td>-9.3</td>
<td>-10.2</td>
<td>-8.6</td>
<td>-9.2</td>
<td>-8.8</td>
<td>-9.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-7.8</td>
<td>-8.3</td>
<td>-9.9</td>
<td>-5.4</td>
</tr>
</tbody>
</table>

Table 4.7: $BE$s of BOC-Tyr docked into BOC-Tyr binding sites after neutralizing individual H-bonding MAAs. The interaction labeled at the top of each column represents the interaction neutralized within binding site prior to docking.
Table 4.8: Effect of each MAA - target interaction on $BE$ and $\alpha_E$. The $\alpha_E$(avg) values were calculated as the average of the enantioselectivities of the R- and S-BOC-Tyr imprinted sites after neutralizing the corresponding interaction. For comparison, the $\alpha_E$(avg) before neutralization is 1.10. It should be noted that the differences in $BE$s between the target and enantiomer did not significantly increase after neutralizing the MAA-O2 and MAA=PhenolOH interactions; all $\Delta BE$ values for the “normal” system, the neutralized MAA-O2 system, and the MAA=PhenolOH system were between 0.9 and 1.2 kcal/mol. However, because the overall $BE$ values are weaker after neutralizing the interactions, the $\Delta BE$ produces a greater $\alpha_E$(avg) relative to the “normal” $\alpha_E$(avg).

<table>
<thead>
<tr>
<th>Monomer – Target Interaction</th>
<th>Avg. bond strength (% Total $BE$)</th>
<th>$\alpha_E$(avg) after neutralization</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAA=COOH</td>
<td>8.46</td>
<td>1.06</td>
</tr>
<tr>
<td>MAA=NH+O1</td>
<td>17.44</td>
<td>1.06</td>
</tr>
<tr>
<td>MAA-O2</td>
<td>4.36</td>
<td>1.12</td>
</tr>
<tr>
<td>MAA=PhenolOH</td>
<td>10.51</td>
<td>1.14</td>
</tr>
</tbody>
</table>

It is also interesting to note that the enantioselectivity is lowest after neutralizing the MAA=COOH or the MAA=NH+O1 bonds. This seems to indicate that it is not bond strength that determines the functional group’s contributions to the enantioselectivity per se, but rather the group’s proximity to the chiral center. Both the carboxylic acid group and the amine group are directly bonded to the chiral center of the molecule, with the double-bonded oxygen and phenol groups separated from the chiral center by several rotatable bonds. It is therefore much more difficult for R-BOC-Tyr to position itself in such a way that it can form strong MAA=COOH or MAA=NH+O1 bonds with the binding site that has been optimized to fit the carboxylic acid and amine orientations of S-BOC-Tyr, and vice-versa. The rotatable bonds between the chiral center and the double-bonded oxygen and phenol groups allow the molecule the freedom to find conformations that still allow for strong MAA-O2 and MAA=PhenolOH interactions to take place (Figure 4.10). To further test the idea that the proximity of the functional group to the chiral center influences its contributions to the overall enantioselectivity, the imprinting, docking, and neutralization steps were repeated after inserting additional rotatable bonds (-CH$_2$-) between the chiral center and the functional groups in question. These modified forms are designated by asterisks, where BOC-Tyr* represents BOC-Tyr with an additional -CH$_2$- between the chiral center
and the carboxylic acid, and BOC-Tyr** represents the molecule with an additional -CH2- between the chiral center and the amine group. Both modified molecules are shown in Figure 4.11. It was found that separating these functional groups from the chiral center by additional rotatable bonds significantly reduced or eliminated the group’s contributions to the final enantioselectivity. These results are summarized in Table 4.9. While the BOC-Tyr* and BOC-Tyr** imprinted sites are still enantioselective, this enantioselectivity is not noticeably reduced after methyl-group substitution of the respective interactions (MAA=COOH for BOC-Tyr* and MAA=NH+O1 for BOC-Tyr**). The unmodified BOC-Tyr docking simulations showed significant decreases in enantioselectivity after neutralizing the MAA=COOH and MAA-NH+O1 bonds. The positioning of these two hydrogen bonding groups, immediately adjacent to the chiral center, means the groups are spatially fixed, unable to move relative to the chiral center. Each additional rotatable bond between the group and the chiral center significantly increases the group’s range of motion, relative to the chiral center. Therefore, the enantiomer of the imprinted molecule will still likely be able to form the H-bonding interactions corresponding to functional groups that are not immediately bonded to the chiral center. Figure 4.12 shows how the additional rotatable bond between the chiral center and the carboxylic acid allows the enantiomer of the imprinted BOC-Tyr* molecule to find a conformation in which the MAA=COOH interaction can still take place.
Figure 4.10: R-BOC-Tyr (magenta) and S-BOC-Tyr (blue) after docking into R-BOC-Tyr imprinted 30 MAA binding site. The MAAs of the MAA=COOH interaction (brown), the MAA=NH+O1 interaction (light blue), the MAA-O2 interaction (pink), and the MAA=PhenolOH interaction (green) are shown as well. All other MAAs are hidden to simplify the image. When S-BOC-Tyr docks into the R-imprinted site, it can be seen that the functional groups that are immediately bonded to the chiral center are less aligned with those R-BOC-Tyr. However, the functional groups separated by multiple rotatable bonds, the phenol OH group and the double bonded oxygen of the urethane group, are better aligned between the two chiral forms. This is because the rotatable bonds allow for the enantiomer to find conformations more suitable for the positioning of the H-bonding sites within the binding pocket.

The last column in Table 4.7 shows the results of docking into a binding site after neutralizing all four H-bonding MAAs. It is worth noting that, after methyl-group replacement, the sites still exhibit some degree enantioselectivity. The steric and hydrophobic effects of the neutralized sites are therefore significant enough to allow for enantioselection based purely on shape.
Figure 4.11: BOC-Tyr molecules modified with additional rotatable bonds (dashed blue circles). Chiral centers are circled in red. Left: BOC-Tyr* has an additional -CH₂- group between chiral center and carboxylic acid group. Right: BOC-Tyr** has an additional -CH₂- group between the chiral center and the nitrogen.

Figure 4.12: The imprinted (blue, S-form) and enantiomer (brown, R-form) BOC-Tyr molecules with an additional -CH₂- between the chiral center (circled in red) and the carboxylic acid group. The MAA that forms the MAA=COOH interaction is shown in pink (other MAAs were hidden to simplify the figure). Figure on the right is the side-view of the figure on the left side. Though optimized with the S-form, the rotatable bond gives the carboxylic acid group of the R-form the freedom of movement to find a conformation that permits the MAA=COOH interaction.
Table 4.9: The $BE$ and $\alpha_E$ values of the modified BOC-Tyr molecules, calculated via docking into the imprinted site. Once modified with additional rotatable bonds separating the functional group from the chiral center, neutralizing the interactions no longer reduces the enantioselectivity of the site.

<table>
<thead>
<tr>
<th>Docking molecule</th>
<th>R-form</th>
<th>S-form</th>
<th>R-form</th>
<th>S-form</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-BOC-Tyr*</td>
<td>-11.9</td>
<td>-10.7</td>
<td>-11.2</td>
<td>-9.6</td>
</tr>
<tr>
<td>S-BOC-Tyr*</td>
<td>-10.8</td>
<td>-12.0</td>
<td>-10.1</td>
<td>-11.2</td>
</tr>
<tr>
<td>$\alpha_E$</td>
<td>1.10</td>
<td>1.12</td>
<td>1.11</td>
<td>1.17</td>
</tr>
<tr>
<td>R-BOC-Tyr**</td>
<td>-10.1</td>
<td>-10.1</td>
<td>-9.0</td>
<td>-9.2</td>
</tr>
<tr>
<td>S-BOC-Tyr**</td>
<td>-9.0</td>
<td>-11.5</td>
<td>-7.8</td>
<td>-10.2</td>
</tr>
<tr>
<td>$\alpha_E$</td>
<td>1.12</td>
<td>1.14</td>
<td>1.15</td>
<td>1.11</td>
</tr>
</tbody>
</table>

4.4 Conclusions

Using MM $\rightarrow$ QM-optimized binding sites, docking algorithms were incorporated to determine the imprinted sites’ selectivities against structural analogues. Some degree of selectivity could be identified and it correlated positively to the amount of surface area contact shared between the imprinted molecule and the target. Idealized, fully internal imprinted sites were created to further test the abilities of the selected docking selectivity method. These models exhibited size-, shape-, and polarity-based preference toward the imprinted molecule, demonstrating the potential of MM docking for QM-optimized MIP binding site selectivity analysis.

We were also able to successfully imprint binding sites in silico that exhibited an enantioselective preference for the computationally imprinted molecules over their opposite chiral forms. The computational imprinting included a multiscale MM $\rightarrow$ QM optimization scheme whereby 30 monomers were energetically minimized surrounding a chiral target molecule. Using a ligand docking technique, we quantified the enantioselectivity of the imprinted binding sites.
Binding sites imprinted with BOC-Tyr and CBZ-Tyr exhibited enantiomeric preference for the imprinted target over its enantiomer, and qualitatively agree with experimental data [4]. The cross-enantioselective properties were also evaluated by docking alternative chiral molecules into the imprinted sites. BOC-Tyr imprinted sites showed cross-enantioselectivity for the structurally similar tyrosine derivatives, CBZ-Tyr and Acetyl-Tyr. As CBZ-Tyr and Acetyl-Tyr share the same functional groups with BOC-Tyr and only differ in the composition of a non-polar non-bonding group, they were able to form the same strong H-bonds with the binding site that had been optimized for CBZ-Tyr. The BOC-Tyr imprinted sites did not show cross-enantioselectivity for the molecules that were not tyrosine derivatives, ClPhEthOH and CBZ-Ser. These molecules did not contain the same arrangement of functional groups as the tyrosine derivatives, and therefore could not fit well with the H-bonding landscape of the binding site. The ability for BOC-Tyr imprinted sites to cross-enantioselect for the chiral forms of other tyrosine derivatives, but not structurally dissimilar chiral molecules, correlates with literature [16].

The monomers formed H-bonds with the target during the imprinting. The effect each H-bonding monomer had on the binding energy and selectivity of the binding site was investigated by neutralizing the H-bonding groups individually. The positioning of the functional group relative to the chiral center, rather than the strength of the bond, most influenced its contributions to the selectivity of the site. Significant decreases in enantioselectivity were observed after neutralizing monomers that H-bond with the functional groups immediately adjacent to the chiral center. Because they were directly bonded to the chiral center, the positioning of the H-bonds within the binding site could be seen as fixed, relative to the chiral center. No change in enantioselectivity was observed after neutralizing the H-bonds with functional groups that were separated from the chiral center by two or more rotatable bonds. We hypothesize that this is because the enantiomer
of the imprinted molecule could rotate until the functional group aligned with the H-bonding monomer. Separating the functional groups that were immediately bonded to the chiral center by rotatable -CH₂- groups eliminated their enantioselective impact. These results indicate the importance of functional group positioning in developing an enantioselective MIP. Chiral targets that do not have functional groups immediately bonded to the chiral center may be less likely to form enantioselective binding sites.

With the ultimate goal of being able to better predict the properties of an imprinted polymer matrix, this work introduces a means of investigating the selectivity of a computationally imprinted MIP binding site, as well as the variance in binding energies and selectivities between binding sites within a QM optimized system. This work introduces a method to both create and characterize computationally imprinted binding sites using a combined MM and QM approach. The novel method of computationally imprinting enantioselective binding sites allows for greater understanding of the mechanisms underlying MIP binding. The methods and results presented in this chapter have been published in the Journal of Molecular Modeling [1] and the Journal of Molecular Recognition [28].

4.5 References


chemical editor, visualization, and analysis platform. J Cheminform. doi: 10.1186/1758-2946-4-17

CHAPTER 5

QM and MD Evaluations of Protonation States, Solvents, and Temperature Effects on Molecular Imprinting

5.1 Introduction to Protonation States, Solvent Effects, and Temperature Effects

5.1.1 Monomer – Target Protonation States

One of the advantages of MIPs over other biological receptors is the ability to sense in acidic and basic environments that would otherwise destroy the biological recognition elements. However, though the structure itself will remain stable, the ability for a MIP to recognize its imprinted target may vary significantly at different pH values [1]. This pH-dependent sensitivity is due to the existence of multiple protonation states for both the recognition element and the target molecule. The complementarity of protonation states between the target and the binding site allows for the formation of the H-bonds that are the basis of MIP interactions.

The most common recognition element, PMAA, has one pKₐ value between 6 and 7 [1]. Above this value, the deprotonated state, PMAA⁻, is more prevalent. HA has two pKₐ values within biologically relevant pH values: pKₐ₁=6.9, pKₐ₂=10.4 [2]. (Histamine also possesses a third pKₐ value at 15-20, but it was not investigated here due to its irrelevance, biologically). In basic environments above pH 10.4, HA exists primarily in its neutral form. As the pH decreases, the aliphatic amino group is increasingly likely to become protonated, creating the HA⁺ ion. HA⁺ is the most prevalent form between pKₐ₁ and pKₐ₂. At pH 6.9, HA⁺ exists in equilibrium with HA⁺⁺, the double protonated state formed by the binding of a proton to the imidazole ring. Below pH 6.9, the double protonated form is the most prevalent. Because both the target and the recognition
element have pK\textsubscript{a} values around 7, there are four potential combinations of target – recognition elements at neutral pH values: HAA\textsuperscript{+}-MAA, HA\textsuperscript{+}-MAA\textsuperscript{-}, HA\textsuperscript{++}-MAA, and HA\textsuperscript{++}-MAA\textsuperscript{-}. The protonation states of histamine and methacrylic acid are shown in Figure 5.1 below.

The compositions of the systems tested, as well as the pH values at which they would likely be found, are shown in Table 5.1 below.

![Figure 5.1: Structures of the protonation states of histamine (HA, HA\textsuperscript{+}, HA\textsuperscript{++}) and methacrylic acid (MAA, MAA\textsuperscript{-})](image)

<table>
<thead>
<tr>
<th>System</th>
<th>Target</th>
<th>Recognition Element</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HA\textsuperscript{++}</td>
<td>MAA</td>
<td>$&lt;6.9$</td>
</tr>
<tr>
<td>2</td>
<td>HA\textsuperscript{++}</td>
<td>MAA\textsuperscript{-}</td>
<td>$\sim7$</td>
</tr>
<tr>
<td>3</td>
<td>HA\textsuperscript{+}</td>
<td>MAA</td>
<td>$\sim7$</td>
</tr>
<tr>
<td>4</td>
<td>HA\textsuperscript{+}</td>
<td>MAA\textsuperscript{-}</td>
<td>6.9-10.4</td>
</tr>
<tr>
<td>5</td>
<td>HA</td>
<td>MAA\textsuperscript{-}</td>
<td>$&gt;10.4$</td>
</tr>
</tbody>
</table>

5.1.2 Solvent and Temperature

To fabricate MIPs, the targets and monomers are dissolved into a solvent, where monomer – target complexes form, primarily through H-bonding between the functional groups. These
complexes are then locked into place via thermal polymerization, typically at 60°C, or photopolymerization, typically at 4°C.

The pre-polymerization solvent is commonly referred to as the porogen. Several factors must be considered when selecting an appropriate porogen; it must be capable of solvating the monomers and targets, but should not interfere with the formation of monomer – target complexes. Today, it is generally accepted that the recognition capabilities of imprinted binding sites can be correlated to the nature of the monomer – target complexes that are formed in the pre-polymerization matrix [3–7]. To prevent disruptive target – solvent and monomer – solvent complexation, solvents with poor H-bonding capabilities are typically preferred. Early studies of solvent effects on MIP performance showed a negative correlation between the relative permittivity of the solvent and the binding capabilities of the resulting imprinted polymer [8]. Relative permittivity has long been used as a rough metric for solvent polarity. Solvent polarity, however, is a complex function of many parameters [9]. More recent studies indicate that relative permittivity, though important, cannot fully predict the degree to which solvents influence monomer – target interactions [10]. A more thorough measure of solvent polarity comes in the form of Hansen solubility parameters (HSPs). The HSPs empirically quantify the solvent with three metrics: (1) energy from the dispersion forces, $\delta_d$, (2) energy from the dipolar intermolecular forces, $\delta_p$, and (3), energy from H-bonds, $\delta_h$ [11]. Together, these parameters more fully describe how solvents may facilitate or disrupt monomer – target complexes.

The porogen is also responsible for the final porosity of the polymer (hence the name). Increasing the porosity increases the surface area to volume ratio, allowing for the exposure of more imprinted binding sites [8]. Porosity facilitates rapid mass transport, thereby making it simpler to extract the analytes from within the matrix and improve response times. As seen in
Table 5.2 below, the porosity cannot easily be predicted based on the chemical properties of the solvent [8]. It appears that an ideal solvent is one that can both solvate the monomers and targets with minimal interference of monomer – target complexation (poor H-bond capability) and produce a high porosity polymer (surface area per mass). It can be reasoned that poor H-bonding solvents will produce stronger individual binding sites, $K_0$, by allowing for the maximum number of monomer – target interactions to form in solution, and high porosity solvents will allow for a greater total number of accessible binding sites, $N_T$.

<table>
<thead>
<tr>
<th>Porogen</th>
<th>Dielectric Constant</th>
<th>Surface area (m²/g)</th>
<th>Pore volume (ml/g)</th>
<th>Avg. pore size (Å)</th>
<th>Enantioselectivity factor</th>
<th>H-bond capability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetic acid</td>
<td>6.2</td>
<td></td>
<td></td>
<td></td>
<td>1.9</td>
<td>Strong</td>
</tr>
<tr>
<td>Dimethylformamide</td>
<td>36.7</td>
<td>127</td>
<td>0.17</td>
<td>52</td>
<td>2.0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Tetrahydrofuran</td>
<td>7.5</td>
<td>194</td>
<td>0.24</td>
<td>52</td>
<td>4.1</td>
<td>Moderate</td>
</tr>
<tr>
<td>Chloroform</td>
<td>4.8</td>
<td>3.5</td>
<td>0.007</td>
<td>91</td>
<td>4.5</td>
<td>Poor</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>37.5</td>
<td>256</td>
<td>0.60</td>
<td>94</td>
<td>5.8</td>
<td>Poor</td>
</tr>
<tr>
<td>Benzene</td>
<td>2.3</td>
<td>216</td>
<td>0.43</td>
<td>75</td>
<td>6.8</td>
<td>Poor</td>
</tr>
</tbody>
</table>

Due to the complications inherent to characterizing imprinted polymers experimentally, computational characterization methods have become increasingly popular [12–14]. Most common are electronic structure based methods, for which molecular interactions are modeled using the principles of quantum chemistry. These methods are considered representative of the interactions in the system and more accurate than molecular mechanics (MM) methods, but the accuracy comes with great computational costs. Due to the computational limitations of these methods, solvent molecules are typically excluded from the models, lest the computational costs become prohibitively high. Attempts have been made to simulate the presence of the solvent using
a polarizable continuum model (PCM). This “implicit solvation” method approximates the effects of the solvent by incorporating a polarizability into the medium between the explicitly modeled molecules, using the empirically determined dielectric constant of the solvent [15]. Though the PCM method has met with some success, it is somewhat limited in its inability to replicate hydrogen bonding between the solvent and the monomers or target molecules [16–18]. This competitive hydrogen bonding affects the monomer – target complexes that will ultimately define the structure of the binding site after polymerization. Therefore, the PCM loses validity except when modeling strictly non-polar solvents [19].

Molecular mechanics (MM) and molecular dynamics (MD) methods can be used to combat the computational limitations of QM methods. MM/MD methods rely on classical mechanics-based force fields and are therefore significantly less computationally expensive than electronic structure methods, allowing for the simulation of much larger systems. MM/MD is an effective technique for modeling solvent interactions with the monomers and targets of the pre-polymerization system [19–26].

With the goal of identifying optimal polymerization conditions, here we present a MM/MD method for simulating and analyzing the monomer – target complexes that will define the imprinted binding site. These interactions were quantified by counting the H-bonds formed in the simulation. Including multiple target molecules into the models allowed for target – target dimerization. Explicit solvation with chloroform, acetonitrile, and water allowed us to study the influence of different solvents on the formation of these complexes. Systems were optimized via simulated annealing. Removing the targets from the annealed system left behind computationally imprinted binding sites. Docking of the target and structural analogues into these imprinted sites allowed us to analyze the selectivity of each binding site. The nature of the MM/MD methods
enabled investigation of temperature effects on binding site formation, which is an inherently dynamic property. Though it has long been suggested that lower temperature polymerization produces more sensitive MIPs because thermal energy disrupts monomer – target complexes in solution, this is, to the best of our knowledge, the first in silico analysis of temperature’s influence on monomer – target interactions for MIPs.

5.2 Methods

5.2.1 Equipment

All simulations were carried out on a Dell Workstation with dual quad-core 5687 Xeons with a NVIDIA Quadro 6000 GPU and NVIDIA Tesla c2075 GPU. It has 48 GB of RAM and is running Oracle Linux 6.5.T7500.

5.2.2 Monomer – Target Protonation State Studies

Geometric optimization was computed for systems consisting of four monomers and one target. This 4:1 ratio was found to be the optimal ratio for small scale, one-target systems [27]. Avogadro, a cross platform molecular editor (avogadro.cc), was used to build the pre-polymerization systems, which include the target(s) and monomers [28, 29]. The simulations were conducted in vacuum in order to maximize the effect of the interactions. First, the small systems consisting of one target and 4 monomers were built in Avogadro. The starting orientations of the molecules were chosen such that the carboxylic acid functional groups of the monomers were in approximate co-location to the polar groups of the target. The hydrophobic moieties of the monomers, which form the backbone of the MIP, were directed away from the target. The systems were then geometrically optimized using the PM6-DH2 [30] basis set with MOPAC [31]. This
method was parameterized to correct for dispersion and H-bonding within the PM6 Hamiltonian, making it suitable for the noncovalent interactions that drive MIP binding [30]. All binding energies, $BE$, are defined:

$$BE = [\Delta H_f(target + monomers) - \Delta H_f(target) - \Delta H_f(monomers)]$$  \hspace{1cm} \text{eq. 5.1}

where $\Delta H_f$ is the total heat of formation.

### 5.2.3 Solvent Effects

MM/MD methods were utilized to simulate the competitive interactions between the pre-polymerization solvent and the monomers and targets. Histamine (HA), Caffeine (CAF), and R-BOC-Tyrosine (R-BOC-Tyr) were used as targets, and three solvents were tested, chloroform, acetonitrile, and water. These solvents were chosen for their prevalence in MIP literature, and as representatives for non-polar, intermediate, and polar solvents, respectively. Table 5.3 shows the relative permittivities and the Hansen solubility parameters for each solvent.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Classification</th>
<th>Relative permittivity</th>
<th>$\delta_d$</th>
<th>$\delta_p$</th>
<th>$\delta_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloroform</td>
<td>Non-Polar</td>
<td>4.81</td>
<td>17.8</td>
<td>3.1</td>
<td>5.7</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>Polar, Aprotic</td>
<td>37.5</td>
<td>15.3</td>
<td>18.0</td>
<td>6.1</td>
</tr>
<tr>
<td>Water</td>
<td>Polar, Protic</td>
<td>80.1</td>
<td>15.5</td>
<td>16.0</td>
<td>42.3</td>
</tr>
</tbody>
</table>

Methacrylic acid (MAA) was selected for the monomer, as it is one of the most commonly used monomers in MIP literature [32]. Figure 5.2 shows the structures of the chemical components included in this study.
Figure 5.2: Molecular structures of the targets (R-BOC-Tyr, HA, and CAF), monomer species (MAA), and solvents (chloroform and acetonitrile), as created with ACD/ChemSketch [33].

The Automated Topology Builder (ATB) v. 2.2 server [34–36] was used to generate topology files for each molecular component compatible with the GROMOS96 force field [37]. A graphical representation of the charge group distribution for BOC-Tyr generated with the ATB server is shown in Figure 5.3 below. PACKMOL [38] was then used to randomly fill a 45x45x45Å box with either one or 20 target molecules, 100 MAAs, and enough solvent molecules to fill the remaining volume to the density of the particular solvent (375 chloroform molecules, density of 1.49 g/ml; 1000 acetonitrile molecules, density of 0.79 g/ml; 2400 water molecules, density of 1.00 g/ml). A consistent volume of solvent was maintained between systems rather than a consistent number of molecules to reflect the experimental protocols for solvent studies, for which solvent volumes were consistent between samples [20, 39]. Systems containing no solvent (one or 20 targets and 100 MAAs) were also investigated. The randomized starting geometry of the 20:100:1000 CAF:MAA:Acetonitrile system is shown in Figure 5.4. Models containing one target
were chosen to investigate the formation of a more idealized binding site, without the potential for disruptive target – target interactions. The 20 target systems contain a 1:5 target-to-monomer ratio, which is common for imprinted polymer formation. These systems allow for target – target interactions that contribute to the heterogeneity of binding sites inherent to imprinted polymers [27].
Figure 5.3: ATB generated charge group distribution for BOC-Tyr
Figure 5.4: Starting geometry of system containing 1000 acetonitrile molecules (white, stick), 100 methacrylic acid molecules (green, ball and stick), and 20 caffeine molecules (colored by atom, van der Waals spheres). Image created with VMD [40].

To ensure that there were no steric clashes or inappropriate geometries, the structure was relaxed through an energy minimization (EM). The EM ran a steepest descent geometry optimization algorithm until the maximum force was less than 10 kJ/(mol*nm). Once converged, the geometry output from the EM was prepared for the MD simulation via a two-phase equilibration. First, a 5 ps NVT ensemble was used to equilibrate the system to 300K, using a leap-frog algorithm for integrating Newton’s equations of motion. Here the number of atoms (N), the volume (V) and the temperature (T) are conserved. Energy is exchanged with a thermostat until the temperature converges at a desired value. Second, a 10 ps NPT ensemble was used to equilibrate the pressure (and therefore the density) of the system, again using a leap-frog integrator.
The system was then considered stabilized for the MD simulation. Here the number of atoms (N), pressure (P), and temperature (T) are conserved. The thermostat is combined with a barostat to converge at the appropriate pressure. The 1.5 ns MD simulations were conducted with time steps of 1 fs (1.5x10^6 steps). Finally, a 500 ps simulated annealing (SA) with 1 fs time steps was used to linearly bring the system from 300K to 0K. SA is a probabilistic technique for approximating the global energy minimum of the system. For the MD and SA simulations, energies, geometries, and velocities were saved every 5000 steps. The EM, NVT, NPT, MD, and SA steps were all carried out with GROMACS 5.1 [41, 42]. Periodic boundary conditions were employed to simulate the conditions of bulk MIP synthesis. The full parameterization files for the EM, NVT, NPT, MD, and SA steps can be found in the Appendix.

H-bond counts were the primary metric for quantifying monomer – target complexes and the solvent effect on said complexes. The number of H-bonds was calculated using the Visual Molecular Dynamics (VMD) Hydrogen Bonds plugin [40]. H-bond donor-acceptor distances were cut off at 4 Å, with an angle cutoff of 30 degrees. Due to the periodic boundary conditions of the simulations, molecules on the edges of the box may form H-bonds with those on the opposite edge. These interactions are ignored by the VMD Hydrogen Bond Plug-In. In order to account for these PBC H-bonds, a 3x3 matrix of periodic systems was built, using the SA-optimized geometry as the unit cell. The 3 x 3 matrices were created with the UCSF Chimera package [43]. The H-bond counts could then be focused on the central unit cell, which was surrounded on all sides by periodic unit cells. For example, counting the H-bonds between the targets of only the central unit and the monomers of the entire 3x3 matrix allows for the inclusion of MAA – target H-bonds across the boundaries. For example, the number of H-bonds could be counted between the 20 targets in the central unit cell and all of the MAA molecules in the entire 3 x 3 matrix (Figure 5.5). This count
would include both the H-bonds present in the single unit cell and the H-bonds formed across boundaries as a result of the periodic conditions. The latter category of H-bonds would have been ignored without the 3x3 matrix. It should be noted that this method counts each H-bond equally rather than weighting each bond by strength. Though simplified, the technique offers a fast and reasonable measure of the MAA – target complex.

Figure 5.5: 3x3 matrix of 20:100 BOC-Tyr:MAA (left) and zoomed-in image of central unit (right). BOC-Tyr molecules in the central unit are shown as atom-colored VDW spheres (with the BOC-Tyr molecules in the surrounding units hidden from view). The MAAs in the central unit are represented by green ball-and-stick molecules. The MAAs in the surrounding matrix are shown as blue sticks.

The final SA geometries of the one-target systems were then used for molecular docking studies. Docking is a computational technique for calculating binding orientation and affinities between receptors and ligands, playing an vital role in modern rational drug screening [44]. Using the SA-optimized binding site and the imprinted target as the receptor and the ligand, respectively, produced the binding energy ($BE$). Structural analogues were also docked into the binding site, allowing us to quantify the selectivity of the site. Selectivity is defined here as the ratio of the $BE$
of the imprinted target to the $BE$ of a structural analogue into the same binding site. Our previous work has shown the application of target docking to investigate the selectivity, enantioselectivity, and heterogeneity of QM-optimized binding sites [27, 45]. AutoDock Vina (a molecular docking program) [46] was used to calculate the $BE$ values of the targets and analogues. AutoDock uses the AMBER force field, which has a proven record with proteins, nucleic acids, and other organics [47]. Grid boxes were centered in the binding site, with dimensions of 20 x 20 x 20 Å, and an exhaustiveness of 8 was used to allow for sufficient sampling of docking orientations. In order to emphasize the role of H-bonds in the binding site, the weight of H-bonds in the scoring algorithm was increased two-fold, from -0.6 to -1.2 [46, 48]. Here, the selectivity factor, $\alpha$, is defined as the ratio between the $BE$s of the target and the analogue. The selectivity of the R-BOC-Tyr imprinted site was calculated using its enantiomer, S-BOC-Tyr, as the analogue. The selectivity of the CAF and HA imprinted sites were calculated using each other as the structural analogue; CAF was docked into the HA imprinted site, and vice versa.

5.2.4 Temperature Effects

To investigate temperature effects on target formation, the number of monomer – target H-bonds was calculated for systems containing 1:100 target:MAA. Temperatures of 277 K (4°C) and 333 K (60°C) were chosen to replicate the cold UV-initiation temperatures and thermal initiation temperatures common to many MIP polymerization protocols, respectively [49, 50]. For this study, the EM, NVT, NPT, and MD steps were executed as described above, but with the thermostats set to 277K or 333K instead of 300K. H-bond counts were extracted from the MD simulations. The geometries, and velocities were saved every 5000 steps with time steps of 1fs, producing 300
frames. The H-bond counts represent the average of each frame for the last half of the simulation’s runtime, the final 750 picoseconds (150 frames), when the systems were more equilibrated.

5.3 Results and Discussion

5.3.1 Protonation States

Table 5.4 below shows the binding energy of each system of 4:1 MAA:HA system at the various protonation states. Figure 5.6 shows experimental results from Bongaers et al. [1], with the “Figure 5.6” column in Table 5.4 identifying which experimental conditions would contain that particular combination of protonation states.

Table 5.4: QM-calculated BEs of 4:1 MAA:HA systems at five potential protonation state combinations.

<table>
<thead>
<tr>
<th>System</th>
<th>Target</th>
<th>Monomers</th>
<th>pH</th>
<th>BE (kcal/mol)</th>
<th>Figure 5.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HA++</td>
<td>MAA</td>
<td>&lt;6.9</td>
<td>-24.27</td>
<td>a,b</td>
</tr>
<tr>
<td>2</td>
<td>HA++</td>
<td>MAA-</td>
<td>~7</td>
<td>-347.65*</td>
<td>b</td>
</tr>
<tr>
<td>3</td>
<td>HA+</td>
<td>MAA</td>
<td>~7</td>
<td>-21.59</td>
<td>b</td>
</tr>
<tr>
<td>4</td>
<td>HA+</td>
<td>MAA-</td>
<td>6.9-10.4</td>
<td>-131.17**</td>
<td>b,c</td>
</tr>
<tr>
<td>5</td>
<td>HA</td>
<td>MAA-</td>
<td>&gt;10.4</td>
<td>-69.33</td>
<td>d</td>
</tr>
</tbody>
</table>

*Highest BE, but low probability
**Strongest experimental binding at pH 9, primarily HA+ and MAA-
Figure 5.6: Relative impedance signal of MIP and NIP in response to addition of histamine under various pH conditions: (a) pH 5, (b) pH 7, (c) pH 9, and (d) pH 12. Additions of 3 nM, 6 nM, and 9 nM histamine are labeled I, II, and III, respectively [1].

The impedance values of the HA-MIP systems show the greatest response to the addition of HA molecules at pH 9 (Figure 5.6, C), with a relative impedance shift of approximately 75%. The second greatest response comes at pH 7 (Figure 5.6, B), with an impedance shift of 15%. Both the pH 5 and pH 12 systems (Figure 5.6, A and D) show negligible impedance shifts in response to the presence of the analyte. The modeling results indicate that System 2, with the doubly protonated target and anionic monomer, has the highest binding energy. However, because HA$^{++}$ transitions to HA$^+$ above pH 7 and MAA transitions to MAA$^-$ above pH 7, this pairing will not exist in great quantity (Figure 5.7). This may explain why the experimental binding results are not highest at pH 7. At pH 9, approximately all of the targets will be protonated and all of the binding sites will be anionic. As seen in Table 5.4, this HA$^+$ and MAA$^-$ combination produces the second strongest $BE$. The high frequency with which this pairing can be found at pH 9, compared to the
low frequency with which the HA$^{++}$ and MAA$^-$ pairing is found at pH7, clarifies why the strongest
impedance shifts occur at the higher pH value. At pH values below 7, the binding sites are more
likely to be neutrally charged, leading to the low binding response at pH 5. At pH values above
10.4, the target molecules are more likely to be neutral, leading to the low binding response at pH
12.

![Speciation graphs for histamine and methacrylic acid](image)

Figure 5.7: Speciation graphs for histamine (left) and methacrylic acid (right) [1]

It should be noted that while this work modeled the binding site as a collection of
monomers, the protonation state behavior for polymers is much more complex than that of
polymers. It is likely that binding sites will have a different pK$_a$ value than the non-selective
regions of the bulk polymer. It is also likely that the pK$_a$ will vary between sites, and even between
functional groups within the same binding. However, we can accurately predict the proportion of
HA, HA$^+$, and HA$^{++}$ present at any pH value. While we cannot accurately determine the
protonation states of individual binding sites experimentally, we can measure broader shifts in the
state of the bulk polymer. Using the methods presented here, we can predict the affinity between
targets and binding sites in different protonation states. Further validation of these models coupled
with experimental MIP pH studies could ultimately offer new insight into the variations in binding site and bulk pKₐ values.

5.3.2 Simulated Annealing of Pre-Polymerization Matrix

Table 5.5 shows the number of monomer – target H-bonds (per target molecule), calculated after the SA step. The boxes containing dashes represent interactions that were considered to be the focus of that particular study. For example, because the majority of MAA molecules in the 1:100 target:MAA systems are not in the vicinity of the target, the number of MAA – MAA H-bonds does not tell us anything about the nature of the MAA – target H-bonds. However, for systems containing 20 targets, every MAA could potentially H-bond with the target, making the MAA – MAA dimerization count relevant to the MAA – target H-bond count.
Table 5.5: Number of H-bonds between various components after SA optimization step. For systems containing 20 targets, MAA-target and target-solvent H-bond counts are presented as the average per target value. MAA-MAA and MAA-solvent counts are presented as the average per MAA value. Boxes containing dashes represent interactions that were deemed irrelevant to the focus of the particular study.

<table>
<thead>
<tr>
<th>System</th>
<th>MAA-target</th>
<th>MAA-MAA</th>
<th>target-target</th>
<th>MAA-solvent</th>
<th>target-solvent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1CAF+100MAA</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1CAF+100MAA+375Chloroform</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>1CAF+100MAA+1000Acetonitrile</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>1CAF+100MAA+2400Water</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>20CAF+100MAA</td>
<td>5.85</td>
<td>1.12</td>
<td>0.15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20CAF+100MAA+375Chloroform</td>
<td>3.55</td>
<td>0.89</td>
<td>0.10</td>
<td>0.42</td>
<td>0.5</td>
</tr>
<tr>
<td>20CAF+100MAA+1000Acetonitrile</td>
<td>1.95</td>
<td>0.47</td>
<td>0.20</td>
<td>3.18</td>
<td>3.15</td>
</tr>
<tr>
<td>20CAF+100MAA+2400Water</td>
<td>0.6</td>
<td>0.15</td>
<td>0.06</td>
<td>5.87</td>
<td>9.8</td>
</tr>
<tr>
<td>1HA+100MAA</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1HA+100MAA+375Chloroform</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>1HA+100MAA+1000Acetonitrile</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>1HA+100MAA+2400Water</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>20HA+100MAA</td>
<td>6.65</td>
<td>1.30</td>
<td>0.58</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20HA+100MAA+375Chloroform</td>
<td>5.3</td>
<td>0.95</td>
<td>0.89</td>
<td>0.44</td>
<td>0.5</td>
</tr>
<tr>
<td>20HA+100MAA+1000Acetonitrile</td>
<td>2.6</td>
<td>0.54</td>
<td>0.22</td>
<td>3.28</td>
<td>5.8</td>
</tr>
<tr>
<td>20HA+100MAA+2400Water</td>
<td>0.8</td>
<td>0.15</td>
<td>0.06</td>
<td>5.6</td>
<td>9.45</td>
</tr>
<tr>
<td>1BOC-Tyr+100MAA</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1BOC-Tyr+100MAA+375Chloroform</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>1BOC-Tyr+100MAA+1000Acetonitrile</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10</td>
</tr>
<tr>
<td>1BOC-Tyr+100MAA+2400Water</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>16</td>
</tr>
<tr>
<td>20BOC-Tyr+100MAA</td>
<td>11.15</td>
<td>0.85</td>
<td>1.06</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20BOC-Tyr+100MAA+375Chloroform</td>
<td>9.95</td>
<td>0.80</td>
<td>1.27</td>
<td>0.43</td>
<td>1.8</td>
</tr>
<tr>
<td>20BOC-Tyr+100MAA+1000Acetonitrile</td>
<td>4.45</td>
<td>0.42</td>
<td>0.30</td>
<td>2.95</td>
<td>9.75</td>
</tr>
<tr>
<td>20BOC-Tyr+100MAA+2400Water</td>
<td>1.35</td>
<td>0.12</td>
<td>0.30</td>
<td>5.98</td>
<td>15.45</td>
</tr>
</tbody>
</table>

Predictably, the number of MAA – target H-bonds is highest when no solvent is present to interfere with the interactions. The number of MAA – target H-bonds decreases as the solvents become increasingly polar. Chloroform has relatively low δ_p and δ_h values, so it does not easily interact with either the target or the monomers, freeing them to bond with each other. Water, on the other hand, has high δ_p and δ_h values, allowing it to interact with both the monomers and the
targets and therefore interfere with the ability of MAA – target complexes to form. Interestingly, acetonitrile has a $\delta_p$ value (even higher than water), but a low $\delta_h$ value (nearly the same as chloroform). The strong dipole energy and the weak H-bonding combine to produce an intermediate amount of MAA – target H-bond disruption.

For the one-target systems, when no solvent molecules were present, the number of MAA – target H-bonds equilibrated at approximately 7 for CAF, 11 for HA, and 15 for R-BOC-Tyr. This number appears to be related not just to the number of functional groups present in the target, but, more importantly, the accessibility of these groups. CAF has more functional groups than HA: four amines and two double-bonded oxygens vs. three amines. However, all of CAF’s amines are tertiary, making them less accessible to the surrounding monomers. HA contains one primary amine and one secondary amine, which can more freely interact with multiple monomers. Looking at the detailed output of the VMD H-bond count confirms these interactions, showing that none of CAF’s nitrogen groups directly participated in H-bonding. HA’s primary amine, on the other hand, formed three H-bonds with surrounding monomers, and the secondary amine formed two H-bonds. The secondary amine in BOC-Tyr also formed two H-bonds with MAA molecules. The detailed information of which atoms participated in the MAA – target bonds can be found in Tables 5.6-5.8, with Figures 5.8-5.11 showing the atom label systems for each of the targets and MAA.
Figure 5.8: CAF with atom labels used in the H-bond counts. Hydrogens were removed for the sake of simplifying the image.

Figure 5.9: HA with atom labels
Figure 5.10: BOC-Tyr with atom labels

Figure 5.11: MAA with atom labels
Table 5.6: Details of MAA – CAF H-bonds in 1:100 CAF:MAA system after SA-optimization

<table>
<thead>
<tr>
<th>Donor</th>
<th>Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAF1-C1</td>
<td>MAA89-C1</td>
</tr>
<tr>
<td>CAF1-C1</td>
<td>MAA89-O3</td>
</tr>
<tr>
<td>MAA19-C5</td>
<td>CAF1-O1</td>
</tr>
<tr>
<td>MAA51-O3</td>
<td>CAF1-O1</td>
</tr>
<tr>
<td>MAA51-O3</td>
<td>CAF1-C5</td>
</tr>
<tr>
<td>MAA61-O3</td>
<td>CAF1-C6</td>
</tr>
<tr>
<td>MAA61-O3</td>
<td>CAF1-O2</td>
</tr>
</tbody>
</table>

Table 5.7: Details of MAA – HA H-bonds in 1:100 HA:MAA system after SA-optimization

<table>
<thead>
<tr>
<th>Donor</th>
<th>Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA1-C4</td>
<td>MAA36-O6</td>
</tr>
<tr>
<td>HA1-N1</td>
<td>MAA24-O6</td>
</tr>
<tr>
<td>HA1-C2</td>
<td>MAA88-O3</td>
</tr>
<tr>
<td>HA1-N1</td>
<td>MAA86-O6</td>
</tr>
<tr>
<td>HA1-C1</td>
<td>MAA57-C1</td>
</tr>
<tr>
<td>HA1-C1</td>
<td>MAA57-O6</td>
</tr>
<tr>
<td>MAA100-C4</td>
<td>HA1-C3</td>
</tr>
<tr>
<td>MAA24-C4</td>
<td>HA1-N1</td>
</tr>
<tr>
<td>MAA101-C4</td>
<td>HA1-N2</td>
</tr>
<tr>
<td>MAA57-O3</td>
<td>HA1-N2</td>
</tr>
<tr>
<td>MAA9-C5</td>
<td>HA1-C5</td>
</tr>
</tbody>
</table>
Table 5.8: Details of MAA – BOC-Tyr H-bonds in 1:100 BOC-Tyr:MAA system after SA-optimization

<table>
<thead>
<tr>
<th>Donor</th>
<th>Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOC-Tyr1-C3</td>
<td>MAA34-O6</td>
</tr>
<tr>
<td>BOC-Tyr1-C3</td>
<td>MAA3-O6</td>
</tr>
<tr>
<td>BOC-Tyr1-N1</td>
<td>MAA38-C1</td>
</tr>
<tr>
<td>BOC-Tyr1-N1</td>
<td>MAA38-O6</td>
</tr>
<tr>
<td>BOC-Tyr1-CB</td>
<td>MAA31-C2</td>
</tr>
<tr>
<td>BOC-Tyr1-CB</td>
<td>MAA31-C5</td>
</tr>
<tr>
<td>BOC-Tyr1-OH</td>
<td>MAA84-C1</td>
</tr>
<tr>
<td>BOC-Tyr1-OH</td>
<td>MAA84-O6</td>
</tr>
<tr>
<td>MAA34-C5</td>
<td>BOC-Tyr1-C3</td>
</tr>
<tr>
<td>MAA76-C5</td>
<td>BOC-Tyr1-C6</td>
</tr>
<tr>
<td>MAA48-O3</td>
<td>BOC-Tyr1-C6</td>
</tr>
<tr>
<td>MAA85-C5</td>
<td>BOC-Tyr1-OXT</td>
</tr>
<tr>
<td>MAA48-O3</td>
<td>BOC-Tyr1-OXT</td>
</tr>
<tr>
<td>MAA99-C4</td>
<td>BOC-Tyr1-CZ</td>
</tr>
<tr>
<td>MAA20-C5</td>
<td>BOC-Tyr1-OH</td>
</tr>
</tbody>
</table>

The MAA – target H-bond numbers decrease significantly when 20 target molecules are present, which is partially due to the presence of target – target interactions. 15% of CAF molecules, 58% of HA molecules, and 106% of R-BOC-Tyr molecules formed target – target H-bonds in the SA-optimized geometries with 20 targets, 100 MAAs, and no solvents. Here 100% would mean an average of one target – target H-bond per target. Values greater than 100% mean that some targets formed multiple H-bonds with another target, making the total number of target – target H-bonds greater than the number of targets. The high percentage of R-BOC-Tyr dimerization is due in part to the size of the molecule; R-BOC-Tyr is much larger than CAF and HA, so there is a greater likelihood of two molecules interacting within a given a volume. A higher ratio of monomers would be required to surround each R-BOC-Tyr and reduce target dimerization.

Another possible contributor to the decrease in MAA – target H-bonds for the 20 target system is the limited availability of monomers. The one-target systems were saturated with monomers, allowing for the maximum number of interactions possible. The 20-target systems may
have adjacent monomers that are H-bonded with a separate target molecule on the opposite side of the MAA. It should also be noted that the CAF – CAF interaction is more prevalent than the H-bond counts indicate, as those values do not take π-π stacking into account. The number of the 20 CAF molecules participating in π-π stacking with other CAF molecules in the SA-optimized systems with no solvent, in chloroform, in acetonitrile, and in water were found to be 7, 12, 6, and 4, respectively. Interestingly, these results seem to indicate that chloroform facilitates the π-π stacking of CAF molecules. This may explain why the number of monomer – target H-bonds decreases more for CAF than for BOC-Tyr or HA when chloroform is included into the 100 MAA, 20 target systems. Figure 5.12 shows the stacking of CAF molecules in the system solvated with chloroform.

![Figure 5.12: π-π stacking of CAF molecules after SA-optimization. MAA and chloroform molecules are hidden to simplify the figure. The dotted lines connect aromatic groups engaging in π-π stacking. Image created with VMD [40].](image)

Including 20 targets in the simulation allowed for a glimpse into the heterogeneity of the system. By the disordered and amorphous nature of pre-polymerization systems and non-covalent imprinting, each binding site will be unique. While Table 5.5 shows the average number of MAA
target bonds per target, a further breakdown of the MAA – target H-bonds in the 20 target systems is shown in the box-and-whisker plot (Figure 5.13). For each 20:100 target:MAA system, we see a distribution of MAA – target H-bond counts. Similar to our QM heterogeneity studies [27], the ranges do not show any noticeable bi-modal discretization, furthering arguments for continuous affinity distribution isotherms for mathematically modeling MIP behavior. The ranges and standard deviations can be found in Table 5.9.

Figure 5.13: Box-and-whisker plot of the individual MAA – target H-bond counts for the 20 Target + 100 MAA systems after SA optimization. Each dot represents the MAA – target H-bond count of an individual target within the system. The boxes represent the middle two quartiles of the H-bond count range, with the “x” marking the mean.
Table 5.9: Average, range, and standard deviation of the number of MAA – target H-bonds in SA-optimized 20 target systems.

<table>
<thead>
<tr>
<th>MAA - Target bonds</th>
<th>System (100 MAA + 20 Targets)</th>
<th>CAF</th>
<th>HA</th>
<th>BOC-Tyr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average per target</td>
<td></td>
<td>5.85</td>
<td>6.65</td>
<td>11.15</td>
</tr>
<tr>
<td>Fewest H-bonds with target</td>
<td></td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Most H-bonds with target</td>
<td></td>
<td>13</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td></td>
<td>3.02</td>
<td>2.63</td>
<td>3.12</td>
</tr>
</tbody>
</table>

The target – solvent H-bonds also follow the expected trend. Non-polar chloroform cannot easily bond with the targets, averaging 0.5, 0.5, and 1.8 H-bonds per CAF, HA, and R-BOC-Tyr, respectively, in the 20 target systems. Target – acetonitrile bonds interfere with the MAA – target bonds, reducing the MAA – target H-bond count. Target – water bonds almost fully replace MAA – target bonds in the 20-target systems, and completely replace MAA – target bonds in the one-target systems. The MAA – MAA bond numbers followed a similar trend as the MAA – target bonds. MAA molecules were less likely to dimerize when the more polar solvents were present.

The analysis of these models shows that the degree to which the solvent interferes with the MAA – target complex is inversely correlated with the polarity of the solvent. The relationship between the solvent relative permittivity and the degree of complex disruption corresponds well with published experimental work on bulk MIPs [20, 39, 51], as well as with QM models utilizing the PCM and combined QM/MD studies [20, 52]. The results also correspond well with more recent experimental work showing that, although the relative permittivity of the solvent produces a reasonable estimate of MAA – target H-bond disruption, solubility parameters may produce a more complete view of the disruption [10]. Both dipole energy and H-bond energy were shown to increase solvent interference. The dispersion energy parameter, $\delta_p$, was ignored for this study, as the MM methods do not account for induced polarizability. Because the dispersion interactions are significantly weaker than the dipole – dipole bonds and H-bonds in imprinted polymers, and
because the $\delta_p$ values for the three solvents are similar, this omission is not expected to change the outcome of the models.

As early as 1993, Sellergren et al. [39] showed experimentally that chloroform produced a significantly higher retention factor per unit of surface area than other, more H-bond capable solvents, including acetonitrile. It should be noted, however, that compared to chloroform, acetonitrile generally produces significantly more porous MIPs. Though the acetonitrile interference weakens individual binding sites, the increased surface area allows for a greater density of these weaker sites per mass of polymer, thus allowing acetonitrile MIPs to perform comparably to chloroform MIPs on a measure of targets bound per mass of MIP. For example, a 1.0 $\mu$g chloroform MIP particle containing 1000 strong binding sites and a 1.0 $\mu$g acetonitrile MIP particle containing 100,000 binding sites that are $1/100^{th}$ as strong will produce similar results using typical bulk MIP binding analysis techniques. This is one reason that it is important to consider the surface area and not just the mass of particles when comparing MIPs.

5.3.3 Docking into Annealed Binding Sites

Table 5.10 below shows the $BE$ and $\alpha$ values of the one-target systems, calculated via docking.
Table 5.10: \( BE \) and \( \alpha \) values for SA-optimized R-BOC-Tyr, CAF, and HA binding sites. Sites were optimized with no solvent, and in the presence of chloroform, acetonitrile, and water. Dashes represent ligand – receptor pairs that were not considered within the scope of this study; analogues were chosen to be structurally similar to the imprinted target.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Imprinted molecule (receptor)</th>
<th>Docked molecule (ligand) (kcal/mol)</th>
<th>R-BOC-Tyr</th>
<th>S-BOC-Tyr</th>
<th>CAF</th>
<th>HA</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Solvent</td>
<td>R-BOC-Tyr</td>
<td>-9.0</td>
<td>- -</td>
<td>- -</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAF</td>
<td>-</td>
<td>- -</td>
<td>- 8.4</td>
<td>-7.4</td>
<td></td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>HA</td>
<td>-</td>
<td>- -</td>
<td>0.0</td>
<td>-7.4</td>
<td></td>
<td>( \infty )*</td>
</tr>
<tr>
<td>Chloroform</td>
<td>R-BOC-Tyr</td>
<td>-7.4</td>
<td>- -</td>
<td>- -</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAF</td>
<td>-</td>
<td>- -</td>
<td>-5.2</td>
<td>-5.5</td>
<td></td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>HA</td>
<td>-</td>
<td>- -</td>
<td>-4.9</td>
<td>-5.7</td>
<td></td>
<td>1.16</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>R-BOC-Tyr</td>
<td>-6.4</td>
<td>- -</td>
<td>- -</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAF</td>
<td>-</td>
<td>- -</td>
<td>-4.7</td>
<td>-5.3</td>
<td></td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>HA</td>
<td>-</td>
<td>- -</td>
<td>-4.9</td>
<td>-5.9</td>
<td></td>
<td>1.20</td>
</tr>
<tr>
<td>Water</td>
<td>R-BOC-Tyr</td>
<td>-4.9</td>
<td>- -</td>
<td>- -</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAF</td>
<td>-</td>
<td>- -</td>
<td>-4.3</td>
<td>-4.2</td>
<td></td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>HA</td>
<td>-</td>
<td>- -</td>
<td>-4.5</td>
<td>-5.4</td>
<td></td>
<td>1.20</td>
</tr>
</tbody>
</table>

*There were no energetically favorable orientations for CAF to dock into the HA binding site.

The \( BE \) values, as the number of MAA – target H-bonds, are highest when no solvent is present, and decrease as the polarity of the solvent increases: \( BE_{\text{No Solvent}} > BE_{\text{Chloroform}} > BE_{\text{Acetonitrile}} > BE_{\text{Water}} \). This inverse relationship between the strength of a binding site and the porogen’s H-bond capability correlates well with both experimental literature [8, 53, 54] and previous theoretical studies [19, 25]. Enantioselectivity is only observed in the no solvent case; here R-BOC-Tyr binds more strongly to the R-BOC-Tyr imprinted binding site than its enantiomer, S-BOC-Tyr. The CAF and HA binding sites also exhibit selectivity when no solvent is present. With no solvent interference, the monomers fully surround the target, creating a tight binding site. HA is smaller than CAF, so the tight HA-imprinted site completely prohibits CAF from finding an
energetically favorable binding mode. HA can fit into the CAF-imprinted site, but the site’s preference for CAF indicates an orientation of H-bonding groups present in the site that better suits the imprinted molecule than the analogue.

The enantioselectivity of the R-BOC-Tyr-imprinted sites against S-BOC-Tyr and the selectivity of the CAF-imprinted sites against HA falls away when sites are formed in the presence of a solvent. The target – solvent H-bonds prevent the monomers from forming all of the MAA – target H-bonds needed to form a tight binding site. Instead, a binding site is formed with the target only partially complexed with the adjacent MAAs. The H-bond positioning within the binding site can therefore be expected to lead to reduced preference for the imprinted molecule over the analogue. The HA-imprinted site manages to maintain selectivity against CAF regardless of the solvent, which is likely still due to size-related steric hindrance.

\subsection*{5.3.4 Temperature Effects}

Table 5.11 below shows the average number of H-bonds formed in the final half of the 1.5ns MD simulations run at 277K and 333K

<table>
<thead>
<tr>
<th>System</th>
<th>Number of H-bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>At 277K</td>
</tr>
<tr>
<td>1CAF+100MAA</td>
<td>6.767</td>
</tr>
<tr>
<td>1HA+100MAA</td>
<td>8.073</td>
</tr>
<tr>
<td>1BOC-Tyr+100MAA</td>
<td>14.773</td>
</tr>
</tbody>
</table>

Table 5.11 shows for all three targets that, at lower temperatures, more H-bonds can form between the targets and the monomers than at higher temperatures, as increasing thermal energy decreases the stability of the complexes. This has been the argument in favor of cold UV-initiated polymerization rather than thermal initiation [49, 51, 55–58]. Though the general trend of the
simulation results is rather intuitive, this computational method represents a novel mean of quantifying just how strongly the MAA – target complexes are influenced by temperature effects.

For experimental studies using non-imprinted polymers (NIPs) as controls, one must be extremely careful to ensure that the MIP and NIP do not significantly differ morphologically after photopolymerization, as the target molecules are likely to interact with initiating photons and alter the progression of the reaction. Thermal initiation produces more morphologically similar MIPs and NIPs, increasing the likelihood that differences in binding between the two are due to the presence/absence of selective binding sites rather than other morphological phenomena.

5.4 Conclusions

Using QM energy optimizations of 4:1 MAA:HA systems at various protonation states, we were able to calculate binding energies for the species at different pH values. These binding energies could be used to explain the binding responses at different pH values from experimental MIP work [1]. This shows that molecular modeling techniques may be applicable for predicting the efficacy of MIPs in a wide range of pH environments.

Further, using MM/MD techniques, we were able to simulate several of the interactions that take place in the pre-polymerization matrix. It is believed that the interactions between the monomers and target molecules in solution will ultimately determine the sensitivity and selectivity of binding sites formed during polymerizations, making it worthwhile to investigate the conditions that facilitate monomer – target H-bond formation. Using H-bond counts as our metric, simulated annealing of systems containing one target, 100 monomers, and no solvents gave us an approximation of an idealized and optimized binding interaction. Increasing the target-to-monomer ratio from 1:100 to 1:5 and explicitly solvating the systems allowed us to quantify the
deviations from the idealized binding interactions brought about by solvent interference and target
dimerization. H-bond counts and molecular docking demonstrated that the strength of the
monomer – target interactions and the strength of the resulting binding site are inversely related to
the complex polarity of the solvent, which correlates with both experimental and theoretical results
in literature [20, 39, 51, 52]. To the best of our knowledge, we believe this to be a novel technique
for computationally imprinting selective binding sites via a simulated anneal and quantifying the
solvent’s effect on the selectivity of the site.

While the dielectric constant serves as a reasonable predictor of a solvent’s ability to
interfere with monomer – target interactions, a more detailed description of the solvent may be
necessary to better understand the mechanisms behind the interference. Comparing solvent HSPs
with the degree to which the solvent influences the monomer – target H-bond counts offered
insight into how the dipole and H-bonding forces influence the solvent interference behavior.

H-bond counts were also used to investigate temperature effects on monomer – target
interactions. It was shown that at 4°C, a temperature common for cold photopolymerization of
MIPs, more monomer – target H-bonds were formed during the simulation than at 60°C, a
temperature common for thermal polymerization of MIPs. This novel computational analysis of
MIP systems adds credence to the experimental work showing that lower polymerization
temperatures produce more sensitive MIP binding sites [49, 51, 55–58]. However, when using
photopolymerization, additional care must be taken to ensure the presence of the target does not
alter the course of polymerization so significantly that comparisons to NIPs are no longer
reasonable.

This work serves as a proof of principal that simulated annealing of large scale MD systems
could produce binding sites that exhibit selectivity in docking studies, thus building on our
previous work creating selective binding sites through small scale QM optimizations. It also introduced a method for measuring the influence of temperature on monomer – target interactions in solution. With the continuing objective of rationally designing, optimizing, and characterizing MIPs in silico, the methods presented above offer new insight into the effects of two key MIP design parameters: the solvent and the temperature.

At the time this chapter was written, the methods and results were under review for publication by the Journal of Molecular Modeling.

5.5 References

Coefficient. Their Importance in Surface Coating Formulation. Technical University of Denmark


47. Pearlman D, Case D, Caldwell J, et al. (1995) Amber, a Package of Computer-Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular-Dynamics and Free-


CHAPTER 6

Conclusions and Future Directions

6.1 Summary

Since they were first developed, the characterization and optimization of MIPs has been hindered by experimental limitations. The amorphous nature of the substrate and the heterogeneity of the binding sites significantly complicates the analysis of binding data. This, in turn, makes it difficult and time consuming to screen for the optimal fabrication for parameters. The motivating hypothesis for this work was that computational chemistry could be used to characterize the monomer – target interactions in ways that would allow for a greater understanding of the MIP binding properties.

6.1.1 Ratio Optimization

One of the parameters that has experimentally proven critical when determining the quality of binding sites is the monomer-to-target ratio. QM methods revealed that the optimum ratio varied between small one-target systems and larger five-target system. We believe this to be the first QM ratio optimization with multiple targets. It was concluded that the variance between the optimal ratios of the two systems was due to the ability of monomers to dimerize and exist in the space between targets without negatively affecting the total energy of the system by steric crowding around the targets.
6.1.2 **Heterogeneity of Binding Sites**

It has long been suspected that a heterogeneity of binding sites was responsible for complicating experimental results. Binding curves did not follow the trends expected for homogenous systems. Because the number of binding sites was unknown, assumptions had to be made in order to mathematically fit isotherms to the binding data.

Our QM analysis of systems containing five targets showed that each binding site possessed a unique binding energy, furthering arguments for the use of continuous distribution models when analyzing MIP binding data. Binding site surface area analysis showed that the amount of contact between the target and the binding site, as well as the degree to which the contact area was “active” with functional groups, contributed to the sensitivity of the site. This was the first *in silico* heterogeneity study at the QM level of theory.

6.1.3 **Binding Site Selectivity**

Binding sites that were imprinted with QM optimization were shown to exhibit selectivity against structural analogues. This selectivity was based on the differences in size, shape, and polarity between the target and the analogue molecules. Chiral targets were imprinted in order to investigate enantioselectivity, which is seen as the purest measure of imprinted selection. Systems imprinted with chiral tyrosine derivatives were shown to bind more strongly to the imprinted molecule than its enantiomer. Further, these binding sites also exhibited a preference for the corresponding chiral forms of other tyrosine derivatives over their enantiomers. By selectively neutralizing H-bonding functional groups in the binding site, we were able to conclude that the functional groups closest to the chiral center contribute more to the enantioselectivity of the site than the functional groups separated from the chiral center by multiple rotatable bonds.
6.1.4 Protonation States

QM optimizations of monomers and histamine in different protonation states allowed for a prediction of the binding energies that would be expected in different pH environments. The modeling results corresponded well with literature results showing that binding is strongest when an anionic binding site can interact with a cationic histamine.

6.1.4 Solvent Effects

Large systems containing one or 20 targets, 100 monomers, and explicit solvation with chloroform, acetonitrile, or water were modeled using MD. It was shown that solvents with a greater complex polarity are more likely to interfere with monomer – target complexation. Targets were imprinted via a simulated anneal and then the targets and structural analogues were docked into the imprinted sites. Selectivity decreased as solvent polarity increased. This represents a novel MD method for modeling whole binding sites.

6.1.5 Temperature Effects

H-bond counts of MD simulations at 4°C and at 60°C showed that monomer – target complexation decreases as temperature increases. These two temperatures are common for MIP photopolymerization and thermal polymerization, respectively. To the best of our knowledge, this was the first attempt at modeling temperature effects in silico.
6.2 Critical Assessment of Modeling Assumptions

By the nature of molecular modeling, many assumptions and simplifications must be made in order to manage the computational expense. The two largest assumptions made in the work presented in this thesis are:

1) The crosslinkers have minimal effect on the monomer – target interactions and can therefore be removed from the models

2) The monomer – target interactions are unchanged by the polymerization process

In 1989, Sellergren probed the crosslinker effects by comparing chromatographic enantioselectivity data for L-phenylalanine imprinted polymers with varying monomer-to-crosslinker ratios (the concentration of target molecules was kept constant) [1]. Figure 6.1 below shows the measured separation factor as a function of the percent functional monomer.

![Optimization of Separation Factor vs. Percent Functional Monomer](image)

Figure 6.1: Determination of the optimal percent functional monomer, adapted from [1, 2]
Figure 6.1 clearly shows that the crosslinker influences the imprinted polymer. Sellergren makes the argument that the crosslinker effect is mechanical in nature, rather than chemical. In his view, the monomer – target interactions are not altered by the presence of the crosslinkers, but there is a minimum crosslinker concentration required to produce a polymer with enough rigidity for the binding sites to remain stable. With standard monomer-to-crosslinker ratios on the order of 1:4, it is difficult to imagine that the abundance of crosslinkers truly has no effect on the monomer – target interactions that ultimately define the binding sites. However, under the Sellergren interpretation that the crosslinkers do not interfere with the formation of binding sites and only serve to uphold the fidelity of the sites, it is believed that the omission of the crosslinkers does not detract from the validity of the results presented in the previous chapters.

The assumption that the interactions between the target and functional groups of the monomers are unaltered by the polymerization is one that we are unable to verify experimentally. To date, the overwhelming majority of MIP theoretical work operates under this assumption [3]. Very recently, methods for modeling the polymerization process around targets using reactive MD simulations appear to show that the selectivity actually increases after the polymerization process [4]. Though polymerization models are in their infancy and limited due to their computational expense, we expect reactive simulations to add another layer of accuracy to the future of MIP modeling. It is for this reason that reactive modeling fits into the potential future directions of the research presented here.

6.3 Future Directions

6.3.1 Reactive Modeling of Polymerization

Reactive molecular simulations, those capable of modeling the formation of bonds, could
potentially prove useful when predicting MIP matrix morphology as a function of polymerization. Polymerization models would provide a greater understanding of the thermodynamics and kinetics of functional matrix – target relationships, rather than relying on the assumption that the monomer – target complex is unaffected by polymerization. Previous studies, including our own, have employed either QM or MM/MD simulations to optimize a single aspect of MIP formulation (monomer selection, monomer/analyte ratio, solvent) [3, 5–9]. Ideally, what is needed is a simulation that can accurately represent both the liquid structure of the pre-polymerization solution and the actual polymerization process itself. Once the structure of the MIP matrix is known, the binding efficacy and specificity for the target analyte can be determined using standard docking methods such as those presented in the previous chapters. The ability to predict the effects of different imprinting and preparation conditions will enable the true “rational design” of novel MIPs.

6.3.2 Preparation and Characterization of Bulk MIPs

Once rationally designed, selected MIPs and corresponding NIPs could be fabricated in bulk as described previously [10–12]. Once prepared, analyte samples could be run, where retention factor (k) and peak widths (tw) are representative of binding affinity and binding site heterogeneity, respectively. The selectivity compares the retention time on the column between two analytes and will be used to compare selectivity for a MIP between target/non-target analytes. This data could feed back into the optimization of the reactive MD simulations to validate the modeling results.
6.3.3 Development of Nanoscale Interdigitated Electrode Transducer

Nanoscale interdigitated electrodes (nano-IDEs) are hypothesized to significantly reduce solution-dependence of the recorded signal, allowing for sensing in complex environments. Nanoscale miniaturization of electrode arrays can significantly improve sensitivity over other impedimetric devices by concentrating the electric field to an elliptical integral within the thickness of the sensing matrix (Figure 6.2) [13].

![Figure 6.2: Field lines over nano-IDEs. Each curve marks the percentage of the field below the line (left). Illustration of how engineering design can control the field path of IDE devices to maximize response from the MIP matrix (middle). COMSOL model of potential field lines over a nano-IDE configuration (right).](image)

6.3.4 Preparation of Surface-Polymerized MIPs.

Surface initiated polymer (SIP) techniques could be used to integrate MIPs on the nano-IDE transducer as described by Piacham et. al. [14]. Typically SIP generates polymer films between 25-150 nm [15], providing a suitable approach to prepare fast-response MIPs integrated on an electronic transducer.

6.3.5 Sensor Analysis

Nano-IDE designs could be evaluated with MIPs prepared via SIP. Sensor response (sensitivity) would be evaluated through capacitance scans at the frequency determined through electrical impedance spectroscopy (EIS). The limit of detection would be determined through the
periodic addition of the analyte into buffer solution until a distinguishable change in capacitance (greater than 3X the systematic noise) is observed. The linear range and saturation point could be determined similarly, by analyzing the capacitance response to increasing analyte concentrations until no further shift is noticed. Though binding site heterogeneity is typical in MIPs, it enables remarkable detection ranges, spanning as much as six orders of magnitude (nanomolar to millimolar) before saturating [16, 17]. These measurements could also be performed in buffer with different pH and increasing ionic strength to assess the performance in the MIP/nano-IDE system in a simulated dynamic environment (cell growth media, etc.). Non-imprinted polymer controls with the same monomer/crosslinker formulation, and non-target analogues would be included and used for comparisons.

6.4 References


A Few Words on Bulk MIP Fabrication and Analysis

Though the focus of this work was on in silico investigations of the properties and MIPs, while scouring the literature I have found three practices common to experimental MIP studies that I believe should be reconsidered or at least critically examined.

First, it is common to present bulk binding data as the number of analytes bound per mass of MIP, and this number is compared to the number of analytes bound per mass of NIP. However, it is not uncommon for studies to make no mention of the porosities of the MIP and the NIP. Because analyte binding is largely a surface interaction, a better measure of the binding capabilities would be a comparison of the number of analytes bound per surface area. It has been shown that addition of analyte molecules to the pre-polymerization complex can dramatically alter the surface area per mass of the polymer (Fig. PS.1) [1]. The non-selective binding found in NIPs is directly proportional to the accessible surface area. In order to say for certain that the differences in binding between the MIP and the NIP are in response to selective imprinted binding sites and not simply extra surface area for non-selective binding, the number of bound analytes must be normalized to the surface area of the polymers rather than the masses.
Second, additional care must be taken to verify the similarities in polymer properties when using photopolymerization rather than thermal polymerization. The analyte molecules in solution will interact with the UV light, absorbing and/or reflecting the photons in ways that alter the polymerization. Photopolymerization of samples containing the monomers, crosslinkers, and porogen (NIP) will produce polymer pellets that are morphologically dissimilar from those with dissolved analytes (MIP) in ways that can be noticed with the naked eye. Under identical conditions, I have found the NIP vial to be fully polymerized and uniform while the MIP vial would be harder near the walls of the vial and more gelatinous or powdery near the center. It is my belief that the analytes near the walls of the vial will absorb and reflect the photons, inducing polymerization of the surrounding monomers and thus reducing the transmittance of UV into the center of the vials. These extreme morphological disparities dramatically reduce the ability to compare the MIP to the NIP. The differences may be less obvious; I have noticed MIP pellets that appear hard and uniform, similar to the NIP formed under the same conditions, but the MIP will strongly adhere to the glass vial while the NIP will easily separate. The advantages of cold UV-
initiation are apparent. Lower temperatures will allow for more stable monomer–target complexes to form in the solution, creating stronger binding sites and reducing heterogeneity. However, if the presence of the target significantly alters the properties of the resulting polymer, thermal polymerization may be a more suitable alternative.

Finally, it seems that many experiments seek to investigate the selectivity of an MIP by comparing the binding of the imprinted molecule to that of a structurally similar analog. However, one cannot say if the preference is, in fact, due to the presence of selective binding sites without also presenting the NIP binding data for both the target and the analog. For example, showing that a histamine imprinted polymer binds more strongly to histamine than L-histidine is not enough verify the presence of selective binding sites. If the NIP also shows a preference for histamine due to general, non-selective binding, as is often the case, the MIP’s “preference” may not be in response to selective binding sites, but rather differences in the amorphous polymer’s van der Waals interactions with the two molecules. To better verify MIP selective binding, one could subtract out each molecule’s non-selective NIP binding affinities. Also, MIPs imprinted with the analog could be used to make the case for selective binding. If the histamine-MIP prefers histamine over L-histidine, AND the L-histidine-MIP prefers L-histidine over histamine, the differences could be attributed to selective binding. It is for these reasons that enantiomers are the preferred method of characterizing MIPs; the target and analog (the enantiomer) should have identical non-selective affinities to the achiral polymer of the NIP.

APPENDIX

Below is the commented content of the .mdp files used for the EM, NVT, NPT, MD, and SA steps. (For the temperature effect studies, ref_t and gen_t were set to 277 or 333.)

**Energy Minimization (EM)**

; Parameters describing what to do, when to stop and what to save
integrator = steep ; Algorithm (steep = steepest descent minimization)
emtol = 10.0 ; Stop minimization when the maximum force < 10.0 kJ/mol
emstep = 0.01 ; Energy step size
nsteps = 50000 ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
nstlist = 1 ; Frequency to update the neighbor list and long range forces
cutoff-scheme = Verlet
ns_type = grid ; Method to determine neighbor list (simple, grid)
list = 1.0 ; Cut-off for making neighbor list (short range forces)
coulombtype = PME ; Treatment of long range electrostatic interactions
rcoulomb = 1.0 ; long range electrostatic cut-off
rvdw = 1.0 ; long range Van der Waals cut-off
pbc = xyz ; Periodic Boundary Conditions
Canonical Ensemble (NVT)

define = -DPOSRES ; position restraints

; Run parameters
integrator = md ; leap-frog integrator
nsteps = 25000 ; 2 * 25000 = 5 ps
dt = 0.0002 ; .2 fs

; Output control
nstxout = 500 ; save coordinates every 1.0 ps
nstvout = 500 ; save velocities every 1.0 ps
nstenergy = 500 ; save energies every 1.0 ps
nstlog = 500 ; update log file every 1.0 ps

; Bond parameters
Continuation = no ; first dynamics run
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
lincs_warnangle = 60 ; maximum angle bond can rotate between steps

; Neighborsearching
cutoff-scheme = Verlet
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 20 fs, largely irrelevant with Verlet
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT

; Temperature coupling is on
tcoul = V-rescale ; modified Berendsen thermostat
tc-grps = system ; two coupling groups - more accurate
tau_t = 0.1 ; time constant, in ps
ref_t = 300 ; reference temperature in K

; Pressure coupling is off
pcoupl = no ; no pressure coupling in NVT
; Periodic boundary conditions
pbc = xyz ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme
; Velocity generation
gen_vel = yes ; assign velocities from Maxwell distribution
gen_temp = 300 ; temperature for Maxwell distribution
gen_seed = -1 ; generate a random seed
Isothermal-Isobaric Ensemble (NPT)

define = -DPOSRES ; position restrain the protein

; Run parameters
integrator = md ; leap-frog integrator
nsteps = 50000 ; .2 * 50000 = 10 ps
dt = 0.0002 ; .2 fs

; Output control
nstxout = 500 ; save coordinates every 1.0 ps
nstvout = 500 ; save velocities every 1.0 ps
nstenergy = 500 ; save energies every 1.0 ps
nstlog = 500 ; update log file every 1.0 ps

; Bond parameters
Continuation = yes ; restarting after NVT
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
lincs_warnangle = 60 ; maximum angle bond can rotate between steps

; Neighborsearching
cutoff-scheme = Verlet
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT

; Temperature coupling is on
tcoul = V-rescale ; modified Berendsen thermostat
tc-grps = system ; two coupling groups - more accurate
tau_t = 0.1 ; time constant, in ps
ref_t = 300 ; reference temperature in K

; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc = xyz ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme
; Velocity generation
gen_vel = no ; Velocity generation is off
Molecular Dynamics (MD)

; Run parameters
integrator = md ; leap-frog integrator
nsteps = 1500000 ; 0.001 * 1500000 = 1500 ps (1.5 ns)
dt = 0.001 ; 1 fs

; Output control
nstxout = 5000 ; save coordinates every 5.0 ps
nstvout = 5000 ; save velocities every 5.0 ps
nstenergy = 5000 ; save energies every 5.0 ps
nstlog = 5000 ; update log file every 5.0 ps

; Bond parameters
continuation = yes ; Restarting after NPT
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
lincs_warnangle = 90 ; maximum angle bond can rotate between steps

; Neighborsearching
cutoff-scheme = Verlet
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT

; Temperature coupling is on
tcoupl = V-rescale ; modified Berendsen thermostat
tc-grps = system ; two coupling groups - more accurate
tau_t = 0.1 ; time constant, in ps
ref_t = 300 ; reference temperature, one for each group, in K

; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1

; Periodic boundary conditions
pbc = xyz ; 3-D PBC

; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme

; Velocity generation
gen_vel = no ; Velocity generation is off
Simulated Annealing (SA)

; Run parameters
integrator = md ; leap-frog integrator
nsteps = 500000 ; .001 * 50000 = 500 ps (0.5 ns)
dt = 0.001 ; 1 fs

; Output control
nstxout = 5000 ; save coordinates every 5.0 ps
nstvout = 5000 ; save velocities every 5.0 ps
nstenergy = 5000 ; save energies every 5.0 ps
nstlog = 5000 ; update log file every 5.0 ps

; Bond parameters
continuation = yes ; Restarting after MD
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
lincs_warnangle = 90 ; maximum angle bond can rotate between steps

; Neighborsearching
cutoff-scheme = Verlet
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)

; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT

tc-grps = system ; two coupling groups - more accurate
tau_t = 0.1 ; time constant, in ps
ref_t = 300 ; reference temperature, one for each group, in K

; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1

; Periodic boundary conditions
pbc = xyz ; 3-D PBC

; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme

; Velocity generation
gen_vel = no ; Velocity generation is off

; Simulated annealing
annealing = single ; single sequence of points for each T-coupling group
annealing_npoints = 2 ; two points - start and end temperatures
annealing_time = 0 500 ; time frame of heating - heat over period of 500 ps
annealing_temp = 300 0 ; start and end temperatures
Permissions

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Mar 27, 2017

This Agreement between Jacob J Terracina ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number 4070940612133
License date
Licensed Content Publisher John Wiley and Sons
Licensed Content Publication Journal of Molecular Recognition
Licensed Content Title In silico characterization of enantioselective molecularly imprinted binding sites
Licensed Content Author Jacob J. Terracina, Susan T. Sharfstein, Magnus Bergkvist
Licensed Content Date Jan 25, 2017
Licensed Content Pages 1
Type of use Dissertation/Thesis
Requestor type Author of this Wiley article
Format Print and electronic
Portion Full article
Will you be translating? No
Title of your thesis / dissertation Computational Characterization and Optimization of Molecularly Imprinted Polymers
Expected completion date Apr 2017
Expected size (number of pages) 170
Requestor Location Jacob J Terracina
303 Cortland St.

ALBANY, NY 12208
United States
Attn: 12208

Publisher Tax ID EU826007151
Billing Type Invoice

148
TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., (“CCC’s Billing and Payment terms and conditions”), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, and any CONTENT (PDF or image file) purchased as part of your order, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.

For STM Signatory Publishers clearing permission under the terms of
the **STM Permissions Guidelines** only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

- **NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.**

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- **IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION**
SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

• Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

• The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

• This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

• Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

• These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

• In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

• WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

• This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

• This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state’s conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in
such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS
Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License
The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License
The Creative Commons Attribution Non-Commercial (CC-BY-NC) License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License
The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations
Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.
Further details can be found on Wiley Online Library http://olabout.wiley.com/WileyCDA/Section/id-410895.html
This Agreement between Jacob J Terracina ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number
License date
Licensed Content Publisher
Licensed Content Publication
Licensed Content Title
Licensed Content Author
Licensed Content Date
Licensed Content Pages
Type of use
Requestor type
Format
Portion
Will you be translating?
Title of your thesis / dissertation
Expected completion date
Expected size (number of pages)
Requestor Location
Publisher Tax ID
Billing Type
Billing Address
TERMS AND CONDITIONS
This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

• The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

• You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, and any CONTENT (PDF or image file) purchased as part of your order, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

• With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the
You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as
nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party’s right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS
Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License
The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License
The Creative Commons Attribution Non-Commercial (CC-BY-NC) License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. (see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License
The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations
Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Further details can be found on Wiley Online Library http://olabout.wiley.com/WileyCDA/Section/id-410895.html