ENANTIOSELECTIVE MECHANISM OF THE WHELK-O1 CHIRAL STATIONARY PHASE: A MOLECULAR DYNAMICS STUDY

by

Chunfeng Zhao

A thesis submitted to the Department of Chemistry

in conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

(September, 2008)

Abstract

The Whelk-O1 chiral stationary phase is widely used in liquid and supercritical chromatography for the separation of enantiomers. The enantioselective mechanism of the Whelk-O1 chiral stationary phase is the main focus of this thesis. Semi-flexible models are developed based on *ab initio* calculations for the Whelk-O1 selector and a series of chiral analytes. Extensive molecular dynamics simulations are then applied to study the solvation, selectivity and *in silico* optimization of the chiral stationary phase.

The solvation of the Whelk-O1 chiral stationary phase has been explored in a normal phase *n*-hexane/2-propanol solvent, a reversed phase water/methanol solvent, and a supercritical CO_2/methanol solvent. We found that, in all three solvents, the Whelk-O1 selectors are open to the bulk, indicating readiness for docking of analyte. Significant solvent partitioning at the interfaces was noticed, which generates a polarity gradient between the stationary phase and the bulk, and may encourage a high analyte concentration at the interface. Hydrogen bonding activities on the amide hydrogen, amide oxygen, and nitro oxygen of the Whelk selector have also been analyzed.

The selectivity of the Whelk selector was studied by molecular dynamics simulations of analyte docking on the chiral stationary phase. The elution orders and the separation factors for a series of analytes were predicted successfully. We found that hydrogen bonding and π-π stacking interactions are essential for the enantioselectivity as they are strong and specific, and they hold analytes to the cleft region of the Whelk selector. Other interactions, both stabilizing interactions such as the CH-π interaction and the edge-to-face π-π interaction, and destabilizing interactions such as steric hindrance and unfavorable conformational changes also contribute to the enantioselectivity.

We identified a dominant docking arrangement for the most retained enantiomers. Other docking arrangements were found to be more frequent for the least retained enantiomers and these involve interactions with alternative selector sites. Based on the identified enantioselective mechanism obtained from the study, an optimization of the Whelk-O1 chiral stationary phase was undertaken and *in silico* evaluation of the modified chiral stationary phases was carried out. It was demonstrated that restriction of the alternative docking arrangements for the least retained enantiomers could possibly improve the enantioselectivity of the chiral stationary phase.

Acknowledgements

I gratefully acknowledge my supervisor Dr. Natalie Cann for her guidance throughout my graduate work. Her insightful suggestions and direct contributions significantly simplified my work. I also thank Drs. Hugh Horton, Axel Becke, David Wardlaw, and Richard Oleschuk for being on my supervisory committee and many helpful suggestions. I would also like to thank the group members in Dr. Natalie Cann's lab: Jenel Vatamanu, Irina Paci, Sorin Nita, Rodica Pecheanu, and Shihao Wang, for many interesting discussions, shared experiences and friendship.

I thank my parents for all the love they offered me in the past years. They have always been supporting and believing in me. I also wish to address many special thanks to my wife for her wisdom, understanding, and full support.

Financial support from NSERC and Queen's University, and the use of computer resources on HPCVL, SHARCNET and WESTGRID are also gratefully acknowledged.

Statement of Originality

I hereby certify that all of the work described within this thesis is the original work of the author. Any published (or unpublished) ideas and/or techniques from the work of others are fully acknowledged in accordance with standard referencing practices.

Chunfeng Zhao

September, 2008

Table of Contents

Abstract ii
Acknowledgements iv
Statement of Originality v
Table of Contents vi
List of Tables ix
List of Figures xii
List of Abbreviations and Symbols xxiii
Chapter 1 Introduction 1
1.1 Chirality and Chiral Resolution 1
1.2 Chiral Liquid Chromatography 3
1.3 Whelk-O1 Chiral Stationary Phase 10
1.4 Theoretical Studies on Chiral Stationary Phases 12
1.5 Research Aims 17
Chapter 2 Methods and Models 20
2.1 Quantum Mechanics Methods 20
2.1.1 *ab initio* Calculations 20
2.1.2 Density Functional Theory Method 23
2.1.3 Basis Functions and Basis Sets 25
2.1.4 Conformational Minimization and Atomic Charges 27
2.2 Molecular Simulation 28
2.2.1 Monte Carlo and Molecular Dynamics Simulations 28
2.2.2 Potentials 29
2.3 Practical Issues Related to Potential Calculations 33
2.3.1 Periodic Boundary Conditions 33
2.3.2 Truncation of the Lennard-Jones Interactions 35
2.3.3 Ewald Summation Method for Electrostatic Potential 36
2.4 Integrating the Equations of Motion 39
2.4.1 Algorithms for Integration of the Equations of Motion 39
2.4.2 Constraint Dynamics 41
2.4.3 Nosé-Hoover Thermostat 43
2.4.4 Molecular Dynamics of Rigid Bodies 45
2.5 Semi-Flexible Models and the Model CSP Interface 47

2.5.1 Model CSP Interface 48
2.5.2 Molecular Representation of the Whelk-O1 Selector 49
2.5.3 Intramolecular Potential for the Whelk-O1 selector 52
2.5.4 Intermolecular Potential of Whelk and Models for Other Components 59
2.5.5 Molecular Simulation Details 60
2.6 Coding and Other Practical Issues 61
2.6.1 Coding for semi-flexible surface simulation in the MDMC program 61
2.6.2 Message-Passing Interface for Nonbonding Force Calculations 62
2.7 Problems Tackled in This Thesis 63
Chapter 3 Solvation of the Whelk-O1 Chiral Stationary Phase 65
3.1 Introduction 65
3.2 Theoretical Details 66
3.2.1 Models 66
3.2.2 Simulation Details 66
3.2.3 Properties 68
3.3 Results 70
3.3.1 Flexibility of Whelk-O1 in Solvated Environments 70
3.3.2 The Solvation of Whelk-O1 in *n*-Hexane/2-Propanol 74
3.3.3 The Solvation of Whelk-O1 in Supercritical CO_2 81
3.3.4 The Solvation of Whelk-O1 in Water/Methanol 86
3.4 Conclusions 92
Chapter 4 Docking of chiral epoxides on the Whelk-O1 stationary phase 94
4.1 Introduction 94
4.2 Theoretical Details 96
4.2.1 Models of the Whelk-O1 Interface 96
4.2.2 Molecular Representation of Chiral Epoxides and Solvent 99
4.2.3 Molecular Dynamics Simulation Details 101
4.2.4 Properties 103
4.3 Results 104
4.3.1 Whelk-O1 Mobility in *n*-Hexane with Styrene Oxide 104
4.3.2 The Selectivity of Whelk-O1 for (R/S)-Styrene Oxide 106
4.3.3 The Selectivity of Whelk-O1 for (R,R/S,S)-Stilbene Oxide 114
4.3.4 The Influence of 2-Propanol on the Selectivity 122

4.3.5 Separation Factors and Other Considerations 127
4.4 Conclusions 130
Chapter 5 Chiral Recognition of the Whelk-O1 Chiral Stationary Phase 132
5.1 Introduction 132
5.2 Theoretical Details 133
5.3 Results and Discussion 136
5.3.1 Overview 136
5.3.2 The Importance of Hydrogen Bonding on the Selectivity of Whelk-O1 142
5.3.3 The Importance of π-π Stack on the Selectivity of Whelk-O1 146
5.3.4 CH-π Bond-Ring Interactions and Edge-to-Face Ring-Ring Interactions 147
5.3.5 Deficiency in the Analyte and Its Interactions 151
5.3.6 Steric Hindrance and Conformational Change in the Analyte 151
5.3.7 Analyte Docking Locations and Conformational Changes in the Selector 155
5.4 Conclusion 158
Chapter 6 Rational Optimization of the Whelk-O1 Chiral Stationary Phase 160
6.1 Introduction 160
6.2 Theoretical Details 164
6.3 Results and Discussions 165
6.3.1 Overview 165
6.3.2 CSP1 169
6.3.3 CSP2 175
6.3.4 CSP3 176
6.4 Conclusions 178
Chapter 7 Conclusions and Outlook 179
Bibliography 183
Appendix A Interaction Potentials for Chiral Selectors and Analytes 197

List of Tables

Table 1.1: Possible mechanisms and primary interactions for different types of CSPs. …… 9

Table 2.1: Details of the model Whelk-O1 selector. Atomic positions and CHELPG charges are extracted from the B3LYP/6-311+G(2d,p) geometry optimization. Values marked with *a* are obtained from geometry optimization of $SiH_3OSi(CH_3)_2CH_2CH_3$. Lennard-Jones energy and length parameters are taken from the OPLS (unmarked) and CHARMM (marked with b). ……. 55

Table 2.2: Details of the bending potentials for the model Whelk-O1 selector. The first column lists the three atoms that define the angle following the numbering in Fig. 2.6(b). The equilibrium angles, extracted from *ab initio* calculations, are given in the second column. The force constants, and their sources, are given in the third and fourth columns, respectively. ……..…… 56

Table 4.1: Torsional potentials for the Whelk-O1 selector. The potential parameters for Eq. 2.25 are presented. The first row identifies the atoms [see Figs. 2.6(b) for atom numbering]. The coefficients c_i are given in kJ/mol and the angles ϕ_i are in degrees. ……. 98

Table 4.2: Improper torsional potentials for the Whelk-O1 selector. The first column identifies the four atoms defining the improper torsional motion [see Figs. 2.6(b) for atom numbering]. The second and third columns define the equilibrium angle and the corresponding force constant (see Eq. 2.26). ……... 99

Table 4.3: Interaction statistics for R/S-styrene oxide in *n*-hexane. The statistics are based on 15425 snapshots, collected every 1.54 ps, and includes 32 Whelk-O1 selectors and 16 epoxides (racemic) for a total of 246800 interactions. The first series of entries is obtained by examining each analyte, in each snapshot, to ascertain whether it is interacting with one or more selectors. If so, the number of interactions is divided into two categories: one interaction, and two or more interactions. In Mechanism 1, a π-π interaction forms with the dinitrophenyl ring of the selector, while Mechanism 2 includes a π-π interaction with the the phenanthryl group. ……. .. 107

Table 4.4: Interaction statistics for (R,R/S,S)-stilbene oxide in *n*-hexane collected. The analysis is based on 22150 snapshots, collected every 1.54 ps, and includes 32 Whelk-O1 selectors and 16 epoxides (racemic) for a total of 354400 interactions. The first series of entries is obtained by examining each analyte, in each snapshot, to ascertain whether it is interacting with one or more selectors. If so, the number of interactions is divided into two categories: one interaction, and two or more interactions. ……... 117

Table 4.5: Interaction statistics for R/S-styrene oxide in *n*-hexane/2-propanol. The statistics are based on 17500 snapshots, collected every 1.54 ps, and includes 32 Whelk-O1 selectors and 16 epoxides (racemic) for a total of 280000 interactions. The first series of entries is obtained by examining each analyte, in each snapshot, to ascertain whether it is interacting with one or more selectors. If so, the number of interactions is divided into two categories: one interaction, and 2 or more interactions. In Mechanism 1, a π-π interaction forms with the dinitrophenyl ring of the selector, while Mechanism 2 includes a π-π interaction with the phenyl rings of the phenanthryl group. …….. 124

Table 5.1: Interaction summary for the ten analytes on Whelk-O1 extracted from MD snapshot analysis. An interaction consists of either a hydrogen bond or a π-π stack (see Section 5.3.1 for details). The first group of columns reports the percentage of analytes that have 0-3 interactions with the Whelk-O1 selector. A "docked" interaction occurs when an H-bond and a π-π stack are present simultaneously. The central column compares the experimental separation factors with the calculated factors, obtained from the ratio of docking probabilities. The last set of columns describe the probability for four types of docking events, given as follows. In M1 and M2 docking, the analyte forms a hydrogen bond with the amide hydrogen [H(26)], and a π-π stack is formed with either the dinitrophenyl moiety (M1) or the phenanthryl group (M2). In M3 and M4, the analyte forms a π-π stack with the dinitrophenyl moiety and a hydrogen bond with either the

amide oxygen[O(28)], for M3 docking, or with a nitro oxygen [O(36),O(37),O(39),O(40)] for M4. …….. 141

Table 5.2: Hydrogen bonding and π-π stacking statistics. The interaction probabilities between specific sites on the Whelk-O1 selector and the ten analytes. …… 145

Table 5.3: A breakdown of the average energy of analyte-selector pairs with M1 and M2 docking arrangement. The "Total" energies include the intramolecular energy of the Whelk-O1 selector (Sel E) and the analyte (Ana E), and intermolecular docking energy (Int E), which is further divided into contributions from the selector phenanthryl group ("phen"), amide linkage ("amide"), dinitrophenyl group ("DNP"). Tether contributions are not listed but they are between -1.1 and -4.2, and can be extracted from the data provided. All energies are listed in kJ/mol and are averages over all the snapshots collected from the simulations. ……... 150

Table 6.1: Interaction summary for the modified CSPs and Whelk-O1 extracted from MD snapshot analysis. The first four columns give the percentage of analytes that have 0-3 interactions with the CSP. A "docked" interaction occurs when an H-bond and a π-π stack are present with a selector simultaneously. The central column compares the calculated and experimental separation factors. The last five columns provide a breakdown of the docking probability according to the docking mode, M1-M4. M1 docking is further subdivided according to the analyte location: "INSIDE" for inside-of-cleft, and "SIDE" for side-of-cleft. ……. 166

Table 6.2: Hydrogen bonding and π-π stacking statistics for the Whelk-O1 and the modified CSPs. ……... 172

List of Figures

Figure 1.1: Illustration of the three-point interaction model (a) and the four-point interaction model (b). *I, II,* and *III* represent enantiomer *I*, enantiomer *II* and an achiral molecule respectively, and *S* represents the chiral selector. Enantiomer *I* matches the selector while enantiomer *II* does not. ……... 5

Figure 1.2: Chemical structures of some CSPs or chiral selectors. ……... 6

Figure 1.3: Illustration of the three-point interaction model proposed for the Whelk-O1 chiral stationary phase is shown in (a) with two views. The three interactions are: a π- π stack formed with the dinitrophenyl group (shown by arrow 1), a hydrogen bond formed with the amide hydrogen (shown by arrow 2), and either an edge-to-face π-π interaction (shown by arrow 3) or a side chain CH-π interaction with the phenanthryl group (shown by arrow 3'). For clarity, the carbon atoms of the analyte are highlighted in yellow. The cleft formed by the 3,5-dinitrophenyl group and the phenanthryl moiety is shown in (b). …….. 11

Figure 2.1: Lennard-Jones potential as a function of the interatomic separation. Key points on the potential and interatomic separation axes are labelled. …… 32

Figure 2.2: A schematic representation of two dimensional periodic boundary conditions. …... 34

Figure 2.3: The treatment of a set of point charges in Ewald summation as a set of screened charges plus the compensating charges is shown in (a). The three contributions to the electrostatic potential at the site of point charge *i* are shown in (b). …… 37

Figure 2.4: Euler angles $\Omega=(\phi, \theta, \chi)$ relating the space-fixed frame ($F=XYZ$) and body-fixed frame ($g=xyz$). ……... 45

Figure 2.5: The rational design of the simulation cell beginning with the chiral chromatography column and ending with the dimensions of a typical pore. The rightmost picture shows a snapshot

of the simulation cell. In the snapshot, the Whelk-O1 selectors are identified by thick grey lines, the *n*-hexane solvent is represented with thin grey lines, and R- and S-1-phenyl ethanol (analytes) are shown in purple and green, respectively. …… 48

Figure 2.6: (a) Structure of the Whelk-O1 chiral stationary phase. (b) The semi-flexible representation of the Whelk-O1 selector with the two rigid units indicated by shaded areas. The atom numbering shown in (b) will be used throughout this thesis. …….. 50

Figure 2.7: Stable conformers of the (3R,4S) Whelk-O1 selector identified from B3LYP/6-311+G(2d,p) calculations: (a) The global minimum, (b) The local minimum corresponding to a roughly 180^{o} rotation about the C(10)-N(25) bond, (c) The local minimum corresponding to a ring flip in the phenanthryl group. (d) The local minimum for both a ring flip and a C(10)-N(25) bond rotation. Energy differences relative to the global minimum are also given. …….. 53

Figure 2.8: Torsional barriers from *ab initio* calculations. The torsional potentials are given for rotation about (a) C(7)-C(8)-C(9)-C(10), (b) N(25)-C(27)-C(29)-C(30), (c) C(9)-C(10)-N(25)-C(27), (d) C(10)-N(25)-C(27)-C(29). *ab initio* points are represented by filled circles. The torsional potentials used in the simulations are represented by solid lines. …….. 58

Figure 3.1: A snapshot from directly above showing the initial Whelk-O1 interface. 16 Whelk-O1 molecules (coverage of 1.10 μmol/m^2) and 48 trimethyl silyl end-caps (coverage of 3.29 μmol/m^2) are bonded to the underlying silicon surface of 24*24 silicon atoms with a Si-Si distance of 0.205 nm. …… 68

Figure 3.2: A snapshot from directly above illustrating the range of conformers explored by Whelk-O1 molecules. ……... 71

Figure 3.3: Distribution of the Whelk-O1 selector in a *n*-hexane/2-propanol solvent. (a) The distribution of tether atoms from the surface. Open circles, filled squares, open squares, filled diamonds, open diamonds, and filled triangles correspond to O(2), Si(3), C(4), C(5) and C(6), C(7), C(8), respectively. (b) The distribution of the phenanthryl atoms [C(9)-H(24)] from the

surface. The curves with filled and open circles identify the two chiral carbons C(9) and C(10), respectively. (c) The distribution of the amide linkage atoms from the surface. Filled circles, open circles, filled squares, and open squares correspond to N(25), H(26), C(27), O(28) respectively. (d) The distribution of the dinitrobenzoyl atoms from the surface. Filled and open circles correspond to C(29)-C(34) and N(35)-O(40), respectively. ……. 72

Figure 3.4: Simulation results from the *n*-hexane/2-propanol solvated Whelk-O1 interface. (a) A snapshot of the simulation cell. Silicon, oxygen, nitrogen, hydrogen atoms are shown in yellow, red, blue, and white respectively. Carbon atoms are shown in grey with the exception of *n*-hexane where green is used. 2-propanol molecules are highlighted. (b) Distribution of 2-propanol (open symbols) and *n*-hexane (filled symbols) atoms above the surface: CH_3, CH, O and H of 2-propanol are identified by circles, squares, diamonds, and triangles respectively; CH_3, 2-CH_2, and 3-CH_2 of *n*-hexane are represented by circles, squares, and diamonds respectively. (c) Radial distributions between H(26)-O(2-propanol) (filled circles), O(28)-H(2-propanol) (open circles), O(40)-H(2-propanol) (filled squares) and N(25)-H(2-propanol) (open squares). …… 75

Figure 3.5: 2D distribution functions and representative snapshots from the *n*-hexane/2-propanol solvated Whelk-O1 interface. Two views of $g(r_c,z_c)$ for H(26)-O(2-propanol) are shown in (a) and (b). Two views of $g(r_c,z_c)$ for O(28)-H(2-propanol) are shown in (c) and (d). Two views of $g(r_c,z_c)$ for O(40)-H(2-propanol) are shown in (e) and (f). Snapshots showing the H-bonding around H(26), O(28) and O(40) are shown in (g), (h) and (i) respectively with distances shown in Å. Blue lines in (a), (c), and (e) identify z_c=-5.625Å, -3.75Å, -1.875Å, 0.0Å, 1.875Å, 3.75Å, 5.625Å, and 7.50Å. Blue lines in (b), (d), and (f) identify r_c=0.094Å, 1.97Å, 3.84Å, 5.72Å and 7.59Å. ……..... 78

Figure 3.6: Simulation results from the sc-CO_2/methanol solvated Whelk-O1 interface. (a) A snapshot of the simulation cell. Silicon, oxygen, nitrogen, carbon, and hydrogen atoms are shown in yellow, red, blue, grey, and white, respectively. Methanol molecules are highlighted. (b)

Distribution of methanol (open symbols) and sc-CO_2 (filled symbols) above the surface. CH_3, O, and H of methanol are identified with circles, squares, and diamonds respectively. C and O of sc-CO_2 are represented with circles and squares respectively. (c) Radial distributions between H(26)-O(methanol), O(28)-H(methanol), O(40)-H(methanol), N(25)-H(methanol), and H(26)-O(CO_2) are represented with filled circles, open circles, filled squares, open squares, and filled diamonds, respectively. …….. 82

Figure 3.7: 2D distribution functions from the sc-CO_2/methanol solvated Whelk-O1 interface. Two views of $g(r_c,z_c)$ for H(26)-O(methanol) are shown in (a) and (b). Two views of $g(r_c,z_c)$ for O(28)-H(methanol) are shown in (c) and (d). Blue lines in (a) and (c) identify z_c=-3.75Å, -1.875Å, 0.0Å, 1.875Å, 3.75Å, 5.625Å, and 7.50Å. In (b) and (d), the blue lines identify r_c=0.094Å, 1.97Å, 3.84Å, 5.72Å and 7.59Å. ……... 85

Figure 3.8: Simulation results from the water/methanol solvated Whelk-O1 interface. (a) A snapshot showing the location of methanol. Silicon, oxygen, nitrogen, carbon, and hydrogen atoms are shown in yellow, red, blue, grey and green respectively. (b) The equivalent snapshot identifying the position of water. (c) The distribution of methanol (filled symbols) and water (open symbols) above the surface. CH_3, O, and H of methanol are identified with circles, squares, and diamonds respectively. H and O of water are represented with circles and squares respectively. (d) Radial distributions between H(26)-O(methanol), O(28)-H(methanol), O(40)-H(methanol), and N(25)-H(methanol) as identified by filled circles, open circles, filled squares, and open squares, respectively. (e) Radial distributions between H(26)-O(water), O(28)-H(water), O(40)-H(water) and N(25)-H(water) are represented by filled circles, open circles, filled squares, and open squares, respectively. ……. 87

Figure 3.9: 2D distribution functions from the water/methanol solvated Whelk-O1 interface. Two views of $g(r_c,z_c)$ for H(26)-O(methanol) are shown in (a) and (b). Two views of $g(r_c,z_c)$ for O(28)-H(methanol) are shown in (c) and (d). Two views of $g(r_c,z_c)$ for H(26)-O(water) are shown

in (e) and (f). Two views of $g(r_c z_c)$ for O(28)-H(water) are shown in (g) and (h). Blue lines in (a), (c), (e), and (g) identify z_c=-3.75Å, -1.875Å, 0.0Å, 1.875Å, 3.75Å, 5.625Å, and 7.50Å. Blue lines in (b), (d), (f), and (h) identify r_c=0.094Å, 1.97Å, 3.84Å, 5.72Å and 7.59Å. …….. 90

Figure 4.1: A side view of the simulation cell showing two interfaces with Whelk-O1 selectors, end-caps, and silanol groups evident. The solvent consists of *n*-hexane (thin grey lines) and 2-propanol (green), R-styrene oxide (yellow), and S-styrene oxide (purple). …….. 97

Figure 4.2: Global energy minima for (a) R-styrene oxide and (c) (S,S)-stilbene oxide. Panels (b) and (d) show the atomic numbering used throughout, and the identification of rigid units (shaded regions). The numbering for the mirror image molecules is identical for all atoms and is generated by exchanging the phenyl group with the hydrogen at the chiral carbon. The energy minima are obtained from B3LYP/6-311++G** full geometry optimizations. ……. 100

Figure 4.3: The distribution of Whelk-O1 selector atoms above the underlying layer of silicon atoms (defined as z=0). The distribution of C(15), C(18), C(21) are represented by filled circles, open circles, and filled squares in (a). C(29) and C(32) surface distributions are represented by filled circles and open circles in (b). In (c), the distribution of the amide linkage atoms N(25), H(26), C(27), and O(28) are represented by filled circles, open circles, filled squares, and open squares, respectively. …….. 105

Figure 4.4: Distributions between the amide hydrogen, H(26), of the Whelk-O1 selector and O(3) of styrene oxide obtained from MD simulations of racemic styrene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for S- and R-styrene oxide, respectively. 2D distributions are shown in (c) and (d), for the S and R enantiomers, respectively. The difference in 2D distributions is shown in panel (b). …… 108

Figure 4.5: Distributions between the dinitrophenyl ring of the Whelk-O1 selector and the phenyl ring of styrene oxide. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic styrene oxide, in

n-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for the S and R enantiomers, respectively. 2D distributions are shown in (c) and (d), for S and R styrene enantiomers, respectively. The difference in 2D distributions is shown in panel (b). …….. 110

Figure 4.6: Radial distributions between the phenyl ring of styrene oxide and the phenyl rings of the selector phenanthryl group. Distributions relative to P1 and P2 are shown in (a) and (b), respectively. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic styrene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. The solid and dotted lines represent the distributions for the S and R enantiomers, respectively. …… 111

Figure 4.7: A selection of snapshots of Whelk-O1 selectors with S- and R-styrene oxide. The snapshot in (a) shows the dominant docking arrangement for S-styrene oxide. The most probable docking arrangement for R-styrene oxide is shown in (b). Panel (c) shows the secondary three-point docking arrangement for R-styrene oxide. Finally, an outside-of-cleft interaction, for R-styrene oxide, is shown in (d). Ring centers are shown in pale blue. …… 114

Figure 4.8: Distributions between the amide hydrogen, H(26), of the Whelk-O1 selector and O(14) of stilbene oxide obtained from MD simulations of racemic stilbene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for S,S- and R,R-stilbene oxide, respectively. 2D distributions are shown in (c) and (d), for the S,S and R,R enantiomers, respectively. The difference in 2D distributions is shown in panel (b). .. 115

Figure 4.9: Distributions between the dinitrophenyl ring of the Whelk-O1 selector and the phenyl rings of stilbene oxide. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic stilbene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for the S and R enantiomers, respectively. 2D distributions are shown in (c) and (d),

for S and R stilbene enantiomers, respectively. The difference in 2D distributions is shown in panel (b). ……. 118

Figure 4.10: Radial distributions between the phenyl rings of stilbene oxide and the phenyl rings of the selector phenanthryl group. Distributions relative to P1 and P2 are shown in (a) and (b), respectively. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic stilbene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. The solid and dotted lines represent the distributions for the S,S and R,R enantiomers, respectively. …….. 119

Figure 4.11: A selection of snapshots of Whelk-O1 selectors with (S,S)- and (R,R)-stilbene oxide. The panels on the left hand side provide an alternate view of the same docking arrangement. The snapshot in (a) shows the dominant docking arrangement for (S,S)-stilbene oxide. The most probable docking arrangement for (R,R)-stilbene oxide is shown in (b). Panel (c) shows the secondary docking arrangement for (R,R)-stilbene oxide. Finally, a bridging interaction, for (R,R)-stilbene oxide interacting with two selectors, is shown in (d). Ring centers are shown in pale blue. …… 121

Figure 4.12: Radial distribution functions for resolution of R/S-styrene oxide on (3R,4S)-Whelk-O1 CSP in a hexane/2-propanol solvent. Solid lines are for S-styrene oxide, dotted lines are for R-styrene oxide. (a)H(26)-O(3), (b)DNP-P, (c)P1-B, (d)P2-P, (e)O(styrene oxide)-H(2-propanol). The radial distribution of 2-propanol relative to the CSP are shown in (f), where the solid line is for H(26)-O(2-propanol), the dashed line is for O(28)-H(2-propanol), and the dotted line is for O(40)-H(2-propanol). ……. 123

Figure 4.13: Snapshots of Whelk-O1 selectors with S/R-styrene oxide in hexane/2-propanol. The panels highlight the role of 2-propanol on solvation and selectivity. Hexane is represented by thin grey lines, while 2-propanol is green (alkyl groups), red (oxygen) and white (hydrogen). The snapshot in (a) shows several 2-propanols hydrogen bonding in the cleft region of a Whelk-O1

selector. The hydrogen bonding between styrene oxide and 2-propanol in the bulk is illustrated in (b). The subsequent four panels (c) illustrate the replacement of docked S-styrene oxide by 2-propanol. ……. 126

Figure 4.14: The dynamics of R/S-styrene oxide docking on Whelk-O1 selectors in *n*-hexane. Results from two simulations are shown. Each row follows the trajectory of a single enantiomer over a 1.15 ns simulation and the upper and lower panels show results for S- and R-styrene oxide respectively. The line colours identify the most strongly interacting Whelk-O1 selectors. For each enantiomer, interaction with the first selector is identified by a red line, and the subsequent selectors are identified by blue, orange, and green. For each enantiomer, the vertical scale is defined by the following interactions: a hydrogen bond is a 0.5, a π-π interaction with the dinitrophenyl group is given a value of 0.3, while a π-π interaction with either the P1 or P2 rings is worth 0.1. In this way, a docking consistent with Mechanism 1 of Table 4.3 corresponds to 0.8 while Mechanism 2 corresponds to 0.6. The dotted lines are placed at 0.5. ……... 129

Figure 5.1: Structures of the ten analytes. the B3LYP/6-311++G** optimized structures are provided for: (S)-Styrene oxide (S-STYO); (S,S)-Stilbene oxide (S,S-STBO); (S)-2-methyloxirane (S-MOXR); (R,R)-1,2-diphenylcyclopropane (R,R-DPCP); (R)-1-phenylethanol (R-PEOL); (S)-1-phenylethanamine (S-PEAM); (S)-1-phenylethane-1,2-diol (S-PEDO); (S,S)-1,2-diphenylethane-1,2-diol (S,S-DPED); (S)-2-hydroxy-1,2-diphenylethanone (S-DPEO); (S)-N-(1-(4-bromophenyl)ethyl)pivalamide (S-PAMD). The shaded areas are treated as rigid units in the simulations. ……... 135

Figure 5.2: The definition of a π-π stack (a) The distance between ring centers (R) must be less than 4.6 Å and the ring tilt angle (θ) is less than 30^{o}. (b) The average ring tilt angle as a function of the center-to-center distance between the DNP ring of the Whelk-O1 selector and the aromatic ring of STYO (S: red solid, R: red dashed) and DPEO (S: blue solid, R: blue dashed). …… 138

Figure 5.3: Four docking arrangements (M1-M4) on Whelk-O1. Snapshots show S-PEOL, docked according to each mechanism. Red arrows identify hydrogen bonding and π-π stacking interactions. Note that other selectors, end-caps, and the solvent are omitted from the snapshots to highlight the docking arrangements. ……... 139

Figure 5.4: A comparison of DPCP and STBO. Radial distributions (a) between H(26) and either the epoxide oxygen of STBO (S,S: Blue solid, R,R: Blue dashed) or the unsubstituted carbon of the DPCP cyclopropane ring (S,S: Red solid, R,R: Red dashed). Snapshots of R,R-DPCP with Whelk-O1 are shown in (b). The dominant docking modes for S,S-STBO (c) and R,R-STBO (d) are shown for comparison. For clarity, analyte carbons are yellow. …….. 144

Figure 5.5: Illustrative snapshots of MOXR interacting with the chiral selector are provided in (a). The docking of S-STYO and R-STYO are shown in (b) and (c), respectively. S-STYO strongly prefers to dock according to M1 while R-STYO docks according to M1 or M2. …… 147

Figure 5.6: Snapshots of M1 and M2 docking of PEDO, DPED, and DPEO. *n*-hexane, end-caps, other selectors, etc. are omitted from the snapshots for clarity. ……. 153

Figure 5.7: Analyte location for M1 docking events. Three locations are identified (a): "S" for on the side of the cleft (Green), "A" for above the cleft (Red), and "B" for below the cleft (Blue). Two angles are used to identify the location: the N(25)-C(27)-C(29)-C(30) dihedral (Φ), and the angle (Θ) between the vector normal to the DNP ring and the vector joining the DNP ring center to the analyte ring center. Snapshots of docked S-PEOL are shown in (b) to illustrate the criteria. Two views of R-PAMD, which docks ONSIDE, are shown in (c) and (d). Two views of S-PAMD, which docks ABOVE, are shown in (e) and (f). The probability distribution for Φ when S-PAMD (red) and R-PAMD (blue) are docked is shown in (g). Analyte location probabilities for eight analytes are given in (h). …….. 156

Figure 6.1: Lowest energy 3D structures, obtained from DFT calculations, and chemical structures of the selectors of the Whelk-O1 CSP and the modified CSPs. For CSP1, the two nearly-degenerate conformations of the saturated ring are shown. …… 163

Figure 6.2: Illustration of docking arrangements. The analyte is composed of an aromatic ring (yellow hexagon), a hydrogen bonding site (yellow circle), and other structural elements(white square). The selector is composed of a tether (white rectangle) and two aromatic regions - phenanthryl group (blue rectangle) and a dinitrophenyl group (green corns for the nitro groups and blue hexagon for the phenyl group) - connected by an amide bridge (green rectangle). The four docking arrangements, M1-M4, on Whelk-O1 are shown. In addition, M1 docking is subdivided according to analyte location. Specifically, inside-of-cleft and side-of-cleft docking is shown above. Red arrows identify hydrogen bonding and π-π stacking interactions. Red dotted arrowd identify secondary CH-π or edge-to-face π-π interaction. …….. 168

Figure 6.3: Surface atomic distributions for selected atoms before and after selector modification. The atoms being compared are shown in (a). The atom colors correspond to the curve colors, and the Whelk-O1 results are represented by solid lines while the modified CSP distributions are given by dashed lines. In (b), atomic distributions for Whelk-O1, CSP1a (dotted), and CSP1b (dashed) are compared. A comparison of CSP2 and Whelk-O1 is given in (d) while CSP3 and Whelk-O1 are compared in (d). …….. 169

Figure 6.4: A comparison of radial distributions for CSP1a, CSP1b, and Whelk-O1 for (R/S)-styrene oxide and (R,R)/(S,S)-stilbene oxide docking. The aromatic regions of Whelk-O1 are identified by P1, P2, and DNP in (a). Radial distributions between the aromatic ring of S-STYO(red) or R-STYO(blue), and the P2 ring of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b(open squares) are shown in (b). Similarly, distributions between the aromatic ring of S-STYO(red) or R-STYO(blue), and the P1 ring of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b(open squares) are shown in (c). Radial distributions for S,S-

STBO(red) and R,R-STBO(blue) are shown in panels (d) and (e). Distributions between the ring centers and P2 of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b (open squares) are shown in (d). Distributions between the ring centers and P1 of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b (open squares) are shown in (e). …… 174

Figure 6.5: Torsional flexibility comparison between the dinitrophenyl ring of the Whelk-O1 selector and the dinitroisoindole ring in the CSP2 selector. The solid line with filled circles shows the torsional potential of N(H)-C(O)-C-C of the Whelk while the dotted line with open circles identify the torsional potential between C-N(H)-C-N of CSP2. …….. 175

List of Abbreviations and Symbols

Abbreviations

2D	2-Dimensional
3D	3-Dimensional
AA	All Atom
AO	Atomic Orbital
CGTF	Contracted Gaussian-Type Functions
CHARMM	Chemistry at HARvard Molecular Mechanics
CHELPG	CHarges from ELectrostatic Potentials using a Grid based method
CI	Configuration Interaction
CMPA	Chiral Mobile Phase Additive
CSP	Chiral Stationary Phase
DFT	Density Functional Theory
GTF	Gaussian-Type Function
H-bond	Hydrogen bond
HF	Hartree-Fock
HPLC	High Performance Liquid Chromatography
LJ	Lennard-Jones
MC	Monte Carlo
MD	Molecular Dynamics
MEP	Molecular Electrostatic Potential
MP	Møller-Plesset
MPI	Message-Passing Interface
NAG	Numerical Algorithms Group

NMR	Nuclear Magnetic Resonance
NSAID	NonSteroidal Anti-Inflammatory Drug
OPLS	Optimized Potential for Liquid Simulations
RU	Rigid Unit
sc-CO_2	Supercritical CO_2
SCF	Self-Consistent Field
SFC	Supercritical Fluid Chromatography
STO	Slater-Type Orbital
UA	United Atom

Symbols of Analytes and Chemical Groups (Fig. 5.1)

DPCP	1,2-Diphenylcyclopropane
DPED	1,2-Diphenylethane-1,2-diol
DPEO	2-Hydroxy-1,2-diphenylethanone
MOXR	2-Methyloxirane
PAMD	(n-(1-(4-bromophenyl)ethyl)pivalamide
PEAM	1-Phenylethanamine
PEDO	1-Phenylethane-1,2-diol
PEOL	1-Phenylethanol
STBO	Stilbene oxide
STYO	Styrene oxide
P	Phenyl group of analytes
P1	Middle ring of the phenanthryl group of Whelk-O1 or modified CSPs (see Fig. 2.6)
P2	Topmost ring of the phenanthryl group of Whelk-O1 or modified CSPs (see Fig. 2.6)
DNP	Dinitrophenyl group

Symbols of Docking Arrangements

M1 Form an H-bond with amide hydrogen and a π-π stack with DNP with a selector simultaneously

M2 Form an H-bond with amide hydrogen and a π-π stack with phenanthryl with a selector simultaneously

M3 Form an H-bond with amide oxygen and a π-π stack with DNP with a selector simultaneously

M4 Form an H-bond with nitro oxygen and a π-π stack with DNP with a selector simultaneously

Chapter 1
Introduction

1.1 Chirality and Chiral Resolution

A chiral molecule is one that is not superimposable on its mirror image. It has no mirror plane, center of inversion, or rotation-reflection axis [1]. Chiral molecules bearing a mirror image relationship are referred to as enantiomers. Enantiomers have essentially identical physical (except for optical activity) and chemical (except in a chiral environment) properties [2]. However, the essential products of life are composed of many chiral molecules, such as amino acids, sugars, proteins, and nucleic acids. Because of this natural asymmetry, the unique handedness of a chiral molecule has both intricate and dramatic influence on how they interact with biological systems. As a result, two enantiomers of a chiral drug may have dramatically different pharmacological effects [3-6]. Development of single-enantiomer drugs is now strongly encouraged by administrations in Canada, Europe, Japan, and United States [2,7]. According to C&EN news [8], in 2006, 80% of small-molecule drugs approved by the United States Food and Drug Administration were chiral and 75% of these were single enantiomers. It is now an essential task for chemists to obtain pure enantiomers.

Pure enantiomers can be obtained in several ways: synthesis starting with materials from a chiral pool, asymmetric synthesis, or chiral resolution of racemates. Two notable examples of asymmetric synthesis are chirally catalyzed hydrogenation reactions [9,10] and chirally catalyzed oxidation reactions [11]. However, tailored asymmetric synthetic routes are very difficult to develop. The pharmaceutical industry usually does not have the luxury of waiting for a perfect synthesis route to come up. Thus, chiral resolution methods are routinely used in the industry [12]. Resolution methods are also widely used by organic chemists for characterization of asymmetric reactions.

Briefly, chiral resolution methods may be divided into three categories based on the use of chiral selectors: spontaneous resolution, indirect resolution, and direct resolution. The first demonstration of chiral resolution was achieved by Louis Pasteur [13] in mid 19th century. Pasteur noticed two forms of crystals of sodium ammonium tartrate and manually sorted them. The crystallization method used by Louis Pasteur is a spontaneous resolution where the chiral selectors are the enantiomers themselves [14]. In spontaneous resolution, packing of one enantiomer with copies of itself is thermodynamically or kinetically more favourable than packing with its mirror image counterpart. This method has been used in the manufacture of L-dopa [15]. Being a fascinating phenomenon, spontaneous resolution of enantiomers, however, is very uncommon [16]. Thus, resolution of enantiomers more often involves interaction with chiral selectors other than themselves.

In indirect chiral resolution the enantiomers are coupled with an auxiliary chiral reagent to form diastereomers [1]. As the diastereomers have different physical and chemical properties even in achiral environments, they can be separated by appropriate achiral separation techniques. Pure enantiomers are obtained from the decomposition of the separated diastereomers [17]. The resolution of bases and acids via diastereomeric salt formation followed by crystallization is an application of indirect resolution. This method is common in industry due to its preparative ability and low cost. For example, Kim *et al.* [18] resolved basic intermediate of flavopyridol, a potent anticancer agent, with the addition of dibenzoyl-D-tartaric acid. Grella *et al.* [19] resolved two aryloxycarboxylic acids using cinchonidine as a resolving agent.

Direct chiral resolution, where the enantiomers to be separated are not modified, is another resolution method. Here the enantiomers are separated due to the different free energy changes upon formation of transient diastereomeric complexes with a chiral selector. Typical examples include chiral liquid chromatography using a chiral stationary phase [20] and chiral simulated moving bed chromatography [21,22]. Chiral liquid chromatography has become one of the most

versatile techniques for resolution of racemates. It can be used both in an analytical scale [23] and in a preparative scale [24]. Fast development in this field in the past three decades enabled a large spectrum of racemates to be separated [25]. In the following section, a brief overview of chiral liquid chromatography will be given.

1.2 Chiral Liquid Chromatography

There are two approaches for chiral resolution with liquid chromatography: using chiral mobile phase additives (CMPAs) and chiral stationary phases (CSPs). In the former approach, a suitable chiral selector is added into the mobile phase to form diastereoisomeric complexes with the two enantiomers, which are then separated on the achiral stationary phase. Debowski *et al.* [26] used β-cyclodextrin as a CMPA to separate racemic mandelic acid. Since then, cyclodextrins have become the most common CMPAs [27]. Other commonly used CMPAs include ligand exchangers [28], and proteins [29]. Chiral chromatography using CMPAs has several severe drawbacks: consumption of chiral additives, necessity to recover enantiomers from diastereomers, and detection problems due to UV absorption by the additives. Thus, the use of CSPs, instead of CMPAs, is nowadays the common approach in chiral liquid chromatography [25]. The CSPs are formed by integrating the chiral selectors into the stationary phases.

Chiral resolution by CSP is through formation of transient diastereomeric complexes between the enantiomers and the immobilized chiral selector [30]. The transient complexes are stabilized by a number of interactions such as hydrogen bonding, π-π interactions, dipole-dipole interactions, and inclusion complexation [25]. Due to their distinct stereo configurations, enantiomers interact differently with the chiral selector when the transient diastereomeric complexes are formed. Chiral resolution results from the difference in free energies between the two diastereoisomers. The enantiomer that fits best is retained most by the CSP and elutes out last from the chromatography column.

Several geometric models have been proposed to explain the difference in molecular fit between the enantiomers and the selector due to the distinct stereo configurations of enantiomers. A three-point attachment model was originally proposed by Easson and Stedman [31] in 1933 to explain the recognition of adrenaline and miotine enantiomers. Over the years, this model has been generalized from three attachments to three interactions [32], in recognition of the fact that some interactions may be repulsive. According to this intuitive model, three interactions – either attractive or repulsive - are identified as the origin of the discrimination. Simultaneous occurrence (or avoidance for repulsive interactions) of the three interactions is regarded as favourable. For example, in Fig. 1.1(a), object *I* interacts with selector *S* better than its mirror-image, object *II*. Object *III* is achiral and its left-front face fits the selector *S* well. At present, three-point interaction models are generally invoked to explain the selectivity of CSPs [30], especially the brush-type CSPs [33-35]. In actual fact, the number of interactions required for chiral selectivity is a contentious issue, with some claiming that two interactions are required [36] while others specify up-to four interactions [37,38]. Fig. 1.1(b) illustrates the four-location model: Object *I* and object *II* can both fit the three locations (A', B', C' of selector *S*) well, but object *I* can also favourably match the fourth location D' while its mirror-image object cannot. Sundaresan and Abrol [39] have recently argued that three- and four- point models are equivalent in biological chiral recognition, with the restriction on direction of approach providing an implicit fourth interaction.

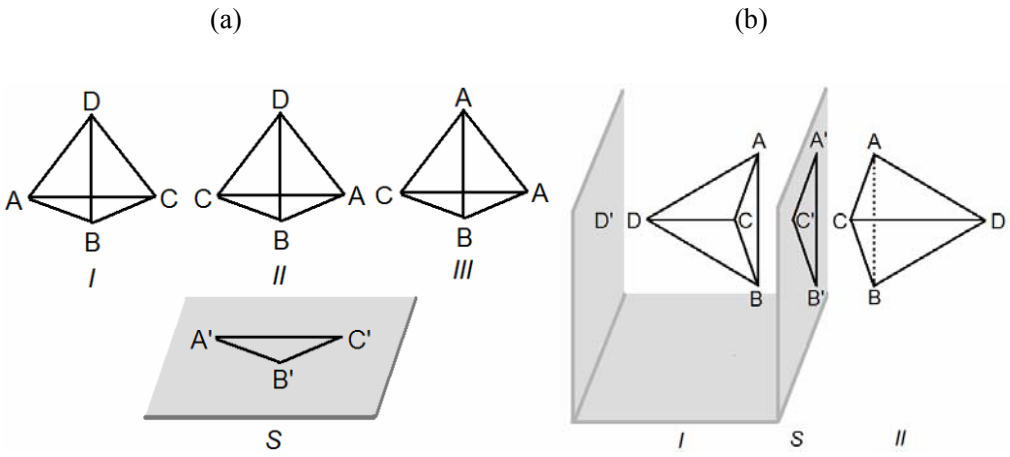


Figure 1.1: Illustration of the three-point interaction model [31] (a) and the four-point interaction model [37,38] (b). *I, II,* and *III* represent enantiomer *I*, enantiomer *II* and an achiral molecule respectively, and *S* represents the chiral selector. Enantiomer *I* matches the selector while enantiomer *II* does not.

The most common types of CSPs [25] include the polysaccharide-based CSPs, cyclodextrin-based CSPs, brush-type CSPs, protein-based CSPs, and macrocylic glycopeptide antibiotics based CSPs. The structure of a CSP for each type, with exception of the complicated protein-based CSPs, is shown in Fig. 1.2. In the rest of this section, a brief description of these CSP types is given. Their enantioselective mechanisms are also discussed.

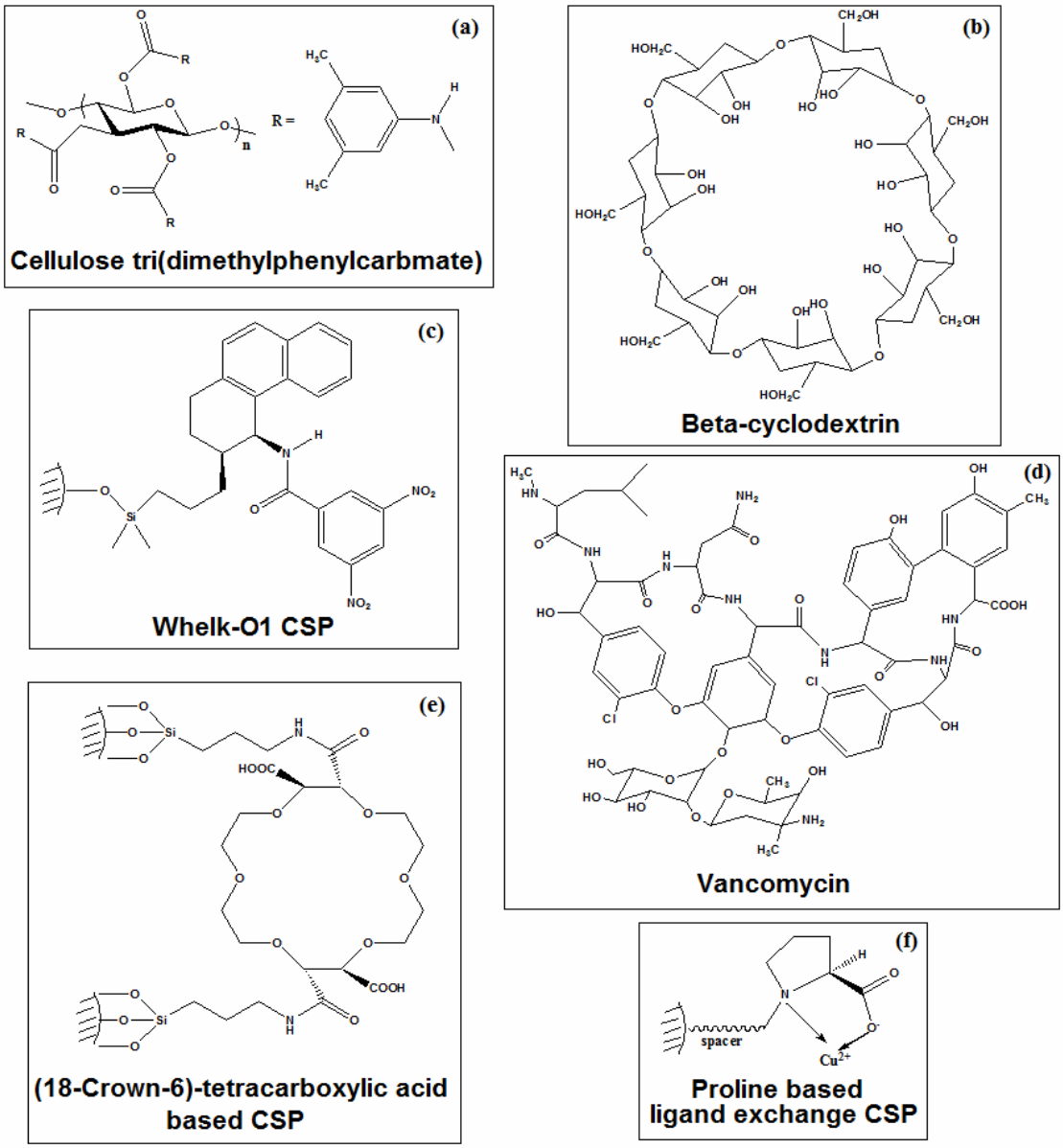


Figure 1.2: Chemical structures of some CSPs or chiral selectors [24,25,40].

Polysaccharides, such as cellulose and amylose, are abundant natural chiral polymers. They can be modified readily to carbamates or esters [41]. The most popular polysaccharide-based selectors are derivatized 3,5-dimethylphenyl carbamates of cellulose [Fig. 1.2(a)] [42,43] and amylose [44]. This type of CSP is broadly discriminating because it possesses multiple chiral centers as well as a long-range helical secondary structure. Hesse and Hagel [45] proposed that the inclusion of enantiomers into the helical chiral grooves is essential for the chiral discrimination by cellulose triacetate. However, for the phenylcarbamate derivatives, hydrogen bonding to the carbamate group may be the most important attractive interaction [41,46]. Aboul-Enein and Ali [47,48] observed that π-π interaction could also be an important binding force for the chiral resolution of aromatic racemates. Therefore, the interaction mechanism has only been partially established [30].

Cyclodextrins (CDs) are small cyclic oligosaccharides with six, seven, or eight glucopyranose units for the α-, β-, or γ- CD, respectively. Fig. 1.2(b) shows the structure of a β-CD. Each cyclodextrin has a cone-shaped cavity formed by its glucopyranose units. Cyclodextrins are widely used in chiral chromatography both as chiral mobile phase additives and chiral stationary phases and vast literature is available [27]. Hamilton and Chen [49] co-crystallized β-CD and racemic fenoprofen, showing the formation of inclusion complexes. The inclusion complexes are stabilized by a number of interactions including hydrogen bonding, dipole-induced dipole interactions, hydrophobic interactions, and π-π interactions (for CSPs based on cyclodextrin aromatic derivatives), and these interactions have been identified as the source of chiral discrimination [49-51]. As the space inside the cavity of cyclodextrin is highly restricted, chiral resolution is affected by even a slight variation in the chromatographic conditions, such as a change in the mobile phase [51] and the derivatization of the CD [52]. As a result, the exact mechanism of chiral resolution on these phases is still not fully understood [25].

Brush-type CSPs originated in 1976 when Mikes *et al.* [53] attached R(−)- and S(+)-2-(2,4,5,7-tetranito-9-fluorenylideneaminooxy)propionic acid and three other chiral charge-transfer agents to silica. This type of CSP has the selectors bonded to a surface via a tether, giving the appearance of a brush. The chiral selector usually contains a π-electron donating group or/and a π-electron accepting group [Fig. 1.2(c), for example]. The three-point interaction model is usually invoked to explain the chiral resolution mechanism of brush-type CSPs [33-35]. Crystallography [54,55] and NMR [56,57] studies indicated π-π stacking and hydrogen bonding to be major interactions for this type of CSP. Other interactions such as dipole-dipole interactions and steric hindrances have also been proposed to be part of the mechanism [58,59]. However, a consensus has not been reached on the enantioselective mechanism of this type of CSP.

Protein-based CSPs, pioneered by Stewart and Doherty [60], and macrocyclic glycopeptide antibiotics-based CSPs, pioneered by Armstrong *et al.* [61], are logical choices since proteins are responsible for the chiral recognitions in the body. Fig. 1.2(d) shows the structure of vancomycin, which has been shown to be a versatile chiral selector [60]. The chiral recognition mechanism for these CSPs is elusive because there are many chiral centers and functional groups and thus many kinds of interactions may be involved. The use of these types of CSPs, especially the macrocyclic glycopeptide antibiotics-based CSPs, for chiral resolution is rapidly increasing [62,63]. A wide range of charged and neutral molecules have been shown to resolve on these CSPs [62,63]. However, as only part of the CSPs are involved in chiral resolution, the disadvantage of these types of CSPs is their limitation on preparative applications [24].

Other major types of CSPs include crown ether-based CSPs [Fig. 1.2(e)] [64,65] and ligand-exchange CSPs [Fig. 1.2(f)] [66,67]. These types of CSPs are limited due to their narrow range of applications. Crown ether-based CSPs have inclusion of NH_3^+ ion into the cavity as the major driving force for docking of enantiomers [68], and thus they are usually only used to discriminate enantiomers with a primary amine group [69]. Ligand-exchange CSPs interact with

enantiomers through the exchange of the chiral ligands and the enantiomers of specific metal ions through coordinated bonds [66,67]. As a result, they are only useful for chiral resolution of enantiomers containing electron donating atoms or groups.

As a summary, the docking mechanism and primary interactions for different types of CSPs are listed in Table 1.1. Two points can be drawn from the above discussion. First, a full understanding of the chiral separation mechanisms is still not available for most CSPs. Thus, the development and application of all these types of CSPs involves a trial-and-error process. For example, in the selection of a CSP to separate a particular racemate, one usually empirically considers the primary interactions available and goes through a trial process [30]. Second, within the most popular CSPs, such as the polysaccharide-based cellulose/amylose tri(3,5-dimethylphenylcarbamate) CSPs, the antibiotics-based CSPs, the protein-based CSPs, and most of the brush-type CSPs, there are common functional groups. Hydrogen bonding, π-π, and dipolar interactions are among the major interactions.

Table 1.1: Possible mechanisms and primary interactions for different types of CSPs.

CSP	**Primary mechanism**	**Primary interaction**
Polysaccharide-based	Insertion into helical structures	Hydrogen bonding, π-π, dipolar, or steric
Cyclodextrin-based	Inclusion	Hydrogen bonding or hydrophobic
Brush-type	Three-point interaction	π-π, hydrogen bonding
Protein-based	Multiple binding sites	Variable
Antibiotics-based	Multiple binding sites	Variable
Crown ether-based	Inclusion	Ion-dipole
Ligand-exchange	Diastereoisomeric selector-metal ion-analyte complex	Coulomb or ion-dipole

1.3 Whelk-O1 Chiral Stationary Phase

As discussed above, brush-type CSPs are characterized by attaching a small chiral selector onto a silica support. Early development and application of these types of CSPs had attracted much attention [25,53,70]. The reciprocity concept put forth by Pirkle *et al.* allowed systematic development of brush-type CSPs [58]. Briefly, if a CSP derived from (S)-A can separate (S)-B and (R)-B, then a CSP derived from (S)-B may separate (S)-A and (R)-A. The reciprocal process helped to discover a series of CSPs in Pirkle's lab, including dinitrobenzoyl-phenylglycine-derived CSP [71], N-(2-naphthyl)-valine derived CSP [72], and phthalide derived CSP [73].

The Whelk-O1 chiral stationary phase was designed in the early 90's [40,74] for the separation of nonsteroidal anti-inflammatory drugs (NSAIDs). Among the NSAIDs, naproxen was sold as a single enantiomer [75], and there had been considerable interest in the resolution of naproxen [76]. Pirkle *et al.* [40,74] immobilized (S)-naproxen and tested a variety of electron-deficient amine derivatives on the naproxen-based CSP. Using the reciprocal idea, they ended up with the Whelk-O1 chiral stationary phase. A separation factor of 2.93, which was significant at that time, was obtained at room temperature for racemic naproxen.

Shown in Fig. 1.2(c), the Whelk-O1 chiral stationary phase is based on 1-(3,5-dinitrobenzamido)-1,2,3,4-tetrahydrophenanthrene. Throughout this thesis, the term "Whelk selector" will refer to a single 1-(3,5-dinitrobenzamido)-1,2,3,4-tetrahydrophenanthrene molecule. The selector consists of an electron-deficient dinitrophenyl group, an electron-rich 1,2,3,4-tetrahydrophenanthrene group (will be referred to as the phenanthryl group), and an amide tether that connects them. The selector is attached to silica through a hydrocarbon tether and a siloxane linkage to the silica surface. There are two chiral centers in the Whelk selector and both the (3R, 4S)-form and (3S, 4R)-form of the chiral stationary phase have been commercialized [77].

It was found that besides NSAIDs, racemates from a number of families including amides, epoxides, esters, ureas, carbamates, ethers, aziridines, phosphonates, aldehydes, ketones,

carboxylic acids, and alcohols can be resolved by the Whelk-O1 chiral stationary phase [77-88]. The phase is compatible with all commonly used mobile phases, including organic solvents [77,81], aqueous systems [89], and supercritical/subcritical CO_2 solvent [90-92]. The Whelk-O1 CSP was also shown to have preparative capacity [93-95]. It has proven to be the most general of the brush-type CSPs developed to-date [30].

Typically, the enantiomers of a compound having an aromatic substituent and a hydrogen bond acceptor near a chiral center can be resolved on the Whelk-O1 CSP. Two three-point interaction models have been proposed to explain the selective mechanism of Whelk-O1. The three interactions are shown in Fig. 1.3(a). First, a π-π stack is formed between the electron deficient dinitrophenyl group of Whelk-O1 and an aryl group in the analyte. Second, a hydrogen bond is formed between the Whelk amide hydrogen and a hydrogen bond acceptor in the analyte. Third, according to different opinions, either CH–π interactions (interaction 3') [96] or edge-to-face π-π interactions (interaction 3) [55,57,77,97] are formed between the phenanthryl moiety and the analyte.

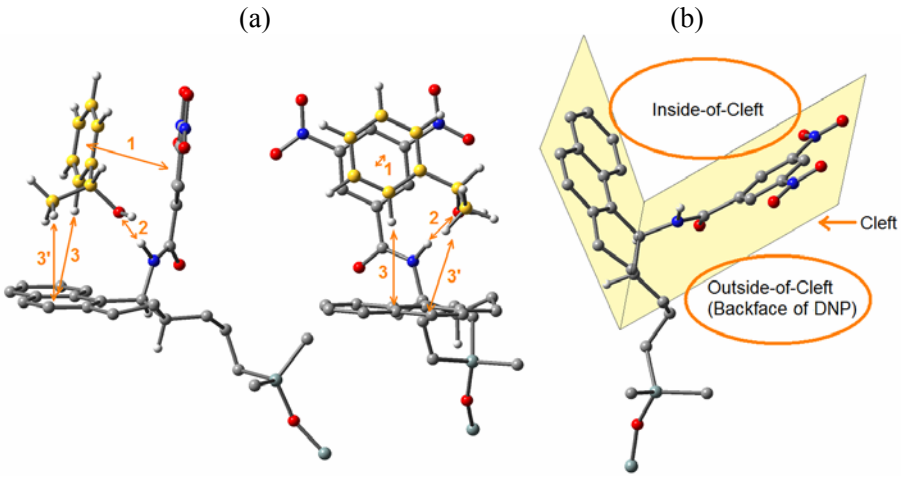


Figure 1.3: Illustration of the three-point interaction model proposed for the Whelk-O1 chiral stationary phase is shown in (a) with two views. The three interactions are: a π- π stack formed with the dinitrophenyl group (shown by arrow 1), a hydrogen bond formed with the amide hydrogen (shown by arrow 2), and either an edge-to-face π-π interaction (shown by arrow 3) [55,57,77,97] or a side chain CH-π interaction with the phenanthryl group (shown by arrow 3') [96]. For clarity, the carbon atoms of the analyte are highlighted in yellow. The cleft formed by the 3,5-dinitrophenyl group and the phenanthryl moiety is shown in (b).

Koscho *et al*. [55,57] further pointed out that the cleft region formed by the phenanthryl group and the dinitrophenyl group plays an essential role in discriminating racemates [Fig. 1.3(b)]. They proposed that the more retained enantiomer docks inside of the cleft, and thus it can form the three interactions simultaneously, while the less retained enantiomer docks outside of the cleft, and cannot form the edge-to-face π-π interaction with the Whelk phenanthryl group. A recent co-crystallization study [55] between the Whelk selector and n-(1-(4-bromophenyl) ethyl)pivalamide supported this mechanism. In the crystal structures of the selector-analyte complexes, the more retained enantiomer docked inside of the cleft while the less retained enantiomer docked outside of the cleft. NMR studies [57,97] of a variety of analytes in the presence of a Whelk-O1 selector are also consistent with this mechanism. The nuclei which are part of, or between, the main interaction sites of the analyte, of the more retained enantiomers are more shielded than those of the less retained enantiomers because of the shielding effect of the cleft formed by the aromatic rings.

1.4 Theoretical Studies on Chiral Stationary Phases

Theoretical studies of separations can generally be divided into three categories: force field exploration of selector-analyte complexes, selectivity predictions via structure-function relationships, and simulations. These will be discussed in this section.

The separation of enantiomers is governed by two equilibria of enantiomers forming diastereomeric complexes with the selector:

$$S + A^R \Leftrightarrow S \bullet A^R \qquad 1.1$$

$$S + A^S \Leftrightarrow S \bullet A^S \qquad 1.2$$

S refers to the chiral selector and A is the analyte. The superscripts *R* and *S* are stereochemical descriptors [98]. The equilibrium constants of these two equilibria are:

$$K^R = \frac{[S \bullet A^R]}{[S][A^R]} \qquad 1.3$$

$$K^S = \frac{[S \bullet A^S]}{[S][A^S]} \quad 1.4$$

The separation factor α is the ratio of the two equilibrium constants:

$$\alpha = \frac{K^{more-retained}}{K^{less-retained}} \quad 1.5$$

Note that, by definition, α is always greater than or equal to 1. The equilibrium constant is related to free energy change ΔG by

$$\Delta G = -RT \ln K \quad 1.6$$

where R and T are the ideal gas constant and the absolute temperature. Comparing Eq. 1.5 and Eq. 1.6 we have:

$$\Delta\Delta G = \Delta G^{more-retained} - \Delta G^{less-retained} = -RT \ln \alpha \quad 1.7$$

Eq. 1.7 shows that the separation factor α is determined by the difference between the free energy changes of the two equilibria.

Thus, the differential free energy change, $\Delta\Delta G$, can be used to predict the separation factor. Recognizing that single solvated enantiomers have the same free energies, one only needs to compute the free energy difference of the diastereomeric complexes to obtain $\Delta\Delta G$. To date, the majority of theoretical studies of chiral separations are through force field exploration of the docked diastereomeric complexes [98]. There have been several major concerns with these calculations.

First, one needs to decide where and how to place the analyte around the selector. A "motif-based" strategy has been used in most early attempts [99]. Here one takes advantage of the known interactions that occur between the selector and the analyte. For example, one constrains the distance between a selector-analyte hydrogen bonding pair and then searches the potential energy surface for docking arrangements. This approach works quite well for focusing on the important part of an otherwise complex potential energy surface. Nonetheless, the results may be biased by using preconceived notions about docking. Some authors have tried to search

the entire potential energy surface by 3-dimensional grid-guided search [100]. In this method, the selector is usually fixed at the origin and the analyte is sequentially placed at each grid point and allowed to rotate to find the lowest-energy docking arrangement.

Second, one needs to decide what structures of selector and analyte should be used. Clearly, using only the lowest energy conformations of each molecule would be inappropriate as the most effective docked conformations of the selector and the analyte need not be the global minima. Hence, all reasonable conformations of selector and analyte should be explored. For example, Lipkowitz *et al.* [101] first searched the conformational shapes of the selector and the analyte, and a grid search for docking conformations was then carried out for all the possible combinations of rigid bodies of the selector and the analyte. Recently, Monte Carlo and molecular dynamics algorithms have been utilized in docking arrangement searches [96,102]. Usually, a selector and an analyte are constrained within certain distances and Monte Carlo or molecular dynamics simulations are carried out to sample the low energy docking arrangements. This way, the selector and the analyte can be fully flexible.

Third, one needs to decide how to treat the surface. With regards to the surface, the tether connecting the chiral selector to the silica, the silanol groups and the end-caps on the silica, and the neighbouring chiral selectors all have effects on the interaction of chiral selectors with analytes. For example, Pirkle *et al.* [40] have shown that shortening the tether that connects the Whelk selector to the silica from $–(CH_2)_{10}Si(CH_3)_2O-$ to $–(CH_2)_3Si(CH_3)_2O-$ improved the separation of naproxen on the CSP, suggesting that the surface functional groups sterically reduce the docking of the analyte under the cleft which is preferred by the less retained analyte. However, in most theoretical studies, the surface effect is neglected.

Fourth, a solvent used in chiral chromatography changes the potential energy surface by solvating the CSP, the analytes and the complexes. Thus, one should also consider solvent effect on chiral selectivity. In most of the published studies, solvent is either neglected or treated implicitly with continuum models [98].

Fifth, force field exploration of diastereomeric complexes often ignores the entropy contribution to the free energy as only the energies of the complexes are calculated. The free energy can be calculated with the free energy perturbation method [103], but the method is very resource demanding, and thus highly restricted. Recently, free energies were calculated directly by some researchers [96,104], applying the mode integration algorithm [105]. However, the method uses a continuum solvent model and needs careful identification of duplicate conformations, which is sometimes arbitrary. These factors could introduce significant errors to the calculated free energies.

Despite all of the above concerns, studies of individual (isolated) docked complexes have nonetheless achieved successes in the understanding of chiral selective mechanisms in chiral chromatography. This type of study has been undertaken mostly for brush-type CSPs [96,99,102,106,107] and for cyclodextrins [46,108]. For the latter, the selector cavity is expected to minimize the contribution of the solvent to the stability of the docked complex. For reviews on earlier work, see refs. [109] and [98].

Recently, as a result of the increasing computer power, more complicated problems can be tackled and resource-demanding calculations can be used in the study of enantioselective mechanisms. For instance, Maier *et al.* [102] studied the chiral discrimination between a cinchona alkaloid carbamate selector and 3,5-dinitrobenzoyl amino acids with force field exploration of the diastereomeric complexes. In their study, stochastic molecular dynamics simulations have been used to sample the configurations of the diastereomeric mixtures in implicit solvent. They found that the average energy difference between the two diastereomeric mixtures was consistent with the chromatographic elution order and the sampled structures of the diastereomeric mixtures were consistent with NMR and X-ray crystallography results. Mihlbachler *et al.* [110] predicted the correct elution order of the (±)-Tröger's base enantiomers on the amylose tri(3,5-dimethylcarbamate) selector coated to silica in a continuum solvent environment using a combination of molecular mechanics and molecular dynamics sampling of the diastereomeric

mixtures. Kasat *et al.* [111] applied DFT calculations to explain the infrared spectroscopy of ten simple solutes absorbed on amylose tris(3,5-dimethylphenylcarbamate). Due to the complexity of the *ab initio* calculation, the sampling has been highly restricted with motif based optimization and only a side chain of the selector has been considered. Temporini *et al.* [112] studied the docking of two aryloxyacetic acids on a Penicillin G Acylase-based CSP. The structure of the selector was fixed based on the known 3D structure of Penicillin G Acylase. They found that π-π interactions and steric effects are important to the chiral separation.

Docked complexes between the Whelk-O1 selector and a series of 1-(4-halogeno-phenyl)-1-ethylamine derivatives were explored by Del Rio *et al.* [96]. They found that calculated enthalpy and free energy differences between the docked complexes were consistent with measured elution orders but the calculated differential free energy changes could be 100% off the experimental values. They also found that in the lowest energy configurations of the diastereomeric complexes, both R- and S- analytes form three interactions with the CSP: a hydrogen bond with the amide hydrogen, a π-π stack with the dinitrophenyl ring and a CH-π interaction with the phenanthryl group. This is in contrast [77] with the three-point interaction model, which states that the more retained enantiomer forms all the three interactions while the less retained enantiomer cannot.

Other than force field exploration of docked complexes, a second category of theoretical study on chiral chromatography is statistical modeling via structure-function relationships. This type of study is especially useful for those protein-based CSPs where the details about the chiral active sites are not known at an atomic level. Instead, a number of analytes are separated and molecular descriptors of those analytes are correlated with the ability of the protein to resolve them [113-115]. Structure-function relationships have been derived for the amylose [116] and cellulose [117] tris(4-methylbenzoate) CSPs, the ULMO CSP [118] and the Whelk-O1 CSP [119].

A training set of 143 diverse chiral compounds associated with a large range of separation factors was used to obtain a statistical model for the separation factors on Whelk-O1 CSP [119].

The model highlighted the importance of π-π interactions and pointed out the possible interference of hydrogen bonding that may determine a decrease of the enantioselectivity.

The third category of theoretical study of chiral chromatography is simulation. Monte Carlo and molecular dynamics simulations have been applied to search energy minima of selector-analyte complexes. However, as noted above, these studies are usually in the gas phase or in an implicit solvent, and constraints are usually imposed on the selector and the analyte. Simulation of the full stationary phase with explicit solvent is relatively recent, due to the complexity of the interface involved. Efforts have typically focused on achiral interfaces [120-122], such as the *n*-octadecane stationary phase [122].

1.5 Research Aims

Since chiral chromatography demonstrates its ability and promise in the analytical and preparative separation of racemates, much effort has been directed towards the discovery, design, optimization, and application of chiral chromatography. A full understanding of the chiral selective mechanism will facilitate these aspects of chiral chromatography. However, most chiral selective mechanisms are obtained indirectly and often incompletely through comparing experimental elution data. NMR and co-crystallization studies of selector-analyte complexes are helpful, but these studies are only suitable to a limited number of analytes. In addition, the co-crystals are in the solid state, and NMR studies are carried out in specific solvents that differ from those used in chiral chromatography. Furthermore, an underlying surface is not present in either NMR studies or crystallography studies. However, it should be noted that in a recent work by Hellriegel *et al.* [123], the enantioselective docking of (R)- and (S)- dinitrobenzoyl leucine on a (tert-butyl carbamoyl) quinine based brush-type CSP was studied using an ^{1}H HR/MAS (high resolution/magic angle spinning) 2D transfer NOESY method. They pointed out that the inclusion of the underlying silica surface helped to correlate the NMR study with chromatography results.

Molecular modeling and related calculation techniques are nowadays practical and valuable tools for obtaining crucial information about the mechanism of chromatographic separations [98]. The goal of this thesis is to study the selective mechanism of the Whelk-O1 chiral stationary phase in chiral chromatography. Molecular dynamics simulations of the Whelk-O1 chiral interface, including details on the silica surface, the solvent molecules in the mobile phase, the Whelk selectors and analytes, are employed to gain an atomistic understanding of the chiral selective mechanism. To the best of our knowledge, our simulation was the first example of MD simulation with full interfacial details of the working environment in the field of chiral chromatography.

The Whelk-O1 chiral stationary phase is chosen because it is the most successful example of brush-type CSPs [30]. This CSP is versatile and robust, yet it is also relatively simple, containing only two chiral centers which make it a perfect object for a mechanism study. As we have reviewed, two different versions of three-point interaction models have been proposed for the Whelk-O1 CSP. One [55,57,77,97] stated that inside-of-cleft docking makes the more retained enantiomers form the third interaction, the edge-to-face π-π interaction with the phenanthryl moiety of the Whelk selector, and outside-of-cleft docking prevents the less retained enantiomers from forming this third interaction. The other [96] stated that both more retained and less retained enantiomers form the third interaction, which is the CH-π interaction with the phenanthryl moiety, but to a different degree. Theoretical studies of the Whelk-O1 chiral stationary phase using docked complex energy minimization method [96] and quantitative structure-function relationship method [119] have been carried out. However, a powerful model with predictive ability has not been achieved. Thus, we decided to apply a different method, molecular dynamics simulation, to study the chiral selective mechanism of the Whelk-O1 chiral stationary phase.

A precise and efficient model of the molecular flexibility of the selector and analyte is essential for the success of the study of chiral recognition. Thus, our first step is to develop a semi-flexible model for the Whelk-O1 selector and the analytes through *ab initio* calculations and numerical fitting. This will be presented in Chapter 2, after a description of theoretical methods used in our study. In Chapters 3 to 6, the solvation, selectivity, and optimization of the Whelk-O1 CSP will be studied with molecular dynamics simulations. A conclusion of our study and an outlook of applying molecular dynamics simulation to chiral chromatography are presented in Chapter 7.

Chapter 2
Methods and Models

2.1 Quantum Mechanics Methods

2.1.1 *ab initio* Calculations

Quantum mechanics calculations are applied in this thesis to obtain details of equilibrium structures and intramolecular flexibility of the relevant molecules. In the postulates of quantum mechanics [124], molecular electronic properties can be obtained by solving the time independent Schrödinger equation:

$$\hat{H}\Psi(\vec{r}_i,\vec{r}_\alpha) = E\Psi(\vec{r}_i,\vec{r}_\alpha) \qquad 2.1$$

where $\hat{H}$ is the Hamiltonian. $\Psi(\vec{r}_i,\vec{r}_\alpha)$ is the wave function of the system and it depends on the coordinates of all the electrons ($\vec{r}_i$) and nuclei ($\vec{r}_\alpha$). The solution of the Schrödinger equation gives the form of the wave function and its eigenvalue, *E*, the total energy of the system.

The Hamiltonian is the quantum mechanical operator corresponding to the classical total kinetic and potential energies of the system. The nonrelativistic Hamiltonian for a polyatomic molecule, in atomic units, is

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_i \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} - \frac{1}{2}\sum_\alpha \frac{1}{M_\alpha}\nabla_\alpha^2 + \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \qquad 2.2$$

where *i*, *j* run over electrons and *α, β* run over nuclei. The first term corresponds to the kinetic energy of all the electrons where the Laplacian ∇_i^2 operator is a sum of second-order differential operators. The second term corresponds to the electron-nucleus potential energy where Z_α and $r_{i\alpha}$ are the charge on nucleus α and the separation distance between nucleus α and electron *i*. The third term corresponds to the electron-electron potential energy where r_{ij} is the distance between

electron *i* and electron *j*. The fourth and fifth terms correspond to the kinetic energy of the nuclei and the potential energy of nucleus-nucleus interaction respectively. With the Born-Oppenheimer approximation, the fourth and fifth terms can be factored out of the Hamiltonian as nuclei move much slower than electrons. The complete Hamiltonian can be reduced to a purely electronic Hamiltonian

$$\hat{H}_{elec} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_i \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} = \hat{T} + \hat{V}_{Ne} + \hat{V}_{ee} \qquad 2.3$$

and we have

$$\hat{H}_{elec}\Psi_{elec}(\vec{r}_i) = E_{elec}\Psi_{elec}(\vec{r}_i) \qquad 2.4$$

The solution of the electronic Schrödinger equation gives the electronic wave function Ψ_{elec} which now depends on the coordinates of electrons and the coordinates of nuclei enter only parametrically and do not appear explicitly. The total energy of the system is then the sum of E_{elec} and the constant internuclear repulsion energy.

From now on, only Eq. 2.4 will be considered and the subscript "elec" will be dropped. From the postulates of quantum mechanics, the wave function is not directly observable but its square gives the probability density of finding the electrons in the corresponding volume element $\prod_i d\vec{r}_i$. The expectation value of an observable O can be obtained through:

$$< \hat{O} >=< \Psi \mid \hat{O} \mid \Psi >= \int ... \int \Psi^* \hat{O} \Psi d\vec{r}_1 d\vec{r}_2 ... d\vec{r}_N \qquad 2.5$$

where $\hat{O}$ is the quantum operator of observable O, and Ψ^* is the complex conjugate of the wave function Ψ . Therefore, the ground state energy, E_0 , of a molecule is given by $< \Psi_0 \mid \hat{H} \mid \Psi_0 >$ as its corresponding quantum mechanical operator is $\hat{H}$.

With a few exceptions, Eq. 2.4 cannot be solved analytically. In this case, other approaches are adopted. One approach is based on the variational principle, which states that given any trial

wave function Ψ_{trial}, the expectation value of the energy E_{trial} will always be higher than, or equal to, the exact ground state energy E_0:

$$< \Psi_{trial} | \hat{H} | \Psi_{trial} > = E_{trial} \geq E_0 = < \Psi_0 | \hat{H} | \Psi_0 > \quad 2.6$$

The variational principle provides a basis for systematically approaching the exact wave function: One only needs to minimize the trial energy by searching through all continuous, quadratic-integrable, and single-valued wave functions.

The Hartree-Fock (HF) approach [124] applies this concept and uses a self-consistent field method to approximately solve the electronic Schrödinger equation. Briefly, for an N-electron system, a Slater determinant is formed by a group of guessed spin orbitals to serve as the trial wave function. A set of derived HF equations is solved. The resulting new spin orbitals are then used in the next iteration and so on until a certain convergence threshold is met.

The only flexibility of the Slater determinant trial wave function is the choice of the spin orbitals. A spin orbital is a product of a spatial molecular orbital (MO) and a spin function. In practice, the MOs are usually expressed as linear combinations of basis functions. These functions will be discussed later in this section. Intuitively, the quality and quantity of basis functions determines the quality of the trial wave function, and thus the calculation results. The HF method is a typical *ab initio* (means from the beginning) calculation as it does not use any empirical parameters. It is a corner-stone of quantum mechanics, providing a consistent way to approach the HF limit, for which an infinite set of basis functions must be used.

Despite its importance, the HF method has major limitations. First, the use of a single Slater determinant is a fairly drastic approximation to the exact N-electron wave function. Second, the HF method uses an average effective potential which neglects the dynamical electron correlation that is caused by the instantaneous repulsion of the electrons. Various schemes to deal with electron correlation have been devised, including Moller-Plesset perturbation theory [125], configuration interactions (CI) [124], and coupled cluster approaches [126]. Unfortunately, these

methods are very computationally expensive [124]. To overcome this disadvantage, density functional theory (DFT) attempts to calculate the ground state energy and other molecular properties from the ground state electron density $\rho_0(x, y, z)$, a function of only three variables.

2.1.2 Density Functional Theory Method

Density functional theory, as it applies to molecules, derives from the Hohenberg and Kohn existence and variational theorems [127]. These two theorems state that, for molecules with a nondegenerate ground state, the ground state molecular energy, wave function and all other molecular electronic properties are uniquely determined by the ground state electron probability density and the true ground state electron density minimizes the energy. Practically, Kohn and Sham [128] devised a method to find the ground state density and ground state energy. They introduced a fictious reference system of N noninteracting electrons that each has the same external potential energy function $v_s(\vec{r}_i)$ which makes the ground state electron probability density the same as that of the real system. The energy of the real system can be calculated as:

$$E(\rho) = \int \rho(\vec{r}) v(\vec{r}) d\vec{r} + \overline{T}_s(\rho) + \frac{1}{2}\int\int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 + \Delta T(\rho) + \Delta \overline{V}_{ee}(\rho) \qquad 2.7$$

where the first term is the exact electron-nucleus attractive electrostatic energy; the second term is the kinetic energy of the reference system; the third term is the classic expression for electrostatic repulsion energy; the fourth and fifth terms are the difference between the real system and the reference system for the kinetic energy and inter-electronic repulsion energy, respectively. The first and third terms are trivial to evaluate. The second term can be calculated by:

$$\overline{T}_s(\rho) = -\frac{1}{2}\sum_i < \theta_i^{KS}(1) \left| \nabla_1^2 \right| \theta_i^{KS}(1) > \qquad 2.8$$

where ∇_1^2 is the Laplacian operator. θ_i^{KS} is the spatial part of Kohn-Sham orbitals and satisfies the Kohn-Sham one-electron eigenfunction of the noninteracting reference system:

$$[-\frac{1}{2}\nabla_i^2 + v_s(\vec{r}_i)]\theta_i^{KS} = \varepsilon_i^{KS}\theta_i^{KS} \quad 2.9$$

where ε_i^{KS} are Kohn-Sham orbital energies. By definition we have

$$\rho_s = \sum_i \left|\theta_i^{KS}\right|^2 = \rho \quad 2.10$$

The functionals $\Delta T(\rho)$ and $\Delta\overline{V}_{ee}(\rho)$ in Eq. 2.7 are unknown. An exchange-correlation energy functional is defined to include both terms:

$$E_{xc}(\rho) = \Delta T(\rho) + \Delta\overline{V}_{ee}(\rho) \quad 2.11$$

By Eqs. 2.7, 2.8, 2.9, and 2.11 we have,

$$[-\sum_\alpha \frac{Z_\alpha}{r_{1a}} - \frac{1}{2}\nabla_1^2 + \int \frac{\rho(\vec{r}_2)}{r_{12}} d\vec{r}_2 + v_{xc}(1)]\theta_i^{KS}(1) = \varepsilon_i^{KS}\theta_i^{KS}(1) \quad 2.12$$

where

$$v_{xc}(\vec{r}) = \frac{\delta E_{xc}[\rho(\vec{r})]}{\delta\rho(\vec{r})} \quad 2.13$$

is the exchange-correlation potential. The Kohn-Sham operator (the operator in the square brackets in Eq. 2.12) considers both the exchange and electron correlation via the exchange-correlation potential, while the Fock operator neglects the latter. However, at this point the exact form of the exchange-correlation potential is still unknown.

Similar to the HF method, the Kohn-Sham orbitals can be obtained through SCF iterations as long as an approximate exchange-correlation potential is used. Much effort has been made to find suitable exchange-correlation energy functionals [129-133]. One very popular exchange-correlation energy functional is the B3LYP, which is a hybrid functional combining the gradient corrected HF definition of exchange energy, the local-spin-density approximation exchange functional [131], the B88 exchange functional [129], and the Lee-Yang-Par correlation functional [130]. The B3PW91 [133] is also very popular, and this functional uses the B88 exchange

functional [129] and the PW91 correlation functional [132]. These functionals usually contain empirical parameters. Thus, DFT, in a sense, cannot be called *ab initio* and is not variational --- unlike HF and CI that can consistently approach the exact ground state energy by improving the basis set, DFT does not always give better results with larger basis sets. However, the performance of DFT is outstanding for most molecular properties and it is much more efficient than CI, perturbation theory, and coupled cluster methods and thus can be applied to larger molecular systems.

2.1.3 Basis Functions and Basis Sets

As discussed above, the spatial molecular orbitals (φ) used in the HF and DFT methods are represented by linear combinations of a set of basis functions:

$$\varphi_i = \sum_r c_{ri} \mathrm{X}_r \qquad 2.14$$

where c_{ri} is the coefficient for basis function X_r in molecular orbital φ_i. An adequate set of basis functions is required for the success of these calculations.

The most popular basis functions are known as the Gaussian-type functions (GTFs) [124]. A Cartesian Gaussian centered on atom *b* is defined as

$$g_{i,j,k} = N x_b^i y_b^j z_b^k e^{-\alpha r_b^2} \qquad 2.15$$

where *i, j, k* are nonnegative integers, α is a positive orbital exponent, x_b, y_b, z_b are Cartesian coordinates with the origin at nucleus b and N is the normalization constant. The GTF is called an s-type, a p-type and a d-type Gaussian when $i+j+k$ equals to 0, 1, and 2 respectively. So, there is 1 s-type Gaussian [(i,j,k)=(0,0,0)] and 3 p-type Gaussians [(i,j,k)=(1,0,0), (0,1,0), or (0,0,1)] as expected. However, there are 6 d-type Gaussians, but 5 linear combination of the 6 d-type Gaussians can be formed to represent the 5 "actual" d AO's ($x_b y_b$, $x_b z_b$, $y_b z_b$, $x_b^2 - y_b^2$, and $3z_b^2 - r_b^2$). The sixth linear combination with $x_b^2 + y_b^2 + z_b^2$ is like a 3s function and is

sometimes omitted from the basis set. Similarly, there are 10 f-type Gaussians and they can be used to form the 7 "actual" f AO's.

The most natural basis functions are actually the Slater orbitals (STOs) that have a cusp at the nucleus and resemble AOs well. But they are impractical for polyatomic, nonlinear molecules where a large number of multi-centered electron-repulsion integrals need to be evaluated. Gaussians, on the other hand, are efficient for the evaluation of the electron-repulsion integrals, but give a poor representation of an AO as there is no cusp at the nucleus. Current practice is to use contracted Gaussian-type functions (CGTF) as basis functions. A CGTF is a linear combination of a few Gaussians (called as the primitive Gaussians) and has the desired shape.

There are several types of basis sets. A minimal basis set consists of one CGTF/STO for each inner-shell AO and for each valence-shell AO. A double-zeta basis set has two CGTFs/STOs for such AO. A split-valence basis set uses two or more CGTFs/STOs for each valence-shell AO but only one CGTF/STO for each inner-shell AO.

One of the most common types of basis sets is the split-valence basis set arising from the group of Pople [134] and has the notation of X-YZG, where X, Y, and Z are positive integers. X indicates that for the inner-shell AO's, one CGTF contracted by X primitive Gaussians is used. The Y and Z indicate that the valence orbitals are composed of two basis functions each (split-valence double zeta), the first one composed of a linear combination of Y primitive Gaussian functions, the other composed of a linear combination of Z primitive Gaussian functions. Split-valence triple- and quadruple-zeta basis sets are also used and are denoted as X-YZWG, and X-YZWVG respectively. As an example, with the 6-31G basis set, each inner-shell AO is represented by a CGTF formed by 6 primitive Gaussians, and each valence-shell AO is represented by two CGTFs, one formed by 3 primitive Gaussians and another by 1 Gaussian.

AOs are distorted in some molecules, and to allow such distortion, polarization functions with higher angular momentum quantum number than the maximum in the valence-shell of the ground state atom may be added. A 6-31G* basis function adds a set of 6 d-type Gaussian

functions to heavy atoms and a 6-31G** also adds a set of 3 p-type Gaussian functions to hydrogen. Some systems, such as anions and compounds with lone pairs, have significant electron density at large distances from the nuclei. Highly diffuse functions, which have very small orbital exponents, can be added to improve the accuracy for such systems. For example, 6-31+G is formed from 6-31G by adding four diffuse functions (*s, p_x, p_y, p_z*) to each non hydrogen atom, whereas 6-31++G is formed from 6-31+G by including a highly diffuse *s* function on each hydrogen atom.

2.1.4 Conformational Minimization and Atomic Charges

DFT calculations are applied in this thesis to obtain molecular equilibrium structures, atomic charges, and conformational energies. In the Gaussian program [135], local minima are found through the location of stationary points on a potential energy surface. Briefly, the potential energy surface is a hyper surface with 3N-6 degrees of freedom. A stationary point has the 3N-6 first partial derivatives of the potential energy all zero. It may be a minimum, maximum, or a saddle point and can be further identified via the second order derivatives of the potential. To start a geometry optimization, one guesses a geometry (Gaussian input), and the DFT calculation is done to obtain the potential energy and its first derivatives. The geometry optimization routine changes the 3N-6 nuclear coordinates to a new set which is likely to be closer to a minimum than the previous structure. The optimization iterates until a stationary point is obtained [124].

Atomic charges can be obtained for the optimized structure through fitting of the molecular electrostatic potential (MEP). The MEP (in atomic units) at a point x_1,y_1,z_1 is given by:

$$\phi(x_1, y_1, z_1) = \sum_{\alpha} \frac{Z_\alpha}{r_{1\alpha}} - \iiint \frac{\rho(x_2, y_2, z_2)}{r_{12}} dx_2 dy_2 dz_2 \qquad 2.16$$

where the first and second term calculates the MEP due to all nuclei and all electrons respectively. With atomic partial charges (Q), the MEP at the same point becomes:

$$\phi^{approx}(x_1, y_1, z_1) \equiv \sum_{\alpha} \frac{Q_\alpha}{r_{1\alpha}} \qquad 2.17$$

It was found that the MEP outside of the van der Waals surface is the most significant for nonbonding interactions. Electrostatic potential charges Q_α's are obtained by minimizing the deviation for all the MEP at each grid point chosen outside of the van der Waals surface. Various ways of choosing the grid points and of including other refinements give different schemes for finding the atomic charges. Two common schemes are the Merz-Singh-Kollman method [136] and the charges from electrostatic potentials using a grid method (CHELPG) [137].

2.2 Molecular Simulation

2.2.1 Monte Carlo and Molecular Dynamics Simulations

Despite their rapid development, quantum mechanics methods are very expensive computationally and are often prohibitive for large many-body systems such as condensed phase systems. Classical molecular simulations based on statistical mechanics are widely employed for this purpose. Computer simulations were first carried out by Metropolis *et al.* [138] in early 1952 and since then, have become bridges between model, theory, and experiment. On one hand, one can evaluate the model by comparing the calculated properties of the model system with those of an experimental system. On the other hand, one can test a theory by comparing the outcome of the theory and the result of simulation applied to the same model. In this sense, computer simulations are experiments on a computer.

Monte Carlo and molecular dynamics simulations are the two most common classical statistical mechanics methods. "Monte Carlo" simulation is named for its heavy use of computer generated random numbers. During a Monte Carlo simulation, a random displacement is given to a randomly selected particle and the displacement is accepted with a probability. The probability depends on the energy change due to the random displacement as well as the acceptance rule that is devised to sample averages in the specified ensemble. The most common Metropolis acceptance rule produces canonical averages [138].

While the Monte Carlo method is very elegant and efficient, it is applied to the study of equilibrium properties. A different method is needed to obtain the dynamics. Molecular dynamics simulation provides an option. Here, we numerically solve Newton's equations of motion for many body systems. Briefly, consider a typical system in a fixed simulation cell with N particles, its classical Hamiltonian is:

$$H(\vec{p}^{\,N},\vec{q}^{\,N}) = \sum_i \frac{\vec{p}_i^{\,2}}{2m_i} + U(\vec{q}^{\,N}) \qquad 2.18$$

where $\vec{p}_i$, $\vec{q}_i$, and m_i are the momentum, position and mass for the particle *i*, respectively. The first term is the total kinetic energy and the second term is the total potential energy of the system. In Cartesian coordinates, the equations of motion for this system are:

$$\frac{d\vec{q}_i}{dt} = \frac{\vec{p}_i}{m_i} \qquad 2.19$$

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial U(\vec{q}^{\,N})}{\partial \vec{q}_i} = \vec{f}_i \qquad 2.20$$

In each time step of a molecular dynamics simulation, the force is calculated and the equations of motion are integrated.

2.2.2 Potentials

The potential describes the interactions within and between molecules. It is the basic input for a molecular simulation as the negative of its first derivative, the force, governs the evolution of the equations of motion. The potential can be molecule-based or atom-based. It can have terms that depend on the coordinates of one, two, three, or more particles [139]. The common atom-based potential usually includes intramolecular terms and pairwise non-bonding terms.

The intramolecular potential dictates the energetic costs for molecular conformations. Common force fields typically include bond stretching, angle bending, improper torsions and torsions.

The bond stretching potential is usually given in a harmonic form:

$$U^{stretch}(r) = k_r (r - r_e)^2 \qquad 2.21$$

where k_r is the stretching force constant, r is the interatomic separation, and r_e is the equilibrium bond length. The right hand side of the equation comes from the first non-zero term of a Taylor expansion about r_e of $U^{stretch}(r)$. A more accurate equation that describes the bond stretching energy would be the Morse potential [140]. However, it is not applied by most force fields due to its computational inefficiency because of the involvement of exponential calculations. Also, since the energetic costs of bond length variation are relatively high, Eq. 2.21 is a good approximation for the Morse potential around the minimum. Practically, bonds are sometimes "frozen" out in simulations: They are usually constrained by SHAKE [141] or RATTLE [142] algorithms and a bond stretching potential is not needed.

The angle bending motion is also usually represented by a harmonic potential:

$$U^{bend}(\theta) = k_\theta (\theta - \theta_e)^2 \qquad 2.22$$

where k_θ and θ are the force constant and the angle, respectively, and θ_e corresponds to the equilibrium angle.

For four atoms connected in sequence, ABCD, the torsional motion is typically modeled by a Fourier series which has the general form [143,144] of

$$U^{torsion}(\phi) = \frac{1}{2}\sum_{\{i\}} V_i [1 + (-1)^{i+1} \cos(i\phi + \phi_i)] \quad \text{with } i=0,1,2,3\ldots \qquad 2.23$$

where ϕ is the dihedral angle and ϕ_i is the corresponding phase shift. The set of periodicities included in the sum are specific to the torsional linkage. The majority of the torsional potentials in the OPLS [143] and CHARMM [144] force fields have less than 4 terms (i.e., i=0,1,2).

In this thesis, two Ryckaert-Bellemans [145] potentials of the form

$$U^{torsion}(\phi) = \sum_{i=0}^{6} c_i \cos^i(\phi + 180) + c_7 \sin^5(\phi + 180) + c_8 \sin^7(\phi + 180) \qquad 2.24$$

$$U^{torsion}(\phi) = \sum_{i=0}^{6} c_i \cos^i(\phi + \phi_i) \qquad 2.25$$

are employed where c_0 to c_8 are the term amplitudes, ϕ is the dihedral angle, and ϕ_i is the corresponding phase shift. These equations, especially Eq. 2.25, are similar to Eq. 2.23 after expansion. The sine terms in 2.24 and the phase shifts in 2.25 are introduced to model the asymmetric torsional potential caused by chirality. Eq. 2.24 is applied in Chapter 3 and a better form in Eq. 2.25 is applied in Chapter 4 and thereafter.

For three atoms ABC surrounding D, either an improper torsion (dihedral ADBC) or an out-of-plane angle (dihedral ABCD) can be defined. The latter definition is sometimes also called an improper torsion. Improper torsion potentials are sometimes included in force fields to constrain planarity or chirality [146]. For the out-of-plane angle definition which we have used, it usually has the form of

$$U^{imp}(\omega) = k_{\varpi}(\omega - \omega_e)^2 \qquad 2.26$$

where k_ω is the force constant, ω is the out-of-plane angle and ω_e is the corresponding equilibrium value. ω is defined by the angle between the vectors normal to the two planes of ABC and BCD.

In forming the mirror-image molecule, the vectors used to define torsion and improper torsion angles are reversed if the atom numbering is conserved. This has the effect of changing the sign on the phase shift in Eq. 2.25 and the equilibrium out-of-plane angle in Eq. 2.26.

Other than intramolecular potentials, a Lennard-Jones (LJ) potential and a partial charge electrostatic potential are usually included in force fields to model the non-bonding interactions

between atoms. Both potentials are pairwise additive. The Lennard-Jones potential can be written in the form of

$$U(r_{ij})^{LJ} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \qquad 2.27$$

where A_{ij} and B_{ij} are constants specific to atoms i and j and r_{ij} is the distance between the two atoms. The inverse 12th power term represents the strong repulsive interaction at short interatomic separation, while the inverse 6th power term models the attractive London dispersion interaction. The potential can also be written as:

$$U(r_{ij})^{LJ} = 4\varepsilon_{ij}\left(\frac{\sigma_{ij}}{r_{ij}^{12}} - \frac{\sigma_{ij}}{r_{ij}^{6}}\right) \qquad 2.28$$

where the constants ε_{ij} and σ_{ij} are directly related to A_{ij} and B_{ij} in Eq. 2.27. As shown in Fig. 2.1 which depicts the LJ potential versus interatomic separation, the parameter σ_{ij} equals to the interatomic separation when $U(r_{ij})^{LJ}$=0 and the parameter ε_{ij} is the well depth of the potential. Usually, in force fields, σ_{ii} and ε_{ii} are parameterized for each type of atom to model certain experimental results. Lorentz–Berthelot combining rules [147,148], $\sigma_{ij} = (\sigma_{ii} + \sigma_{jj})/2$ and $\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}$, are then used for all pairs of distinct atoms. Without the combining rules, N(N-1)/2 sets of σ_{ij} and ε_{ij} would need to be parameterized for N types of atoms.

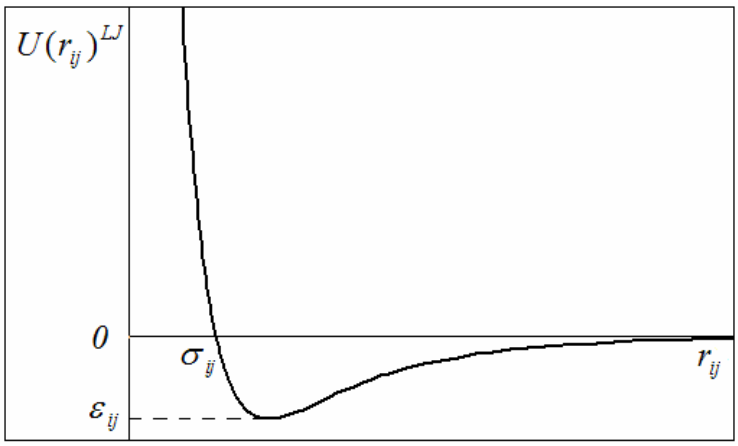


Figure 2.1: Lennard-Jones potential as a function of the interatomic separation. Key points on the potential and interatomic separation axes are labelled.

In molecular systems, the LJ potential can be supplemented by other interactions such as ion-ion, ion-dipole, dipole-dipole and higher multipole interactions. Typically, a simple partial charge electrostatic potential, with the form of Eq. 2.29, is used to model these interactions.

$$U(r_{ij})^{EL} = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \qquad 2.29$$

ε_0 is the dielectric constant and r_{ij} is the distance between atom i and atom j. q_i and q_j are the partial charges on the atoms. In most force fields, the charges are "fixed": all atoms carry permanent charges according to their atom types. Alternatively, fluctuating charge models [149,150] in which atomic charges fluctuate with the change of environment can be developed.

2.3 Practical Issues Related to Potential Calculations

2.3.1 Periodic Boundary Conditions

Periodic boundary conditions are usually applied in molecular simulations to overcome the artificial surface effect at the boundary of the simulation cell [139]. When the conditions are applied, the simulation cell is replicated throughout space to form an infinite lattice. A two-dimensional periodic system is shown in Fig. 2.2. During the simulation, as a particle moves in

the primitive (central in the figure) cell, its periodic images in all the replicated cells move the same way. If one particle leaves the primitive cell at one side, one of its images will enter through the opposite side from a replicate cell. Thus, by using periodic boundary conditions, there is no surface effect at the boundary of the simulation cell.

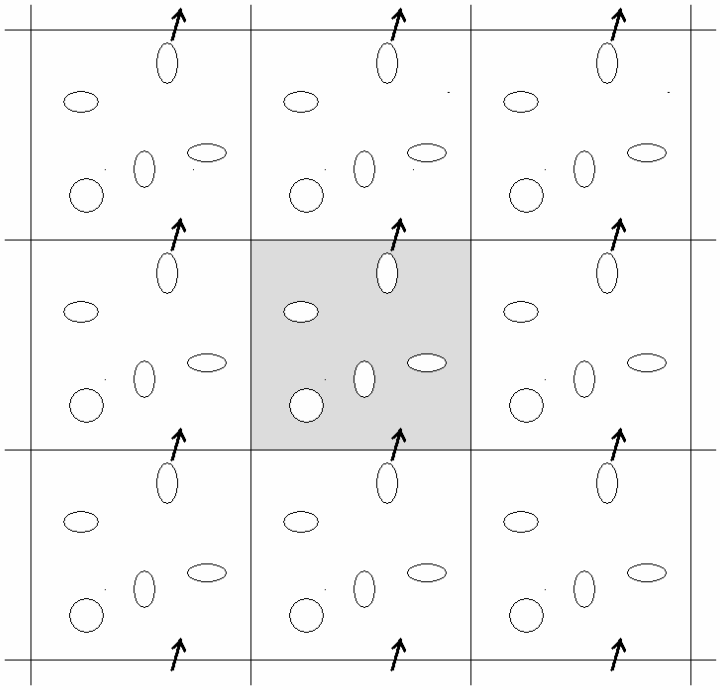


Figure 2.2: A schematic representation of two dimensional periodic boundary conditions.

Consider a system with N particles in the primitive simulation cell. With periodic boundary conditions, a given particle now interacts with all other particles in this infinite periodic system. In our case, all the intermolecular interactions are pairwise additive, so the total non-bonding potential energy of all the particles in the primitive cell (assume cubic for convenience) is:

$$U_{tot} = \frac{1}{2}\sum_{\vec{n}}\sum_{i=1}^{N}\sum_{j=1}^{N}{}' u(|\vec{r}_{ij} + \vec{n}L|) \qquad 2.30$$

where *i* and *j* represent any two particles in the primitive cell, while L is the diameter of the periodic box, and $\vec{n}$ is an arbitrary vector of three integer numbers. Thus, the sum over $\vec{n}$ is sum

over all the lattice points (*i.e.* all the cells). Since the sums over *i* and *j* each includes all particles, the factor ½ is required to correct for double counting of the intermolecular interactions. The prime over the sum indicate that the term with *i=j* is to be excluded when $\vec{n}$ =(0,0,0): a particle interacts with all its images, but not itself.

In this general form, periodic boundary conditions seem to be impractical as we now have to calculate the potential energy as an infinite sum rather than a finite one. In practice, however, we usually truncate the short-range LJ interaction beyond a certain cutoff distance and apply Ewald summations [151,152] for the electrostatic interaction.

2.3.2 Truncation of the Lennard-Jones Interactions

For short-range interactions, such as the LJ potential, the total potential energy for a particle is dominated by interactions with neighbouring particles that are closer than some cut-off distance r_c. A truncation and shift scheme is commonly used in molecular dynamics simulations [138]:

$$u^{tr-sh}(r) = u(r) - u(r_c) \quad \text{when} \quad r \le r_c \tag{2.31}$$

$$u^{tr-sh}(r) = 0 \quad \text{when} \quad r > r_c \tag{2.32}$$

This way, the potential vanishes at the cut-off radius. Practically, the cut-off distance r_c is chosen according to two criteria: First, it should be sufficiently large such that the error resulting from ignoring the interactions with particles at beyond r_c is small; Second, it should be less than the distance of the half of the length (L/2) of the simulation cell such that one does not need to count any atomic pair more than once due to the application of periodic boundary conditions.

For our case, we have a rectangular simulation cell with $L_z > L_x = L_y$. The cut-off radius is chosen to be L_x/2.05, which is sufficient to make the error from truncation and shifting of the Lennard-Jones potential small and we do not need to count any atomic pair more than once.

2.3.3 Ewald Summation Method for Electrostatic Potential

Consider a system of N point charges q_i at position $\vec{r}_i$ which satisfy the charge neutrality condition in a rectangular simulation cell with lengths L_x, L_y, and L_z. Similar to Eq. 2.30, when applying the periodic conditions, the total Coulombic potential can be written as:

$$U(r_{ij})^{Elec} = \frac{1}{2}\sum_{i=1}^{N} q_i \phi(r_i) \qquad 2.33$$

where $\phi(r_i)$ is the electrostatic field at position of point charge *i*:

$$\phi(r_i) = \sum_{\vec{n}} \sum_{j=1}^{N}{}' \frac{q_j}{|\vec{r}_{ij} + \vec{n}|} \qquad 2.34$$

where the sum over $\vec{n}$ is the sum over all lattice points, $\vec{n}$ =(n_xL_x, n_yL_y, n_zL_z) with n_x, n_y, n_z integers. The prime on the summation symbol indicates the omission of the *i*=*j* term when $\vec{n}$ = (0, 0, 0). Reduced units are employed with simulations so that the constant factor $1/(4\pi\varepsilon_0)$ is omitted in Eq. 2.34. The electrostatic field decays as r^{-1} and thus it is a long-range interaction. For this kind of interaction, truncation of the potential (at *L*/2) can lead to artificial behaviour [153]. To treat this potential accurately and efficiently, we use the Ewald summation technique [151,152].

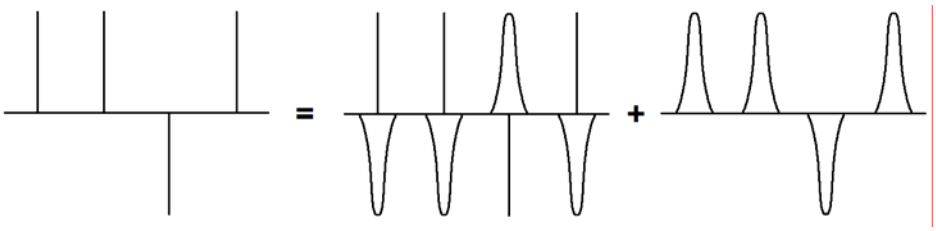
In Ewald summations, a diffuse charge cloud of opposite sign is introduced to surround every partial charge q_i. The total charge of this screening cloud exactly cancels q_i. The screening charge is conveniently chosen to be a Gaussian with the form:

$$\rho(r) = -q_i(\alpha/\pi)^{\frac{3}{2}} \exp(-\alpha r^2) \qquad 2.35$$

where the value of α determines the width of the screening charge. In order to obtain the potential due to point charges, a set of compensating charges which has the same distribution but opposite sign to the screening charges (Fig. 2.3) is introduced. Now consider the potential at the site of

point charge *i*, we have three contributions: the one due to the point charges, the one due to the screening charge clouds, and the one due to the compensating charge clouds.

(a)



(b)

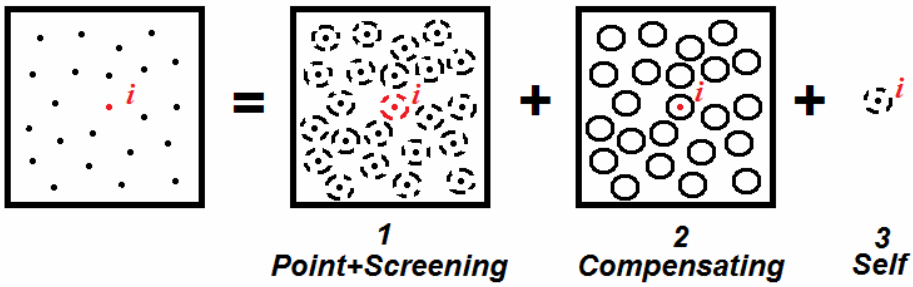


Figure 2.3: The treatment of a set of point charges in Ewald summation [151,152] as a set of screened charges plus the compensating charges is shown in (a). The three contributions to the electrostatic potential at the site of point charge *i* are shown in (b).

The electrostatic field due to the screened point charges [Box 1 in Fig. 2.3(b)] is short-ranged and is calculated by summing over all the screened charges in the original cell and their images in the real space lattice of image cells. The electrostatic field caused by the compensating charge clouds [Box 2 in Fig. 2.3(b)] is calculated in the Fourier space, as the compensating charge distribution is a smoothly varying function and is also periodic if we include the spurious field from the compensating charge cloud around site *i* in the original cell. This spurious self-interaction field should be corrected [Box 3 in Fig. 2.3(b)]. A detailed derivation is given in Ref. [138]. For brevity, only the final form of the sum of the electrostatic field at site *i* is given here:

$$\phi(r_i) = \sum_{\vec{n}} \sum_{j=1}^{N}{}' \frac{q_j}{\left|\vec{r}_{ij} + \vec{n}\right|} erfc(\sqrt{\alpha}\left|\vec{r}_{ij} + \vec{n}\right|) + \frac{1}{V} \sum_{\vec{k} \neq 0} \sum_{j=1}^{N} \frac{4\pi q_j}{k^2} \exp(i\vec{k} \bullet \vec{r}_{ij}) \exp(\frac{-k^2}{4\alpha}) + 2(-q_i)(\frac{\alpha}{\pi})^{\frac{1}{2}} \quad 2.36$$

The three terms in the sum correspond to the three "boxes" respectively. Function *erfc(x)* in the first term is the complementary error function which decays to zero with increasing *x*. In the second term, *V* is the volume of the simulation cell while $\vec{k}$ is the lattice vector in Fourier space. The value of α, which determines the size of the charge clouds, affects the convergence of the first two terms. A large α corresponds to a sharp distribution of screening charge and makes the first term converge fast, but the second term converge slowly (i.e., need a large number of k-vectors). Practically, the aim is to choose a value of α, such that the first term can be truncated at $\vec{n}$ = (0, 0, 0) but a modest number of *k*-vectors are needed for the evaluation of the second term.

The total electrostatic potential energy for a 3-dimensional (3D) system can then be calculated according to Eq. 2.33. However, to simulate the chiral interfaces, a simulation cell with slab geometry (two surfaces facing each other with solvent in between) is applied in our study. Such slab systems have a finite length along one of the three dimensions (assume the z direction) and the conventional Ewald summation technique cannot be used. 2D-Ewald methods have been developed but these methods are computationally expensive. An obvious and convenient way is to use the 3D-Ewald summation by placing empty space above and below the slab. However, if the slab has a net dipole moment along the *z* direction, a spurious depolarizing field will be generated between the periodic images of the slab. Yeh and Berkowitz [152] added a shape-dependent correction term to account for the electrostatic potential due to the depolarizing field. This term has the form of:

$$J(\vec{M}, P) = \frac{2\pi}{V} \vec{M}^2 \quad 2.37$$

where *P* is the shape of the summation geometry and $\vec{M}$ is the dipole moment of the unit cell:

$$\vec{M} = \sum_{i=1}^{N} q_i \vec{z}_i \qquad 2.38$$

Other algorithms that treat the long range electrostatic interaction include the fast multipole method and particle mesh approaches. The fast multipole method, introduced by Appel [154] and further developed by Banes *et al.* [155] and Greengard *et al.* [156], is based on the idea that a group of particles at a large distance can be considered as one big cluster and one can avoid individually calculating particle-particle interactions by clustering the system into groups. The particle-mesh approach was first developed by Hockney and Eastwood [157] to improve the efficiency of the Ewald summation. This approach is based on the fact that the interpolation of the charges on a grid improves the efficiency of the calculation of the Fourier part of the Ewald summation.

2.4 Integrating the Equations of Motion

2.4.1 Algorithms for Integration of the Equations of Motion

Once all the potentials and forces between the particles have been calculated, Newton's equations of motion can be integrated by various algorithms. The normal Verlet algorithm [158] is one of the simplest yet also practical and reliable [138]. The Taylor expansions of the positions around time t for the coordinates at time $t+\Delta t$ and time $t-\Delta t$ are:

$$\vec{r}(t+\Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{\vec{f}(t)}{m}\frac{\Delta t^2}{2} + \vec{r}'''\frac{\Delta t^3}{3!} + \mathrm{O}(\Delta t^4) \qquad 2.39$$

$$\vec{r}(t-\Delta t) = \vec{r}(t) - \vec{v}(t)\Delta t + \frac{\vec{f}(t)}{m}\frac{\Delta t^2}{2} - \vec{r}'''\frac{\Delta t^3}{3!} + \mathrm{O}(\Delta t^4) \qquad 2.40$$

Summing these two equations and rearranging, we get:

$$\vec{r}(t+\Delta t) \approx 2\vec{r}(t) - \vec{r}(t-\Delta t) + \frac{\vec{f}(t)}{m}\Delta t^2 \qquad 2.41$$

Thus, with the normal Verlet algorithm, the new position at time $t+\Delta t$ can be estimated by the old (time $t-\Delta t$) position, current (time t) position, and current forces, with an error of the order of Δt^4, where Δt is the time step in MD simulations. Note that the trajectory is obtained without calculating the velocity. The velocity can be derived explicitly for the evaluation of the kinetic energy.

The normal Verlet algorithm has several equivalent versions that generate the same trajectory including the Leap Frog algorithm [159], the velocity Verlet algorithm [160], and the Beeman algorithm [161]. In our simulation, the velocity Verlet algorithm, where the velocity is needed during the integration, has been used to enable the application of RATTLE [142] and the Nosé-Hoover thermostat [162,163]. The velocity Verlet algorithm takes the form

$$\vec{r}(t+\Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{\vec{f}(t)}{2m}\Delta t^2 \qquad 2.42$$

$$\vec{v}(t+\Delta t) = \vec{v}(t) + \frac{\Delta t}{2m}\left[\vec{f}(t) + \vec{f}(t+\Delta t)\right] \qquad 2.43$$

In practice, the integration is carried out in two stages. In the first stage, the position is advanced a full time step (Eq. 2.42) and the velocity is advanced a half time step:

$$\vec{v}(t+\frac{\Delta t}{2}) = \vec{v}(t) + \frac{\Delta t}{2m}\vec{f}(t) \qquad 2.44$$

In the second stage, the force at the new position is calculated and the velocity is evolved fully:

$$\vec{v}(t+\Delta t) = \vec{v}(t+\frac{\Delta t}{2}) + \frac{\Delta t}{2m}\vec{f}(t+\Delta t) \qquad 2.45$$

Higher-order (employing higher order of the derivatives of the particle coordinates) predictor-corrector algorithms [164] have also been developed. A desirable integration algorithm should be time reversible (*i.e.* the system will trace back its trajectory in the phase space if one reverses the momenta of all particles at a given time), area preserving and energy conserving. Although the short-term energy conservation is very good, most of the higher-order algorithms, such as the predictor-corrector algorithms, are neither time reversible nor area preserving, which

usually leads to a long-time energy drift. The Verlet algorithms, on the other hand, are time reversible and area preserving. They also exhibit little long-time energy drift despite the fact that they have only moderate energy conservation in a short time [138].

2.4.2 Constraint Dynamics

As discussed in the previous subsection, the error associated with integrating the equations of motion is correlated to the time step used. The smaller the time step, the smaller the error. However, as long as the error is acceptable, a large time step is desired so that the simulation is completed in a reasonable time. In a molecular dynamics simulation, the time step should be chosen such that it is appreciably shorter than the period of the highest-frequency motion, which is the intramolecular bond stretching for a typical molecular system. Several schemes have been developed to tackle the time-scale problem. The multiple time scale approach [165] applies different time steps for high-frequency and low-frequency degrees of freedom. An alternative is to apply constraint dynamics in which the high-frequency degrees of freedom are frozen. This is commonly done by the SHAKE approach proposed by Ryckaert *et al.* [141] and its velocity version, the RATTLE algorithm, devised by Andersen [142].

In the approach proposed by Ryckaert *et al.*, the equation of motion with a constraint force on each atom is:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{f}_i + \vec{g}_i \qquad 2.46$$

where $\vec{f}_i$ and $\vec{g}_i$ are the real force included in the potential and the constraint force. Using the Verlet algorithm (Eq. 2.41) to integrate the equation of motion, we have:

$$\vec{r}_i(t+\Delta t) = \vec{r}_i'(t+\Delta t) + \frac{\Delta t^2}{m_i}\vec{g}_i(t) \qquad 2.47$$

where $\vec{r}_i'(t+\Delta t)$ is the position which would have been reached without the constraint.

Consider an example of fixing bond lengths. Since the role of the constraint force is to keep the bond lengths constant, they must be directed along the bonds and must conform to Newton's third law. We have,

$$\vec{g}_i = \sum_j \lambda_{ij} \vec{r}_{ij} \quad \text{and} \quad \lambda_{ij} = \lambda_{ji} \tag{2.48}$$

where i is the atom in consideration and j represents the atoms around and forming a bond with atom i. λ_{ij} are undetermined time-dependent Lagrange multipliers. Combining Eqs. 2.47 and 2.48, we have:

$$\vec{r}_{ij}(t+\Delta t) = \vec{r}_{ij}'(t+\Delta t) - \sum_{h \neq j} \frac{\Delta t^2}{m_h} \lambda_{hi} \vec{r}_{hi}(t) + \frac{\Delta t^2 (m_i + m_j)}{m_i m_j} \lambda_{ij} \vec{r}_{ij}(t) - \sum_{k \neq i} \frac{\Delta t^2}{m_k} \lambda_{jk} \vec{r}_{jk}(t) \tag{2.49}$$

where h runs over all the atoms, except atom j, that connect to atom i. k runs over all the atoms, except atom i that connect to atom j. $\vec{r}_{ij}'(t+\Delta t)$ represents the bond in unrestrained condition and can be obtained with trivial integration of the unrestrained equations of motion. We take the square modulus for each side in Eq. 2.49, and we should have

$$\left|\vec{r}_{ij}(t+\Delta t)\right|^2 = \left|\vec{r}_{ij}(t)\right|^2 = {d_{12}}^2 \tag{2.50}$$

where d_{12} is the bond distance. We end up with a series of quadratic equations for λ_{ij}. In practice, these equations are solved numerically in an iterative fashion, and the resulting values for λ_{ij} are used in Eq. 2.48. This SHAKE approach can be generalized to constrain the angles in a molecule [141].

SHAKE is most easily applied to the Verlet algorithm, in which only positions and accelerations (forces) appear. Andersen [142] proposed a modified RATTLE scheme built around the velocity Verlet algorithm. Thus, in RATTLE, we have two stages. In the first stage, adding constraints to Eqs. 2.42 and 2.44 we have:

$$\vec{r}_i(t+\Delta t) = \vec{r}_i'(t+\Delta t) + \frac{\Delta t^2}{2m_i}\vec{g}_i^{\,r}(t) \qquad 2.51$$

$$\vec{v}_i(t+\frac{\Delta t}{2}) = \vec{v}_i'(t+\frac{\Delta t}{2}) + \frac{\Delta t}{2m_i}\vec{g}_i^{\,r}(t) \qquad 2.52$$

The constraint forces in this stage are directed along the bond vectors $\vec{r}_{ij}(t)$. Similar to SHAKE, they are solved iteratively to constrain the bond lengths as Eq. 2.50. The velocity at the half-time step is then calculated with Eq. 2.52. In the second stage, the velocity is evolved fully:

$$\vec{v}_i(t+\Delta t) = \vec{v}_i'(t+\Delta t) + \frac{\Delta t}{2m_i}\vec{g}_i^{\,v}(t+\Delta t) \qquad 2.53$$

Note that the constraint forces here are different than the constraint forces that will be used to fix the bond length in the next time step and they are chosen such that the velocities satisfy the constraints at time $t+\Delta t$:

$$\frac{d(r_{ij}^2)}{dt} = 2\vec{r}_{ij} \bullet \frac{d\vec{r}_{ij}}{dt} = 0 \qquad 2.54$$

Constraint dynamics, such as SHAKE and RATTLE, provide a neat and simple approach to fix high-frequency degrees of freedoms. In the next subsection, we will proceed to the final issue faced in integrating the equations of the motion: the constraint of constant temperature.

2.4.3 Nosé-Hoover Thermostat

The temperature should be kept constant during the simulation for a NVT ensemble. This is usually done by putting the system in contact with a thermal reservoir. Isokinetic [166] and stochastic [167] methods had been designed to achieve this goal but these methods do not reproduce Newtonian Dynamics. An alternative method is the extended system method. The idea is to include a degree of freedom which represents the reservoir in contact, and simulate the resulting "extended system". [163]

Following Nosé's work, Hoover [162] implemented constant temperature dynamics for a fixed time step. The Hamiltonian of this extended system is:

$$H_{Nosé} = \sum_{i}^{N} \frac{\vec{p}_i^{\,2}}{2m_i s^2} + U(\vec{r}^{\,N}) + \frac{p_s^2}{2Q} + g\frac{\ln s}{\beta} \qquad 2.55$$

where s is the additional degree of freedom. Q can be regarded as its effective mass and p_s is its momentum. g is a parameter. Hoover [163] showed that when g was set to be the number of degrees of freedom +1, the ensemble average of the extended system reduces to the canonical average.

$$< A(\vec{p}/s, \vec{r}) >_{Nosé} = < A(\vec{p}/s, \vec{r}) >_{NVT} \qquad 2.56$$

In the ensemble average, the phase space is spanned by the coordinates $\vec{r}$ and the scaled momenta $\vec{p}/s$. In a real simulation, we would rather refer to the scaled momenta as real momenta and interpret $\vec{p}$ as virtual momenta. Hoover [162] wrote the equation of the motions for the real variables as:

$$\frac{d\vec{r}_i}{dt} = \frac{\vec{p}_i}{m_i} \qquad 2.57$$

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial U(\vec{r}^{\,N})}{\partial \vec{r}_i} - \zeta \vec{p}_i \qquad 2.58$$

$$\frac{d\zeta}{dt} = \frac{(\sum_i \frac{p_i^2}{m_i} - \frac{g}{\beta})}{Q} \qquad 2.59$$

$$\frac{ds}{sdt} = \frac{d\ln s}{dt} = \zeta \qquad 2.60$$

Note that Eq. 2.57 to Eq. 2.59 form a closed set and the last equation is redundant. It is usually solved to check the conservation of the Nosé Hamiltonian. These equations of motion are solved iteratively as the momentum appears on the right and left hand sides of Eq. 2.58.

2.4.4 Molecular Dynamics of Rigid Bodies

Up to this point, we have discussed the molecular dynamics of flexible molecular systems. However, it is sometimes economical to treat a molecule or part of a molecule as a rigid body: it translates and rotates about its center of mass. This approach has been applied in numerous situations including disconnected rigid bodies [168-170] and linked semi-flexible molecular models [171-173]. In this subsection, a short overview of the equations of motion involving rotation of rigid bodies in molecular dynamics is provided. [139]

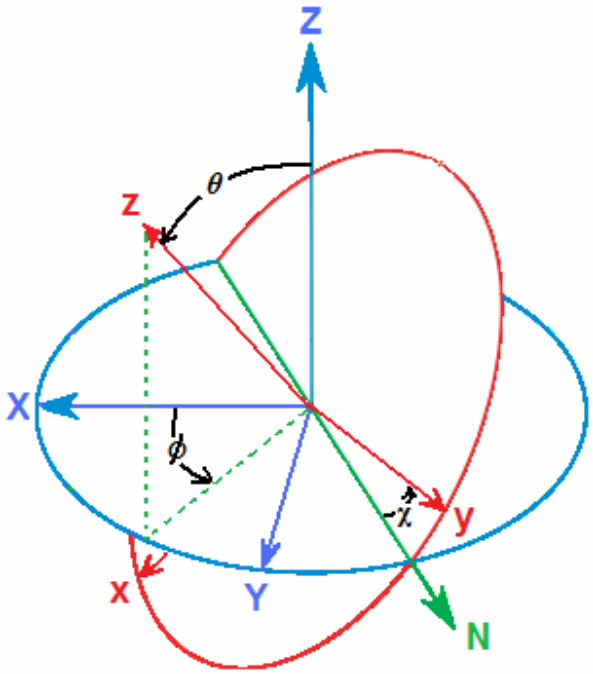


Figure 2.4: Euler angles $\Omega=(\phi, \theta, \chi)$ relating the space-fixed frame ($F=XYZ$) and body-fixed frame ($g=xyz$).

A set of Euler angles, $\Omega=(\phi, \theta, \chi)$, can be used to relate the body-fixed frame (g=*xyz*) with the space-fixed frame (*F=XYZ*) (Fig. 2.4) by three successive rotations: a counter-clockwise

rotation ϕ about Z, a counter-clockwise rotation θ about the line of nodes N, and a counter-clockwise rotation χ about z. With this definition, one can convert the space-fixed coordinates to body-fixed coordinates by a rotation matrix A. It was recognized that with three Euler angles, the equations of motion may be divergent due to singularity problems involving in first derivatives of trigonometric functions of the Euler angles. Evans [174] suggested using quaternions as generalized coordinates. The quaternions are a set of four scalar quantities whose squared sum equals to unity:

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \tag{2.61}$$

So, a set of four quaternions has three degrees of freedom and the last three elements (q_1, q_2, q_3) can be regarded as constituting a vector. Goldstein [139] discussed the way in which this vector may represent the orientation of a rigid body. It is most convenient to define

$$q_0 = \cos\frac{\theta}{2}\cos\frac{\phi+\chi}{2} \tag{2.62}$$

$$q_1 = \sin\frac{\theta}{2}\cos\frac{\phi-\chi}{2} \tag{2.63}$$

$$q_2 = \sin\frac{\theta}{2}\sin\frac{\phi-\chi}{2} \tag{2.64}$$

$$q_3 = \cos\frac{\theta}{2}\sin\frac{\phi+\chi}{2} \tag{2.65}$$

This way, we have,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{2.66}$$

and the rotation matrix A is determined by simple combination of the quaternions. The quaternion time derivatives are:

$$\begin{pmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \begin{pmatrix} 0 \\ \omega_x^b \\ \omega_y^b \\ \omega_z^b \end{pmatrix} \qquad 2.67$$

where ω^b is the body-fixed angular velocity. The equations of motion for rigid bodies can thus be trivially integrated.

To maintain the rotational temperature, a Nosé-Hoover thermostat [162,163] is used to generate canonical (NVT) properties. One could apply the same thermostat to rotations and translations, but we have chosen to apply independent thermostats to avoid the risk of hot-translation/cold-rotation or *vice versa*. The conserved quantity is written as:

$$H_{NH} = E^{full} + \frac{Q_t}{2}\zeta_t^{\,2} + g_t kT \ln s_t + \frac{Q_r}{2}\zeta_r^{\,2} + g_r kT \ln s_r \qquad 2.68$$

where E^{full} is the total kinetic and potential energy of the system, the quantities marked with t and r are for translation and rotation, respectively. The two additional degrees of freedom, s_t for translation and s_r for rotation, have masses of Q_t and Q_r, respectively. g_t and g_r are the translational and rotational degrees of freedom. The equations of motion are integrated following the algorithm of Martyna *et al*. [175].

2.5 Semi-Flexible Models and the Model CSP Interface

Now that we have briefly reviewed the theoretical methods, we use these methods to study the Whelk-O1 chiral stationary phase. In this section, we start with an introduction of the model interface and then proceed to the development of models for the molecules involved.

2.5.1 Model CSP Interface

Computer simulations have been applied to study many kinds of interfacial systems, including but not limited to, aqueous/metal [176], atom/graphite [177], thin-film devices [178], crystal growth [179], self-assembled monolayers on gold surfaces [180,181], peptide/protein on model biomaterials [182,183], and cold-adapted proteins on ice surfaces [184]. Simulations of interfaces present many challenges. Stationary phase interfaces are particularly difficult to represent in simulations because they require an atomic level description of the selector, usually a large molecule, along with some approximation of its distribution on the surface and a representation of the underlying substrate. Nonetheless, some simulations of achiral selectors for HPLC have been published [120-122]. Most notably, the entanglement and solvation of long chain alkanes tethered to silica has been examined in some detail [120-122]. In comparison, only a few simulations of chiral interfaces have been published [185,186]. We note Nita *et al.*'s previous work [185,186] on the solvation of N-(1-phenylethyl)-*N'*-[3-(triethoxysilyl)propyl]-urea, where the chiral molecules were interconnected by hydrogen bonds and consequently closely packed on the surface. Thus, the model surface was highly organized and rigid. The surface coverage of Whelk-O1 does not allow such an approximation to be made but, by virtue of the semi-rigidity of the selective molecule, a realistic model interface is still feasible.

Fig. 2.5 illustrates the rationale for the model Whelk-O1 chiral interface. Experimentally, porous silica beads are coated with the chiral selector. The exposed surface of the bead is dominated by the pores [187], with widths typically between 50-100 Å. The model Whelk-O1 interface is representative of the pore surface: the interface includes (3R,4S)-Whelk-O1 selectors, trimethylsilyl end-caps, and silanol groups corresponding to coverages of 1.07 μmol/m^2, 3.20 μmol/m^2 and 4.26 μmol/m^2, respectively, within the range of experimental values [40,188,189]. An underlying layer of Si atoms, to which end-caps, silanol groups, and selectors are attached, is stationary and represents the solid substrate. Such a simple representation is acceptable since the

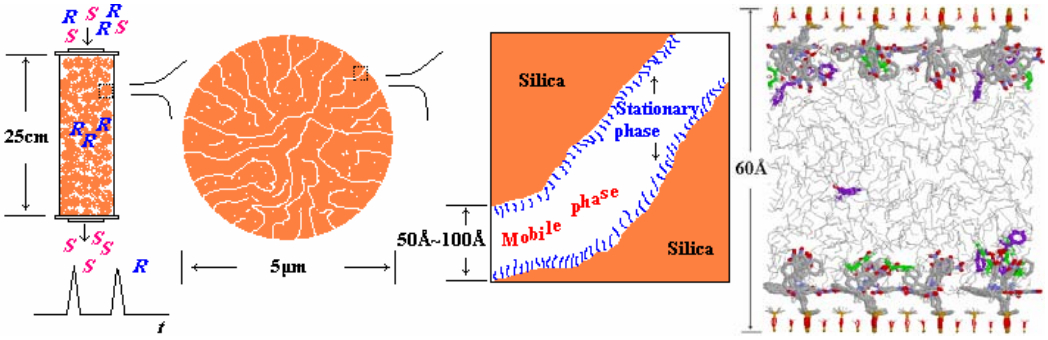


Figure 2.5: The rational design of the simulation cell beginning with the chiral chromatography column and ending with the dimensions of a typical pore. The rightmost picture shows a snapshot of the simulation cell. In the snapshot, the Whelk-O1 selectors are identified by thick grey lines, the *n*-hexane solvent is represented with thin grey lines, and R- and S-1-phenyl ethanol (analytes) are shown in purple and green, respectively.

Si layer in the simulation cell is inaccessible to the solvent and far from the selective portion of Whelk-O1.

The simulation cell includes two chiral interfaces, with solvent and racemate in-between, as the side-view in Fig. 2.5 shows. In actual fact, the figure does not show the entire simulation cell: as discussed in Section 2.3.3, empty space is included above and below the two underlying Si layers. This extra space is required [152] to minimize interactions between pairs of surface-fluid-surface slabs when the simulation cell is replicated in three dimensions.

2.5.2 Molecular Representation of the Whelk-O1 Selector

As shown in Fig. 2.6(a), the selector of the Whelk-O1 CSP consists of two rigid moieties, a dinitrophenyl group and a 1,2,3,4-tetrahydrophenanthrene group, connected by an amide linkage, and is attached to the silica surface by an alkyl tether and a siloxyl group. The selector includes two chiral carbons that connect the 1,2,3,4-tetrahydrophenanthrene group to the alkyl tether and the amide linkage, respectively. The commercially available Whelk-O1 CSP are based on the (3R,4S) form or the (3S,4R) form of the isomers [77]. We will study the (3R,4S) version of the CSP. The phrases “selective portion” and “phenanthryl group” will be used to refer to the non-tether portions of the selector and to the 1,2,3,4-tetrahydrophenanthrene moiety, respectively. P1,

P2, and DNP will be used to represent the middle ring of the phenanthryl group, topmost ring of the phenanthryl group, and the dinitrophenyl group, respectively. Fig. 2.6(b) shows the atom numbering which will be used to identify atoms throughout this thesis.

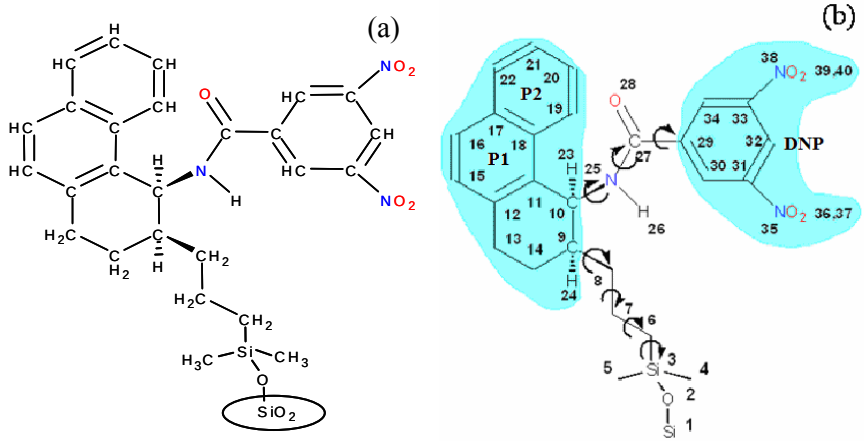


Figure 2.6: (a) Structure of the Whelk-O1 chiral stationary phase. (b) The semi-flexible representation of the Whelk-O1 selector with the two rigid units indicated by shaded areas. The atom numbering shown in (b) will be used throughout this thesis.

A semi-flexible representation of the Whelk-O1 molecule is employed in our simulations. Semi-flexible models are often used to study molecules containing stable structural elements [172,173]. The rings in the Whelk selector are relatively rigid at low temperature, justifying the use of a rigid representation. A semi-flexible representation can effectively reduce internal degrees of freedom and can also avoid the slow convergence problem involved in applications of constraint methods, such as SHAKE [141] and RATTLE [142], on ring systems or triangulated units [172,173,190]. Gay and Berne [191] pioneered the Gay-Berne potential for aspherical groups of atoms treated as a single site and this approach has gained much attention [192,193] because it effectively reduces intermolecular potential pairs and intramolecular degrees of freedom. Alternatively, semi-flexible models bearing linked rigid bodies formed with explicit atoms of invariant relative positions were also developed [171-173]. Kneller and Hinsen [171-173] generalized the Euler equations of motion to the case of linked rigid bodies using Lagrange mechanics. Forester and Smith [172] derived the QSHAKE iterative constraint algorithm for

semi-flexible models. Wu *et al.* [173] developed a method using rigid fragment constraints for simulation of semi-flexible macromolecules. We apply the same [171-173] treatment of the linked rigid bodies, but we have chosen to use intramolecular potentials to constrain the motion of the rigid bodies. Our approach is more straight-forward and can also generate physical trajectories with a moderate time step.

In our semi-flexible representation, the alkyl tether is flexible, with allowance for torsion and bending. The amide linkage is also flexible. The selective portion of the molecule is represented as two rigid units joined by a flexible amide linkage [Fig. 2.6 (b)]. A rigid unit (RU) translates and rotates consistent with the forces on the atoms within the unit. However, the *relative* atomic positions within the unit do not vary with time. In defining an interaction potential, we do not count the intramolecular potentials within the RU since these potentials are constant. But the RU atoms still interact with atoms outside of the unit, including atoms within the same selector, from other selectors, and from other particles in the system. The phenanthryl RU also interacts with the alkyl tether and amide-linkage via bond stretching, angle bending, and dihedral torsion while the dinitrophenyl RU shares bond stretching, angle bending, and dihedral torsions with the amide linkage. These potentials can properly position and orient the RUs.

The energy of the interfacial system is given by:

$$E^{full} = \frac{1}{2}\sum_i m_i \vec{v}_i^{\,2} + \frac{1}{2}\sum_{RU} m_{RU} \vec{v}_{RU}^{\,2} + \frac{1}{2}\sum_{RU} I_{RU} \vec{\omega}_{RU}^{\,2} + \sum U^{LJ} + \sum U^{elec} + \sum U^{\text{intra}} \quad 2.69$$

The first term is the translational kinetic energy of all the atoms, except those involved in RUs. The second and third terms are the translational and rotational kinetic energies, respectively, of the rigid units. m_{RU} is the mass of the rigid unit and is the sum of the masses of all the atoms in the unit. I_{RU} and $\vec{\omega}_{RU}$ are the moment of inertia tensor and angular velocity, respectively, in the principal axis system of the unit. The last three terms are the Lennard-Jones potential (U^{LJ}), the electrostatic potential (U^{elec}) between atoms bearing partial charges, and the intramolecular

potential. The latter consists of three parts, as discussed in Section 2.2.2 and shown in Equation (2.70): bond stretching, angle bending, and dihedral torsion.

$$U^{\text{intra}} = U^{stretch} + U^{bend} + U^{torsion} \qquad 2.70$$

For the studies after Chapter 3, we also add a U^{imp} term to restrict out-of-plane motion.

2.5.3 Intramolecular Potential for the Whelk-O1 selector

A correct representation of the molecular flexibility of the Whelk selector and analytes is essential for study of chiral recognition as induced conformational changes are possible upon docking of analytes onto selectors. However, torsional potentials from available force fields are not specifically designed for our system. We thus undertook a comprehensive series of *ab initio* geometry optimizations to examine the structure and flexibility of the Whelk-O1 selector. The calculations have been carried out using the Gaussian 03 [135] electronic structure program, using density functional theory with the B3LYP functional [133], and the 6-311+G(2d,p) basis set [194,195]. An extensive basis set is required since the molecule is large and branched: Polarization and diffuse functions are required to allow for interactions between atoms that are relatively far apart.

Our modeling starts with a full geometry optimization of an isolated selector. Since the silyl group is not expected to directly impact upon CSP solvation and chiral discrimination, we simply replace the group with a single H for the *ab initio* calculations. Several stable conformers have been identified for the resulting molecule and these are shown in Fig. 2.7. Briefly, the cyclohexene ring in the phenanthryl moiety can adopt two conformers but the barrier is sufficiently high that interconversion will rarely occur at 298K. Within the amide linkage, rotation about the C(10)-N(25) [see Fig. 2.6(b) for atom numbering] bond leads to two minima separated by 22.50 kJ/mol. This energy difference is substantial and population of the local minimum will not be significant at 298 K. The combination of the cyclohexene ring flip with the two minima identified for rotation about the C(10)-N(25) bond leads to four stable conformers.

As shown in Fig. 2.7, the global energy minimum in (a) is substantially lower in energy than the other three conformers.

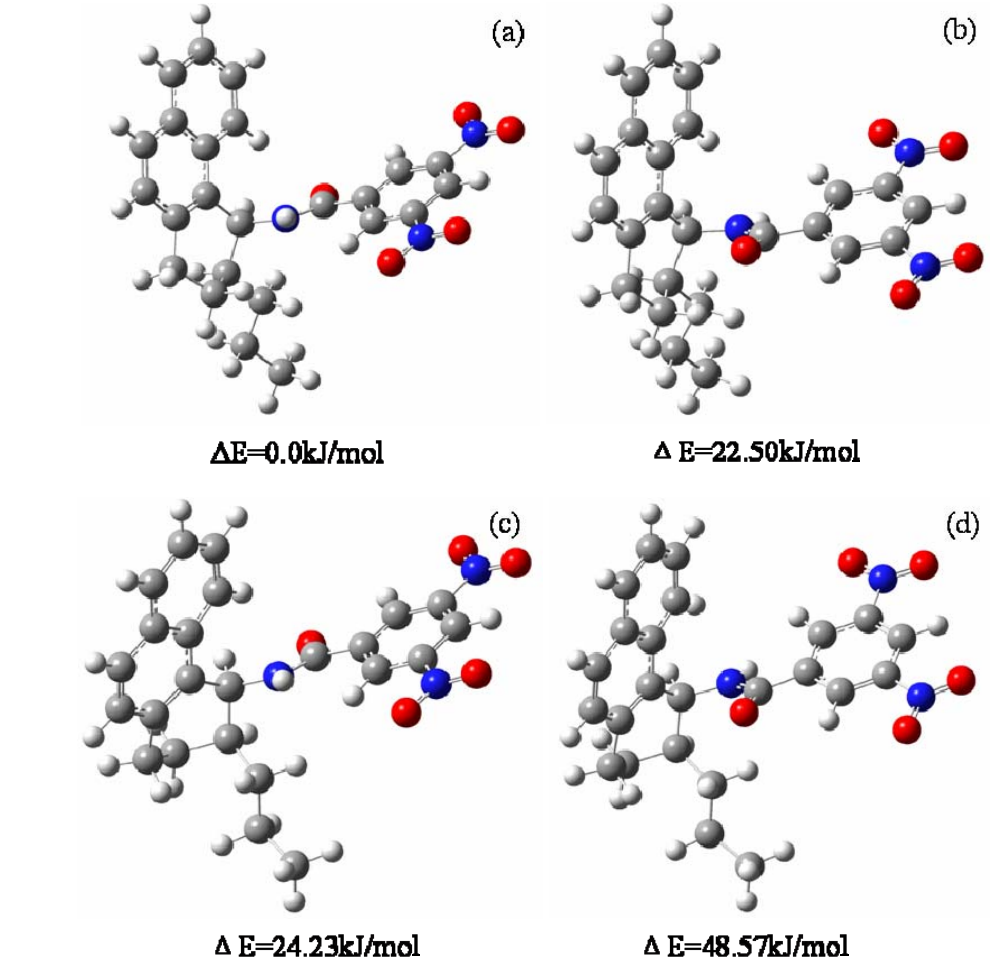


Figure 2.7: Stable conformers of the (3R,4S) Whelk-O1 selector identified from B3LYP/6-311+G(2d,p) calculations: (a) The global minimum, (b) The local minimum corresponding to a roughly 180^{o} rotation about the C(10)-N(25) bond, (c) The local minimum corresponding to a ring flip in the phenanthryl group. (d) The local minimum for both a ring flip and a C(10)-N(25) bond rotation. Energy differences relative to the global minimum are also given.

Since the motion of the rigid unit is governed by translation and rotation about the center of mass, it is necessary to properly position and orient the unit relative to the flexible portions of the molecule. This is accomplished by defining potentials between RU atoms and atoms within the flexible units. For instance, a stretch between the end atom of the tether, C(8), and the associated atom of the RU, C(9), prevents the RU from moving C(9) away from the tether. Bends and

torsions between RU atoms and flexible unit atoms encourage proper orientation of the RU. For example, a bend between C(9)-C(8)-C(7) and a C(10)-C(9)-C(8)-C(7) torsion properly orient the phenanthryl unit relative to the tether.

Consider first bond stretching. The stretching is given by a harmonic potential shown in Eq. 2.21. The C(8)-C(9), C(10)-N(25), and C(27)-C(29) bonds that join flexible moieties to RUs are allowed to stretch and the corresponding force constants are 8.899×10^4 kJ·mol^{-1}·nm^{-2}, 1.017×10^5 kJ·mol^{-1}·nm^{-2}, and 1.100×10^5 kJ·mol^{-1}·nm^{-2}. The corresponding equilibrium bond lengths are from the global energy minimum and may be extracted from Table 2.1. All other bonds are constrained to be rigid by application of the RATTLE algorithm [142].

In the selector, bending motion occurs within the tether, within the amide linkage, and between flexible moieties and RUs. The bending potential is represented by Eq. 2.22. The equilibrium angle is obtained from the global energy minimum and the corresponding force constant is obtained from a least squares fit to ten B3LYP/6-311+G(2d,p) restricted optimization calculations for angles within $\theta_e \pm 5^0$, or from a known force field [196-198]. The full set of bending potential parameters is given in Table 2.2.

Table 2.1: Details of the model Whelk-O1 selector. Atomic positions and CHELPG charges are extracted from the B3LYP/6-311+G(2d,p) geometry optimization. Values marked with *a* are obtained from geometry optimization of $SiH_3OSi(CH_3)_2CH_2CH_3$. Lennard-Jones energy and length parameters are taken from the OPLS [143] (unmarked) and CHARMM [198] (marked with b).

ATOM NO.	ATOM/ GROUP	ε (kJ/mol)	σ (nm)	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)
1	Si	2.5104[b]	0.392[b]	0.0	0.000[a], 0.000[a], 0.000[a]
2	O	0.63639[b]	0.315[b]	0.0	0.000[a], 0.000[a], 1.640[a]
3	Si	2.5104[b]	0.392[b]	0.0	0.326[a], 0.923[a], 2.985[a]
4	CH_3	0.7322	0.3905	0.0	-0.767[a], 2.450[a], 2.964[a]
5	CH_3	0.7322	0.3905	0.0	2.136[a], 1.423[a], 2.964[a]
6	CH_2	0.49371	0.3905	-0.063116	-0.062, -0.175, 4.461
7	CH_2	0.49371	0.3905	0.080299	0.163, 0.462, 5.833
8	CH_2	0.49371	0.3905	0.015478	-0.172, -0.488, 6.988
9	C	0.27614	0.350	0.000603	0.030, 0.121, 8.384
10	C	0.27614	0.350	0.450510	-0.700, -0.678, 9.485
11	C	0.29288	0.3550	-0.287480	-0.467, -0.118, 10.881
12	C	0.29288	0.3550	0.135576	0.636, 0.672, 11.145
13	CH_2	0.49371	0.3905	0.000408	1.618, 1.092, 10.073
14	CH_2	0.49371	0.3905	0.042355	1.499, 0.300, 8.773
15	CH	0.46024	0.3750	-0.062882	0.856, 1.156, 12.462
16	CH	0.46024	0.3750	-0.08132	0.000, 0.860, 13.482
17	C	0.29288	0.3550	0.157075	-1.148, 0.062, 13.251
18	C	0.29288	0.3550	0.167167	-1.392, -0.436, 11.934
19	CH	0.46024	0.3750	-0.086607	-2.556, -1.224, 11.733
20	CH	0.46024	0.3750	0.014195	-3.416, -1.501, 12.768
21	CH	0.46024	0.3750	0.037776	-3.166, -1.010, 14.064
22	CH	0.46024	0.3750	-0.102292	-2.053, -0.244, 14.296
23	H	0.12552	0.2500	-0.029142	-1.764, -0.663, 9.251
24	H	0.12552	0.2500	-0.038570	-0.442, 1.113, 8.387
25	N	0.71128	0.3250	-0.521480	-0.324, -2.107, 9.473
26	H	0.0000	0.0000	0.214072	0.482, -2.364, 10.021
27	C	0.4389	0.3750	0.479756	-1.152, -3.094, 9.049
28	O	0.87864	0.2960	-0.503862	-2.260, -2.895, 8.571
29	C	0.29288	0.3550	0.102854	-0.638, -4.508, 9.199
30	CH	0.46024	0.3750	-0.033172	0.713, -4.832, 9.278
31	C	0.29288	0.3550	0.112896	1.090, -6.164, 9.393
32	CH	0.46024	0.3750	-0.014081	0.168, -7.194, 9.434
33	C	0.29288	0.3550	0.129061	-1.170, -6.843, 9.346
34	CH	0.46024	0.3750	-0.012581	-1.589, -5.529, 9.219
35	N	0.50208	0.325	0.627677	2.530, -6.493, 9.469
36	O	0.71128	0.296	-0.400465	3.322, -5.560, 9.421
37	O	0.71128	0.296	-0.380313	2.835, -7.671, 9.575
38	N	0.50208	0.325	0.612602	-2.186, -7.921, 9.387
39	O	0.71128	0.296	-0.385666	-1.780, -9.071, 9.482
40	O	0.71128	0.296	-0.377331	-3.360, -7.589, 9.324

Table 2.2: Details of the bending potentials for the model Whelk-O1 selector. The first column lists the three atoms that define the angle following the numbering in Fig. 2.6(b). The equilibrium angles, extracted from *ab initio* calculations, are given in the second column. The force constants, and their sources, are given in the third and fourth columns, respectively.

Angle	**θ_e (degree)**	**k_θ ($kJ \cdot mol^{-1} \cdot rad^{-1}$)**	**Source**
3,2,1	143.954	142.256	Ref. [198]
4,3,2	109.117	209.10	Ref. [197]*
5,3,2	109.117	209.10	Ref. [197]*
6,3,2	105.647	209.10	Ref. [197]*
5,3,4	110.161	259.81	Ref. [197]*
6,3,4	111.318	259.81	Ref. [197]*
6,3,5	111.318	259.81	Ref. [197]*
7,6,3	115.513	259.81	Ref. [197]*
8,7,6	112.698	259.81	Ref. [197]
9,8,7	114.228	259.81	Ref. [197]
10,9,8	112.343	427.84	This work
14,9,8	113.787	356.68	This work
24,9,8	107.668	237.63	This work
25,10,9	112.023	450.41	This work
25,10,11	108.913	347.91	This work
25,10,23	105.086	317.60	This work
10,25,26	116.503	158.99	Ref. [196]
10,25,27	123.392	209.20	Ref. [196]
25,26,27	118.225	146.45	Ref. [196]
25,27,28	123.898	334.72	Ref. [196]
25,27,29	116.211	292.88	Ref. [196]
28,27,29	119.890	334.72	Ref. [196]
27,29,30	123.606	399.84	This work
27,29,34	117.042	455.01	This work

*: Si atom is replaced by CH_n, n=0-3

A modified Ryckaert-Bellemans [145] potential, as described by Eq. 2.24, is adopted for torsions. Torsions within the tether were represented by well-known literature potentials for hydrocarbon chains [199]. Published potentials [143] are not of sufficient accuracy for torsions within the amide linkage, and between flexible portions of the selector and the RUs, so we have extracted potentials from *ab initio* data. Specifically, following the numbering in Fig. 2.6(b), we have considered the following torsions: motion about the C(8)-C(9), C(10)-N(25), N(25)-C(27), and C(27)-C(29) bonds. For these motions, the torsional potentials are shown in Fig. 2.8. The filled circles represent B3LYP/6-31G* calculations with the exception of Fig. 2.8(a), where B3LYP/6-311+G(2d,p) results are presented. Ten-to-fifteen restricted geometry optimizations, where the dihedral angle is given a fixed value while the other parameters are fully optimized, were performed to characterize each potential.

The solid lines in Fig. 2.8 show the potential curves used in the simulations. With the exception of Fig. 2.8(d), the potential parameters have been obtained from least-squares fits [200] to the *ab initio* data. The fitted potentials in Figs. 2.8(a) and (b) reproduce the *ab initio* potentials very well. In Fig. 2.8(c), the fitted curve is somewhat shifted but the shift is not critical since the curve reproduces the *ab initio* data very well near the global minimum. Also, the local minimum at -100 degrees would be rarely sampled due to its high energy during the simulation. For the torsion about the N(25)-C(27) bond, we have chosen literature values [201] of the potential parameters to represent this motion, since the curve [Fig. 2.8(d)] is accurate for the angles around the global minimum.

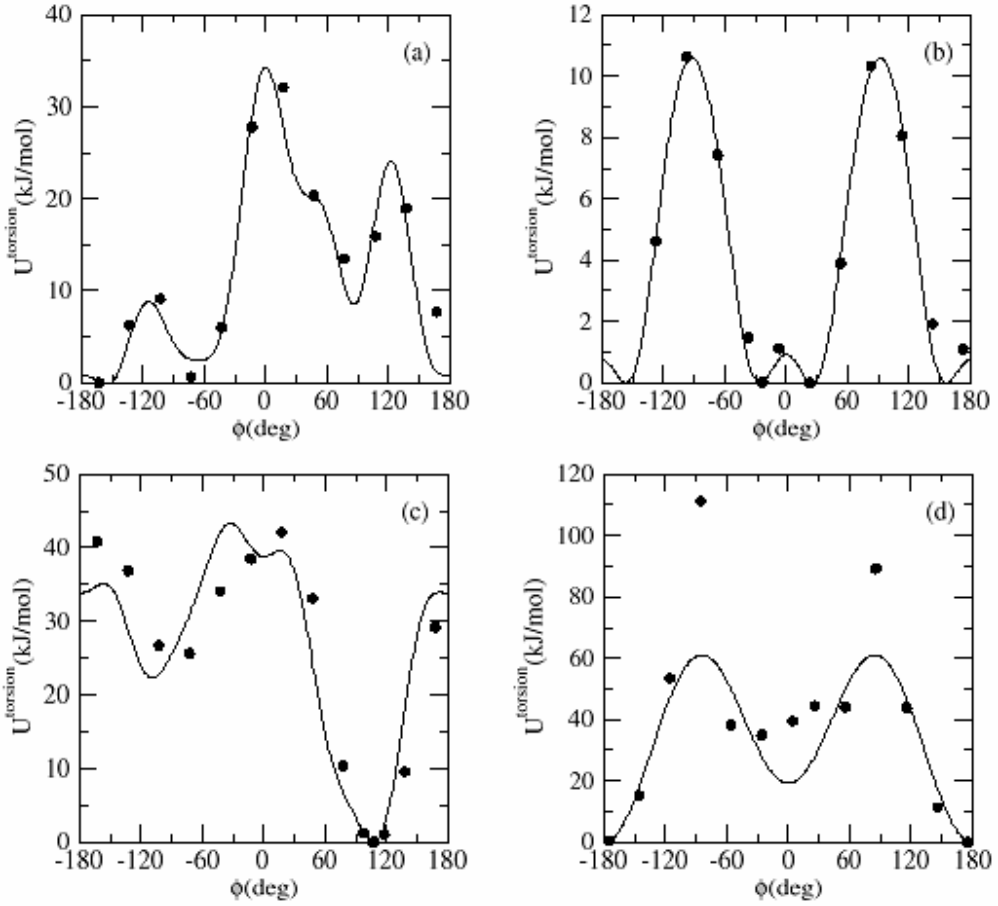


Figure 2.8: Torsional barriers from *ab initio* calculations. The torsional potentials are given for rotation about (a) C(7)-C(8)-C(9)-C(10), (b) N(25)-C(27)-C(29)-C(30), (c) C(9)-C(10)-N(25)-C(27), (d) C(10)-N(25)-C(27)-C(29). *ab initio* points are represented by filled circles. The torsional potentials used in the simulations are represented by solid lines.

Within the amide linkage, the N(25)-C(27) torsion leads to a global energy minimum, corresponding to a dihedral angle of 180 degrees, which keeps the carbonyl and amide bonds in-plane. In contrast, torsion about C(27)-C(29) leads to two nearly degenerate energy minima, separated by a very low energy barrier: the energy difference between the minima is less than 0.15 kJ/mol, and the barrier is less than 1 kJ/mol, which is well within the error of the B3LYP functional [133]. The maximum between the two minima corresponds roughly to the C=O group lying in the plane of the dinitrobenzoyl ring. At room temperature, this dihedral angle will easily vary between -40 and 40 degrees. Thus, the dinitrobenzoyl moiety can reorient itself to

accommodate, and interact favorably, with analytes. The introduction of a torsional potential for the C(10)-N(25) bond means that interconversion between Figs. 2.7(a) and 2.7(b) during a simulation is possible, at least in principle. Torsion about the C(10)-N(25) bond is somewhat flexible, with a range of roughly ±10 degrees and a minimum at 108 degrees. This motion provides the nitrogen with some flexibility such that it may point its hydrogen toward a hydrogen bond acceptor, with very little energy penalty.

2.5.4 Intermolecular Potential of Whelk and Models for Other Components

The intermolecular potential, including the Lennard-Jones potential (U^{LJ}) and electrostatic potential (U^{elec}), are calculated for all pairs of atoms except when the two atoms belong to the same molecule and are less than 4 bonds apart, or when the two atoms belong to a given rigid unit. Electrostatic potentials are evaluated with Ewald summations, as discussed above in Section 2.3.3. The atomic charges in Table 2.1 are obtained using the CHELPG algorithm [137] as applied to the global energy minimum [Fig. 2.7(a)] of the Whelk-O1 selector.

All LJ parameters are taken from the literature, but a single parameter set could not be used due to the diversity of the atoms in the interfacial system. The OPLS force field [199,201,202] is used for the Lennard-Jones potential of the Whelk-O1 molecules and the values are listed in Table 2.1. For the Si atoms, OPLS parameters are not available, and CHARMM [198] values have been used instead.

The intramolecular potentials for the end-caps and the solvents have the same form as for the Whelk-O1: Eqs. 2.21, 2.22, and 2.24 for stretches, bends, and torsions, respectively. The trimethylsilyl end-cap corresponds directly to a truncated Whelk-O1 tether. As a result, the end-cap force fields are the same as the Whelk-O1 potentials applied to the tethers. Specifically, we have chosen OPLS and CHARMM LJ parameters and published potentials for bends and torsions [197,198,203]. The RATTLE algorithm [142] is employed to keep the bond lengths constant. Since the atomic charges on the solvent-accessible portion of the end-cap (the methyl groups) are

expected to be quite small, we have set the atomic charges to zero. Solvent models are adapted from available force fields and will be discussed later in each chapter. All solvent models include bends and torsions, with a consistent functional form for all solvents, and all solvent bond lengths are kept constant using RATTLE [142].

2.5.5 Molecular Simulation Details

As discussed above, a typical simulation cell includes two parts [185]: a slab with two surfaces and solvent and/or analyte in-between (see Fig. 2.4 for example); and empty space above and below the slab. The surface dimensions (L_x and L_y) are adjusted so that a solvent molecule located in the center of the simulation cell is far from either surface. The distance between the two surfaces is determined by the experimental solvent densities. Specifically, in order to achieve the correct solvent density in the simulation cell, we account roughly for the volume of the Whelk-O1 selectors and end-caps in setting the distance between the Si surfaces. Following this, the bulk density in the center of the simulation cell is examined during the simulation to ensure that it is within 5% of the experimental density. Thus, because of inexact estimates of surface volume, the solvent density in the simulation will be close, but not exactly the same, as the experimental density. Ewald summations [151] have been used to treat the long–ranged Coulombic forces between charged atoms. Yeh and Berkowitz [152] have shown that the empty space, along with a correction for the elongated shape of the simulation cell, allow for the use of 3D Ewald sums for systems with a slab geometry. The Ewald convergence parameter of $\alpha = 1.756 / L_x$, with a reciprocal space cutoff of $k^{2*} \leq 27$. Take the simulation cell of the *n*-hexane/2-propanol solvated CSP as an example, the full simulation cell roughly has dimension of 49.20Å x 49.20Å x 150.0Å and 62.7% of this volume is empty space beyond the slabs.

For initialization, the Whelk selector is attached to the surface in its global minimum conformation from DFT calculations. The end-caps, silanol groups, and underlying silicon atoms are positioned sequentially, as will be discussed in Chapters 3 and 4, on a silica (1,1,1) atom grid.

The analyte molecules, if present (Chapters 4, 5, and 6), are attempted to be positioned randomly in the middle of the simulation cell in their global minimum structures. For each attempt, an evaluation of the interacting energy of the all analytes and the surface molecules is carried out to make sure that no analyte atoms overlap. After an attempt is accepted, the centers of mass of all the rigid units are calculated and a stepwise search of the Euler angles for each rigid unit in the principal axis system is carried out. The found Euler angles are transferred to starting quaternions for the rigid units. Excessive numbers of solvent molecules are added in the middle of the simulation cell on a cubic lattice of grid points. Solvent molecules are then removed randomly to match the desired number. Monte Carlo moves are generated for solvent molecules to relax the possible high potential due to solvent-solvent, solvent-analyte and solvent-surface interactions.

After the initialization process, extensive MD simulations were performed in the NVT ensemble. Two Nosé-Hoover thermostats [162,163] are used to generate canonical (NVT) properties. In our simulations, we employ independent translational and rotational thermostats. The equations of motion are integrated following the algorithm of Martyna *et al*. [204] to preserve the RATTLED [142] positions and velocities.

2.6 Coding and Other Practical Issues

2.6.1 Coding for semi-flexible surface simulation in the MDMC program

All MD simulations are run with the locally developed MDMC program [205]. Prior to this thesis, this program had been developed and applied in the simulation of bulk condensed phases [206], liquid crystals with single-site rigid bodies [207], and interfacial systems with rigid surfaces [185]. To accomplish the objectives of this thesis, extensive FORTRAN coding was required to enable simulation of interfacial systems with semi-flexible surface.

Generally, the MDMC program treats the surface and bulk molecules separately by using two sets of variables. Thus, in order to allow simulation of interfacial systems, the subroutines in the program were extended and further developed. For instance, the intermolecular force and

energy calculations are extended from bulk-bulk to bulk-surface, and surface-surface. The integration methods were also extended to accommodate surface molecules.

In other sections of the MDMC program, more significant modifications were required. For initialization, codes were developed to generate the starting configuration of the simulation for an interfacial system with rigid units, as described in Section 2.5.5. Codes have been developed for the treatment of improper torsions (Eq. 2.26) and modified Rickaert-Bellemans [145] torsions (Eq. 2.25). To enable the simulation of rigid bodies formed by multiple atoms, routines were developed to accumulate forces and torques on the rigid unit centers of mass, and to obtain space-fixed atomic positions from the center of mass positions of the rigid units. Subroutines were modified to apply two Nosé-Hoover thermostats for use in translational dynamics of RATTLED [142] flexible surface and bulk molecules and rotational dynamics of the rigid units respectively. For property sampling, codes were developed to collect two-dimensional radial distributions of the bulk atoms about surface atoms. Routines were also modified to collect radial distributions and surface distributions of the surface atoms. To be discussed later in Section 2.6.2, the whole program is parallelized with message-passing interface [208] to enable efficient use of multiple CPU processors.

Other than the simulation programs, various programs and scripts are also developed to process calculation data. Such programs include the fit of *ab initio* data for torsional potentials to Eq. 2.24 and Eq. 2.25 using the NAG [200] least square fit, the energy minimization for semi-flexible molecules using a combined sequential and Monte Carlo searching procedure, and the identification of docking poses of analyte-selector complexes and the evaluation of their energetics.

2.6.2 Message-Passing Interface for Nonbonding Force Calculations

Force calculations, especially non-bonding force calculations are typically the most resource demanding part of a molecular dynamics simulation program. The message-passing interface

(MPI) has been employed in this thesis to parallelize the nonbonding force calculations. The message-passing model of parallization posits a set of computer processors that have only local memory but communicate with other processors by sending and receiving messages [208]. The simulation program uses a master/slave, replicated data model. Briefly, the master node does all the initialization, property sampling, and the Input/Output. It calculates all the intramolecular forces and integrates the equations of motion as well. The calculation of the LJ forces and Ewald summation is accomplished in parallel. Prior to the MD simulation, the task is distributed evenly on all the master and slave processors and the topological information including LJ parameters, partial charges, and atomic connectivity is broadcast from the master processor to the slave processors. During the MD simulation at each time step, communication is done to update the coordinates (MPI_BCAST) and accumulate the nonbonding forces (MPI_REDUCE). We have chosen to run simulations mostly with 16 processors, which provide a good combination of efficiency and speed. For a typical simulation, the scaling factor is 0.75 for 16 AMD Opteron CPUs: it is about 12 times faster to run with 16 processors in parallel than with a single processor in serial.

2.7 Problems Tackled in This Thesis

Using these theoretical tools, we systematically study the Whelk-O1 chiral stationary phase. In Chapter 3, we study the solvation structure of the Whelk-O1 CSP and solvent partitioning in the stationary phase and the bulk phase. In Chapter 4, styrene oxide and stilbene oxide are added into the system and their docking behaviour at the CSP interface is analyzed through snapshot analysis and atomic distributions. Their three-point interaction mechanism with Whelk-O1 is examined. The solvent effect on enantioselectivity is briefly studied by simulating the docking of styrene oxide with different solvent compositions. In Chapter 5, the resolution of 10 analytes on the Whelk-O1 CSP is simulated. The elution order and separation factors are predicted and are consistent with available experiments. The enantioselective mechanism for the Whelk-O1 CSP is

studied systematically. In Chapter 6, under the guidance of the obtained enantioselective mechanism, rational optimizations of the Whelk-O1 selector are proposed, and *in silico* synthesis (model development) and evaluation (simulation) are carried out to virtually screen the modifications.

Chapter 3
Solvation of the Whelk-O1 Chiral Stationary Phase

3.1 Introduction

In this chapter, we consider the solvation of the selective interface of the (3R,4S)-Whelk-O1 chiral stationary phase. Whelk-O1 has been widely used in industry and scientific research [209,210] because of its separation efficiency, versatility, and stability. In particular, there has been a number of experimental studies of the enantioselectivity of Whelk-O1, including separations in a normal-phase hexane/alcohol solvent [77,81], a supercritical CO_2(sc-CO_2) /methanol solvent [90-92], and reversed-phase HPLC in a water/methanol solvent [89]. Thus, we will consider these solvent mixtures.

In liquid or supercritical chromatography, the solvent may participate actively in the enantioselective process. The role of the solvent is difficult to include in the modeling of selectivity, but is clearly required to understand fully the mechanism of selection. In particular, gas phase dimer energies are not an ideal basis upon which to identify the most important docking interactions in a fluid environment since the relative importance of various geometries is significantly modified by the solvent. In many instances, such effects will not be well represented by a continuum solvent either since a certain number of solvent molecules may be directly involved in the complex. In this chapter, we employ molecular dynamics simulations to analyze the solvent distribution in the vicinity of the Whelk-O1 interface. We are also interested in the hydrogen bonding interactions between the solvent and Whelk-O1 selectors, since these could interfere with hydrogen bonds formed by the analyte, which, as noted in Chapter 1, play an essential role in the proposed three-point interaction models for enantioselective mechanism of the CSP.

This chapter begins with a description of details of the molecular dynamics simulations for the Whelk-O1 interfaces in Section 3.2. Solvent distributions and hydrogen bonding at the Whelk-O1 interface are considered in Section 3.3. Selector mobility is also examined in Section 3.3. The chapter concludes with a brief discussion of the results in Section 3.4.

3.2 Theoretical Details

3.2.1 Models

The semi-flexible representation of the (3R,4S)-Whelk-O1 selector is employed in our simulations. The full details of the model have been described in Chapter 2. Briefly, the selective portion of the molecule is represented as two rigid units joined by a flexible amide linkage [Fig. 2.6(b)]. Intramolecular bond stretching, angle bending and dihedral torsional potentials are derived from a series of B3LYP/6-31G* calculations. LJ parameters are adopted from the OPLS [143] and CHARMM [198] force fields. Atomic charges are obtained from the CHELPG charges [137] of the equilibrium structures. The model of end-caps has been discussed in Chapter 2. Siepmann's TraPPE-UA model [211] was chosen for the *n*-hexane, 2-propanol, and methanol solvents. This model was developed from simulations of the alcohols and their mixtures with alkanes. The F3C water model [212], and the EPM2 sc-CO_2 model [213] have also been chosen. All solvent bond lengths are kept constant using RATTLE [142].

3.2.2 Simulation Details

Extensive MD simulations were performed for the Whelk-O1 interface in the presence of three solvents. First, a typical normal-phase solvent is considered: *n*-hexane/2-propanol 80/20 (volume). Second, a reversed-phase solvent is examined: water/methanol 70/30 (mass). Finally, the solvation of Whelk-O1 in supercritical CO_2, with 13.73% (mole fraction) methanol is examined. These particular binary solvents have been chosen since the Whelk-O1 stationary phase has been used to separate enantiomers in each of these three diverse environments [81,89-92,97].

As discussed in Chapter 2, a typical simulation cell includes two parts: a slab with two surfaces and solvent in-between (see Fig. 2.5 for example); and empty space above and below the slab. Ewald summations with a correction for slab systems [152] have been used to treat the long–ranged Coulombic forces between charged atoms. The surface dimensions (L_x and L_y) are adjusted so that a solvent molecule located in the center of the simulation cell is far from either surface. The distance between the two surfaces is determined by the experimental solvent densities: for *n*-hexane/2-propanol, at 25°C, the density is 8.681 mol/L [214]; for sc-CO_2/methanol the density, at 50°C, is 18.982 mol/L [215]; for water/methanol the density, at 25°C, is 45.672 mol/L [216]. The solvent densities are set to be within 5% of experimental values by the method described in Chapter 2.

The full details of the model interface are described in Chapter 2. Briefly, the model surface for *n*-hexane/2-propanol is shown in Fig. 3.1, and 16 of the 64 molecules on the surface are Whelk-O1, the remaining molecules are trimethoxysilyl end-caps. For water/methanol and sc-CO_2/methanol, there are 9 Whelk-O1 molecules on each model surface. The surface dimensions are thus reduced accordingly, requiring fewer solvent molecules to ensure there are only negligible interactions between the two opposite surfaces.

For each of the three solvents, two 1,500,000 time step molecular dynamics simulations have been performed. Our method of generating the starting configuration has been discussed in detail in Chapter 2. Once the starting configuration is obtained, the fluid equilibrates for the first 50,000 time steps leaving a 1,450,000 step collection period. The time steps are 0.71, 0.82, and 0.96 fs for sc-CO_2/methanol, water/methanol, and hexane/2-propanol, respectively. These time steps are sufficient to limit the change in H_{NH} within reasonable limits (0.6% to 1.7% of the total energy fluctuations over 10 ps). The equilibrium properties reported in Section 3.3 are averages over four surfaces (2 surfaces/simulation and 2 simulations per solvent), reflecting a total collection period of 4.35 nanoseconds. All simulations are performed with the MDMC program [205].

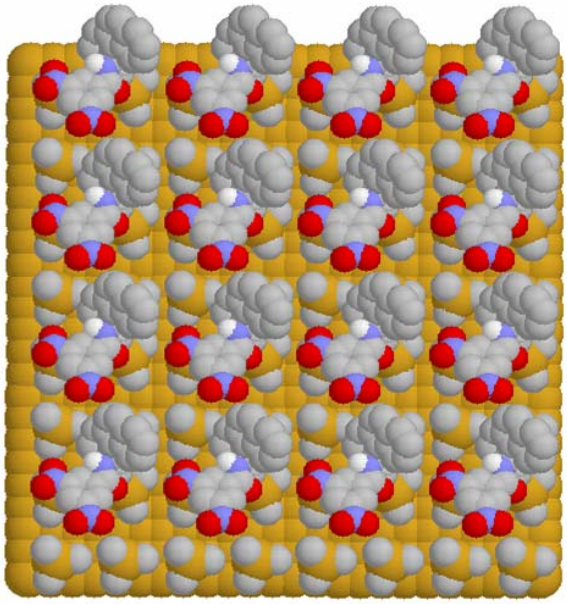


Figure 3.1: A snapshot from directly above showing the initial Whelk-O1 interface. 16 Whelk-O1 molecules (coverage of 1.10 μmol/m^2) and 48 trimethyl silyl end-caps (coverage of 3.29 μmol/m^2) are bonded to the underlying silicon surface of 24*24 silicon atoms with a Si-Si distance of 0.205 nm.

3.2.3 Properties

In order to explore the solvent partitioning and hydrogen bonding about the selective molecules, we collect several distributions of the solvent atoms. These distributions include the radial distribution, the surface distribution, and the 2-dimensional cylindrical distribution.

Radial distributions, *g(r)*, between solvent and Whelk atoms have been evaluated. Several radial distributions are of interest, but we note that these do not decay to unity, as one might expect, because of the surface atoms, the underlying silicon, and the empty space beyond the surface.

Surface distribution functions are collected for the Whelk-O1 atoms and the solvent molecules. These distributions are obtained from

$$g(z) = \frac{n^{act}(z)}{n^{ideal}(z)} = \frac{n^{act}(z)L_z}{N\Delta z} \qquad 3.1$$

where $z=0$ for the underlying, immoveable, layer of silicon, $n^{act}(z)$ is the actual number of atoms found at a distance between z and $z+dz$ above the underlying silicon and n^{ideal} is the number expected for an unstructured fluid. Thus, $g(z)$ identifies areas of concentration and depletion for each atom above the surface. In Eq. 3.1, N is the total number of atoms in the simulation cell and L_z is the distance between the two surfaces. In our case, we have $\Delta z=(1/200)L_z$, which means we divide the intersurface distance into 200 slices. Since the surfaces are identical, the distributions above both surfaces are averaged in the results reported below.

Hydrogen bonding between the solvent and Whelk-O1 can be explored in several ways. First, we apply a geometric criterion [217,218] to define a hydrogen bond: the distance between hydrogen and H-bond acceptor should be less than 2.6Å and the angle formed between donor-H-acceptor should be larger than 150°. We report average numbers of hydrogen bonds, based on applying the geometric criterion to the selector-solvent and selector-selector pairs, for every 2000 iterations of the simulations. Second, snapshots of important hydrogen bonding arrangements about the amide linkage are also provided. Finally, 2-dimensional (2D) cylindrical distributions $g(r_c,z_c)$ are used to study the relative positions of the solvent atoms around the selector. For a 2D distribution the interatomic distance is separated into two components, r_c and z_c, according to $r = \sqrt{(x_c^2 + y_c^2) + z_c^2} = \sqrt{r_c^{\,2} + z_c^{\,2}}$, where we have added the subscript "c" to emphasize that the coordinates are cylindrical. A 2D distribution is calculated from

$$g(r_c, z_c) = \frac{n^{act}(r_c, z_c)}{n^{ideal}(r_c, z_c)} = \frac{n^{act}(r_c, z_c)}{2\pi\, r_c\, \Delta r_c\, \Delta z_c\, \rho} \qquad 3.2$$

where a positive z_c means the solvent atom is above the reference atom. That is, $z_c > 0$ when the solvent atom is further from the underlying silicon layer than the reference atom. In Eq. 3.2, ρ is the number density of the solvent, and $\Delta r_c = \Delta z_c = 0.094$ Å. Thus, we analyze areas of solvent concentration and depletion above the surface and along the surface. This analysis is particularly

instructive for hydrogen bonding and we examine 2D distributions to identify the most probable positions for hydrogen bond donors/acceptors about the Whelk-O1 selectors.

3.3 Results

3.3.1 Flexibility of Whelk-O1 in Solvated Environments

At the beginning of a simulation, each Whelk-O1 molecule shares the same orientation relative to the underlying silicon (Fig. 3.1). Specifically, each molecule is positioned according to the global energy minimum shown in Fig. 2.6(a). One might assume that the molecules do not move much during the course of a simulation by virtue of the size of the Whelk-O1 molecule. This assumption is incorrect, as shown in Fig. 3.2, where the snapshot shows the CSP after 1.35 nano seconds. The degree of reorientation is extensive although most of the Whelk-O1 CSP molecules keep the amide hydrogen, the phenanthryl moiety, and the dinitrobenzoyl group open to the solvent. The open conformation implies that, during separation, analytes have relatively easy access to the chiral region and to the three structural elements identified in the proposed three-point binding model. In Fig. 3.2, two Whelk molecules have reoriented to hydrogen bond with each other (lower left hand corner) although this is a very rare occurrence. In fact, the highest rate of Whelk-Whelk hydrogen bonding occurs in the supercritical solvent, where only 1.5% of the surface molecules participate in a Whelk-Whelk hydrogen bond. In the other solvents, less than 0.1% of the Whelk molecules participate in Whelk-Whelk hydrogen bonds.

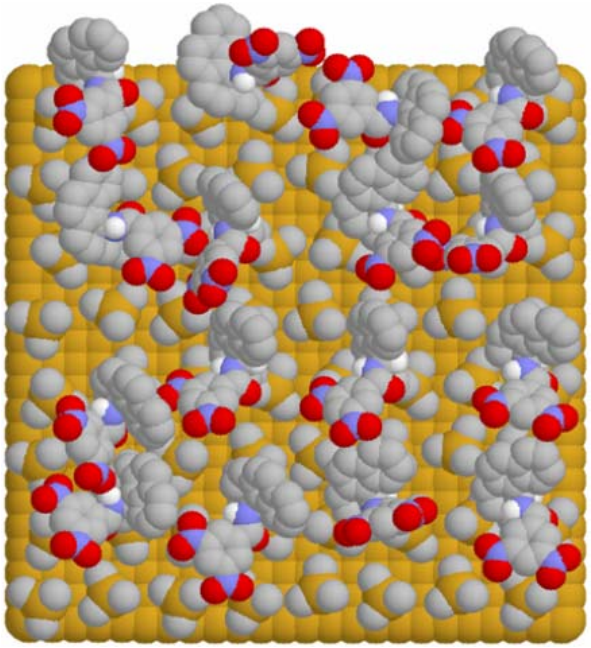


Figure 3.2: A snapshot from directly above illustrating the range of conformers explored by Whelk-O1 molecules.

The distribution of Whelk-O1 selector atoms above the underlying silicon is shown in Fig. 3.3 for the *n*-hexane/2-propanol solvated Whelk-O1 interface. From Fig. 3.3(a), the tether atoms are found between 0 and 8Å above the surface, with more localization for atoms closest to the surface. Specifically, the first 6 atoms are quite localized while the last two atoms, C(7) and C(8) [see Fig. 2.6(b) for numbering], of the tether occupy a much larger range of positions. The C(7) distribution also has 2 local maxima, corresponding to trans and gauche conformations of the tethers. The end-caps are found between 0 and 5Å of the surface, so the topmost atom of the Whelk tether, C(8), is 0-3 Å above the end-caps.

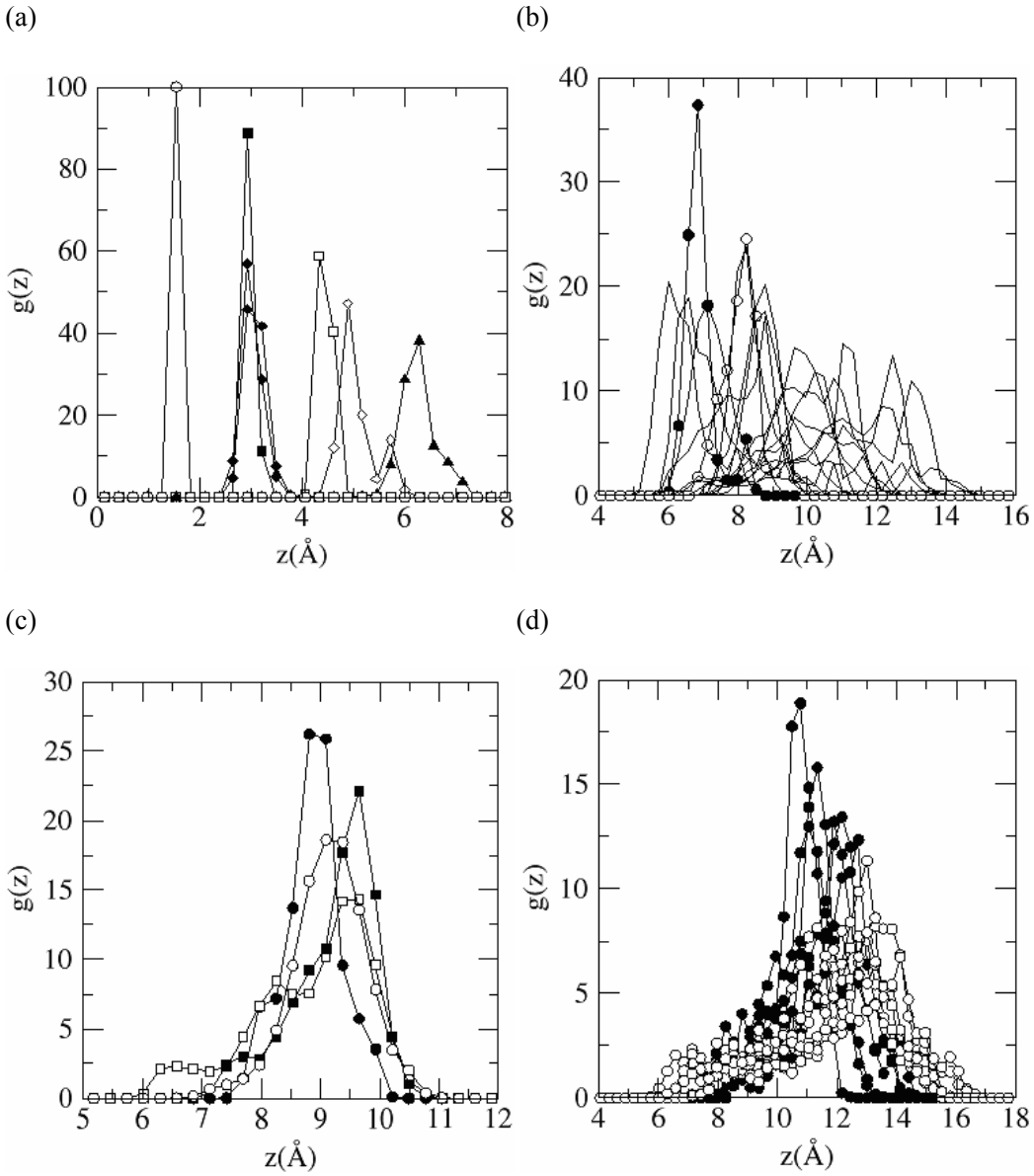


Figure 3.3: Distribution of the Whelk-O1 selector in a *n*-hexane/2-propanol solvent. (a) The distribution of tether atoms from the surface. Open circles, filled squares, open squares, filled diamonds, open diamonds, and filled triangles correspond to O(2), Si(3), C(4), C(5) and C(6), C(7), C(8), respectively. (b) The distribution of the phenanthryl atoms [C(9)-H(24)] from the surface. The curves with filled and open circles identify the two chiral carbons C(9) and C(10), respectively. (c) The distribution of the amide linkage atoms from the surface. Filled circles, open circles, filled squares, and open squares correspond to N(25), H(26), C(27), O(28) respectively. (d) The distribution of the dinitrobenzoyl atoms from the surface. Filled and open circles correspond to C(29)-C(34) and N(35)-O(40), respectively.

The atomic distribution for the phenanthryl group, the amide bridge, and the dinitrobenzoyl group are shown in Figs. 3.3(b), (c), and (d), respectively. The phenanthryl atoms are located within 5-15 Å above the underlying silicon layer, with the chiral carbons located between 6 and 10 Å and showing the sharpest distributions. The atoms in the amide bridge are found between 6-11 Å above the surface, with the atoms closest to the dinitrobenzoyl group being the most delocalized. The carbonyl group in the tether is, on average, 0.5 Å further away from the surface than the nitrogen. The atoms in the dinitrobenzoyl group are found within 6-16 Å of the surface, a very broad range of positions due in part to the flexibility of the torsion about the C(27)-C(29) bond. From Fig. 3.3(d), the nitro groups are very broadly distributed and may be found directly above the end-caps or pointing directly into the fluid (from 5-16.5 Å above the underlying silicon layer). Despite this, most dinitrobenzoyl groups lie roughly parallel to the surface, or point slightly towards the bulk (Fig. 3.2).

Pirkle *et al.* [40] examined the enantioselectivity of Whelk-O1 in *n*-hexane/2-propanol for a relatively short propyl tether and a longer undecyl tether. They found that the shorter tether afforded more enantioselectivity. The Whelk-O1 distributions in Fig. 3.2 and Fig. 3.3 suggest that the shorter tether serves to limit the orientational range displayed by the selective portion of the Whelk-O1 molecule. Specifically, some degree of orientational similarity between individual Whelk-O1 selectors is evident in Fig. 3.2. A long tether would broaden the distribution for each of the atoms in the selective portion of the molecule. Thus, many Whelk molecules could be oriented away from the bulk, or adopt other conformations that diminish their selectivity. As well, the probability of Whelk-Whelk hydrogen bonding would increase for longer chains.

The Whelk distribution above the surface has also been examined for the water/methanol and supercritical solvents. We find that the water/methanol interface shows similar distributions for the Whelk-O1 atoms. In contrast, the *sc*-CO_2/methanol solvated Whelk-O1 interface shows more gauche defects in the tether and more split peaks for the dinitrobenzoyl group, indicating more motion for this moiety of the Whelk-O1 selector under these conditions.

3.3.2 The Solvation of Whelk-O1 in *n*-Hexane/2-Propanol

A snapshot of the *n*-hexane/2-propanol solvent interface with Whelk-O1 is shown in Fig. 3.4(a). Clearly, the alcohol prefers the interfacial region, and the *n*-hexane is found predominantly in the bulk. From the snapshot, molecules do not enter the end-capping region: There is insufficient room to accommodate them between end-caps or between Whelk tethers and end-caps. However, the selectors are surrounded by solvent with solvent appearing between neighbouring selectors, and beneath the selective portions of individual selectors. The distribution of solvent can be seen more readily from Fig. 3.4(b). In the end-capping region, 0-5 Å, the *g(z)* value is less than 0.2 for both solvents showing that they are kept away from the underlying silicon by the end-caps. At large separations, far away from the surface, *g(z)* for *n*-hexane exceeds unity while the corresponding distribution for 2-propanol is less than one. These values reflect the important partitioning of the solvent mixture at the interface, and consequently also in the bulk. In terms of chiral separations, the polarity of the Whelk-O1 phase is increased by the concentration of alcohol at the interface.

Consider the distribution of solvent within 5-9 Å above the underlying silicon layer, between the end-caps and the selective portion of the Whelk-O1. From Fig. 3.4(b), the atoms for *n*-hexane have a peak at around 7.5 Å, at separations between the end-caps and the selective portions of the Whelk-O1. This suggests a confinement of *n*-hexane molecules by the bulky Whelk-O1 selector. The distributions for individual carbons in the six carbon chain have similar peak heights and positions in this region, indicating that the *n*-hexane molecules prefer to lie parallel to the surface. For the distribution of 2-propanol atoms above the surface, the *g(z)* show strong peaks around 8.5 Å. Thus, 2-propanol also has a high probability of being found beneath the selective portion of the Whelk-O1. The concentrated layer of *n*-hexane and alcohol between the end-caps and the selective portions of the Whelk is also evident in the snapshot [Fig. 3.4(a)].

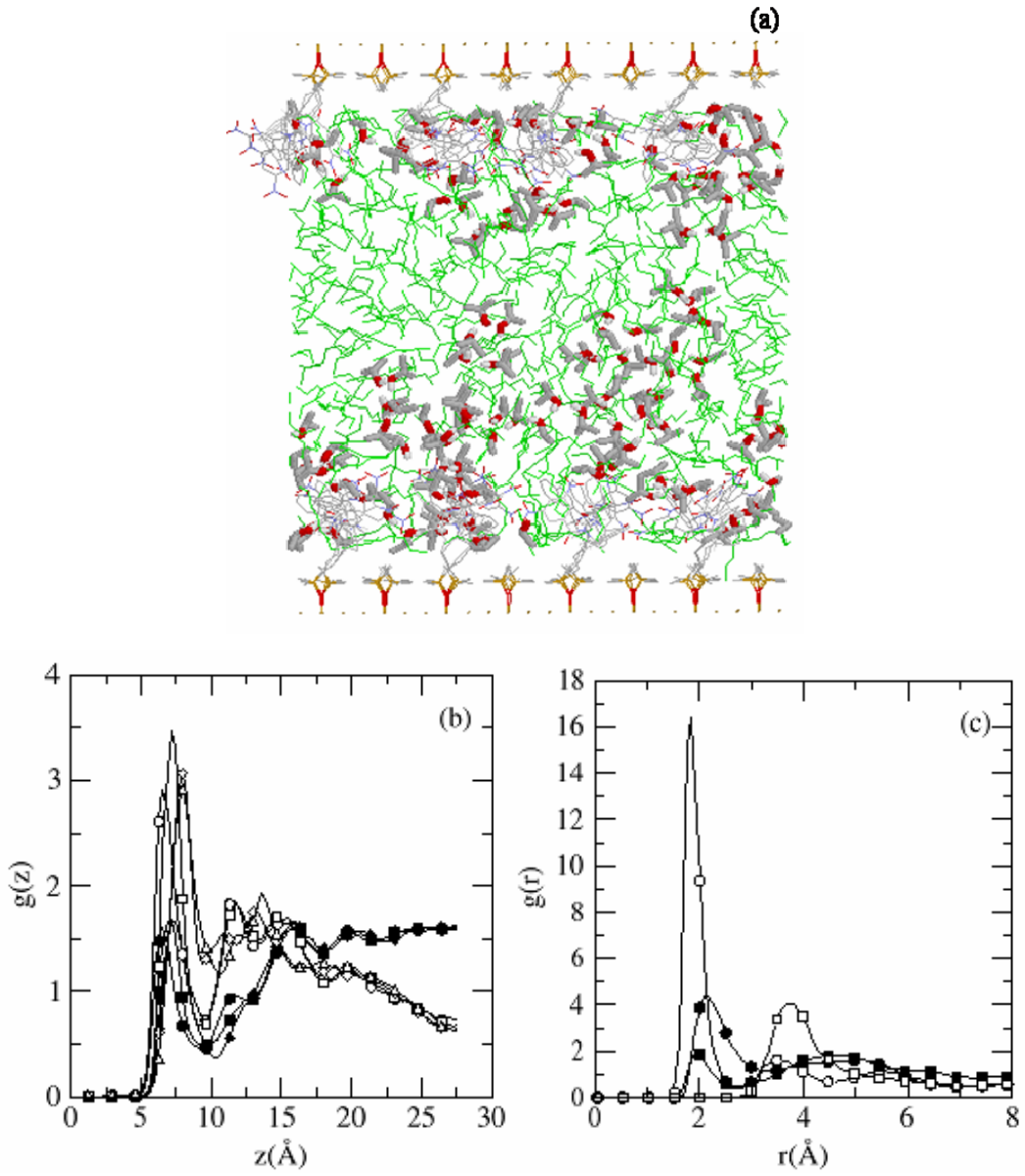


Figure 3.4: Simulation results from the *n*-hexane/2-propanol solvated Whelk-O1 interface. (a) A snapshot of the simulation cell. Silicon, oxygen, nitrogen, hydrogen atoms are shown in yellow, red, blue, and white respectively. Carbon atoms are shown in grey with the exception of *n*-hexane where green is used. 2-propanol molecules are highlighted. (b) Distribution of 2-propanol (open symbols) and *n*-hexane (filled symbols) atoms above the surface: CH_3, CH, O and H of 2-propanol are identified by circles, squares, diamonds, and triangles respectively; CH_3, 2-CH_2, and 3-CH_2 of *n*-hexane are represented by circles, squares, and diamonds respectively. (c) Radial distributions between H(26)-O(2-propanol) (filled circles), O(28)-H(2-propanol) (open circles), O(40)-H(2-propanol) (filled squares) and N(25)-H(2-propanol) (open squares).

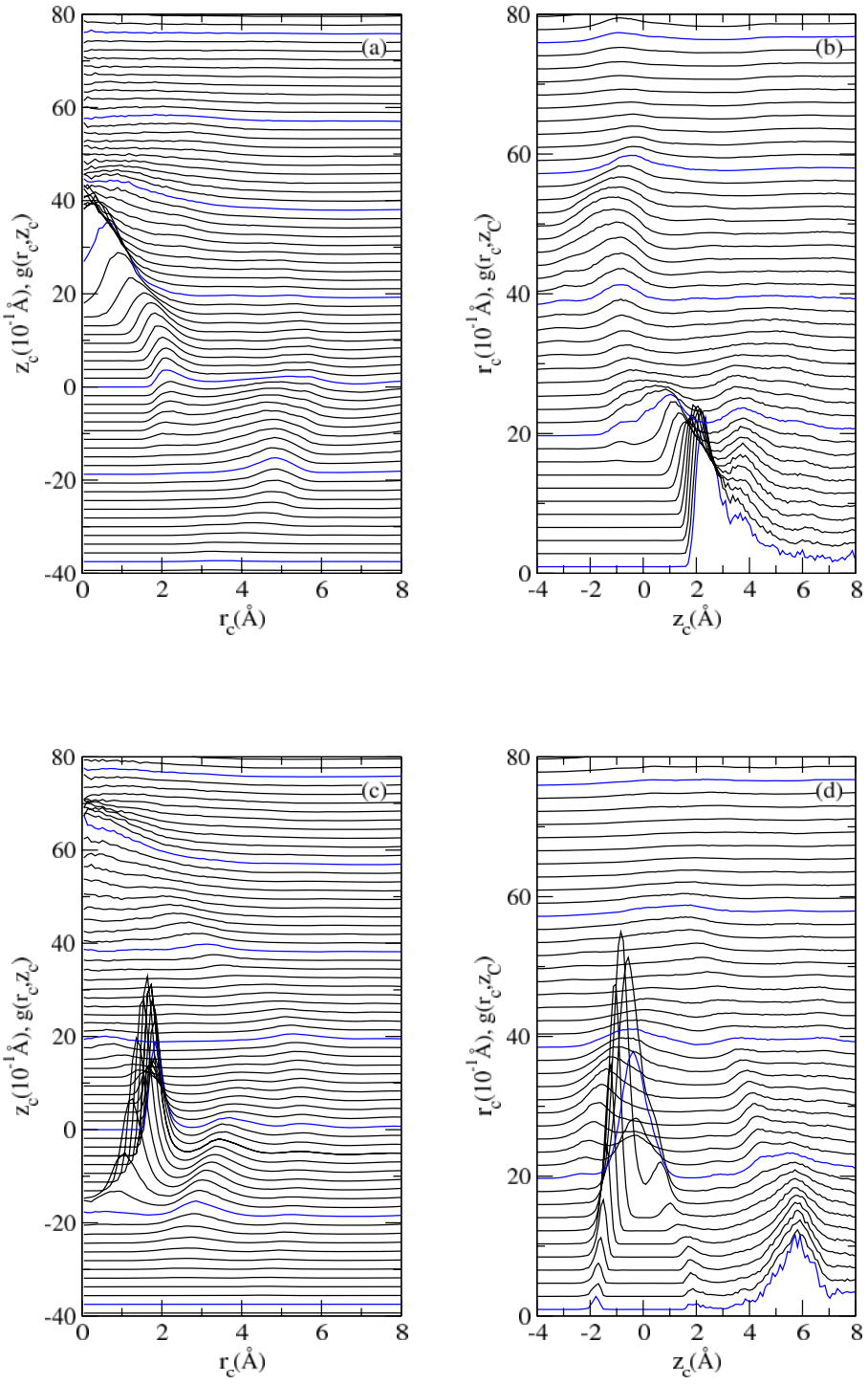
Between 9 and 15Å above the underlying silicon layer, the Whelk-O1 selector takes considerable volume and thus there are fewer solvent molecules found within this region. For the distribution of 2-propanol, however, the depletion is far less severe [see Fig. 3.4(b)]. This indicates that interactions between the alcohol and Whelk-O1 are strong enough to keep the alcohol in this crowded region. As we will discuss later, it is the hydrogen bonding of 2-propanol oxygen and hydrogen with the H(26) and O(28) of Whelk-O1, respectively, that keeps 2-propanol in this region. This can be deduced from Fig. 3.4(b) where most of the 2-propanol molecules close to the surface orient with the CH_3 and CH pointing towards the end-caps while the O and H are directed toward the polar Whelk region.

A concentration of alcohol at the interface may bring analyte into closer proximity with the surface. On the other hand, if the alcohol-selector interactions are too strong, the solvent can directly impede selectivity by sterically obstructing the formation of a selector-analyte pair or by forming energetically competitive hydrogen bonds. Consider first the hydrogen bonds between 2-propanol and the selector, identified by analysis of snapshots. We find that, on average, 15% of H(26) and 36% of O(28) are hydrogen bonded. The hydrogen bonding ratio for H(26) is not very high, indicating that most of the Whelk-O1 molecules have the amide hydrogen available for hydrogen bonding with analytes. On the other hand, the carbonyl oxygen is not usually invoked in the three-point binding model yet our results show that this atom forms a significant number of hydrogen bonds with solvent.

The presence of hydrogen bonds between 2-propanol and Whelk-O1 can also be examined from radial distribution functions between the relevant atomic pairs. In Fig. 3.4(c), distributions between potential hydrogen bonding pairs are shown. Consider donor-acceptor distances less than 2.6 Å. The distribution between O(28) and H(2-propanol) shows a very high enhancement in probability with a peak of about 16 at roughly 1.9 Å. Thus, a 2-propanol hydrogen is 16 times more likely to be found at this separation than expected based on the bulk density. The

distribution also shows a secondary peak at around 3.6 Å and a weak third peak 5.5 Å, and these are due to hydrogen bonding between 2-propanol molecules. We will analyze this in detail below. The distribution between H(26) and O(2-propanol) shows a peak at 2.2 Å. The hydrogen bonding distance for this pair is larger than for O(28)-H(2-propanol) and the peak height is significantly lower. Both of these factors suggest that the amide hydrogen forms weaker hydrogen bonds with 2-propanol than the carbonyl oxygen can form. From Fig. 3.4(c), there are also signs of hydrogen bonding between the nitro oxygen, as represented by O(40), and H(2-propanol), but it is much weaker. As for N(25), the radial distribution function indicates that there is insignificant hydrogen bonding to this atom. This is probably due, in part, to steric hindrance around the N(25), which renders this atom largely inaccessible to the solvent.

The arrangement of the alcohol about the Whelk-O1 gives some indication of the ease by which the solvent can be displaced by analyte. 2D distributions show the relative orientations, and $g(r_c,z_c)$ for the H(26)-O(2-propanol), O(28)-H(2-propanol), and O(40)-H(2-propanol) are shown in Fig. 3.5(a,b), Fig. 3.5(c,d), and Fig. 3.5(e,f), respectively. For each distribution, two views are shown. In one view [Figs. 3.5(a), (c), and (e)], $g(r_c,z_c)$ is shown as a function of r_c and the y-axis shows the magnitude of $g(r_c,z_c)$. However, the series of $g(r_c,z_c)$ with different z_c are shifted and presented in a single figure so that the vertical axis also corresponds roughly to the z_c value. In the other view [Figs. 3.5(b), (d), and (f)], the $g(r_c,z_c)$ are shifted according to r_c, and plotted as a function of z_c.



(Figure 3.5. To be continued on next page)

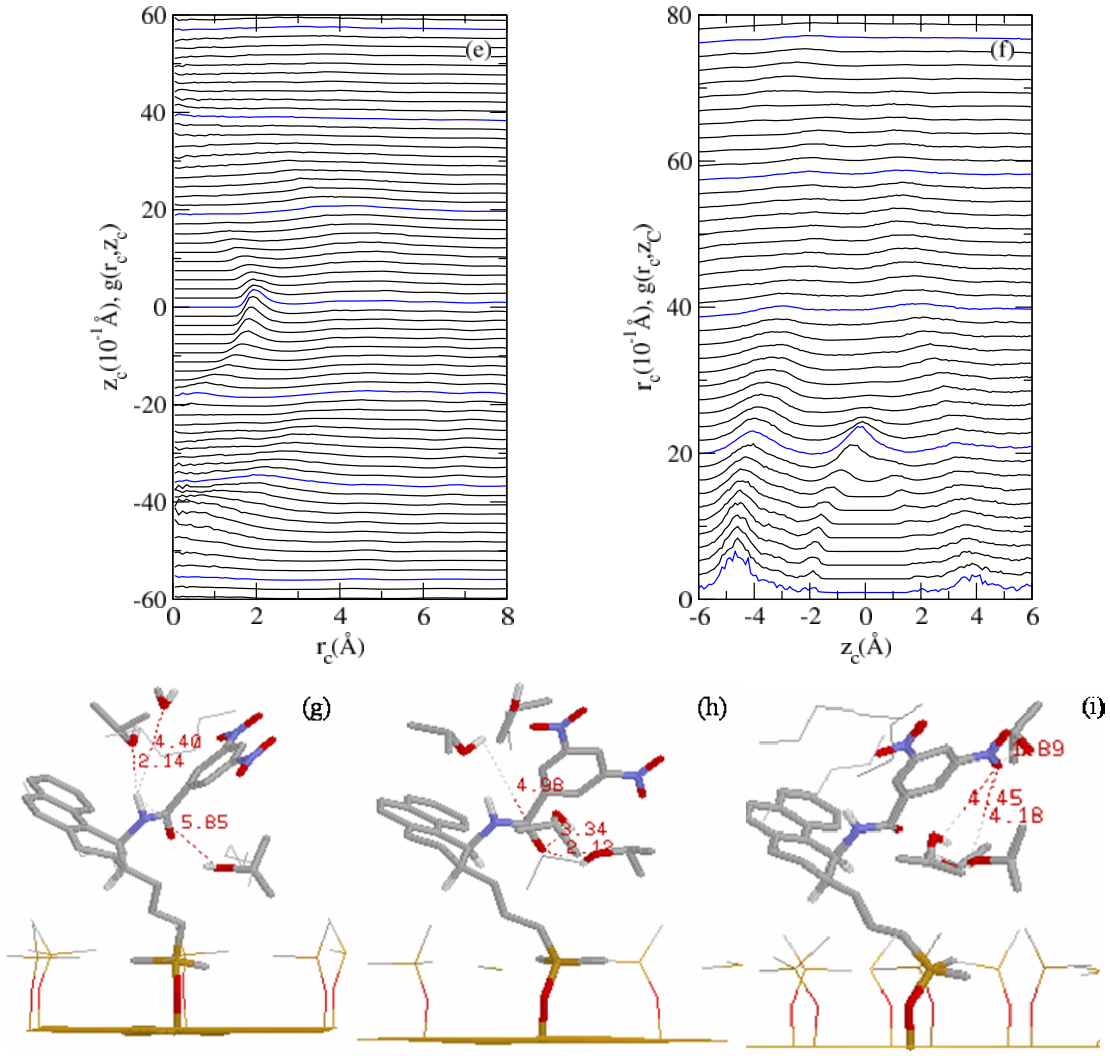


Figure 3.5 (Continued): 2D distribution functions and representative snapshots from the *n*-hexane/2-propanol solvated Whelk-O1 interface. Two views of $g(r_c,z_c)$ for H(26)-O(2-propanol) are shown in (a) and (b). Two views of $g(r_c,z_c)$ for O(28)-H(2-propanol) are shown in (c) and (d). Two views of $g(r_c,z_c)$ for O(40)-H(2-propanol) are shown in (e) and (f). Snapshots showing the H-bonding around H(26), O(28) and O(40) are shown in (g), (h) and (i) respectively with distances shown in Å. Blue lines in (a), (c), and (e) identify z_c=-5.625Å, -3.75Å, -1.875Å, 0.0Å, 1.875Å, 3.75Å, 5.625Å, and 7.50Å. Blue lines in (b), (d), and (f) identify r_c=0.094Å, 1.97Å, 3.84Å, 5.72Å and 7.59Å.

The 2D distribution of O(2-propanol) about H(26) [Figs 3.5(a) and 3.5(b)] shows a series of peaks at a distance about 2.2Å, as expected from the radial distribution in Fig. 3.4(c). However, the peaks in the region around r_c=0Å, z_c=2.2Å are much more pronounced than the peaks in the region around r_c=2.2Å, z_c=0Å, indicating that the 2-propanol oxygen prefers to approach the

H(26) from directly above. The peaks at a distance of 4.4Å (r_c=0-2Å, z_c=3.6-4.0Å) are mainly from the oxygen of 2-propanol hydrogen bonding to the H (2-propanol). That is, a second shell of 2-propanol hydrogen bonding from above. A broad series of peaks in the region around r_c=5.0Å, z_c=-1.0Å, is due to hydrogen bonding to the carbonyl oxygen. Specifically, the hydrogen bond to O(28) is probable enough that the oxygen from this 2-propanol appears in the distribution of oxygen about H(26). The snapshot in Fig. 3.5(g) shows a typical solvation structure around H(26). Three hydrogen bonding alcohols are shown along with their distance to H(26). Notice that the 2-propanol molecule that hydrogen bonds to H(26) is directly above the hydrogen, the second shell hydrogen bonding 2-propanol is also above, but shifted towards the dinitrobenzoyl ring, and the molecule hydrogen bonding to O(28) is 5.8 Å away and lies closer to the surface.

The 2D distribution of the 2-propanol hydrogen about O(28) is shown in Figs. 3.5(c) and 3.5(d). Three series of peaks, at distances of roughly 1.9 Å, 3.8 Å and 6 Å are evident. These originate from a direct hydrogen bond to O(28), a second shell hydrogen bond between 2-propanol molecules, and direct hydrogen bonding to H(26), respectively. The snapshot in Fig. 3.5(h) shows a typical hydrogen bonding arrangement about O(28). In contrast with the structure around H(26), the hydrogen bonding of 2-propanol to O(28) is mainly from the side and slightly below, as emphasized from the snapshot and the pronounced peaks in the region around r_c=1.9Å, z_c=0.0Å.

The radial distributions in Fig. 3.4(c) suggest some hydrogen bonding to the oxygen of the nitro group. The orientation of the alcohol involved in this hydrogen bonding is particularly relevant since the dinitrobenzoyl group is predicted, by the three-point binding model, to form a π-π interaction. If the hydrogen-bonding solvent lies above the oxygen, then it may hinder the relative orientation of the rings. The 2D distribution of H(2-propanol) about O(40) is shown in Figs. 3.5(e) and 3.5(f). Clearly, the hydrogen distribution is weaker than for the two other H-bonding scenarios considered in Figs. 3.5(a)-(d). The direct hydrogen bonding peak is found at a distance of about 2.0Å, with the highest probability at z_c=0 Å, r_c=2.2 Å. The series of peaks at

around 4.5Å (r_c=0.0 to 4.0Å, z_c=-4.6 to -3.0Å) is quite strong relative to the direct hydrogen bonding peak. As shown in Fig. 3.5(i), the distance between O(40) and the hydrogens from the alcohols that hydrogen bond to O(28) are 4.45Å and 4.18Å, so that the solvent structure about O(28) is the source of these peaks. We conclude that the single 2-propanol that H-bonds to O(40) does not lie above the ring.

The hydrogen bonding between Whelk-O1 and 2-propanol is complex. However, the most prevalent interaction, the H-bond to O(28), occurs below the amide linkage and is unlikely to interfere with the approach of analyte. As well, the less probable H-bond to the nitro oxygen is oriented away from the ring and is not likely to interfere with π-π stacking. Hydrogen bonding to the amide H does occur from above and would be broken by a selector-analyte interaction. However, a relatively small fraction of Whelk molecules have an H(26)-O(2-propanol) hydrogen bond. Overall, the alcohol in the *n*-hexane/2-propanol solvent will encourage polar analytes to aggregate at the interface without undue interference in analyte-selector hydrogen bonding or π-π stacking.

3.3.3 The Solvation of Whelk-O1 in Supercritical CO_2

Chiral separation using supercritical fluid chromatography (SFC) has the advantage of being highly efficient, fast with a high throughput, and environmentally friendly [90]. CO_2 is the supercritical fluid of choice in SFC [219], but polar modifiers, such as methanol or acetonitrile, are commonly used to increase the solubility [220]. We examine the Whelk-O1 interface solvated by sc-CO_2/methanol mixture. A snapshot of this interface is shown in Fig. 3.6(a). The methanol molecules tend to prefer the interfacial region over the bulk. Unlike the *n*-hexane/2-propanol interface where the end-cap region is inaccessible to solvent molecules, some of the small, linear CO_2 molecules enter the end-cap region. However, there is virtually no methanol penetration into the region.

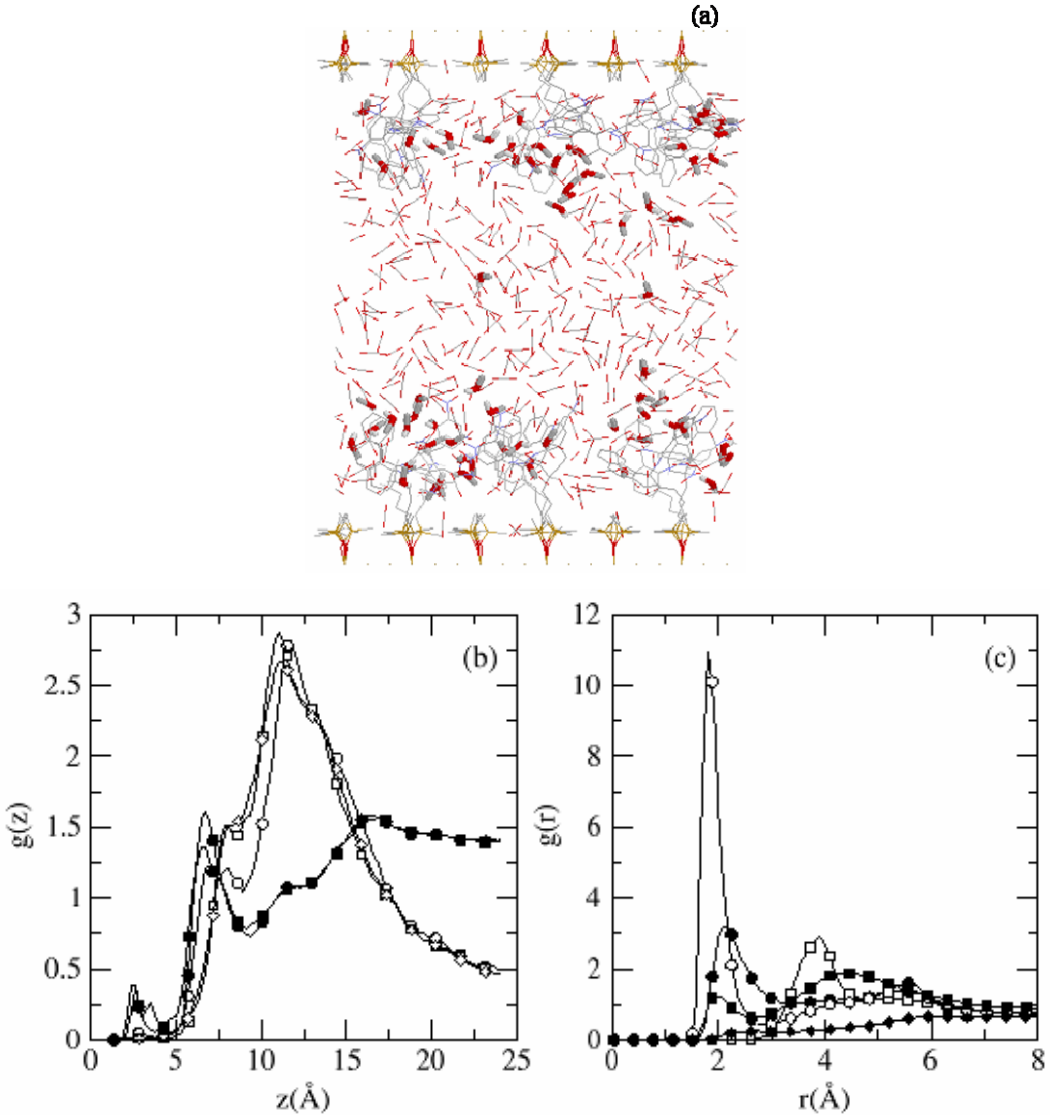


Figure 3.6: Simulation results from the sc-CO_2/methanol solvated Whelk-O1 interface. (a) A snapshot of the simulation cell. Silicon, oxygen, nitrogen, carbon, and hydrogen atoms are shown in yellow, red, blue, grey, and white, respectively. Methanol molecules are highlighted. (b) Distribution of methanol (open symbols) and sc-CO_2 (filled symbols) above the surface. CH_3, O, and H of methanol are identified with circles, squares, and diamonds respectively. C and O of sc-CO_2 are represented with circles and squares respectively. (c) Radial distributions between H(26)-O(methanol), O(28)-H(methanol), O(40)-H(methanol), N(25)-H(methanol), and H(26)-O(CO_2) are represented with filled circles, open circles, filled squares, open squares, and filled diamonds, respectively.

The distributions of methanol and CO_2 from the surface are shown in Fig. 3.6(b). The small peaks at z<5.0Å in the distribution of CO_2 show the limited extent to which these molecules penetrate the end-cap region. Otherwise, the distribution of CO_2 molecules is similar to the *n*-hexane distribution [Fig. 3.4(b)]: a concentrated layer beneath the selective portion of the Whelk-O1, a zone of depletion around 10 Å, and an overall preference for the bulk. In Fig. 3.6(b), the distribution of methanol atoms shows that, in the region close to the end-caps, the methanol orients itself by pointing the CH_3 towards the end-cap region and the OH towards the Whelk-O1. However, in contrast to 2-propanol, the peaks at about z=8.0Å for methanol are much smaller than those for the 2-propanol in the *n*-hexane/2-propanol solvated Whelk-O1 interface. This indicates that there are fewer methanol molecules confined between the end-caps and the selective portions of Whelk-O1. However, methanol has a high probability of being found within and slightly above the amide linkage, as shown by the large broad peak between z=8Å to z=15Å in Fig. 3.6(b).

Hydrogen bond statistics obtained by analysis of snapshots show that, on average, 19% of H(26) and 32% of O(28) are hydrogen bonded in this solvent. Thus, a somewhat larger number of amide hydrogens form hydrogen bonds with the supercritical solvent. Consider the radial distributions between potential selector-solvent hydrogen bonding pairs shown in Fig. 3.6(c). The methanol hydrogen has a high probability of being found within 2.6 Å of O(28), emphasizing that hydrogen bonding to the carbonyl oxygen of the selector is important in this solvent. However, unlike the O(28)-H(2-propanol) distribution for the *n*-hexane/2-propanol solvated interface, the radial distribution does not have well-defined second or third peaks, indicating that second shell hydrogen bonding is less pronounced in this solvent. The distribution of H(methanol) about O(40) is very similar to that of the 2-propanol hydrogen [Fig. 3.6(c)] about this oxygen: a weak peak at 1.8Å and a stronger broad peak from 3.5Å to 6.0Å. This indicates that direct hydrogen bonding to nitro oxygens occurs in the supercritical solvent, but only to a minor extent. For other potential

hydrogen bonding pairs, such as H(26)-O(sc-CO_2) and N(25)-H(methanol), we find that hydrogen bonding events are extremely rare with $g(r)$<0.2 when r < 2.6Å.

Two views of the 2D distribution of the methanol oxygen about H(26) are shown in Figs. 3.7(a) and 3.7(b). The hydrogen bonding interactions, at a distance at about 2.2Å, are somewhat more pronounced around r_c=1.0Å, z_c=1.8Å. This indicates that the methanol oxygen approaches the H(26) from the top, but will also approach from the side, to some extent. Unlike the *n*-hexane/2-propanol solvated Whelk-O1 interface, a second shell of methanol, expected at distance of roughly 4Å, is not formed in this solvent. This is due, in part, to the low methanol concentration. The series of peaks corresponding to a methanol hydrogen bond to O(28) is evident, around the region of r_c=4.5Å to 6.0Å, z_c=-2.0Å to 3.0Å. The peaks in the positive z_c region show that, unlike the *n*-hexane/2-propanol system, the methanol can also hydrogen bond to O(28) from above, which is shown more directly in Figs. 3.7(c) and 3.7(d). From these two figures, the hydrogen bonding peak series at roughly 1.9Å shows that the H(methanol) hydrogen bonds to O(28) mainly from the side since the most pronounced peaks occur around r_c=1.9Å, z_c=0.0Å, partially from the top and seldom from below (slight peaks around r_c=1.0Å, z_c=-1.5Å). Similar to the case of H(26), a second shell of methanol, expected at a distance of about 4.0Å, is not evident.

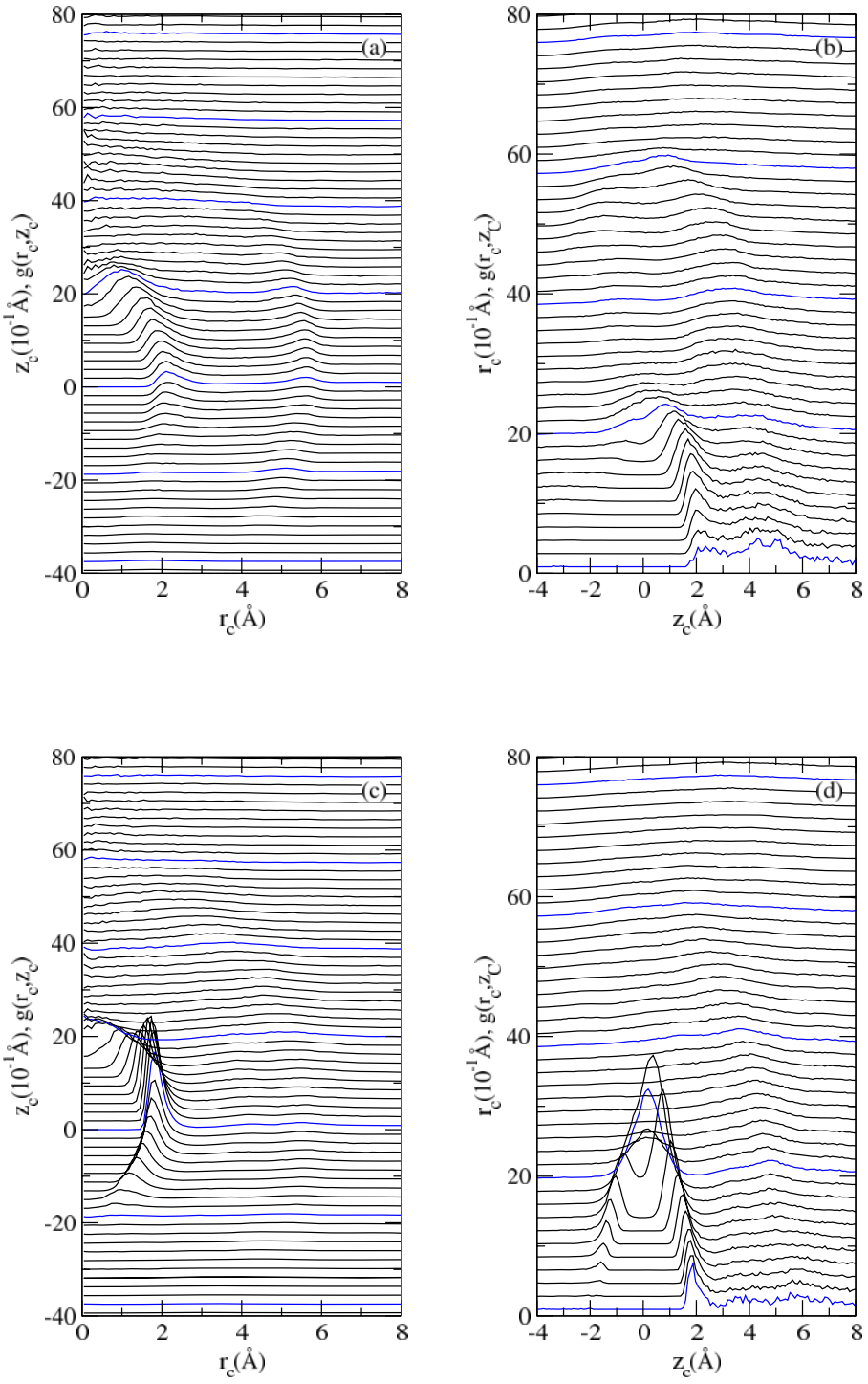


Figure 3.7: 2D distribution functions from the sc-CO_2/methanol solvated Whelk-O1 interface. Two views of $g(r_c,z_c)$ for H(26)-O(methanol) are shown in (a) and (b). Two views of $g(r_c,z_c)$ for O(28)-H(methanol) are shown in (c) and (d). Blue lines in (a) and (c) identify z_c=-3.75Å, -1.875Å, 0.0Å, 1.875Å, 3.75Å, 5.625Å, and 7.50Å. In (b) and (d), the blue lines identify r_c=0.094Å, 1.97Å, 3.84Å, 5.72Å and 7.59Å.

Hydrogen bonding between methanol and Whelk-O1 in a supercritical CO_2 solvent is less well-defined than H-bonding in *n*-hexane/2-propanol. Although methanol strongly prefers the interfacial region and forms more hydrogen bonds to the amide H than does 2-propanol, the methanol can approach from above or below. This leads to a less well-defined solvent structure about H(26) and O(28) and the 2D distributions do not show any second solvation shells.

3.3.4 The Solvation of Whelk-O1 in Water/Methanol

Relative to normal-phase and supercritical fluid HPLC, reversed-phase HPLC is seldom used for the Whelk-O1 CSP [89]. In this section, we explore the solvation of the Whelk-O1 interface in a polar water/methanol solvent. The solvent is predominantly water, with 4.15 water molecules for every methanol. Two snapshots, one showing the methanol and the other showing water, are given in Figs. 3.8(a) and (b). At first glance, one might conclude that the strong solvent partitioning observed for *n*-hexane/2-propanol and sc-CO_2/methanol is not present in the polar solvent. A closer inspection of the interface reveals that this is only partially correct: water does not descend far beneath the Whelk-O1 amide linkage, while Fig. 3.8(a) shows several methanol molecules located within the end-cap region. Fig. 3.8(c) provides a clearer depiction of the interfacial region: methanol has a strong tendency to stay between the end-cap region and the selective portion of the Whelk-O1 molecules, corresponding to the strong peaks in *g(z)* around 7.5 Å, whereas water is more likely to be found in the bulk (z>15 Å). Similar to the normal-phase and supercritical fluid solvents, methanol orients such that the CH_3 points towards the end-caps and OH points toward the selector.

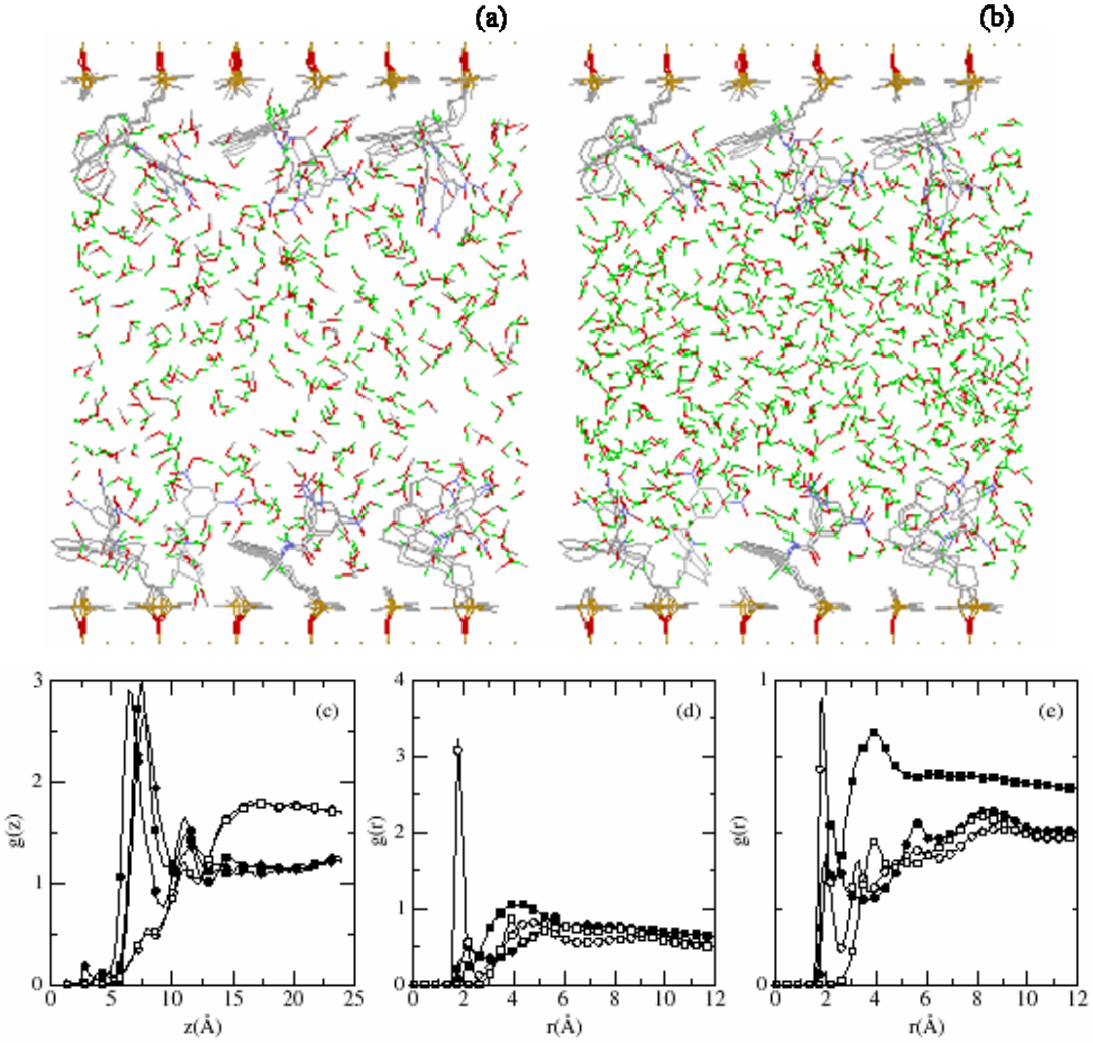


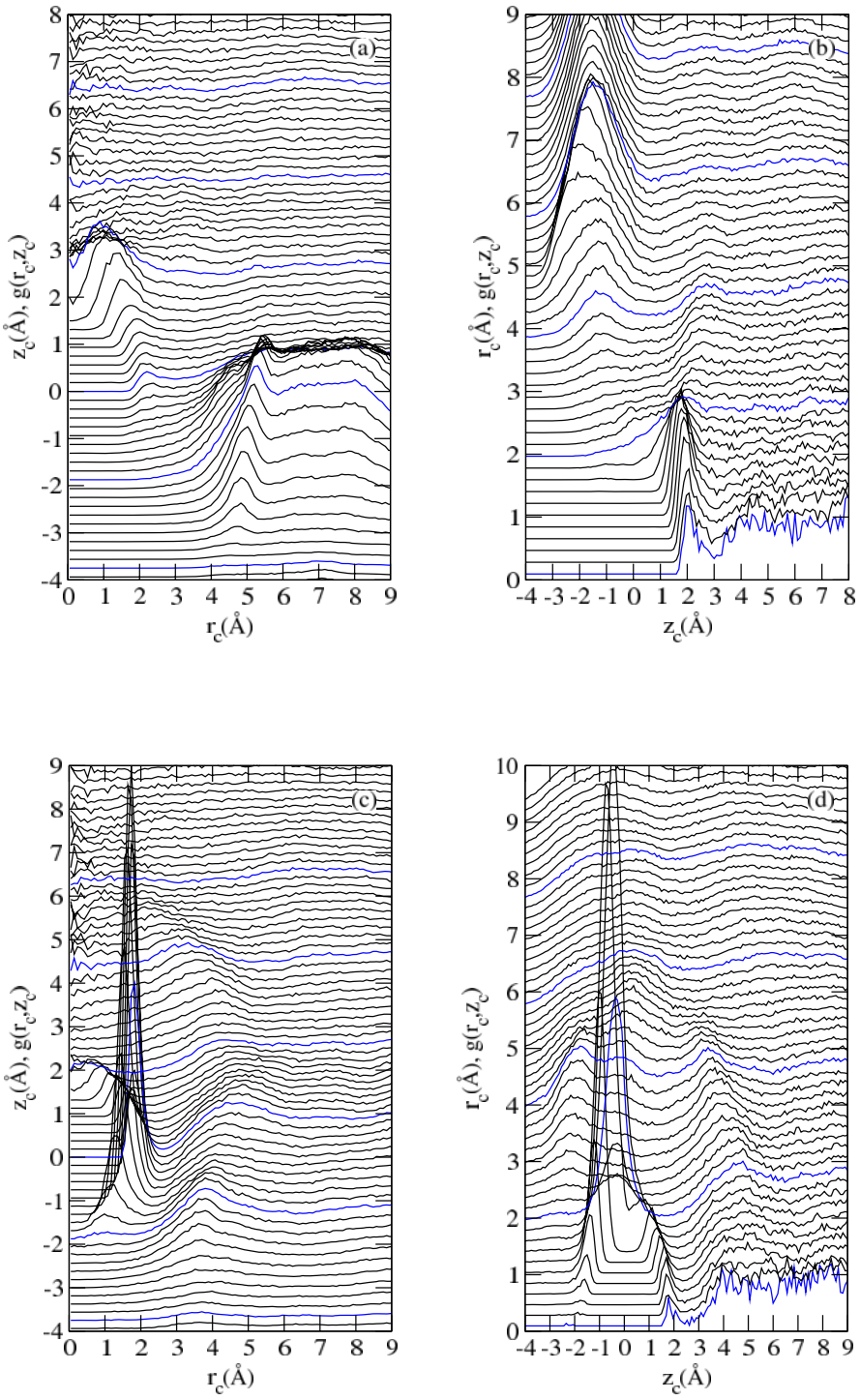
Figure 3.8: Simulation results from the water/methanol solvated Whelk-O1 interface. (a) A snapshot showing the location of methanol. Silicon, oxygen, nitrogen, carbon, and hydrogen atoms are shown in yellow, red, blue, grey and green respectively. (b) The equivalent snapshot identifying the position of water. (c) The distribution of methanol (filled symbols) and water (open symbols) above the surface. CH_3, O, and H of methanol are identified with circles, squares, and diamonds respectively. H and O of water are represented with circles and squares respectively. (d) Radial distributions between H(26)-O(methanol), O(28)-H(methanol), O(40)-H(methanol), and N(25)-H(methanol) as identified by filled circles, open circles, filled squares, and open squares, respectively. (e) Radial distributions between H(26)-O(water), O(28)-H(water), O(40)-H(water) and N(25)-H(water) are represented by filled circles, open circles, filled squares, and open squares, respectively.

The concentration of the less polar solvent, methanol, in the interfacial region is due, in part, to the distinct hydrogen-bonding preferences of water and methanol. To be precise, bulk water forms a particularly dense network of hydrogen bonds and evidently prefers to hydrogen bond with itself rather than with methanol or the Whelk-O1 selectors. Several studies of bulk water/alcohol mixtures show partial segregation, or percolation, despite the complete miscibility of the two components [221,222]. A recent study of water/alcohol mixtures near a surface of N-(1-phenylethyl)-N'-[3-(triethoxysilyl)propyl]-urea [186] showed that the alcohol excludes water, to some extent, from the interface. Thus, the partial exclusion of water at the Whelk-O1 interface is consistent with previous work [186].

Application of the geometric definition of a hydrogen bond to snapshots shows that 31.7% of H(26) and 62.2% of O(28) form hydrogen bonds with water. In contrast, 6.5% of H(26) and 12.3% of O(28) hydrogen bond to methanol. Consequently, water forms roughly 5 times more hydrogen bonds than methanol, a ratio that exceeds their bulk density ratio. As discussed above, methanol tends to concentrate at the interface and partially exclude water. It is, however, important to realize that the distributions in Fig. 3.8(c) are relative to the bulk density, so that the amide linkage, found roughly 8-10 Å above the surface, is surrounded by more water molecules than methanol despite the fact that *g(z)* for methanol is larger than for water at these separations [Fig. 3.8(c)]. The relative hydrogen bonding ability of the solvents is another factor influencing the hydrogen bond statistics. In particular, water can accommodate more hydrogen bonds than methanol. Thus, the hydrogen bond statistics are not in contradiction with the molecular distribution at the interface: they simply reflect the number of each molecule in the vicinity of the selector and also the overall hydrogen bonding ability of the solvents. However, they do suggest that the nature of the water/methanol interface with Whelk-O1 may be highly sensitive to the methanol concentration. At the present composition, only about 60% of the H(26) are free to bind with analytes.

The radial distribution functions between the selector and methanol atom pairs are shown in Fig. 3.8(d). The radial distribution for O(28)-H(methanol) has a sharp peak at r=1.8 Å, and a broad secondary peak from r=3.8Å to r=6.0Å. The latter includes contributions from a second solvation shell with methanol or with water. The intensity of the first peak is roughly 3 times smaller than those of the normal-phase and supercritical fluid system, due in part to the competition for hydrogen bonds by the water. In this regard, Fig. 3.8(e) shows a sharp peak in the O(28)-H(water) distribution at the same separation, r=1.8 Å. The amide hydrogen, H(26), also forms hydrogen bonds with both solvents. From Fig. 3.8(d), methanol has a strong preference for the amide oxygen, O(28), as was observed for other alcohols. However, water is far less selective, as illustrated by the relative peak heights in Fig. 3.8(e): the H(26)-O(water) peak height is only half the O(28)-H(water) peak. Both solvents also show a small propensity for hydrogen bond donation to the nitro group oxygens.

The 2D distribution for H(26)-O(methanol) is given in Figs. 3.9(a) and (b). Methanol hydrogen bonds to H(26) mainly from the top (r_c=0Å to 1.5Å, z_c=1.8Å to 2.2Å), but in a less selective way than 2-propanol in the normal-phase solvent. A second solvation shell of methanol is not evident in the 2D distributions. The series of peaks at roughly 5.4Å (r_c=4.8Å to 5.4Å, z_c=-3.0Å to 0.0Å) corresponds to the methanol molecules that hydrogen bond to O(28). The underlying series of broad peaks between 5Å and 9Å (r_c=5.5Å to 9.0Å, z_c=-3.0Å to 0.0Å) is induced by methanol molecules localized between the end-caps and the selective portions of the selectors. The 2D distribution of O(28)-H(methanol), shown in Figs. 3.9(c) and (d), has a very pronounced peak corresponding to a hydrogen bond beside and slightly below the amide oxygen. Between the two most possible hydrogen bonding sites in the Whelk-O1 amide linkage, H(26) and O(28), methanol clearly prefers to hydrogen bond to O(28). In Figs. 3.9(c) and (d), the complicated series of peaks at around 4.5Å, may mainly come from two sources: second shell hydrogen bonding to methanol and second shell hydrogen bonding to water.



(Figure 3.9. To be continued on next page)

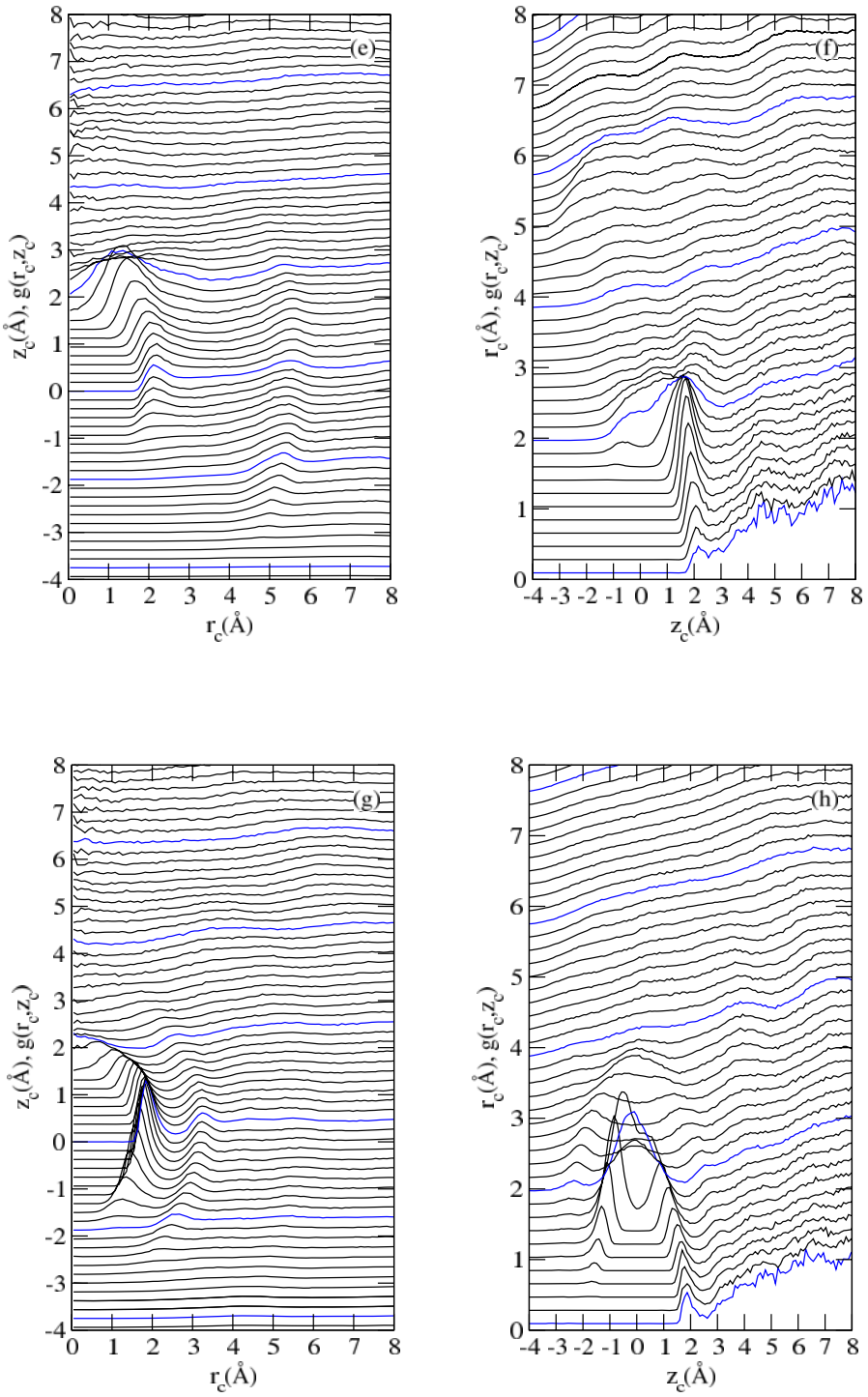


Figure 3.9 (Continued): 2D distribution functions from the water/methanol solvated Whelk-O1 interface. Two views of $g(r_c,z_c)$ for H(26)-O(methanol) are shown in (a) and (b). Two views of $g(r_c,z_c)$ for O(28)-H(methanol) are shown in (c) and (d). Two views of $g(r_c,z_c)$ for H(26)-O(water) are shown in (e) and (f). Two views of $g(r_c,z_c)$ for O(28)-H(water) are shown in (g) and (h). Blue lines in (a), (c), (e), and (g) identify z_c=-3.75Å, -1.875Å, 0.0Å, 1.875Å, 3.75Å, 5.625Å, and 7.50Å. Blue lines in (b), (d), (f), and (h) identify r_c=0.094Å, 1.97Å, 3.84Å, 5.72Å and 7.59Å.

Interestingly, the 2D distribution for H(26)-O(water), shown in Fig. 3.9(e) and 3.9(f), is very similar to that of H(26)-O(methanol) in the sc-CO_2/methanol solvated Whelk-O1 interface. Specifically, direct hydrogen bonding peaks at 2.2 Å and peaks from hydrogen bonding to O(28) at 5.5Å are evident. The primary difference is that the values of $g(r_c,z_c)$ tend to increase with the increase of z_c [Fig. 3.9(f)], which is due to the increased water density in the bulk. This trend is also evident in the 2D distribution of O(28)-H(water), Figs. 3.9(g) and 3.9(h). For O(28)-H(water), the hydrogen bonding peak series is again very similar to the corresponding distribution in the sc-CO_2/methanol solvated Whelk-O1 interface: the water hydrogen is mainly beside O(28) with the most pronounced peaks occurring for r_c=1.9Å and z_c=0.0Å. But water also approaches O(28) from the top and below. In contrast to methanol, there is a well-defined series of peaks in Figs. 3.9(g) and 3.9(h) at around 3.3Å, indicating second shell hydrogen bonding for water.

3.4 Conclusions

In this chapter, we have examined the Whelk-O1 chiral stationary phase in three distinct solvent environments: a normal-phase *n*-hexane/2-propanol solvent; a reversed-phase water/methanol solvent; and a supercritical-phase of carbon dioxide and methanol. Molecular dynamics simulations have been performed to examine the mobility and solvation of the Whelk-O1 interface.

In terms of selector mobility, we find broad atomic distributions for the selectors, in particular for the dinitrobenzoyl moiety. However, snapshots of the surface indicate that most selectors adopt an "open" configuration with the phenanthryl moiety roughly perpendicular to the surface and the dinitrophenyl ring lying roughly parallel to the underlying silicon layer. The end result of this preference is that, at any point in time, many selectors have the potential to bind with analyte.

Our simulations indicate significant solvent partitioning at the interface for all three solvents. In the predominantly apolar *n*-hexane/2-propanol solvent and the supercritical CO_2/methanol solvent, the alcohol is found primarily near the surface. In the reversed-phase water/methanol solvent, the alcohol also prefers the interface. Thus, the three solvents exhibit a polarity difference between the bulk and the interfacial region. The former two solvents lead to an increased polarity at the interface whereas the latter shows a decrease. This polarity gradient may be controlled, and used to advantage, by varying the alcohol concentration.

Simulations of the Whelk-O1 interface reveal important distinctions between solvent-selector hydrogen bonding in the three solvent environments. In the normal-phase solvent, 2-propanol has a very well-defined hydrogen bonding structure with H-bonds to the amide hydrogen primarily from above and hydrogen bonds to the carbonyl oxygen from the side and below the amide linkage. The supercritical solvent is similar except that methanol adopts a wider range of positions about the amide linkage. In a polar water/methanol solvent, the hydrogen bonding structure about Whelk-O1 is complex, and the overall number of hydrogen bonds to Whelk-O1 is higher. This is significant since analyte will need to replace the solvent in forming hydrogen bonds with the Whelk-O1 selectors.

The carbonyl oxygen in the amide linkage participates in a much larger fraction of hydrogen bonds than the amide hydrogen. However, these hydrogen bonds tend to occur from the side or below the amide linkage and consequently should not sterically hinder the approach of analyte. The nitro oxygens also participate in a small number of hydrogen bonds. However, the hydrogen bond donor tends to position itself away from the dinitrobenzoyl ring, thereby leaving this moiety available for π-π stacking. Finally, the amide hydrogen, which appears in the proposed three-point binding model for Whelk-O1, forms a small but significant number of hydrogen bonds. These bonds tend to occur from above and must be broken in the interaction between Whelk-O1 and analyte. Our simulations show that the largest fraction of available H(26) occurs for *n*-hexane/2-propanol.

Chapter 4
Docking of chiral epoxides on the Whelk-O1 stationary phase

4.1 Introduction

The (3R,4S)-Whelk-O1 chiral stationary phase (CSP) has successfully separated many racemates [77] including chiral carboxylic acids, aryl sulfoxides, aryl epoxides, and aliphatic and heterocyclic amines. As a general rule, enantiomers that exhibit a hydrogen bond acceptor and an aromatic moiety close to the stereogenic center tend to be well-resolved on this CSP [77]. The selective molecule of Whelk-O1 is unique in that it adopts a semi-rigid cleft configuration: it consists of two rigid aromatic moieties connected by an amide bridge of limited flexibility.

As previously discussed, Whelk-O1 has been the subject of several theoretical studies in the past few years. In Chapter 3, we employed molecular dynamics simulations to examine the solvation of the (3R,4S)-Whelk-O1 chiral stationary phase. We found that the selective molecules tended to orient the phenanthryl group perpendicular to the surface while the dinitrophenyl groups were roughly parallel to the surface. The solvated interface was examined in three binary solvents: a normal phase *n*-hexane/2-propanol solvent; a reversed phase water/methanol solvent; and a supercritical CO_2/methanol solvent. We found that the alcohol preferred the interfacial region in all three solvents, and hydrogen bonding between the solvent and the amide bridge was favourable. This was particularly true for the carboxyl oxygen of the amide bridge but the interaction places the solvent outside of the cleft region of Whelk-O1 and should not interfere with selectivity. Interactions with the amide hydrogen, on the other hand, tended to place solvent in the cleft region.

In this chapter, we examine the selectivity of Whelk-O1 for two epoxides: (R/S)-styrene oxide, and (R,R/S,S)-stilbene oxide. We have selected these epoxides for several reasons: experimental data is available [77,223] and their separation factors are reasonable and distinct; they are relatively small with limited flexibility; and they are closely related to each other,

allowing an assessment of steric effects. For each racemic epoxide, we employ lengthy molecular dynamics simulations of model Whelk-O1 interfaces in the presence of a hexane solvent and a small amount of the racemate. The interactions between each enantiomer and the selective portions of Whelk-O1 are monitored to ascertain the major docking conformations for each. The mechanism of selectivity is assessed based on multiple distribution functions and an analysis of snapshots. The effect of a polar additive is considered for (R/S)-styrene oxide. In particular, we examine the selectivity in 80:20 *n*-hexane:2-propanol.

Despite the well-documented [77,223] effectiveness and versatility of the Whelk-O1 chiral stationary phase, and recent work at modelling this interface [96,119,224], little is directly known about the mechanism of selectivity. As noted in Chapter 1, according to the three-point binding model, a π-π stack is formed between the electron deficient dinitrophenyl group of Whelk-O1 and an aryl group in the analyte, the amide H of Whelk-O1 forms a hydrogen bond, and either CH-π interactions [96,119,209] or edge-to-face π-π interactions [77,97,225,226] are formed with the phenanthryl moiety. A co-crystallization experiment involving a Whelk-O1 derivative [55] also provides information of the selectivity: outside-of-cleft interactions occurred for the heterochiral crystals. Recent NMR studies [57] of a variety of analytes, differing in their hydrogen bonding abilities and the number and character of their rings, in the presence of a Whelk-O1 derivative is consistent with this mechanism. Undoubtedly, the details of the interaction between enantiomers and Whelk-O1 CSP are more varied and complex than a simple three-point model implies. In this chapter, the mechanism of selectivity is examined in detail, with particular emphasis on the three-point binding model. We also examine outside-of-cleft interactions and bridging interactions involving more than one selector, although these are expected to be rare.

This chapter begins with a brief description of the model Whelk-O1 interface, details of the 1-(3,5-dinitrobenzamido)-1,2,3,4-tetrahydrophenanthrene molecular model, and the molecular models for the epoxides, which have been derived from *ab initio* calculations. Details of the

molecular dynamics simulations are also briefly provided in Section 4.2. The distribution of the enantiomers at the interface, the nature and number of specific interactions, the applicability of the three-point model, and the impact of a polar co-solvent are discussed in Section 4.3. The chapter concludes with a brief discussion of the results in Section 4.4.

4.2 Theoretical Details

4.2.1 Models of the Whelk-O1 Interface

Suitable selector and interface models have been developed in previous chapters for solvation studies. However, details of the selector and interface models become more critical for a docking study. Here we made several small changes to these models.

A model of the Whelk-O1 CSP was derived in previous chapters. First, the Whelk-O1 surface coverage is 1.07 μmol/m^2 and trimethylsilyl end-caps are included at a density of 3.20 μmol/m^2, within the range of experimental end-cap coverage [189]. This model is suitable for solvation studies where the solvent molecule is small in size. In a docking study, however, larger analytes could approach the selector from the back face of the dinitrophenyl group, close to the surface [40]. Silanol groups are known [227] to remain on a "fully" end-capped surface and they may interfere with some docking modes of the analytes. In this chapter, we have added silanol groups to the model surface, at a surface density of 4.26 μmol/m^2. Fig. 4.1 presents a snapshot of the model simulation cell. Each surface consists of 16 Whelk-O1 selective molecules, 48 trimethylsilyl end caps, 64 silanol groups, and an underlying layer of 128 silicon atoms.

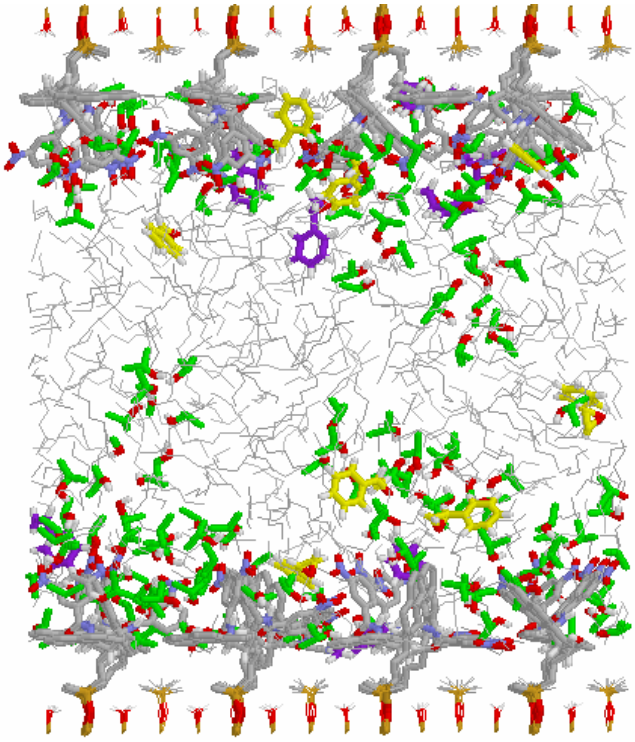


Figure 4.1: A side view of the simulation cell showing two interfaces with Whelk-O1 selectors, end-caps, and silanol groups evident. The solvent consists of *n*-hexane (thin grey lines) and 2-propanol (green), R-styrene oxide (yellow), and S-styrene oxide (purple).

In previous chapters, a suitable molecular model for the Whelk-O1 selector was developed based on an extensive series of density functional calculations. We have adopted the atomic charges, Lennard-Jones parameters, and the stretch and bending potentials from this earlier work, but we have modified the torsional potentials and added improper torsions. Consider first the torsional potentials. Literature potentials [143] were employed in Chapter 3 except for torsions about C(8)-C(9), C(10)-N(25), C(27)-C(29) where potentials were derived from *ab initio* calculations. Using the same set of *ab initio* calculations, these torsions were modified as follows. First, following others [228], we have corrected the torsional potentials for non-bonding interactions. That is, a Rickaert-Bellemans [145] potential (Eq. 2.25) is fitted to *corrected* torsional potentials, where the corrections consist of subtracting U^{LJ} and U^{Elec} for some intramolecular atomic pairs. Second, Eq. 2.25 differs slightly from the previous torsional

potential (Eq. 2.24) where all the phase shifts were fixed at 180^o and the potential included two additional sinusoidal terms to compensate. Eq. 2.25 is adopted here since it provides more flexibility than the previous potential and the latter can be represented in the form of Eq. 2.25. Improper torsions have been added for the amide linkage since we have noted too much flexibility in out-of-plane bending motions in this region. Such flexibility could influence analyte docking as analyte molecules are usually much larger than solvent molecules. To correct for this, we fit improper torsional potentials (Eq. 2.26) with the equilibrium out-of-plane angle and the corresponding force constants extracted from least squares fits to twenty B3LYP/6-31G* calculations with values within 5 degrees of the equilibrium value. The revised torsion and improper torsion parameters are presented in Tables 4.1 and 4.2, respectively.

The end-cap potentials were discussed in Chapter 2. A simple apolar potential with OPLS values [143] chosen for the Lennard-Jones parameters and CHARMM [198] for the bending potential about the oxygen has been applied for the silanol group.

Table 4.1: Torsional potentials for the Whelk-O1 selector. The potential parameters for Eq. 2.25 are presented. The first row identifies the atoms [see Figs. 2.6(b) for atom numbering]. The coefficients c_i are given in kJ/mol and the angles ϕ_i are in degrees.

Torsion potential parameters	10,9,8,7	27,25,10,9	30,29,27,25
c_0	13.66	20.59	12.91
c_1	-2.41	19.84	-1.18
c_2	-11.27	59.97	-24.87
c_3	14.02	-26.51	0.73
c_4	19.86	-35.92	6.09
c_5	-16.73	20.86	0.08
c_6	-5.37	-31.41	10.03
ϕ_1	191.95	58.99	154.31
ϕ_2	17.40	14.77	3.23
ϕ_3	9.28	40.90	114.28
ϕ_4	24.84	46.93	28.77
ϕ_5	46.22	43.75	0.00
ϕ_6	39.88	4.88	175.96

Table 4.2: Improper torsional potentials for the Whelk-O1 selector. The first column identifies the four atoms defining the improper torsional motion [see Figs. 2.6(b) for atom numbering]. The second and third columns define the equilibrium angle and the corresponding force constant (see Eq. 2.26).

Improper Torsion	ω_e (degree)	k_ω (kJ mol^{-1} rad^{-2})
29,34,30,27	1.05	231.85
27,25,28,29	-0.16	274.26
25,10,27,26	-8.54	42.01

4.2.2 Molecular Representation of Chiral Epoxides and Solvent

Flexible molecular models for styrene oxide and stilbene oxide are required to properly represent the conformational changes associated with the docking of these molecules. The epoxide models begin with full geometry optimizations at the B3LYP/6-311++G** level. The global energy minima are shown in Fig. 4.2 along with the atom numbering which will be used throughout.

For Lennard-Jones interatomic potentials, we employ the OPLS-AA parameter set [143] for the well depth and length parameters. Atomic charges for the epoxides have been assigned based on the CHELPG algorithm [137] applied to the global energy minimum.

Consider the intramolecular motion of the epoxides. The rings are treated as rigid units (RU) but it is important to note that all atoms, including hydrogens, appear explicitly in the rigid unit. The intramolecular motion for styrene oxide [Fig. 4.2(b)] consists of the following: a bond stretch between the rigid units; five bending potentials to orient the rings relative to the C(1)-C(7) bond; a torsional potential for the relative orientation of the rings; and an improper torsional potential for out-of-plane motion at C(7). The intramolecular potential for stilbene oxide consists of the same elements, except that a second set of potentials is required for the additional phenyl ring.

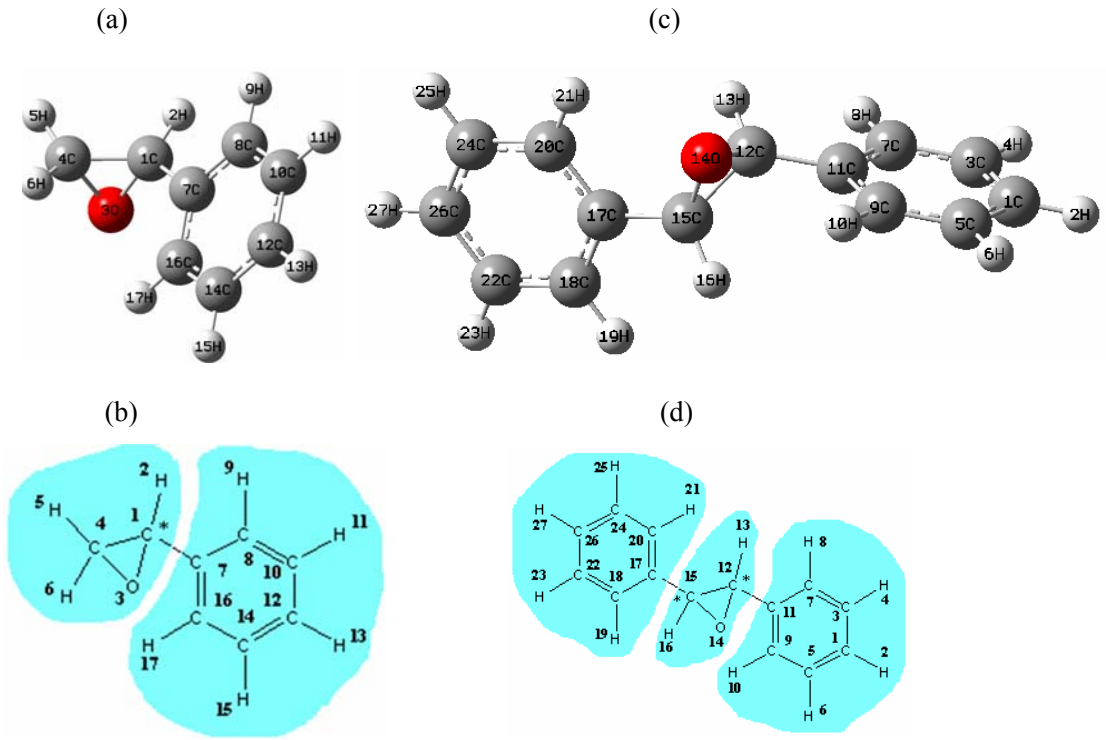


Figure 4.2: Global energy minima for (a) R-styrene oxide and (c) (S,S)-stilbene oxide. Panels (b) and (d) show the atomic numbering used throughout, and the identification of rigid units (shaded regions). The numbering for the mirror image molecules is identical for all atoms and is generated by exchanging the phenyl group with the hydrogen at the chiral carbon. The energy minima are obtained from B3LYP/6-311++G** full geometry optimizations.

The potentials for intramolecular motion are extracted from a series of B3LYP/6-31G* calculations. Specifically, for stretching motion, the force constant and equilibrium bond length are extracted directly from the geometry optimization. The styrene-oxide C(1)-C(7) bond which joins the two RU has a length of 1.491Å and a force constant of 1.160×10^{5}kJ mol^{-1} nm^{-2}. In stilbene oxide, the C(11)-C(12) and C(15)-C(17) bonds are equivalent with a bond length of 1.489Å and a force constant of 1.160×10^{5}kJ mol^{-1} nm^{-2}. For bending, the force constant is obtained from a least squares fit to 20 restricted geometry optimizations for angles within $\pm5^{o}$ of the equilibrium angle. The resulting parameters are given in Appendix A. For torsional potentials, 20 restricted geometry optimizations are performed over the full 360^{o} rotation. The potentials are corrected by subtracting the intramolecular Lennard-Jones and electrostatic energies, as discussed above. For convenience, the resulting potentials are shifted to give a minimum energy of zero and

Rickaert-Bellemans potentials (see Eq. 2.25) are fitted by least-squares. The resulting parameters are given in Appendix A. Improper torsions, of the form shown in Eq. 2.26, have been fitted to restrict out-of-plane motion for the epoxides. Specifically, restricted geometry optimizations have been performed for 20 angles within $\pm 5^{o}$ of the equilibrium angle. The parameters are also given in Appendix A.

The TraPPE-UA model [211] was chosen for *n*-hexane and 2-propanol. This model was employed previously (Chapter 2) for the solvation study of Whelk-O1 and is consistent with the bending and torsion potential forms we employ (Eqs. 2.22 and 2.25).

4.2.3 Molecular Dynamics Simulation Details

Extensive molecular dynamics simulations were performed for the Whelk-O1 interface. We examine the selectivity of Whelk-O1 for both styrene oxide and stilbene oxide in pure *n*-hexane. We also examine the selectivity of Whelk-O1 for styrene oxide in a 80:20 (volume) *n*-hexane:2-propanol. Pure *n*-hexane and *n*-hexane/2-propanol solvents have been chosen since experimental information is available for styrene oxide in a 99:1 hexane:2-propanol and for stilbene oxide in 90:10 *n*-hexane:2-propanol. Specifically, Pirkle and Welch [77] report retention factors for the initially eluted enantiomers of $k_1^{'} = 1.37$ and $k_1^{'} = 0.45$ for styrene and stilbene oxide, respectively, and the corresponding separation factors are 1.37 and 2.00. Thus, stilbene oxide is eluted more quickly and separated more effectively than styrene oxide on Whelk-O1. Schaus *et al.* have examined a series of epoxides [223], including styrene oxide. They report that the S enantiomer elutes first on (3S,4R)-Whelk-O1 in 99:1 hexanes/2-propanol. Our model selector is the (3R,4S) enantiomer and R-styrene oxide should be less retained on this CSP.

A side-view of the simulation cell is shown in Fig. 4.1. The figure does not show the empty space included above and below the surfaces. The underlying layer of Si atoms, to which end-caps, silanol groups, and selectors are attached, is stationary throughout the simulation. This layer of silicon is arranged perpendicular to the *z*-axis of the simulation cell and defines the boundary

between the interfacial system and the empty space beyond the slab. The electrostatic potential appearing in Eq. 2.29 is evaluated using Ewald summations [151] with a correction for the shape of the simulation cell [152]. The Ewald convergence parameter of α=0.1478 $Å^{-1}$ was tested and chosen for the simulations along with a reciprocal space cutoff of $k^{2*} \leq 27$.

For styrene oxide in the pure *n*-hexane solvent, the fluid consists of 470 hexane molecules, 8 R-styrene oxide molecules, and 8 S-styrene oxide molecules. For stilbene oxide, the fluid consists of 450 *n*-hexane molecules and 8 of each enantiomer. The number of *n*-hexane has been adjusted to yield roughly the same solvent density. In the 80:20 *n*-hexane:2-propanol solvent, the fluid consists of 376 *n*-hexane, 161 2-propanol, 8 R-styrene oxide, and 8 S-styrene oxide. We have verified that the hexane density in the center of the simulation cell is consistent with the experimental hexane density of 7.859 $molL^{-1}$ at 273 K [214]. All simulations are performed in the canonical ensemble at 273 K. The simulation temperature is somewhat lower than usual [77,83], although separation factors for Whelk-O1 have been measured at 273K [210,226,229]. At this temperature, the selectivity factor is slightly higher than for 298K [210,229], but retention factors increase more significantly. The temperature in the simulation cell is maintained using two independent Nosé-Hoover thermostats [162,163], one for translational motion and one for rotational motion.

For styrene oxide, twenty 1,500,000 time step molecular dynamics simulations have been performed; this corresponds to a simulation time of over 20 ns. Equilibrium properties reported in Section 4.3 correspond to averages over 40 surfaces (2 surfaces per simulation * 20 simulations). Distribution functions are monitored during the simulations and were generally converged for styrene oxide after 13 ns. For stilbene oxide, we have performed 30 simulations of 2,000,000 time steps, yielding over 40 ns of simulation time. The latter analyte is larger and requires more time to explore the full range of docking arrangements.

(R,R/S,S)-stilbene oxide is chiral, but it possesses a C_2 axis of symmetry. This symmetry provides an important confirmation of our simulations: structural information involving groups

exchanged by rotation about the symmetry axis must be the same. After 20 ns of simulation time, all distributions which are equivalent by symmetry are qualitatively similar but the intensities differ somewhat. Some quantitative differences remain after 40 ns but these are generally small and will not impact the conclusions in this chapter. The results presented in Section 4.3 are averages over the equivalent distributions. We have chosen to perform a large number of simulations, rather than a small number of longer simulations, to allow for statistical averaging over the various docking arrangements. Snapshots are collected every 2,000 iterations (1.54 ps).

The initial configurations are generated according to the method discussed in Chapter 2. Here, we note that the enantiomers and the solvent are placed randomly within the fluid region of the cell and Monte Carlo cycles are applied to eliminate strong overlaps. The first 50000 time steps are used to equilibrate the fluid. The time step in the simulations is 0.77 fs. With this time step, fluctuations in H_{NH} (Eq. 2.68) are within 5% of the total energy fluctuations over a 10 ps simulation. All simulations are performed with the MDMC program [205].

4.2.4 Properties

An analysis of equilibrium structure begins with radial distributions. We report distributions between atoms from the Whelk-O1 selector and from each enantiomer. These distributions reflect the average number of atomic pairs found for a given separation, relative to an unstructured fluid. We also report distributions between rings. In this case, the geometric center of the ring is used in the distribution function. While radial distributions are instructive, the averaging over all orientations eliminates information about the relative positions of the selector and the analyte. To partially eliminate this averaging, we consider 2D distributions (see Chapter 3 for details). The structural distribution of the Whelk-O1 selectors at the surface is analyzed based on a surface distribution, $g(z)$, with Δz=0.301 Å in Eq. 3.1.

The interactions between the epoxides and the Whelk-O1 selector can also be assessed from the snapshots. In particular, snapshots are examined to identify major docking arrangements.

We also target particular analyte-Whelk interactions and examine snapshots for occurrences of these interactions. For example, we apply the same geometric criterion as Chapter 3 to define a hydrogen bond: the distance between H and the hydrogen bond acceptor should be less than 2.6Å and the angle formed between donor-H-acceptor should be larger than 150^{o}. We report average numbers of hydrogen bonds, based on applying the geometric criterion to analyte-selector pairs in the snapshots.

4.3 Results

4.3.1 Whelk-O1 Mobility in *n*-Hexane with Styrene Oxide

ab initio calculations (Chapter 2) show that the Whelk-O1 selector in the global energy minimum orients the dinitrophenyl and phenanthryl groups roughly perpendicular to each other. In Chapter 3, we examined the conformations of the Whelk-O1 selector in *n*-hexane/2-propanol, water/methanol, and supercritical CO_2/methanol solvents. The selector was found to preferentially place the dinitrophenyl ring roughly parallel to the surface. In the present chapter, we consider pure *n*-hexane as a solvent, and the selector mobility may differ somewhat from the previous chapter with binary solvents. Also, the model surface in this chapter includes silanol groups, in addition to the trimethoxysilyl end-caps, and these may result in more flexibility for the selectors since the silanol groups are smaller than the end-caps. With these considerations in mind, we begin with a brief examination of the selector distribution at the surface. In *n*-hexane with (R/S)-styrene oxide, Whelk-O1 presents a range of conformers to the fluid, as one expects, but two dominant conformations are observed. The first of these, "phenanthryl up", corresponds roughly to the dinitrophenyl ring parallel to the surface, while the phenanthryl group is oriented perpendicular. This is the dominant configuration observed previously (Chapter 3). The other preferred conformation, "dinitrophenyl up", places the phenanthryl group roughly parallel to the surface while the dinitrophenyl points toward the fluid. The surface distributions, $g(z)$, in Fig. 4.3 illustrate the importance of these two conformers. In Fig. 4.3(a), $g(z)$ for C(15), C(18), and

C(21) [see Fig. 2.6(b) for atom numbering] in the phenanthryl group are shown. The distributions for the latter two show a clear two peak character. The first peak, at $g(z)$ of roughly 8 Å for all, corresponds to the "dinitrophenyl up" conformation where the phenanthryl group is nearly parallel to the surface. The second peak in Fig. 4.3(a) corresponds to the "phenanthryl up" conformer but the difference in peak positions indicates that on average the phenanthryl group is tilted up by 40 degrees. C(15) is closer to the surface and the pivot point for the rotation. As such, its *g(z)* is broader. However, the first peak also appears at 8 Å. Fig. 4.3(b) shows g(z) for C(29) and C(32) of the dinitrophenyl group. As with Fig. 4.3(a), two peaks are observed and the first peak is consistent with the dinitrophenyl lying nearly parallel to the surface while the second set of peaks places the ring at 60 degrees above the surface.

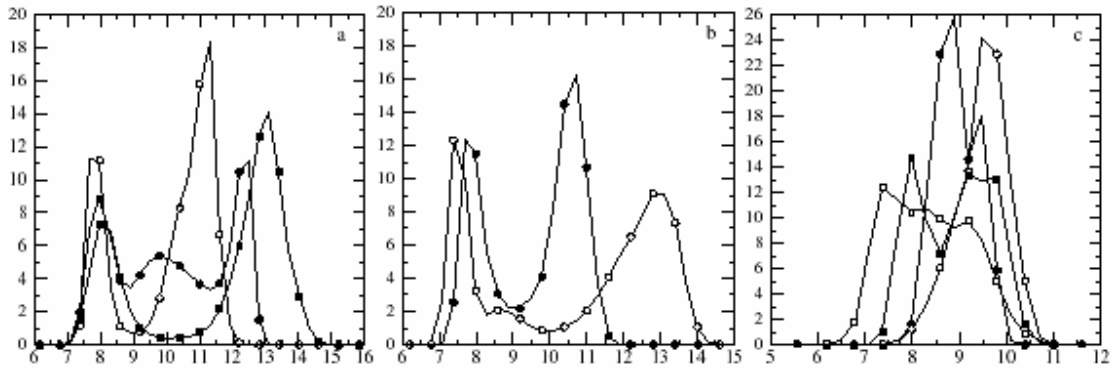


Figure 4.3: The distribution of Whelk-O1 selector atoms above the underlying layer of silicon atoms (defined as *z*=0). The distribution of C(15), C(18), C(21) are represented by filled circles, open circles, and filled squares in (a). C(29) and C(32) surface distributions are represented by filled circles and open circles in (b). In (c), the distribution of the amide linkage atoms N(25), H(26), C(27), and O(28) are represented by filled circles, open circles, filled squares, and open squares, respectively.

Fig. 4.3(c) illustrates the distribution of the amide linkage atoms above the surface. Interestingly, the N(25) and H(26) distributions show a single strong peak indicating that these atoms maintain their position above the surface despite the movement of the dinitrophenyl and phenanthryl groups. This result suggests that rotational pivoting occurs near the top of the tether. The peak positions in Fig. 4.3(c) indicate that the CO group is generally closer to the surface than

the NH group. As well, the hydrogen is clearly above N(25) and is therefore directed towards the cleft region of Whelk-O1.

4.3.2 The Selectivity of Whelk-O1 for (R/S)-Styrene Oxide

Hydrogen bonding between styrene oxide and the Whelk-O1 selector is expected to be an important contributor to the docking interaction. The radial distributions between H(26) of the Whelk-O1 selector and O(3) of R- and S-styrene oxide are compared in Fig. 4.4(a). The peak heights indicate that the oxygen is 20-30 times more likely to be near the amide hydrogen than expected based on its density in the fluid. Clearly, the distributions are qualitatively similar but the S enantiomer is more likely to have its oxygen atom close to H(26). Additional information on the interaction is provided by snapshot analysis. From Table 4.3, and using the structural definition of a H-bond [217,230], S-styrene oxide forms 35% more hydrogen bonds with H(26) than the R enantiomer forms. The 2D distributions shown in Figs. 4.4(c) and 4.4(d) are remarkably simple: each is dominated by a strong peak at z_c of 1.4-1.5 Å and r_c of 1.4 Å. Using the peak positions in Fig. 4.3(c) as a guide, the peak position in the 2D distributions in Figs. 4.4(c) and 4.4(d) are consistent with the epoxide oxygen positioned such that the N-H-O angle is roughly 180 degrees. In other words, the specificity of the H-bond, coupled with the lack of mobility in the amide linkage, restricts the positions from which an epoxide oxygen can form a successful H-bond. Evidently, S-styrene oxide is better at meeting these spatial constraints. The difference in 2D distributions is presented in Fig. 4.4(b). The differences show that the oxygen of the R enantiomer approaches H(26) from slightly further away of surface (z_c larger)

Table 4.3: Interaction statistics for R/S-styrene oxide in *n*-hexane. The statistics are based on 15425 snapshots, collected every 1.54 ps, and includes 32 Whelk-O1 selectors and 16 epoxides (racemic) for a total of 246800 interactions. The first series of entries is obtained by examining each analyte, in each snapshot, to ascertain whether it is interacting with one or more selectors. If so, the number of interactions is divided into two categories: one interaction, and two or more interactions. In Mechanism 1, a π-π interaction forms with the dinitrophenyl ring of the selector, while Mechanism 2 includes a π-π interaction with the phenyl rings of the phenanthryl group.

Structural Properties	**S-styrene oxide**	**R-styrene oxide**	**S/R ratio**
Interaction modes			
No interactions	75349	80049	0.94
1 interaction (H-bond or π-π)	34334	32929	1.04
2 or 3 interactions (at least 2 of the three must be H-bond or π-π)	13395	10155	1.32
Bridging	322	267	1.21
Total	123400	123400	1.00
Mechanism 1			
1. H-O<2.6 Å N-H-O >150^o	18340	13622	1.35
2. DNP-P<4.6 Å	38010	27518	1.38
3. 4.6 Å <P1/P2-P<6.5 Å	57573	47065	1.22
1+2	13139	5018	2.62
1+2+3	11215	2344	4.78
Mechanism 2			
1. H-O<2.6 Å N-H-O >150^o	18340	13622	1.35
4. P1/P2-P<4.6 Å	4778	12109	0.39
5. 4.6 Å <DNP-P<6.5 Å	22142	33571	0.70
1+4	243	5112	0.048
1+4+5	109	3918	0.028
Total contributions from Mechanisms 1 and 2			
1+2+3 or 1+4+5	11324	6262	1.81

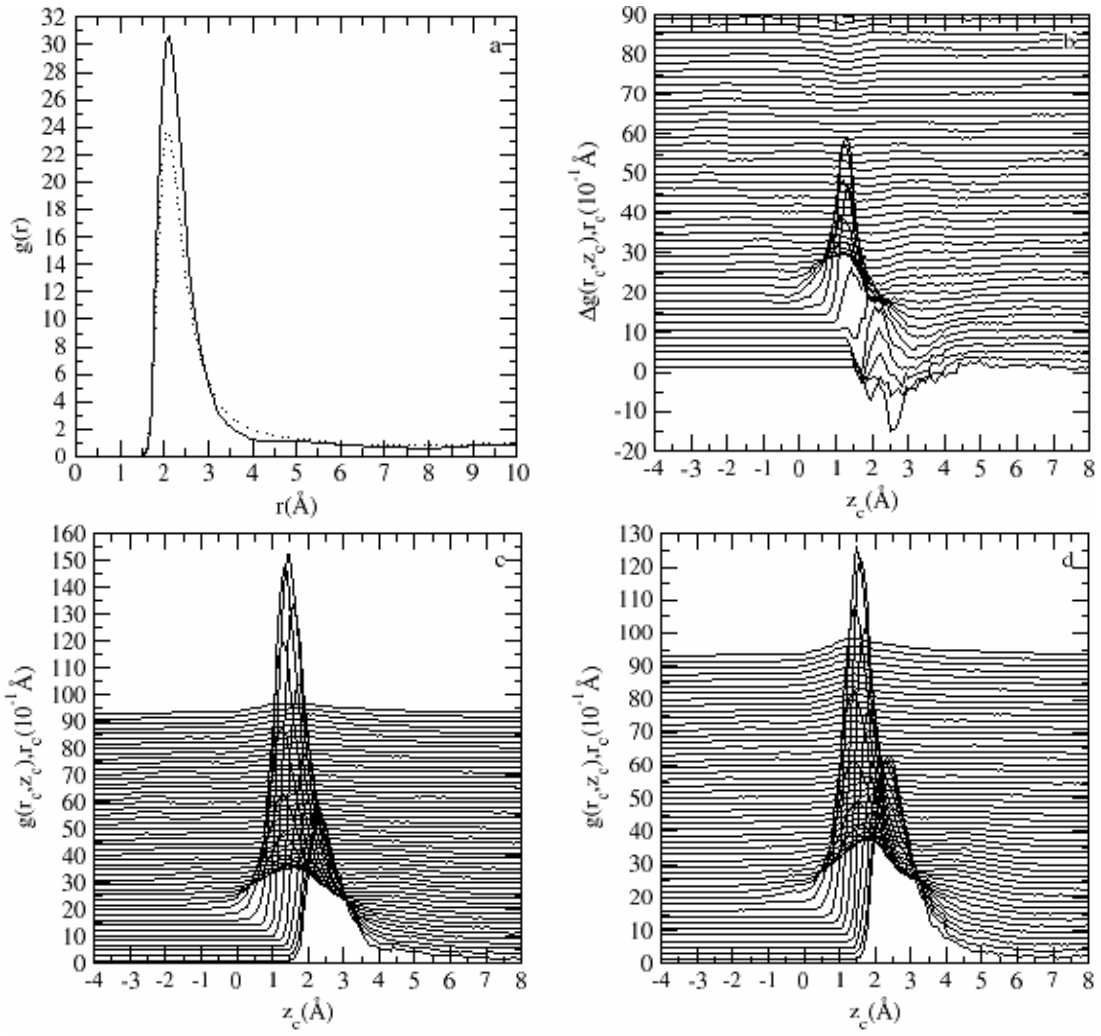


Figure 4.4: Distributions between the amide hydrogen, H(26), of the Whelk-O1 selector and O(3) of styrene oxide obtained from MD simulations of racemic styrene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for S- and R-styrene oxide, respectively. 2D distributions are shown in (c) and (d), for the S and R enantiomers, respectively. The difference in 2D distributions is shown in panel (b).

According to the three-point binding model, the electron deficient dinitrophenyl group of the selector is expected to form a π-π stack with the phenyl rings of the epoxides. Benzene-benzene interaction potentials are particularly challenging to evaluate from *ab initio* methods. These potentials were explored recently by Tsuzuki [231], using coupled cluster theory, including triples corrections, and large basis sets. Overall, slipped parallel configurations were most energetically favourable but edge-to-face arrangements were also very stable. Snapshots

have been analyzed to evaluate the angle between the rings as a function of the center-to-center distance. We find that rings separated by less than 4.6 Å have an angular distribution that is sharply peaked at around 10 degrees, as expected for face-to-face interactions. Beyond this separation, the angular distribution flattens out considerably but the average angle between the rings rises to a maximum of around 70 degrees at a center-to-center separation of roughly 7 Å. Thus, when the center-to-center distance is less than 4.6 Å the interaction is considered to be a π-π stack.

The radial distribution between the center of the dinitrophenyl ring and the aromatic ring of the styrene oxide is presented in Fig. 4.5(a). From the figure, the handedness of the epoxide has a pronounced effect on the ring-ring distributions. In particular, the S enantiomer is much more likely to have its aromatic ring within 4.6 Å of the DNP center. Beyond this region, there is a higher probability for the R enantiomer over an extended region up-to 9.0 Å. The 2D distributions shown in Figs. 4.5(c) and 4.5(d) provide an indication of the nature of the difference between R- and S-styrene oxide. Regardless of the epoxide enantiomer, the 2D distributions show ridge-like features corresponding to a roughly constant center-to-center ring distance. The ridges originate from the Whelk-O1 selector motion, which results in a distribution of positions for the DNP ring center. However, the relative ring-ring positions remain constant along the ridge. The most prominent ridge corresponds to a π-π stack at a separation of roughly 4 Å. This feature is dominant for S-styrene oxide. In Figs. 4.5(c) and 4.5(d), a second ridge at larger separations is particularly prominent for R-styrene oxide. This second ridge is consistent with the formation of an edge-to-face π-π interaction. Thus, we find that the ring interactions are selective and S-styrene oxide has a more pronounced tendency to form a π-π stack with the dinitrophenyl group.

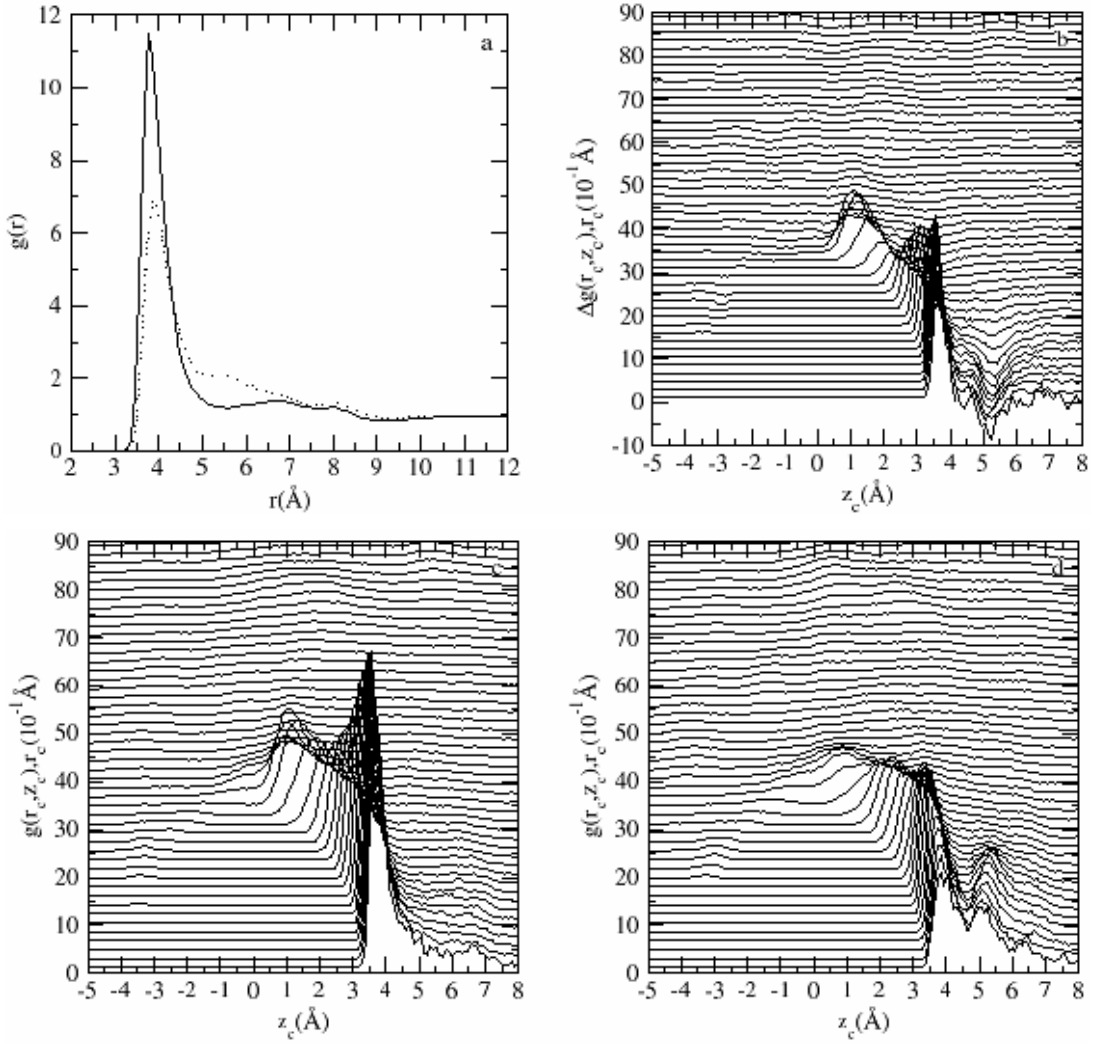


Figure 4.5: Distributions between the dinitrophenyl ring of the Whelk-O1 selector and the phenyl ring of styrene oxide. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic styrene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for the S and R enantiomers, respectively. 2D distributions are shown in (c) and (d), for S and R styrene enantiomers, respectively. The difference in 2D distributions is shown in panel (b).

On the other hand, R-styrene oxide forms more interactions with the dinitrophenyl group at intermediate separations.

In Figs. 4.5(c) and 4.5(d), the most prominent ridge persists into the region of negative z_c. That is, the analyte phenyl ring can be closer to the surface than the dinitrophenyl ring of the

selector although the probability for such arrangements is small. This type of "outside-of-cleft" interaction has been invoked to explain the selectivity of Whelk-O1 [40]: the most retained enantiomer is assumed to form a three-point interaction inside of the cleft while the first enantiomer eluted is assumed to form interactions outside of the cleft. Our simulations show that outside-of-cleft docking is limited. However, it is interesting to note from Fig. 4.5(d) that R-styrene oxide forms more of these interactions than does S-styrene oxide, but the number of interactions is small and insufficient to impact the overall preference for S-styrene oxide "docking" to the Whelk-O1 selector.

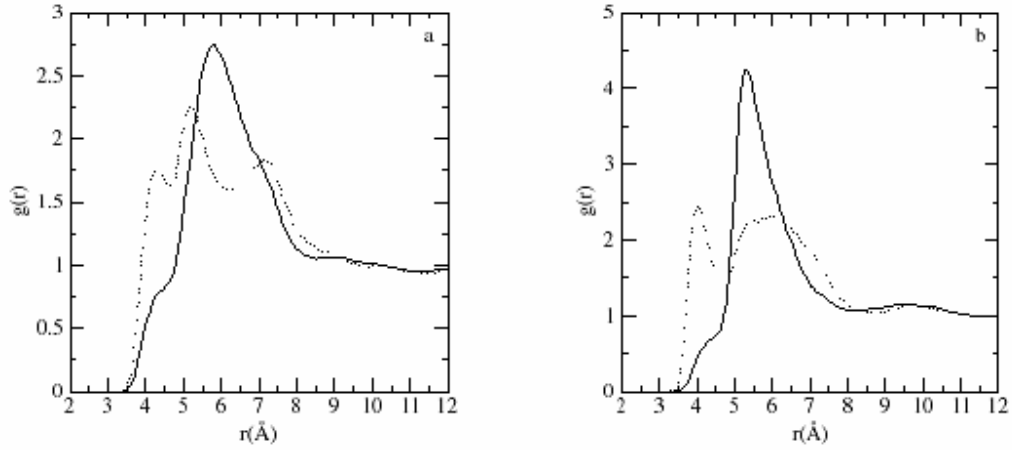


Figure 4.6: Radial distributions between the phenyl ring of styrene oxide and the phenyl rings of the selector phenanthryl group. Distributions relative to P1 and P2 are shown in (a) and (b), respectively. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic styrene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. The solid and dotted lines represent the distributions for the S and R enantiomers, respectively.

Fig. 4.6 presents radial distributions between the aromatic rings of the phenanthryl group where P1 represents the middle ring formed by C(11),C(12), C(15)-C(18) and P2 represents the topmost ring formed by C(18)-C(22) [see Fig. 2.6(b) for atom numbering], and the aromatic ring of the chiral epoxides. Both distributions indicate that the R enantiomer forms significantly more π-π stacking interactions than the S-enantiomer, which prefers π-π interactions with the dinitrophenyl ring. These results are consistent with Fig. 4.5 and suggest two important inside-of-

cleft interaction mechanisms. The first mechanism, preferred by S-styrene oxide, involves a π-π stack with the dinitrophenyl ring of the selector. The second mechanism, displayed preferentially by the R enantiomer, involves a π-π stack with the phenanthryl group. Both mechanisms include a hydrogen bond with the amide hydrogen.

A third interaction, in addition to hydrogen bonding and π-π stacking, is required to complete the three site binding model. In the proposed mechanism for Whelk-O1, CH-π or edge-to-face π-π interactions are assumed to provide this interaction. However, other interactions can form the third point of contact. These include steric repulsions between selector and enantiomers and conformational costs involved in simultaneously forming a hydrogen bond and a face-to-face interaction between rings. Edge-to-face interactions were discussed above. Consider the characteristics of CH-π interactions. Tsuzuki *et al.* [232] have recently examined the strength of CH-π interactions between benzene and small hydrocarbons. From an exhaustive series of CCSD(T) calculations, extrapolated to the basis set limit, the minimum in the potential energy curve occurs when the H points directly toward the center of the ring. For methane, the minimum energy is -6.07 kJ/mol, while for ethane the minimum is -7.61 kJ/mol. These potentials reach their minimum when the H is roughly 2.6 Å above the ring center but they vary slowly with distance. Even with the H about 4.6 Å above the ring, Tsuzuki et al. [232] found an interaction energy of -1.26 kJ/mol for methane. The potential is strongest when the CH-π bond is perpendicular to the ring, but the potential is expected to be slowly varying with the orientation of the CH-π bond [232]. In other words, CH-π interactions are very likely to stabilize the docking of both enantiomers to some extent, since these interactions are slowly decaying and non-specific in terms of orientation. We include edge-to-face π-π interactions in the discussion below but note that there is no well-defined third "point of contact" for the docking process. Rather, the third interaction encompasses multiple contributions.

Table 4.3 presents a summary of interactions based on the analysis of 15425 snapshots of the simulation cell collected every 1.54 ps. For each snapshot, the epoxide molecules are examined to ascertain the nearest Whelk-O1 selector. From this information, the number and type of interactions is evaluated. We have applied a structural definition of H-bonds[217,230] in the table. We also adopt the spatial differentiation between π-π stacking and edge-to-face π-π interactions: ring centers within 4.6 Å are considered to form a stack and ring centers from 4.6 Å to 6.5 Å are considered to form an edge-to-face π-π interaction. With these definitions in hand, consider the statistics in Table 4.3. Over 35% of the chiral epoxides interact with the surface. Most of these interactions involve a single hydrogen bond or a π-π stack, but 9.5% [(13395+10155)/(123400*2)] of the enantiomers interact with multiple points of contact. Focus on the two inside-of-cleft mechanisms noted above. According to our statistics 75% of the interactions involve multiple points of contact involve docking according to Mechanisms 1 and 2. The first three-point mechanism, involving a hydrogen bond with H(26), a π-π stack with the dinitrophenyl group, and an edge-to-face π-π interaction with the phenanthryl group is roughly 5 times more likely for S-styrene oxide although R-styrene oxide does also form this type of interaction. Figs. 4.7(a) and 4.7(c) show representative snapshots for S- and R-styrene oxide respectively. The second mechanism, involving a hydrogen bond with H(26), a π-π stack with the phenanthryl group, and a CH-π interaction with the dinitrophenyl group is the most common three-point interaction for R-styrene oxide but is very rarely observed for S-styrene oxide. Fig. 4.7(b) shows a representative snapshot for R-styrene oxide. A comparison of the total number of inside-of-cleft three-point interactions formed by the epoxides shows that S-styrene oxide is roughly twice as likely to form this type of interaction. Outside-of-cleft interactions, although rare, do occur. Fig. 4.7(d) shows a snapshot for R-styrene oxide. Note, from the snapshot, that the docking involves a hydrogen bond and a π-π stack.

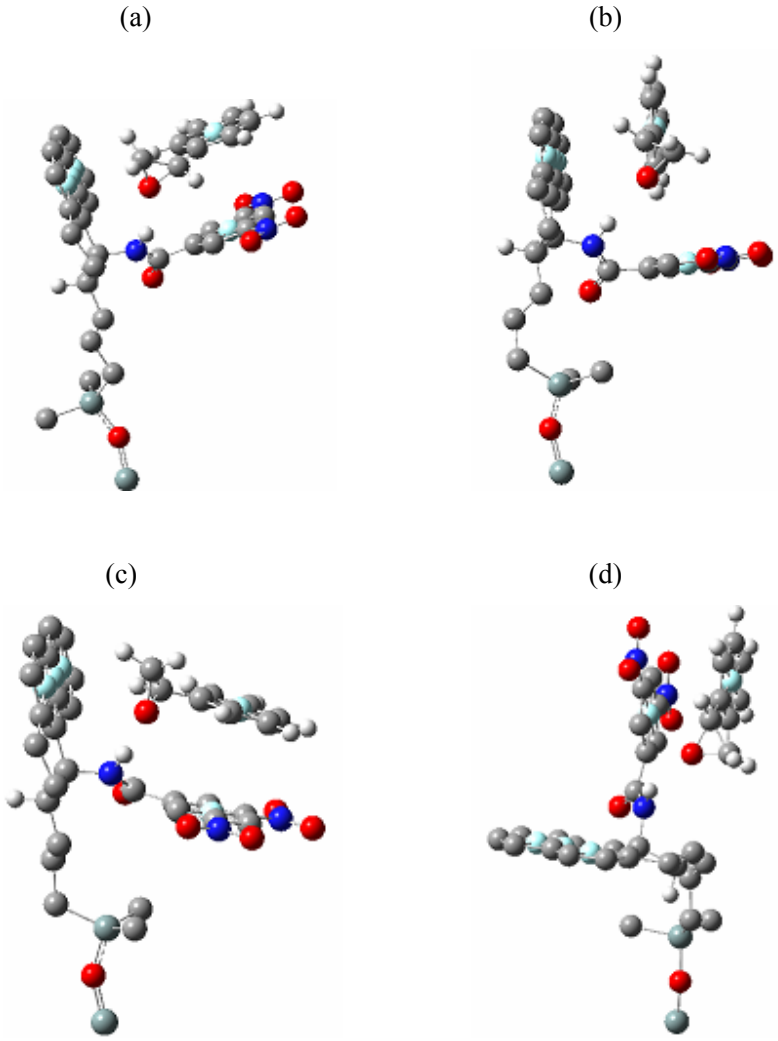


Figure 4.7: A selection of snapshots of Whelk-O1 selectors with S- and R-styrene oxide. The snapshot in (a) shows the dominant docking arrangement for S-styrene oxide. The most probable docking arrangement for R-styrene oxide is shown in (b). Panel (c) shows the secondary three-point docking arrangement for R-styrene oxide. Finally, an outside-of-cleft interaction, for R-styrene oxide, is shown in (d). Ring centers are shown in pale blue.

4.3.3 The Selectivity of Whelk-O1 for (R,R/S,S)-Stilbene Oxide

(R,R/S,S)-stilbene oxide differs from R/S-styrene oxide by the replacement of a hydrogen atom with a phenyl group. (S,S)-stilbene oxide is shown in Fig. 4.2(c) with the numbering used in this article identified in Fig. 4.2(d). The additional ring may introduce steric repulsions which will impact the docking arrangements.

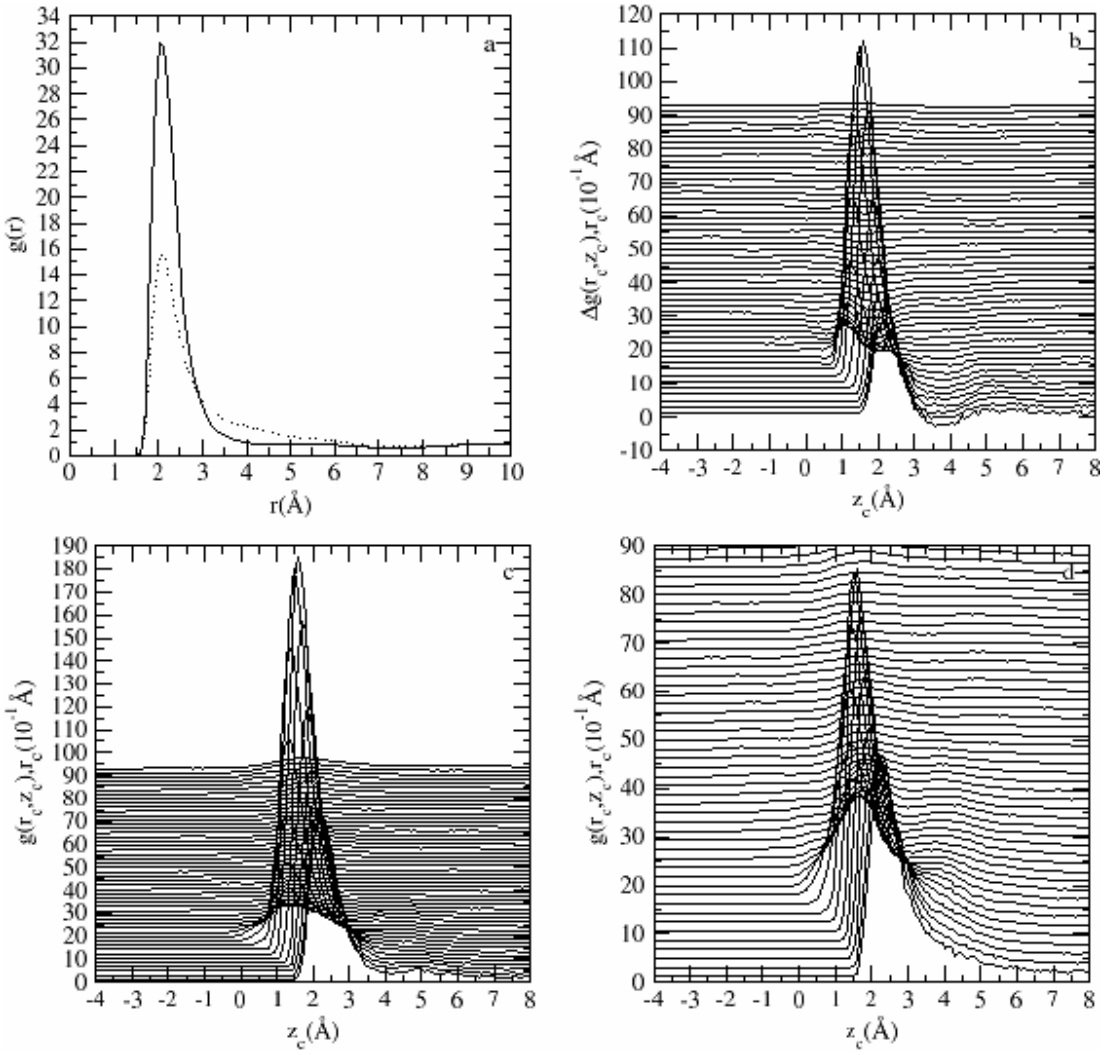


Figure 4.8: Distributions between the amide hydrogen, H(26), of the Whelk-O1 selector and O(14) of stilbene oxide obtained from MD simulations of racemic stilbene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for S,S- and R,R-stilbene oxide, respectively. 2D distributions are shown in (c) and (d), for the S,S and R,R enantiomers, respectively. The difference in 2D distributions is shown in panel (b).

The distribution between the amide hydrogen, H(26), and the oxygen from the epoxide ring of (R,R)- and (S,S)-stilbene oxide is shown in Fig. 4.8. For (S,S)-stilbene oxide the distribution is very similar to S-styrene oxide. In contrast, (R,R)-stilbene oxide is much less likely to position its oxygen near the amide hydrogen than is R-styrene oxide. 2D distributions are also shown in Fig. 4.8 and these have the same features as for styrene oxide: a single well-defined peak resulting from the relative rigidity of the Whelk-O1 selector and the specificity of the H-bond. An

analysis of H-bonds from snapshots (Table 4.4) reveals that (S,S)-stilbene oxide forms twice as many H-bonds as its mirror image. Thus, in terms of the oxygen distribution about H(26), the addition of an aromatic ring to styrene oxide has served to destabilize one enantiomer while leaving the other largely unaffected.

The positioning of the aromatic rings of stilbene oxide, relative to the dinitrophenyl ring, is explored in Fig. 4.9. Radial distributions show that (S,S)-stilbene oxide has a significantly higher probability of positioning its ring within 4.6 Å of the DNP. This indicates that (S,S)-stilbene oxide preferentially forms π-π stacks with the DNP. A secondary peak at around 7.8 Å is consistent with the average positions of the second ring in stilbene oxide, when the first ring forms a π-π stack. That is, only one ring can interact effectively with the DNP group and this leaves the second ring some distance away. For (R,R)-stilbene oxide, the radial distribution is less structured. In particular, π-π stacking does occur but is far less probable than for the mirror-image molecule. The radial distribution shows a broad shoulder between 5 and 7 Å, where (R,R)-stilbene oxide rings are more probable. Overall, (R,R)-stilbene oxide does place one of its rings in the general vicinity of the DNP but without a strong preference displayed in the radial distribution. 2D distributions are also shown in Fig. 4.9. For (S,S)-styrene oxide, a peak at r_c=0.4 Å and z_c=3.7 Å corresponds to a stilbene oxide ring directly above the DNP. The ridge following the peak is due to other arrangements of the Whelk-O1 selector, maintaining the relative orientation of the stacked rings. The broad peaks at larger separations are likely due to secondary interactions. That is, they appear because the molecule includes two rings, only one of which interacts with the DNP. As one expects, the 2D distribution for (R,R)-stilbene oxide is less structured, but shows a ridge corresponding to π-π stacking followed, at larger separations, by a broad ridge at $z_c \approx$ 5.5Å. Distributions between the rings of the phenanthryl group and the aromatic rings of stilbene oxide are shown in Fig. 4.10. The radial distributions are qualitatively very similar to the corresponding distributions for styrene oxide (Fig. 4.6). The similarity is particularly striking for P2, the top-

most ring of the phenanthryl group, and suggests that steric effects are less important for interactions with this ring.

Table 4.4: Interaction statistics for (R,R/S,S)-stilbene oxide in *n*-hexane collected. The analysis is based on 22150 snapshots, collected every 1.54 ps, and includes 32 Whelk-O1 selectors and 16 epoxides (racemic) for a total of 354400 interactions. The first series of entries is obtained by examining each analyte, in each snapshot, to ascertain whether it is interacting with one or more selectors. If so, the number of interactions is divided into two categories: one interaction, and two or more interactions.

Structural Properties	**S,S-stilbene oxide**	**R,R-stilbene oxide**	**S,S/R,R ratio**
Interaction modes			
No interactions	97527	110379	0.88
1 interaction (H-bond or π-π)	45419	48286	0.94
2 or 3 interactions (at least 2 of the three must be H-bond or π-π)	28143	14630	1.92
Bridging	6111	3905	1.56
Total	177200	177200	1.00
Mechanism 1			
1. H-O<2.6 Å N-H-O >150^o	26633	12765	2.09
2. DNP-P<4.6 Å	60346	33644	1.79
3. 4.6 Å <P1/P2-P<6.5 Å	83856	57397	1.46
1+2	22118	2802	7.89
1+2+3	21414	1229	17.42
Mechanism 2			
1. H-O<2.6 Å N-H-O >150^o	26633	12765	2.09
4. P1/P2-P<4.6 Å	7163	25478	0.28
5. 4.6 Å <DNP-P<6.5 Å	34352	62868	0.55
1+4	111	8225	0.013
1+4+5	85	7637	0.011
Total contributions from Mechanisms 1 and 2			
1+2+3 or 1+4+5	21499	8866	2.42

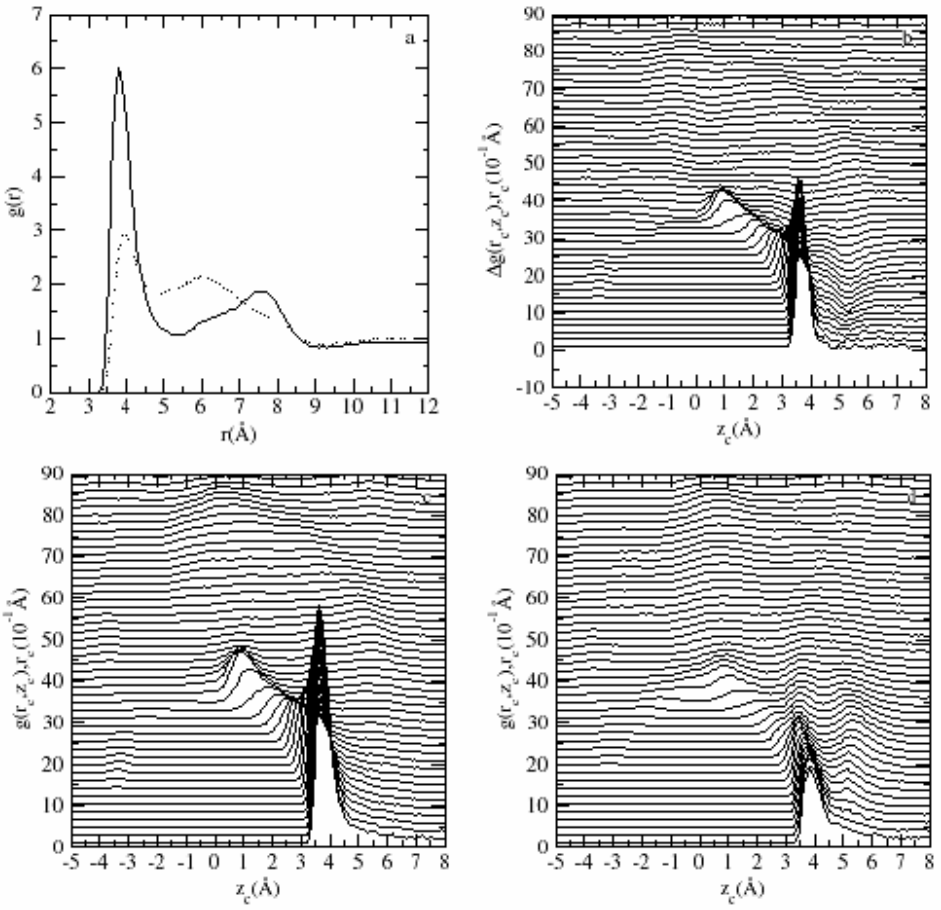


Figure 4.9: Distributions between the dinitrophenyl ring of the Whelk-O1 selector and the phenyl rings of stilbene oxide. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic stilbene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. In (a), the solid and dotted lines represent the distributions for the S and R enantiomers, respectively. 2D distributions are shown in (c) and (d), for S and R stilbene enantiomers, respectively. The difference in 2D distributions is shown in panel (b).

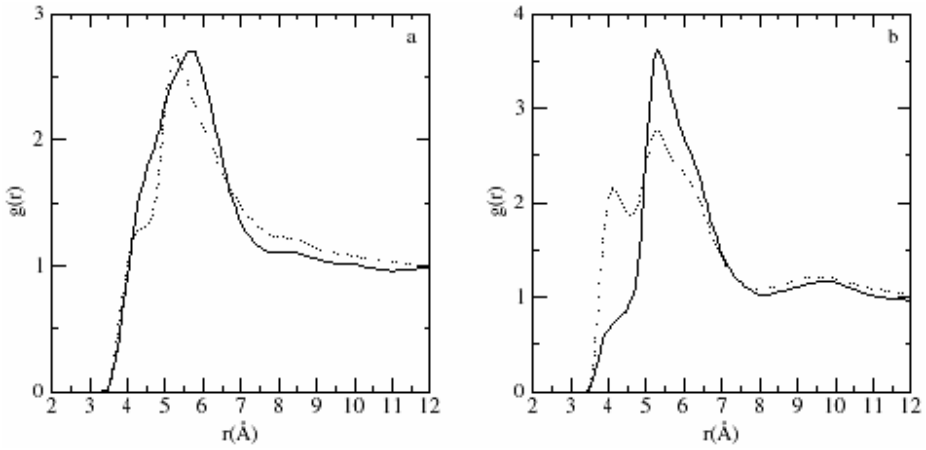


Figure 4.10: Radial distributions between the phenyl rings of stilbene oxide and the phenyl rings of the selector phenanthryl group. Distributions relative to P1 and P2 are shown in (a) and (b), respectively. The distributions are obtained by placing a site at the geometric center of each ring and collecting center-to-center data from MD simulations of racemic stilbene oxide, in *n*-hexane, exposed to the Whelk-O1 interface. The solid and dotted lines represent the distributions for the S,S and R,R enantiomers, respectively.

Stilbene oxide is clearly a bulkier analyte than styrene oxide. For the (S,S)-enantiomer, the additional ring only has a small impact on the docking and allows the formation of an H-bond and a π-π stack. This is not the case for (R,R)-stilbene oxide where the probability of forming an H-bond has dropped markedly relative to R-styrene oxide. As well, interactions with the DNP ring are less probable. Table 4.4 compares the two three-point docking arrangements discussed previously. In this case, the tabulation of edge-to-face π-π and π-π stacking interactions based on center-to-center distances proceeds with either ring in the epoxide. From the table, hydrogen bonds are more than twice as likely for (S,S)-stilbene oxide. Including the simultaneous presence of a π-π stack with DNP and an edge-to-face π-π interaction with P1 or P2, (S,S)-stilbene oxide is 17 times more likely to form this type of three-point interaction. Mechanism 2 is strongly favoured by (R,R)-stilbene oxide although it should be noted that this molecule still exhibits Mechanism 1 to some extent. Overall, (S,S)-stilbene oxide is 2.4 times more likely to form one of these three-point docking arrangements. Fig. 4.11 presents snapshots corresponding to these interaction mechanisms. For (S,S)-stilbene oxide, if a H-bond is present, then the majority of the

snapshots are very similar to the one shown in Fig. 4.11(a). That is, one docking arrangement is dominant. In this arrangement, one of the aromatic rings is effectively outside of the cleft with the remainder of the molecule interacting with the Whelk-O1 selector. In this way, the additional aromatic ring only has a slight impact on the docking. The two three-point docking arrangements for (R,R)-stilbene oxide are also shown in Fig. 4.11. Here, the secondary docking arrangement analogous to Fig. 4.11(a) requires a conformational change in the analyte that introduces torsional strain. Thus, it is less favoured than the primary docking arrangement shown in Fig. 4.11(b). The "cost" associated with docking according to Mechanism 1 can be quantified to some extent by evaluating the bend and torsional energies for docked (R,R)-stilbene oxide. We find that on average docking according to Mechanism 1 involves 20.3 kJ/mol energy for bends and torsions whereas Mechanism 2 requires only 14.2 kJ/mol. This cost is reflected in the balance between the two mechanisms: For R-styrene oxide, Mechanism 1 accounts for 37% of three-point docking events while only 14% of the (R,R)-stilbene oxide three-point docking events are consistent with Mechanism 1. The decrease is attributed to the additional strain (as reflected by high bend and torsion energies) required to place an (R,R)-stilbene oxide molecule into the cleft region with a hydrogen bond to H(26) and a π-π stack with DNP.

Larger analytes may interact with multiple selectors on the surface. This "bridging" interaction is much more probable for stilbene oxide than for styrene oxide (Tables 4.3 and 4.4). Nonetheless bridging interactions are still quite rare for stilbene oxide, and when bridging occurs it almost always involves simultaneous π-π stacks with rings from two different selectors. A snapshot of a bridging interaction is shown in Fig. 4.11(d). In the figure, a π-π stack has been formed with the DNP ring of one selector and a second π-π stack is formed with the phenanthryl rings.

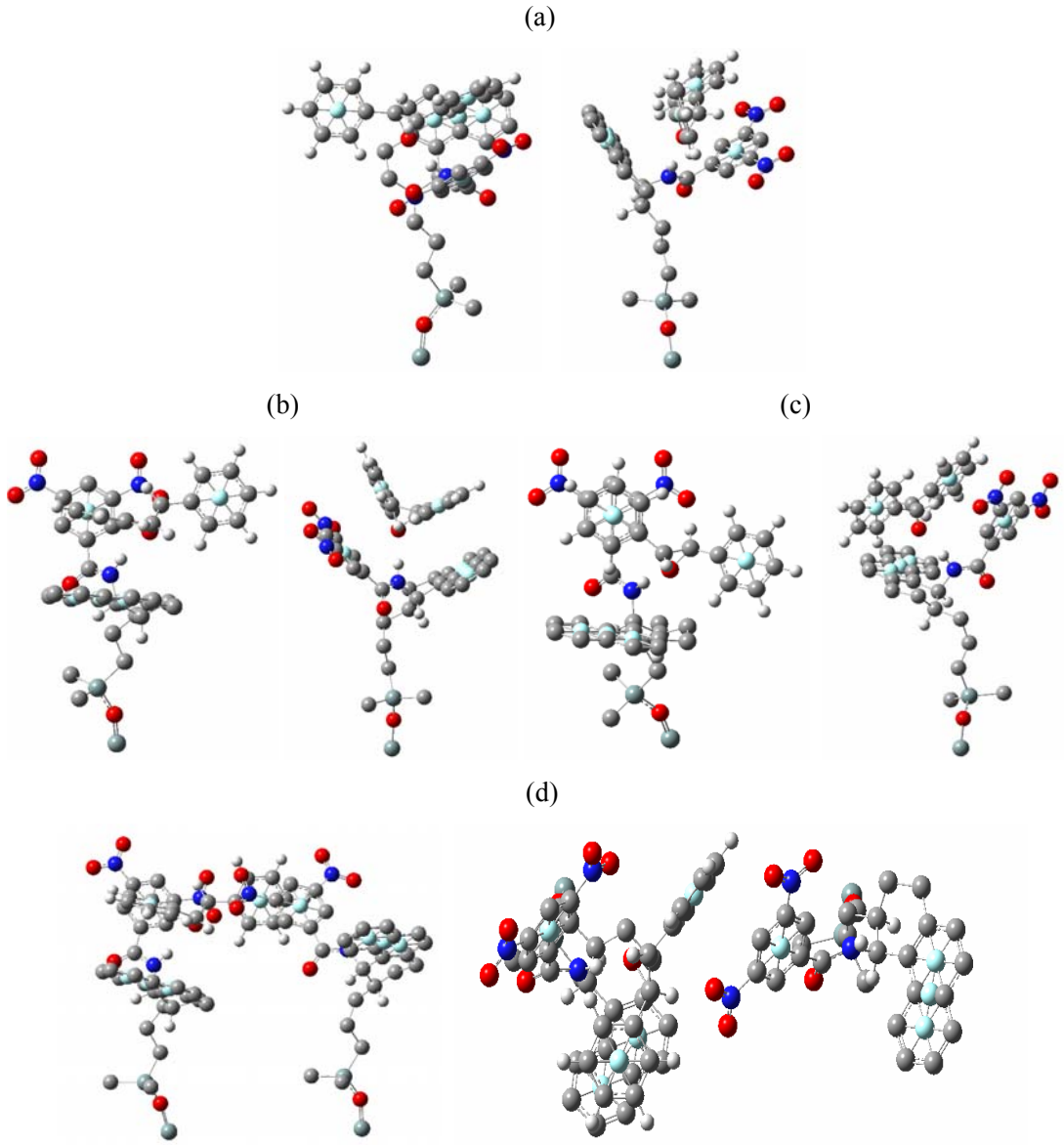


Figure 4.11: A selection of snapshots of Whelk-O1 selectors with (S,S)- and (R,R)-stilbene oxide. The panels on the left hand side provide an alternate view of the same docking arrangement. The snapshot in (a) shows the dominant docking arrangement for (S,S)-stilbene oxide. The most probable docking arrangement for (R,R)-stilbene oxide is shown in (b). Panel (c) shows the secondary docking arrangement for (R,R)-stilbene oxide. Finally, a bridging interaction, for (R,R)-stilbene oxide interacting with two selectors, is shown in (d). Ring centers are shown in pale blue.

Experiments [77] show that the retention factor for stilbene oxide is significantly smaller than for styrene oxide. Our results indicate that (R,R)-stilbene oxide does not interact particularly

well with Whelk-O1 selectors due to steric interference and we predict that it will be eluted first. The separation factor for stilbene oxide is larger than for styrene oxide and this is also consistent with our results: we find that the enantiomers of stilbene oxide show a large difference in the probability for three-point docking, relative to the difference for styrene oxide (Table 4.3).

4.3.4 The Influence of 2-Propanol on the Selectivity

Fig. 4.12 presents radial distributions for (R/S)-styrene oxide about Whelk-O1 selectors when the solvent includes 20% 2-propanol (by volume). The addition of 2-propanol leads to a decrease in the peak intensities. Consider, for example, the distribution between the oxirane oxygen and H(26) of the Whelk-O1 selector. Without the polar solvent modifier, the peak heights (Fig. 4.4) are 30 for S-styrene oxide and 23 for R-styrene oxide. In the presence of 2-propanol, these peaks decrease to 19 and 9, respectively. It is important to note that the peak for R-styrene oxide has dropped by slightly more, thereby potentially increasing the selectivity. Table 4.5 examines this point in more detail by snapshot analysis. The fraction of hydrogen bonds between analyte and Whelk-O1 selectors has dropped by nearly 50% when 2-propanol is present. However we find that, relative to R-styrene oxide, in the presence of 2-propanol S-styrene oxide forms almost twice as many H-bonds with the selector. Thus, the polar modifier has sharply decreased the number of analyte-selector hydrogen bonds but increased selectivity.

The ring distributions in Fig. 4.12 also show that the radial distributions between analyte and selector are unchanged qualitatively by the presence of 2-propanol (see Figs. 4.5 and 4.6 for comparison). However, the peaks are less intense, indicating a decreased propensity for docking. The docking probability for R-styrene oxide, the least interacting and most rapidly eluted enantiomer, is impacted more severely.

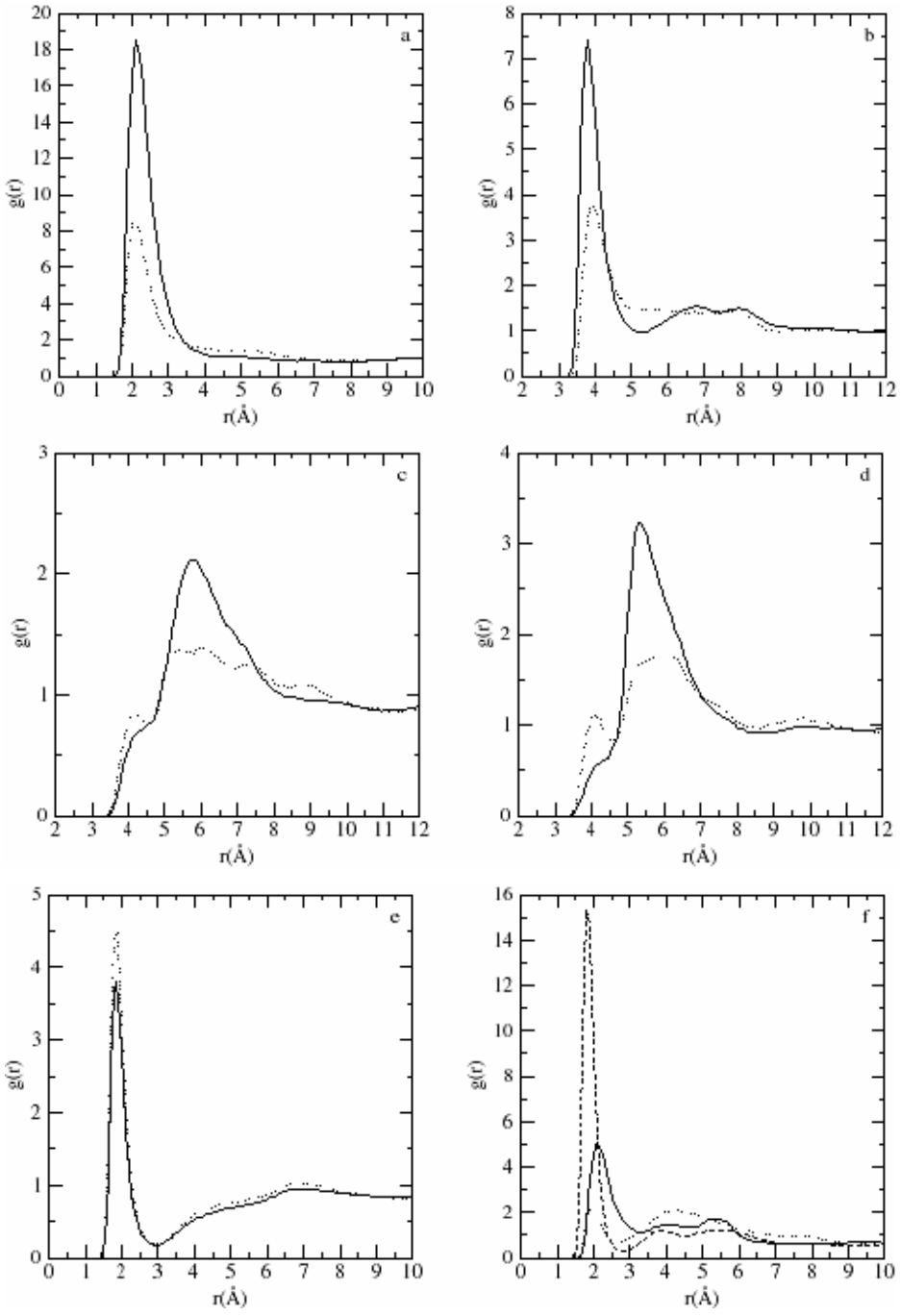


Figure 4.12: Radial distribution functions for resolution of R/S-styrene oxide on (3R,4S)-Whelk-O1 CSP in a hexane/2-propanol solvent. Solid lines are for S-styrene oxide, dotted lines are for R-styrene oxide. (a)H(26)-O(3), (b)DNP-P, (c)P1-B, (d)P2-P, (e)O(styrene oxide)-H(2-propanol). The radial distribution of 2-propanol relative to the CSP are shown in (f), where the solid line is for H(26)-O(2-propanol), the dashed line is for O(28)-H(2-propanol), and the dotted line is for O(40)-H(2-propanol).

Table 4.5: Interaction statistics for R/S-styrene oxide in *n*-hexane/2-propanol. The statistics are based on 17500 snapshots, collected every 1.54 ps, and includes 32 Whelk-O1 selectors and 16 epoxides (racemic) for a total of 280000 interactions. The first series of entries is obtained by examining each analyte, in each snapshot, to ascertain whether it is interacting with one or more selectors. If so, the number of interactions is divided into two categories: one interaction, and 2 or more interactions. In Mechanism 1, a π-π interaction forms with the dinitrophenyl ring of the selector, while Mechanism 2 includes a π-π interaction with the phenyl rings of the phenanthryl group.

Structural Properties	**S-styrene oxide**	**R-styrene oxide**	**S/R ratio**
Interaction modes			
No interactions	102487	111448	0.92
1 interaction (H-bond or π-π)	28114	24011	1.17
2 or 3 interactions (at least 2 of the three must be H-bond or π-π)	9207	4456	2.07
Bridging	192	85	2.26
Total	140000	140000	1.00
Mechanism 1			
1. H-O<2.6 Å N-H-O >150^o	12760	6513	1.96
2. DNP-P<4.6 Å	28431	18831	1.51
3. 4.6 Å <P1/P2-P<6.5 Å	54778	43657	1.25
1+2	8946	1922	4.65
1+2+3	7838	823	9.52
Mechanism 2			
1. H-O<2.6 Å N-H-O >150^o	12760	6513	1.96
4. P1/P2-P<4.6 Å	5339	7583	0.70
5. 4.6 Å <DNP-P<6.5 Å	23238	29882	0.78
1+4	250	2520	0.10
1+4+5	100	1906	0.05
Total contributions from Mechanisms 1 and 2			
1+2+3 or 1+4+5	7938	2729	2.91

The polar modifier influences the docking of analyte in several ways. Fig. 4.13 presents snapshots that illustrate the impact of the polar solvent modifier. First, the polar analyte will be stabilized by hydrogen bonding to the alcohol. Fig. 4.13(b) illustrates this solvation effect. Note the hydrogen bond to the epoxide oxygen and the network of hydrogen bonded 2-propanols near the polar region of the epoxide. As a result of this interaction, styrene oxide will have less impetus to interact with the selector. Second, the alcohol may form "in the cleft" hydrogen bonds with the selector and these will reduce the analyte docking probability by sterically interfering with the docking process. This is illustrated in Fig. 4.13(a) where several 2-propanols occupy the cleft region and hydrogen bonds to H(26) and O(28) are evident. As in Chapter 3, we find that 2-propanol hydrogen bonds to H(26) from above. That is, the alcohol is directly in the cleft region. The alcohol may also replace a docking analyte. This is illustrated by the panels in Fig. 4.13(c) but we note that this replacement is a rare process. Finally, the solvent may stabilize a docked analyte-Whelk-O1 complex. We find little evidence of this process from our simulations but this process may play a role for larger analytes.

The radial distributions and snapshot analysis indicate that the overall effect of 2-propanol on the selectivity of Whelk-O1 for styrene oxide is primarily non-selective. That is, the dominant impact is shown in Figs. 4.13(a) and (b), where the alcohol stabilizes the analyte in the bulk and blocks the cleft region of the Whelk-O1 selectors, respectively. Both of these effects are non-selective. That is, they impact both enantiomers equivalently. This effect is illustrated by the decrease in the number of hydrogen bonds observed for R- and S-styrene oxide. From Tables 4.3 and 4.5, and accounting for the different number of total interactions, S-styrene oxide would lose 2874 H-bonding interactions (out of a total of 100000) with the Whelk-O1 selectors by the addition 2-propanol whereas R-styrene oxide would lose 3193 interactions. The net drop is comparable, although slightly larger for R-styrene oxide. The secondary role of the alcohol is to replace docked analyte, as shown in Fig. 4.13(c). This is a selective process and R-styrene oxide will be more strongly affected since its docking is less favourable and is, as a result, more

susceptible to interference by the alcohol. This secondary effect accounts for the more significant loss of H-bonds for R-styrene oxide. The final result, summarized in the last row of Table 4.5, is a sharp drop in the total number of 3-point docking events with the overall S/R ratio of 1.8 observed in pure *n*-hexane increased to 2.9 by the addition of 2-propanol.

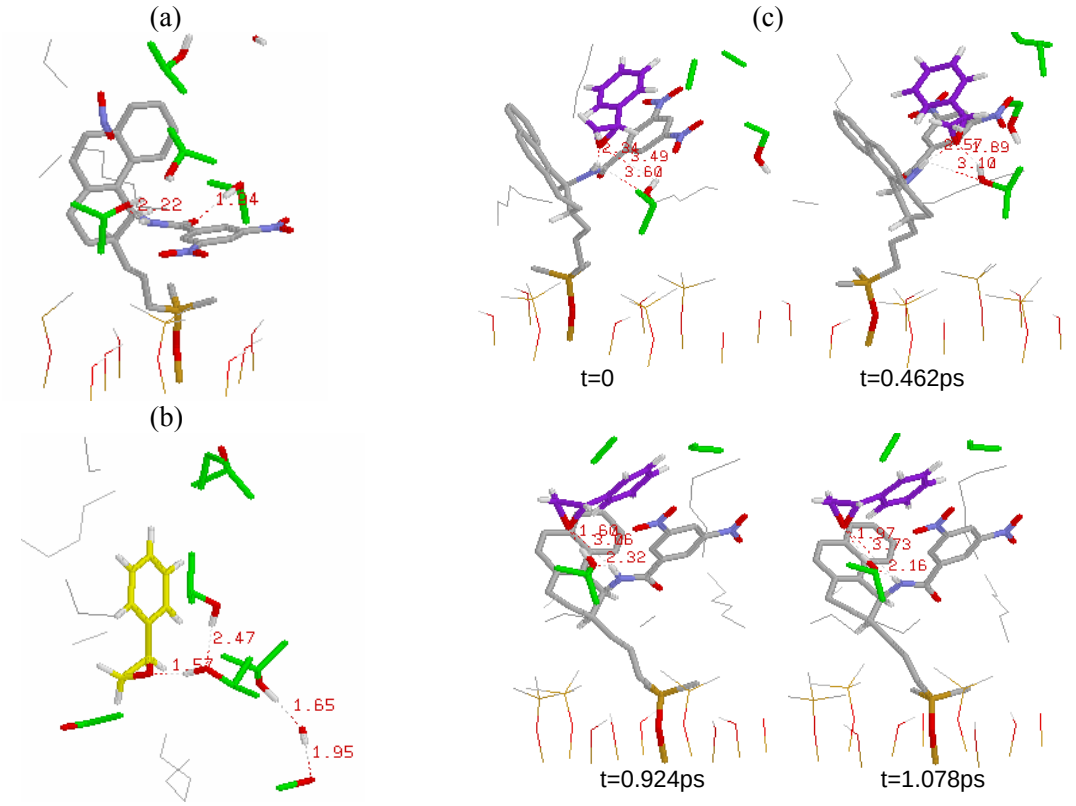


Figure 4.13: Snapshots of Whelk-O1 selectors with S/R-styrene oxide in hexane/2-propanol. The panels highlight the role of 2-propanol on solvation and selectivity. Hexane is represented by thin grey lines, while 2-propanol is green (alkyl groups), red (oxygen) and white (hydrogen). The snapshot in (a) shows several 2-propanols hydrogen bonding in the cleft region of a Whelk-O1 selector. The hydrogen bonding between styrene oxide and 2-propanol in the bulk is illustrated in (b). The subsequent four panels (c) illustrate the replacement of docked S-styrene oxide by 2-propanol.

In summary, the theoretical results are consistent with experiment, where the amount of polar modifier is typically optimized. Our results show that the addition of too much polar modifier would lead to an excessive drop in the total number of analyte-selector interactions. On the other hand, for this particular case, the alcohol does increase the selectivity.

4.3.5 Separation Factors and Other Considerations

The docking process can be summarized by two equilibria: The docking/undocking of the enantiomers (Eqs. 1.1 and 1.2). The experimentally measured separation factor α corresponds to the ratio of the two equilibrium constants where the equilibrium constant for the most docked enantiomer appears in the numerator (Eq. 1.5). Following Eq. 1.5, we have:

$$\alpha = \frac{K^{more-retained}}{K^{less-retained}} = \frac{[Most\ retained]_{docked}/[Most\ retained]_{undocked}}{[Least\ retained]_{docked}/[Least\ retained]_{undocked}} \qquad 4.1$$

Using the data from Tables 4.3-4.5, we can estimate the separation factors from the fraction of docked and undocked enantiomers. A significant difficulty in this estimate arises from the definition of a docked enantiomer. That is, what interactions should be present before an enantiomer is identified as docked? The docking invariably proceeds with a hydrogen bond to H(26), so this bond could be used. If this is done, we predict α=1.4 for styrene oxide in pure *n*-hexane, α=2.3 for stilbene oxide in pure *n*-hexane, and α=2.0 for styrene oxide in 80:20(vol) *n*-hexane:2-propanol. In contrast, the measured separation factor is 1.37 for styrene oxide in 99:1(vol) hexane:2-propanol and 2.00 for stilbene oxide in 90:10 (vol) hexane:2-propanol. The agreement is good. An alternative definition is to define a docked enantiomer by 2 or more interactions. Again, using the information from Tables 4.3-4.5, and considering π-π and H-bonding interactions, the theoretical predictions are in close agreement with experiment. For instance, the predicted α, for styrene oxide in *n*-hexane, is 13395*(123400-10155)/[10155*(123400-13395)]=1.4, and for stilbene oxide in *n*-hexane, is 2.1. This definition gives slightly better predictions than the first one. Finally, one can use the incidences of three-point docking, as defined by the simultaneous occurrence of hydrogen bonding, π-π stacking, and edge-to-face π-π interactions. The lower right hand entries in Tables 4.3-4.5 present this ratio. Although the theoretical predictions are systematically high, the qualitative trends are correct. This latter definition is appealing but problematic since the third interaction mechanism is broader than implied by the mechanism. As noted above, the edge-to-face π-π interaction is

slowly decaying and not particularly specific in terms of orientations. In addition, the emphasis on this one particular mode of contact leaves important steric considerations aside and does not include energetic penalties due to conformational changes required to form the first two interactions.

Fig. 4.14 provides styrene oxide trajectories from two 1.15 nanosecond simulations. The figure summarizes a wealth of information by identifying individual Whelk-O1 selectors by a color and the interaction type by the vertical scale ("score"). In particular, Mechanism 1 corresponds to a score of 0.8 while Mechanism 2 scores 0.6. From the figure, some enantiomers (R6 and S6, for example) remain in the bulk for the duration of the simulation and do not interact with any selectors. Other analytes, such as S7 and R16, interact only weakly with selectors. These interactions typically consist of isolated π-π stacks and the interaction is often short-lived. On the other hand, many of the analytes interact with a single Whelk-O1 selector for an extended period of time (R10 and S14 are good examples). However, even when high scores above 0.5 occur, thermal fluctuations are evident so that scores may return to zero temporarily as the analyte and the selector move in the fluid. In some instances, the scores are zero over an extended period of time.

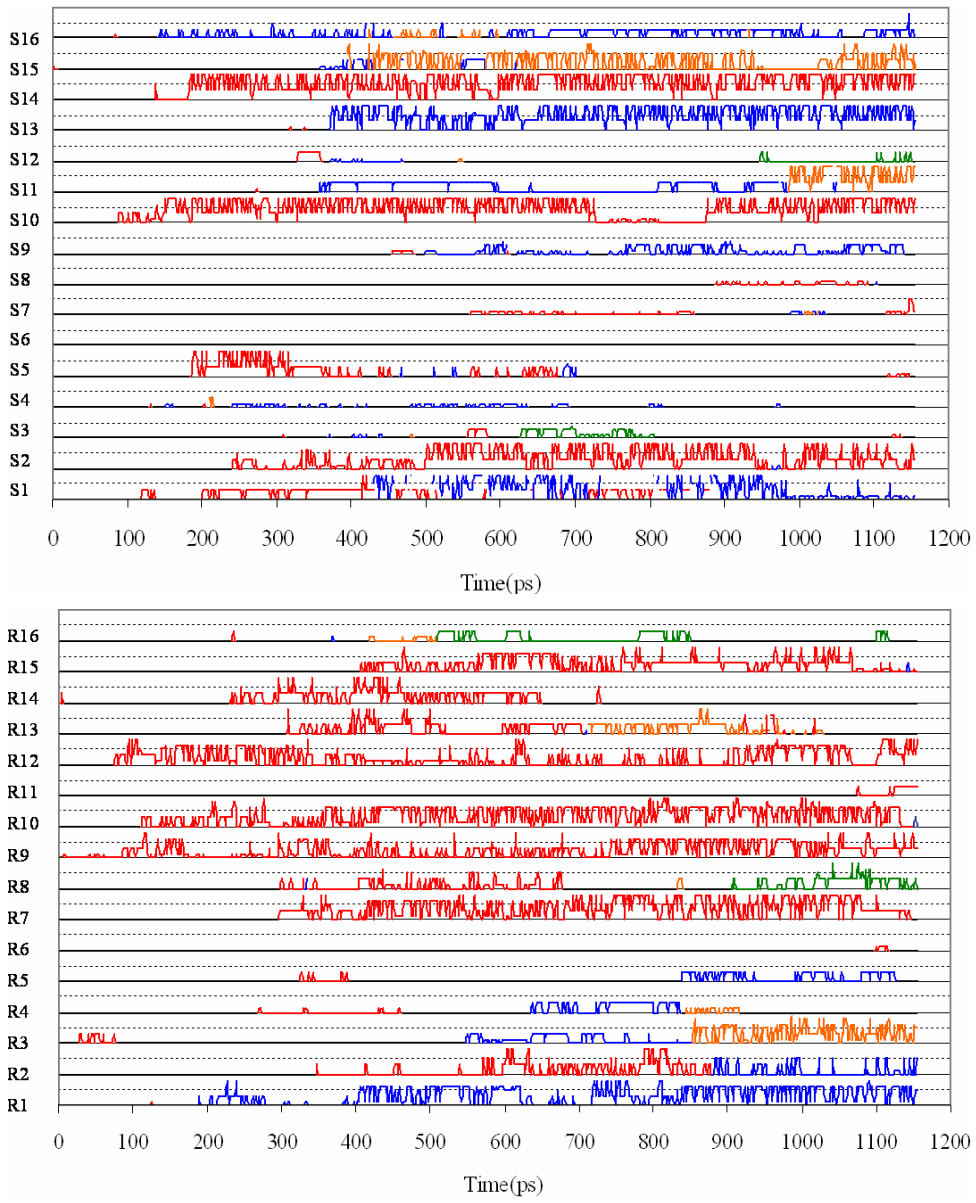


Figure 4.14: The dynamics of R/S-styrene oxide docking on Whelk-O1 selectors in *n*-hexane. Results from two simulations are shown. Each row follows the trajectory of a single enantiomer over a 1.15 ns simulation and the upper and lower panels show results for S- and R-styrene oxide respectively. The line colours identify the most strongly interacting Whelk-O1 selectors. For each enantiomer, interaction with the first selector is identified by a red line, and the subsequent selectors are identified by blue, orange, and green. For each enantiomer, the vertical scale is defined by the following interactions: a hydrogen bond is a 0.5, a π-π interaction with the dinitrophenyl group is given a value of 0.3, while a π-π interaction with either the P1 or P2 rings is worth 0.1. In this way, a docking consistent with Mechanism 1 of Table 4.3 corresponds to 0.8 while Mechanism 2 corresponds to 0.6. The dotted lines are placed at 0.5.

4.4 Conclusions

Simulations of racemic styrene oxide in *n*-hexane show that the S-enantiomer forms roughly twice the number of three-point interactions with the Whelk-O1 selectors. The three points of interaction found in the simulation are consistent with the proposed interaction mechanism for Whelk-O1: a hydrogen bond with H(26), a π-π stack, and an edge-to-face π-π interaction. However, we find two competing mechanisms which differ in the placement of the styrene ring. In the first mechanism, the ring is close to the dinitrophenyl group of the selector while the second mechanism places the ring near the phenanthryl group. Thus, the selectivity of Whelk-O1 results from the predominance of the former mechanism. It is important to note that both of these interactions are inside of the cleft. We have identified outside-of-cleft interactions, which are more prevalent for R-styrene oxide, but these are not common due to steric constraints.

(R,R)- and (S,S)-stilbene oxide interact with the Whelk-O1 selector in distinct ways. In particular, the (S,S) enantiomer is much more likely to form a hydrogen bond with the amide hydrogen of the selector. As well, this enantiomer has a definite preference for one particular three-point docking mechanism that involves the placement of one ring to form a π-π stack with the dinitrophenyl group while the other ring is outside of the cleft region. In effect, the interaction mechanism is very similar to S-styrene oxide. For (R,R)-stilbene oxide, docking is much more difficult. As a result, the radial distributions for ring-ring interactions are broader, suggesting a range of competing mechanisms. Overall, this enantiomer cannot form as many three-point interactions with the Whelk-O1 selector. Based on our results, we expect the (R,R) enantiomer to be eluted first and our analysis based on snapshots is in qualitative agreement with the experimental retention rate and separation factor [77].

The addition of 2-propanol to the *n*-hexane solvent has a dramatic impact on the selectivity of Whelk-O1. We find that, for styrene oxide, the radial distributions are systematically decreased. This arises from three factors: stabilization of styrene oxide in the bulk, interference by 2-propanol in the docking process (2-propanol is docked), and replacement of docked analyte by 2-

propanol. The least retained enantiomer, R-styrene oxide, is more impacted by the 2-propanol so that selectivity is also increased by the polar modifier.

The separation factor α can be estimated from our simulations. We find that the calculated values are in excellent agreement with simulations.

Chapter 5
Chiral Recognition of the Whelk-O1 Chiral Stationary Phase

5.1 Introduction

The Whelk-O1 chiral stationary phase (CSP) was designed [40,74] in the 1990's to isolate S-naproxen, a non-steroidal anti-inflammatory drug. As it turns out, this stationary phase is selective for a broad range of racemates [77,83] and can be used in various solvent environments [77,89,90]. Today, the Whelk-O1 stationary phase is the most successful [30] charge transfer CSP. The overall shape of the selective molecule is a cleft, and this shape is viewed as key to its selectivity [55,57,81,233]. As a general rule, enantiomers with a hydrogen bond acceptor and an aromatic group close to a chiral center tend to be well resolved on the Whelk-O1 CSP [83].

As noted in Chapter 1, two three-point mechanisms have been proposed for Whelk-O1. Both agree on the first two interactions: a hydrogen bond forms between the amide hydrogen and a hydrogen bonding acceptor in the analyte; and a π-π stack forms between the dinitrophenyl group and an aromatic ring in the analyte. For the third point, Del Rio *et al.* [96] proposed that the mechanism includes side-chain CH-π interactions. In contrast, Pirkle *et al.* [77,97] emphasized an edge-to-face π-π interaction formed by the phenanthryl moiety of the Whelk and an aromatic ring in the analyte. Koscho and Pirkle further highlighted the role of the cleft of the selector on enantioselectivity with spectroscopic techniques [55,57]. A comprehensive ^{1}H-NMR study of 22 analytes in the presence of a modified Whelk-O1 selector [57] found that most of the analytes show non-equivalent chemical shifts and these were interpreted as evidence that the favoured enantiomer docks inside of the cleft while the other docks outside of the cleft. An X-ray co-crystallization study [55] of a modified Whelk-O1 selector with N-(1-(4-bromophenyl)ethyl)pivalamide indicates that the favoured enantiomer docks inside of the cleft, while the least retained enantiomer is positioned outside of the cleft region.

In Chapter 4, we have carried out MD simulations to explore the selectivity of the Whelk-O1 CSP for epoxides. We found that hydrogen bonding with the amide hydrogen, and ring-ring stacking were common elements in the docked complexes. These are expected based on the proposed three-point interaction models [77,97] but other factors, such as torsional strain, were also identified as contributors to the overall selectivity. In addition, competition between multiple inside-of-cleft docked complexes was observed and this is not consistent with the idea that only one enantiomer enters the cleft to form three specific interactions with the Whelk-O1. These studies highlight the need for a comprehensive exploration of the factors that contribute to the chiral selectivity of Whelk-O1. In this chapter, the docking of ten analytes is explored, with each analyte chosen to highlight the various aspects of the docking process.

This chapter begins with a brief review of the model of the (3R,4S)-Whelk-O1 selector, the ten analytes, and the chiral interface. Details of the molecular dynamics simulations are also provided in Section 5.2. The simulation results will be presented in Section 5.3. The chapter concludes with a brief discussion of the results in Section 5.4.

5.2 Theoretical Details

The Whelk-O1 selector is shown in Fig. 2.6(b). The amide linkage is somewhat flexible and this is relevant to docking since flexibility in this region changes the cleft. *ab initio* studies (Chapter 2) show that torsion about the C(27)-C(29) bond is particularly flexible with a $\pm 40^{o}$ range within 2.5 kJ/mol of the minimum. In this way, the dinitrophenyl ring can twist, with little energy penalty, to form a ring-ring interaction with the analyte. Limited torsion, of roughly $\pm 10^{o}$, about the C(10)-N(25) bond may also occur so that the amide hydrogen can reorient somewhat to accommodate the analyte. A suitable semi-flexible model for the selector, based on extensive *ab initio* calculations, was discussed in Chapters 3 and 4.

The model Whelk-O1 interface is representative of the pore surface and has been discussed in previous chapters. Briefly, the interface includes 16 (3R,4S)-Whelk-O1 selectors, 48

trimethylsilyl end-caps, and 64 silanol groups corresponding to coverages of 1.07 μmol/m^2, 3.20 μmol/m^2 and 4.26 μmol/m^2 respectively, within the range of experimental values [40,188,189]. An underlying layer of 128 Si atoms, to which end-caps, silanol groups, and selectors are attached, is stationary and represents the solid substrate. The end-cap and silanol models have been discussed in Chapter 4.

Fig. 5.1 shows the ten analytes under consideration. Racemates of (n-(1-(4-bromophenyl)ethyl)pivalamide) (PAMD) are very well resolved by Whelk-O1, while styrene oxide (STYO), stilbene oxide (STBO), 1,2-diphenylethane-1,2-diol (DPED), and 2-hydroxy-1,2-diphenylethanone (DPEO) are well resolved, 1-phenylethane-1,2-diol (PEDO) is poorly resolved, and 1,2-diphenylcyclopropane (DPCP), 2-methyloxirane (MOXR) and 1-phenylethanamine (PEAM) are not resolved at all, and conflicting results [234,235] have been reported for 1-phenylethanol (PEOL). Comparisons among these analytes provide a direct assessment of the factors that influence selectivity. For instance, a comparison between STYO and MOXR highlights the impact of ring-ring interactions. In the simulations, the analyte potentials must accurately represent molecular flexibility. Similar to the STYO and STBO models developed in Chapter 4, the models for the other 8 analytes are also based on extensive series of *ab initio* calculations. Full details are provided in Appendix A.

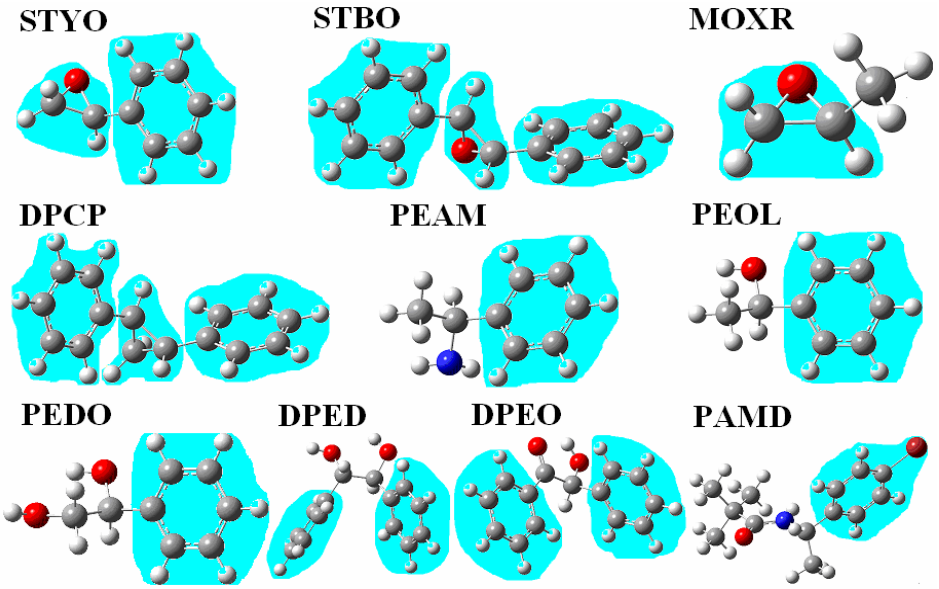


Figure 5.1: Structures of the ten analytes. the B3LYP/6-311++G** optimized structures are provided for: (S)-Styrene oxide (S-STYO); (S,S)-Stilbene oxide (S,S-STBO); (S)-2-methyloxirane (S-MOXR); (R,R)-1,2-diphenylcyclopropane (R,R-DPCP); (R)-1-phenylethanol (R-PEOL); (S)-1-phenylethanamine (S-PEAM); (S)-1-phenylethane-1,2-diol (S-PEDO); (S,S)-1,2-diphenylethane-1,2-diol (S,S-DPED); (S)-2-hydroxy-1,2-diphenylethanone (S-DPEO); (S)-N-(1-(4-bromophenyl)ethyl)pivalamide (S-PAMD). The shaded areas are treated as rigid units in the simulations.

Extensive MD simulations were performed for the Whelk-O1 interface. As noted in Chapter 4, the simulation cell includes two chiral interfaces, with solvent and racemate in-between and empty space above and below the two underlying Si layers. All simulations are performed in the NVT ensemble at 273K, with the temperature maintained by Nosé-Hoover thermostats [162,163]. The equations of motion are integrated following the algorithm of Martyna *et al*. [175] to preserve the RATTLED [142] positions and velocities. The time step in the simulations is 0.77 fs. The solvent consists of 450-470 *n*-hexane molecules, consistent with the experimental hexane density [214] of 7.859 mol/L at 273 K. Each simulation also includes 16 analytes: 8 of each enantiomer. An *n*-hexane solvent has been chosen for the simulations by appeal to its ubiquitous use as the major solvent in normal phase chromatography [77]. In addition, several experimental studies [84,210] report only small changes in separation factors

with modest additions of 2-propanol. In accordance with solvation studies for Whelk-O1 (Chapter 3), the TraPPE-UA [211] model for *n*-hexane has been selected.

For each analyte, fifteen to thirty six 1,500,000-1,600,000 time step MD simulations have been performed; this corresponds to a total simulation time of 17-41 ns. The lower limit is appropriate for analytes with low retention times while the upper limit applies to strongly interacting analytes. In all cases, the racemate distribution is assessed over shorter time periods and simulations continue until satisfactory convergence is achieved. We have chosen to perform a large number of ≈1 ns simulations, rather than a small number of longer simulations, to improve statistical averaging over the various docking arrangements. This choice is based, in part, on a dynamical study of STYO (Chapter 4) which showed that analytes typically visited several selectors during this time. Initial configurations are generated as discussed in Chapter 2. The first 50,000-100,000 time steps are used for equilibration. Following this, distributions are assessed every 20 iterations and a snapshot of the simulation cell is collected every 2000 iterations. All simulations are performed with the MDMC program [205].

5.3 Results and Discussion

5.3.1 Overview

The simultaneous presence of a hydrogen-bond (H-bond) and a face-to-face ring-ring interaction ("π-π stack") brings the analyte into close proximity with the Whelk-O1 selector, allowing other interactions to be discerned. A recent NMR study [57] of analytes in the presence of a modified Whelk-O1 selector showed that analytes that lack an H-bond acceptor were not discriminated by the solvating agent while those without an aromatic ring only interacted weakly. The simulation study in Chapter 4 of chiral epoxides also revealed that the recognition process involved these two specific interactions. These results support the idea that H-bonding and π-π stacking are essential and we begin with an analysis of analyte-selector interactions based on these two modes of interaction. Further justification for our emphasis on these interactions will be based on

comparisons between the ten analytes. Specifically, the necessity for hydrogen bonding will be evident in the comparison between STBO and DPCP in Section 5.3.1. Likewise the comparison between STYO and MOXR, presented in Section 5.3.2, will illustrate the importance of π-π stacking.

At this point, working definitions for a hydrogen bond and a π-π stack are required. As described in Chapter 4, geometric criterion [217,230] is used to define a H-bond: the distance between the hydrogen and the H-bond acceptor should be less than 2.6Å and the donor-H-acceptor angle should be larger than 150^{o}. A definition for a π-π stack is more problematic. The benzene dimer has been the subject of several recent *ab initio* studies [231,236]. These studies reveal that the slipped parallel configuration is slightly more stable than a T-configuration. Within the simulations, the rings will have a broad range of relative orientations but, unfortunately, few *ab initio* [231] studies include any consideration of ring tilt. We have proceeded by evaluating the average angle between the analyte rings and the DNP group. Illustrative results are shown in Fig. 5.2(b). The average angle between the two rings increases sharply from less than 20^{o} to nearly 70^{o}, and the transition occurs when the ring centers are separated by roughly 4.6 Å. Based on this, we define the following: Two aromatic rings form a π-π stack when they have a center-to-center distance of less than 4.6 Å and a ring-ring tilt angle of less than 30^{o}. This criterion, illustrated in Fig. 5.2(a), is more stringent than we adopted previously (Chapter 4).

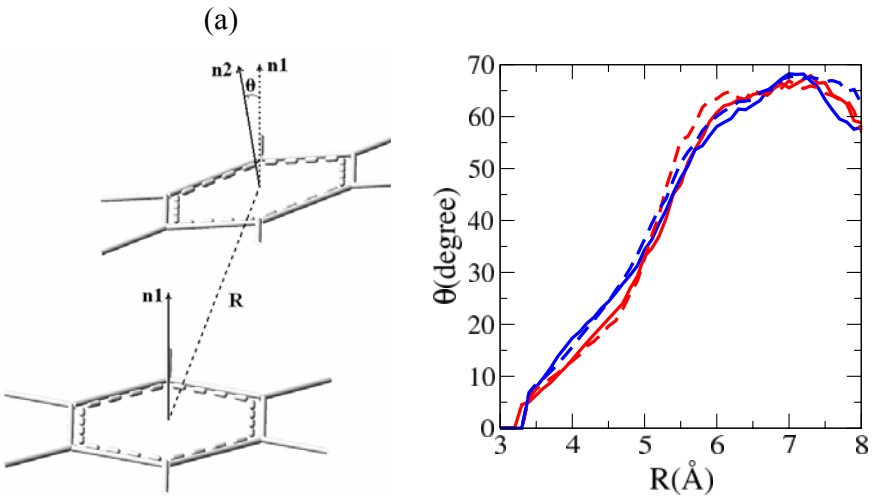


Figure 5.2: The definition of a π-π stack (a) The distance between ring centers (R) must be less than 4.6 Å and the ring tilt angle (θ) is less than 30°. (b) The average ring tilt angle as a function of the center-to-center distance between the DNP ring of the Whelk-O1 selector and the aromatic ring of STYO (S: red solid, R: red dashed) and DPEO (S: blue solid, R: blue dashed).

An analyte is "docked" when it simultaneously forms an H-bond and a π-π stack, and the presence of these interactions is determined by the application of the geometric criteria described above. The Whelk-O1 selector has three potential H-bonding sites: the amide hydrogen [H(26)]; the amide oxygen [O(28)]; or one of the nitro oxygens [O(36), O(37), O(39), O(40)]. Likewise, two π-π stacking interactions are possible since the Whelk-O1 selector has two aromatic regions. The Pirkle [77] and del Rio [96] binding models consider only H-bonding to the H(26) and π-π stacking with the DNP ring, but we consider a total of six possible docking modes. The four probable docking arrangements are illustrated in Fig. 5.3.

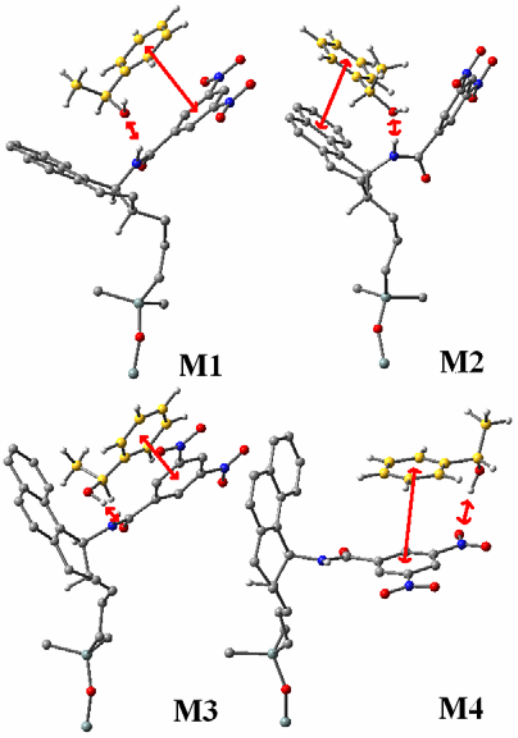


Figure 5.3: Four docking arrangements (M1-M4) on Whelk-O1. Snapshots show S-PEOL, docked according to each mechanism. Red arrows identify hydrogen bonding and π-π stacking interactions. Note that other selectors, end-caps, and the solvent are omitted from the snapshots to highlight the docking arrangements.

The analysis of the selector-analyte interactions begins with an evaluation of the frequency of interactions. This analysis is summarized in the first four columns of Table 5.1. At any instant in time, roughly 60-80% of the analytes do not participate in either a π-π stack or an H-bond with a selector. However, 10-30% of the enantiomers form one of these interactions while 2-15% form two interactions. The latter category is dominated by docked analytes but also includes analytes which interact with multiple selectors, form multiple H-bonds, or multiple π-π stacks. At first glance, the interaction frequencies in Table 5.1 suggest that most analytes are "far" from the selector. In fact this is not the case. The H-bond and π-π stack definitions adopted here are stringent. By relaxing these conditions slightly, for example by increasing the distance criteria for H-bonding and π-π stacking to 3.0 Å and 5.0 Å respectively, the number of 1- and 2-point interactions increase by a few percent. Thus, a significant proportion of the analytes are in the

general vicinity of a selector but many are not quite close enough to the H-bonding sites and the aromatic rings of the selector.

From Table 5.1, the frequency of docked interactions varies greatly among the analytes. R-PAMD is the least likely to dock with, on average, only 1.4% of the molecules simultaneously forming an H-bond and a π-π stack at any given time. In contrast, an average of 14.1% of the S,S-DPED enantiomers will be docked at any point in time. Although six docked arrangements are possible, snapshot analysis reveals that only four are probable and these are shown in Fig. 5.3 for S-PEOL. The final four columns in Table 5.1 identify the relative probabilities for these four docking arrangements. Note that the M1 docking arrangement is consistent with the Pirkle [77] and del Rio [96] interaction models. M2 was identified as an important docking mode for STYO and STBO in Chapter 4 but, to our knowledge, this is the first evidence for M3 and M4 docking. For all analytes, the most frequently docked enantiomer has a preference for the M1 mechanism. In contrast the other enantiomer adopts multiple docking arrangements. A similar result [96] has been observed from the molecular mechanics optimization of docked 1-(4-halogeno-phenyl)-1-ethylamines on Whelk-O1 where the binding of the most retained enantiomer was highly specific, while the other showed multiple low-energy conformations. Similar trends have been noted for other CSPs [123,237].

Table 5.1: Interaction summary for the ten analytes on Whelk-O1 extracted from MD snapshot analysis. An interaction consists of either a hydrogen bond or a π-π stack (see Section 5.3.1 for details). The first group of columns reports the percentage of analytes that have 0-3 interactions with the Whelk-O1 selector. A "docked" interaction occurs when an H-bond and a π-π stack are present simultaneously. The central column compares the experimental separation factors with the calculated factors, obtained from the ratio of docking probabilities. The last set of columns describe the probability for four types of docking events, given as follows. In M1 and M2 docking, the analyte forms a hydrogen bond with the amide hydrogen [H(26)], and a π-π stack is formed with either the dinitrophenyl moiety (M1) or the phenanthryl group (M2). In M3 and M4, the analyte forms a π-π stack with the dinitrophenyl moiety and a hydrogen bond with either the amide oxygen[O(28)], for M3 docking, or with a nitro oxygen [O(36),O(37),O(39),O(40)] for M4.

Analyte		**Interaction frequency**					α^{calc}	**Docking Arrangement**			
		0	**1**	**2**	**3**	**Docked**	(α^{exp})	**M1**	**M2**	**M3**	**M4**
ST YO	S	63.6	26.3	10.1	<0.1	9.9	1.3	98.5	1.5	N/A	N/A
	R	66.8	25.4	7.8	<0.1	7.7	(1.37[77])	48.8	51.2	N/A	N/A
ST BO	S,S	62.5	23.5	13.0	1.0	9.6	1.8	98.5	1.5	N/A	N/A
	R,R	65.6	24.7	9.1	0.5	5.7	(2.00[77])	17.6	82.4	N/A	N/A
DPCP	S,S	66.6	30.1	3.3	0	N/A	1.0	N/A	N/A	N/A	N/A
	R,R	69.0	28.7	2.3	0	N/A	(1.00[57])	N/A	N/A	N/A	N/A
MOXR	S	88.0	12.0	0	0	N/A	1.0	N/A	N/A	N/A	N/A
	R	88.2	11.8	0	0	N/A		N/A	N/A	N/A	N/A
PEOL	S	63.8	26.5	9.0	0.7	7.0	1.2 (1.00[234],	50.7	15.5	11.1	19.7
	R	63.7	25.8	9.9	0.7	8.3	1.68[235])	57.7	1.9	10.4	28.1
PEAM	S	66.4	22.7	9.9	1.0	8.5	1.1	48.9	8.9	7.1	34.6
	R	69.0	20.7	9.0	1.0	8.1	(1.00[57])	35.0	4.9	8.2	51.0
PEDO	S	65.7	23.0	10.0	1.4	5.5	1.1	72.4	6.1	11.5	8.2
	R	64.0	25.0	9.3	1.7	5.0	(1.08[77])	25.3	46.7	14.0	10.6
DPED	S,S	56.7	25.3	15.1	3.0	14.1	1.2	88.0	4.2	2.1	5.5
	R,R	59.1	24.6	13.3	3.1	12.1	(1.32[77])	82.6	7.3	3.9	5.9
DPEO	S	66.6	23.8	8.9	0.7	7.6	1.8	68.4	28.2	0.2	0.4
	R	62.8	23.2	13.2	0.9	12.6	(1.96[77])	88.3	10.8	0.4	0.3
PAMD	S	79.7	14.8	5.4	0.2	4.6	3.4	82.9	3.0	0.9	13.2
	R	82.7	14.7	2.5	0.2	1.4	(10.11[57])	17.1	16.6	10.1	55.9

Theoretical estimates of the separation factor, α^{calc}, can be obtained from Eq. 4.1 via the docking frequencies since they reflect the equilibria in Eqs. 1.1 and 1.2. The agreement between theory and experiment in Table 5.1 is outstanding, considering that the experiments were performed with a range of co-solvents, in different proportions. The only exception is PAMD, where the simulations significantly underestimate the separation factor. To confirm that docking modes have been well explored, several 3-5ns simulations were carried out for PAMD and these also underestimate the separation factor. Since PAMD is flexible and branched, relative to the other analytes, some docking modes may require even longer times and may not be sampled sufficiently in the simulations. On the other hand, PAMD is particularly complex since the enantiomers have very distinct docking preferences and these involve analyte and selector torsion: docking probabilities may be sensitive to details of the potential. Regardless, simulations correctly predict a PAMD separation factor that is much larger than for the other analytes (Table 5.1).

Absolute elution orders have been determined from experiment for STYO [223] and PAMD [238] and the most retained enantiomers are S-STYO and S-PAMD. Elution orders can be extracted from the simulations by identifying the enantiomer that interacts most frequently with the Whelk-O1. Table 5.1 reveals that the docking counts are indeed higher for S-STYO, and S-PAMD. For the other analytes, Table 5.1 provides a prediction of the elution order: The R-enantiomer is predicted to elute first, except for PEOL and DPEO.

5.3.2 The Importance of Hydrogen Bonding on the Selectivity of Whelk-O1

The importance of hydrogen bonding is well illustrated by an example. STBO and DPCP are equivalent in all regards except that the oxygen in the former has been replaced by a carbon in the latter. Experimentally, STBO can be well separated [77] on Whelk-O1 while DPCP cannot [57]. The radial distribution between H(26) of the selector and either the oxygen of STBO or the corresponding carbon of DPCP is shown in Fig. 5.4(a). Both enantiomers of STBO have a

significant probability of having their oxygen near the amide hydrogen. The distributions for DPCP are, by comparison, structureless showing that the cyclopropane ring is not localized near H(26). Snapshots of DPCP interacting with the Whelk-O1 selector are shown in Fig. 5.4(b). The predominant docking arrangement involves the placement of one analyte ring above the DNP ring but the other ring does not adopt any preferred position. In contrast, S,S-STBO docks almost exclusively according to M1 while R,R-STBO prefers M2 [see Figs. 5.4(c) and (d)]. That is, the H-bond "anchors" the analyte in the cleft region.

Table 5.2 provides H-bonding statistics for all the analytes. The probabilities are obtained by analyzing each snapshot to identify H-bonding events. Consider first the amide oxygen O(28). This oxygen forms relatively few H-bonds, in part due to the fact that it is located slightly below the cleft region and may be inaccessible to the analyte. Of all the analytes, PEDO, which has two alcohol groups and significant conformational freedom forms the most H-bonds with the carbonyl oxygen. Nonetheless, less than 5% of the PEDO enantiomers form these bonds. H-bonding to the nitro oxygens is highly variable between the analytes, with a minimum of 1.7% for R-DPEO and a maximum of 15.9% for R-PEDO. This interaction is favored by the location of the nitro oxygens at the edge of the cleft region and the fact that there are four oxygens available. Even so, most analytes prefer to H-bond with the amide hydrogen. Note that DPEO and PAMD, the two analytes with the highest predicted selectivity factors, have a strong preference for the amide hydrogen.

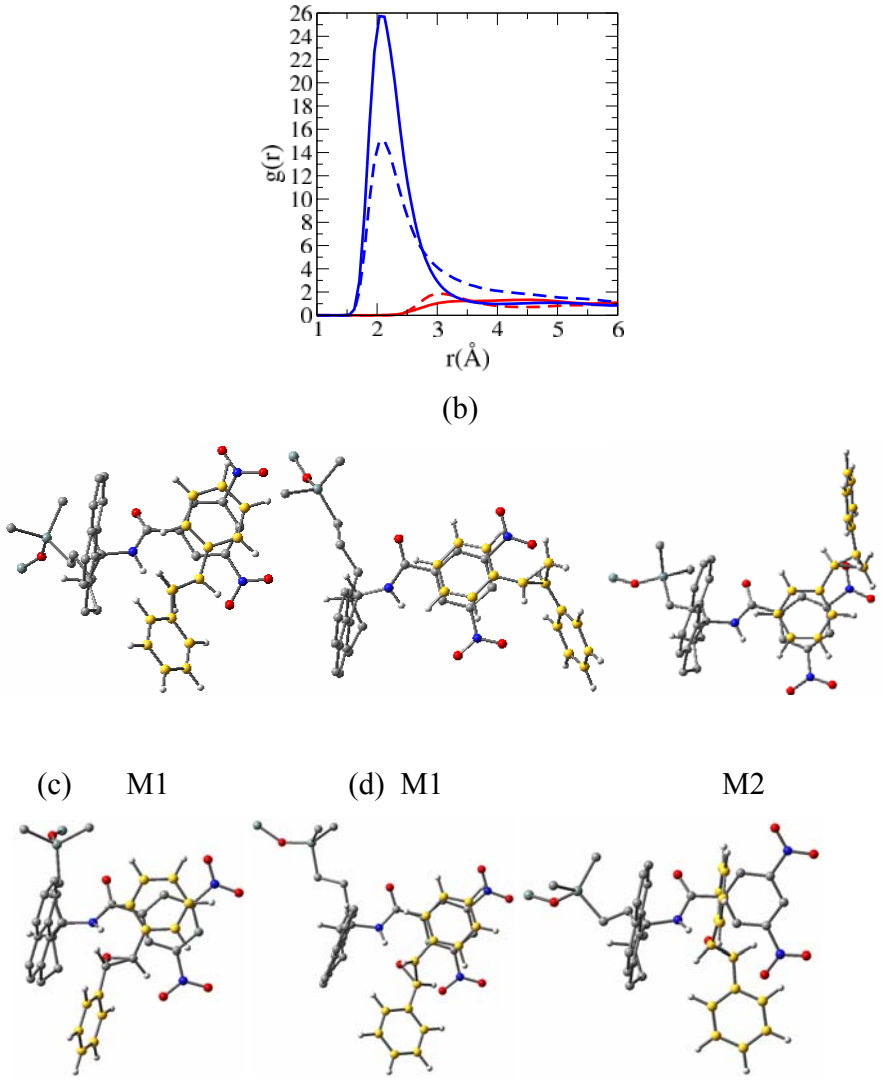


Figure 5.4: A comparison of DPCP and STBO. Radial distributions (a) between H(26) and either the epoxide oxygen of STBO (S,S: Blue solid, R,R: Blue dashed) or the unsubstituted carbon of the DPCP cyclopropane ring (S,S: Red solid, R,R: Red dashed). Snapshots of R,R-DPCP with Whelk-O1 are shown in (b). The dominant docking modes for S,S-STBO (c) and R,R-STBO (d) are shown for comparison. For clarity, analyte carbons are yellow.

Table 5.2: Hydrogen bonding and π-π stacking statistics. The interaction probabilities between specific sites on the Whelk-O1 selector and the ten analytes.

Analyte		H-bond Probability			π-π stack probability	
		Amide hydrogen [H(26)]	Amide oxygen [O(28)]	Nitro oxygen [O(36), O(37), O(39),O(40)]	π-π stack with Phenan-thryl	π-π stack with dinitro-phenyl
STYO	**S**	14.6	N/A	N/A	3.2	28.8
	R	11.2	N/A	N/A	8.6	21.2
STBO	**S,S**	11.9	N/A	N/A	9.3	30.8
	R,R	6.6	N/A	N/A	17.5	19.4
DPCP	**S,S**	N/A	N/A	N/A	6.5	29.3
	R,R	N/A	N/A	N/A	4.7	27.8
MOXR	**S**	12.0	N/A	N/A	N/A	N/A
	R	11.8	N/A	N/A	N/A	N/A
PEOL	**S**	9.0	2.7	9.5	5.9	19.5
	R	10.4	3.3	10.5	3.6	19.9
PEAM	**S**	8.0	1.9	10.8	4.4	20.1
	R	6.9	1.8	12.3	4.5	17.0
PEDO	**S**	13.2	3.7	15.1	2.6	12.1
	R	11.9	5.0	15.9	5.4	9.7
DPED	**S,S**	16.9	1.3	9.5	5.7	30.3
	R,R	14.7	1.8	9.6	4.1	29.5
DPEO	**S**	14.1	0.2	3.0	7.9	18.3
	R	18.1	0.1	1.7	5.9	26.1
PAMD	**S**	10.7	0.2	2.4	3.2	9.5
	R	7.7	0.2	3.1	4.6	4.4

The quality of the H-bond acceptor in the analyte has a well-known impact on selectivity [80]. For example, PEOL and PEAM differ in that an alcohol group is replaced by an amine, but an alcohol oxygen is a better H-bond acceptor than an amine nitrogen [239]. The oxygen of

PEOL is more likely, relative to the nitrogen of PEAM, to be near the amide hydrogen [H(26)] of the selector and this is evident from the H-bonding statistics in Table 5.2. However, the docking arrangements in Table 5.1 indicate an important shift from PEOL to PEAM. M1 is the dominant docking arrangement in the former, followed by M4 while, for the latter, M1 is less dominant with a shift towards more M4 docking events. The transition from M1 to M4 signals that the amine has a stronger preference to donate a hydrogen bond rather than accept one. The transition from M1 to M4 is accompanied by a drop in selectivity due, in part, to cancellation between the M1 and M4 probabilities for R- and S-PEAM (Table 5.1).

5.3.3 The Importance of π-π Stack on the Selectivity of Whelk-O1

The Whelk-O1 CSP has a π-electron accepting (π-acidic) aromatic region, the DNP group, while the phenanthryl group is π-electron donating (π-basic). In principle, analytes that have aromatic rings can form a π-π stack with either group and mechanisms M1 and M2 differ precisely in this regard: both have an H-bond to the amide hydrogen, but differ in the aromatic group chosen for π-π stacking. *ab initio* studies [240] indicate that the presence of an electron withdrawing group lowers the interaction energy for the parallel ring configuration. Based on this, the analytes should favor π-π stacking with the DNP group, all other things being equal. This preference is best shown by DPCP, which is free to form a π-π stack with either the dinitrophenyl ring or the phenanthryl group. Snapshot analysis reveals that roughly 30% of the DPCP form π-π stacks with the dinitrophenyl group while only about 5% stack with the phenanthryl group. This preference may be changed by other interactions such as hydrogen bonds or steric hindrance. In R,R-STBO for example, roughly 18% of the enantiomers form a π-π stack with the phenanthryl group. The probability of π-π stacking events is given in Table 5.2. The DNP ring is strongly favored by most analytes, and the stacking probability is as much as nine times higher than for the phenanthryl group.

A comparison between STYO and MOXR illustrates the overall impact of π-π stacking. Both analytes have an epoxide ring, but the aromatic ring in STYO is replaced by a methyl group in MOXR. Fig. 5.5 presents snapshots of MOXR interacting with the selector. MOXR H-bonds with H(26) but the enantiomers display a range of positions relative to the selector. For STYO, the ring-ring interactions restrict the docking arrangements so that the enantiomers are more localized.

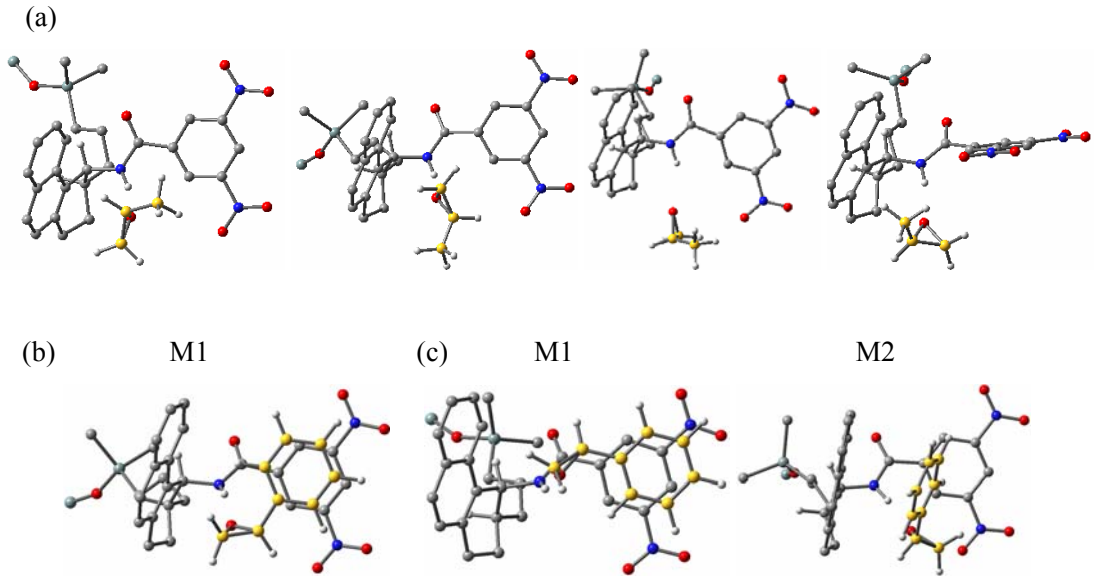


Figure 5.5: Illustrative snapshots of MOXR interacting with the chiral selector are provided in (a). The docking of S-STYO and R-STYO are shown in (b) and (c), respectively. S-STYO strongly prefers to dock according to M1 while R-STYO docks according to M1 or M2.

5.3.4 CH-π Bond-Ring Interactions and Edge-to-Face Ring-Ring Interactions

Edge-to-face ring-ring interactions have been invoked by Pirkle *et al.* as the third point of contact in the proposed three-point mechanisms [77,97]. *ab initio* studies indicate that these interactions are strongest when the ring centers are separated by 5.0-5.2 Å, but unsubstituted rings separated by 7 Å may still benefit from a 2 kJ/mol energy lowering [231]. Consider first the structural feasibility of simultaneous π-π stacking and edge-to-face interactions. Fig. 5.3 shows snapshots for S-PEOL docked according to the four mechanisms (M1-M4). In each snapshot, the aromatic ring is π-π stacked with one of the aromatic groups of Whelk-O1, and is roughly perpendicular

and within 7 Å of the other aromatic group. In other words, the formation of a π-π stack is simultaneously accompanied by an edge-to-face interaction. This result follows because of the dimensions of the cleft, the rigidity of the selector, and the nearly perpendicular positioning of the aromatic groups in the Whelk-O1 selector.

CH-π interactions have been the subject of several recent studies [232,241]. Bond-ring interaction energies depend on many factors: the orientation of the bond relative to the ring, whether the bond points towards the ring center or towards a carbon atom, and the identity of the parent molecule. These interactions decay slowly with separation and, even with the hydrogen 4 Å above the ring center, the energy may be lowered by a few kJ/mol. The possibility for bond-ring docked complex stabilization is evident from Fig. 5.3 where the M1 docking mechanism results in the α-hydrogen of PEOL pointing towards the phenanthryl group. The other docking mechanisms do not point this bond towards any of the aromatic rings (the image for M2 suggests an interaction but this is not the case). Thus, CH-π interactions may play a role in stabilizing docked complexes.

Table 5.3 presents docking interaction energies, Int E, for all ten analytes. To be precise, Int E is the average intermolecular energy, for docked analyte-selector pairs, subdivided into contributions from the four distinct regions of the Whelk-O1 selector (see Fig. 5.1). In physical terms, Int E is the energetic impetus for analyte docking. For each analyte, the lowest intermolecular energy corresponds to the most frequently docked enantiomer, except for DPED where the R,R- and S,S-energies are very close. Regardless of the docking mechanism, M1 or M2, the interaction with the tether is small since the analyte is relatively far from the tether. Consider the energies for docking according to M1. The analyte forms a π-π stack with the DNP and the energy is lowered considerably by this interaction (see Column 8 in Table 5.3). CH-π and edge-to-face interactions with the phenanthryl group occur also and these lower the energy by a significant amount (see Column 6 in Table 5.3). When the analyte forms a π-π stack with the phenanthryl group (M2 docking), the interaction energy is lowered substantially by the π-π

stacking interaction (see Column 12 of Table 5.3). However, the interaction with the DNP is significantly smaller, relative to the stabilizing contribution of the phenanthryl group for M1 docking (see Column 14 of Table 5.3). This indicates that edge-to-face π-π and CH-π interactions are less favorable for the electron-deficient DNP group. Interactions with the amide bridge always favor docking and the energies are typically lower for M1 docking. With only one exception (S-DPEO), M1 docking corresponds to a lower interaction energy (Int E) than M2 and the pivotal energy contribution is the secondary edge-to-face π-π and CH-π interactions.

The components of Int E illustrate clearly that interactions with the amide linkage, and with both aromatic regions of the Whelk-O1, provide a significant incentive to dock into the cleft region. The energetic contributions from π-π stacks, edge-to-face ring interactions, CH-π interactions, and hydrogen bonding occur *simultaneously* and *both* enantiomers benefit from them. In fact, the enantiomers differ only in the extent to which they benefit from these energy-lowering interactions, but both do achieve a significant energy lowering.

Table 5.3: A breakdown of the average energy of analyte-selector pairs with M1 and M2 docking arrangement. The "Total" energies include the intramolecular energy of the Whelk-O1 selector (Sel E) and the analyte (Ana E), and intermolecular docking energy (Int E), which is further divided into contributions from the selector phenanthryl group ("phen"), amide linkage ("amide"), dinitrophenyl group ("DNP"). Tether contributions are not listed but they are between -1.1 and -4.2, and can be extracted from the data provided. All energies are listed in kJ/mol and are averages over all the snapshots collected from the simulations.

Analyte		M1						M2					
		Total	Sel E	Ana E	Int E			Total	Sel E	Ana E	Int E		
					phen	amide	DNP				phen	amide	DNP
ST YO	**S**	**-19.7**	39.2	6.4	-28.5	-4.0	-31.3	**-1.2**	40.3	8.3	-32.6	-2.3	-13.4
	R	**-12.9**	40.6	7.8	-23.5	-5.3	-31.0	**-11.3**	40.0	5.5	-37.8	-2.3	-15.1
STBO	**S,S**	**-25.1**	38.1	12.2	-36.2	-5.2	-32.3	**-12.8**	42.9	12.5	-37.4	-4.4	-24.9
	R,R	**-5.0**	40.9	22.5	-28.5	-4.6	-32.4	**-4.5**	41.1	18.9	-39.5	-3.9	-19.2
PEOL	**S**	**-0.8**	39.0	26.5	-26.4	-5.4	-32.8	**4.2**	39.0	22.7	-35.2	-3.5	-17.3
	R	**-3.7**	38.1	25.4	-27.2	-5.2	-33.4	**11.5**	38.4	25.6	-32.3	-3.7	-15.1
PEAM	**S**	**-10.0**	38.7	18.1	-26.0	-6.5	-32.7	**5.8**	37.3	23.7	-29.1	-6.7	-18.3
	R	**-2.5**	37.9	25.9	-25.7	-6.6	-32.6	**8.6**	37.7	25.4	-30.7	-5.9	-16.7
PEDO	**S**	**-6.4**	38.2	25.5	-29.8	-5.4	-33.2	**10.2**	40.9	35.7	-37.1	-2.4	-25.2
	R	**4.3**	42.1	27.6	-25.1	-5.3	-33.1	**3.2**	40.4	24.6	-37.4	-3.4	-19.4
DPED	**S,S**	**-7.9**	38.6	27.7	-34.2	-5.2	-32.9	**5.5**	41.2	30.0	-35.8	-2.3	-25.8
	R,R	**-5.5**	39.0	29.2	-33.7	-5.0	-33.1	**4.8**	39.8	28.1	-37.7	-1.9	-21.7
DPEO	**S**	**-10.1**	40.1	21.5	-32.8	-4.2	-32.9	**-6.0**	39.5	29.1	-43.3	-3.4	-25.9
	R	**-18.3**	38.6	19.4	-36.0	-7.0	-31.6	**-6.0**	39.5	24.7	-39.3	-2.3	-26.7
PAMD	**S**	**-27.6**	40.0	27.5	-44.1	-7.2	-41.5	**-10.3**	44.8	26.4	-45.8	-7.2	-25.9
	R	**-22.4**	41.3	27.7	-35.9	-5.7	-45.6	**-8.2**	40.2	32.0	-47.7	-7.8	-22.8

5.3.5 Deficiency in the Analyte and Its Interactions

DPCP and MOXR form one primary interaction with the Whelk-O1 selector. The former π-π stacks with the DNP while the latter H-bonds to H(26). However, neither of these analytes can simultaneously form a hydrogen bond and a π-π stack and they are not well localized in the cleft region. Table 5.1 shows that the enantiomers of DPCP and MOXR have slight differences in the number of single interactions with the selector. In the case of DPCP, the simultaneous formation of two π-π stacks could give rise to some selectivity but less than 0.9% of the DPCP interacts with the selector in this manner. Thus, the selectivity is very small (but not zero) for analytes that do not include an H-bonding site and an aromatic ring.

Selectivity can also be lost when both enantiomers dock to comparable extents. PEAM is a good example of an analyte that docks well, but is not well separated. From Table 5.1, both enantiomers dock according to all four mechanisms M1-M4. This adaptability of the analyte acts against selectivity: if both enantiomers have a variety of docking options available, retention time may be high but selectivity could be low.

A high degree of conformational flexibility may also reduce selectivity. For example, PEDO has two alcohol groups that are relatively free to reorient for optimal H-bonding. As with PEAM, docking according to all four mechanisms is observed and selectivity is low.

5.3.6 Steric Hindrance and Conformational Change in the Analyte

Analytes can change their conformation to better interact with the Whelk-O1 selector and the energetic costs for this change will offset some of the energy gained by forming a docked complex. As illustrated in Table 5.3, M1 docking typically offers the largest energy benefit. However if the conformational costs associated with M1 docking are too high, then the enantiomer may dock according to a different mechanism. R,R-STBO is a good example: it prefers to dock according to M2 since M1 docking introduces significant intramolecular strain

(see Column 5 in Table 5.3). If the "costs" of docking according to M1 are not high, then M1 will be preferred with other docking arrangements occurring to a lesser extent.

A clear illustration of the impact of molecular size is provided by a comparison between STYO and STBO (Chapter 4), where a hydrogen atom in STYO has been replaced by an aromatic ring to yield STBO. Steric hindrance forces a rotation of the added ring when R,R-STBO docks according to M1. In contrast, M2 docking of R,R-STBO is not accompanied by torsional strain. As a result, R,R-STBO docks preferentially according to M2 even though R-STYO prefers M1 docking. The additional ring does not interfere with the docking of S,S-STBO [see Fig. 5.4(c)] and the docking characteristics of S-STYO and S,S-STBO are very similar.

A second example is provided by PEDO and DPED, which also differ by the replacement of a hydrogen atom with a phenyl ring. Here, the addition of the ring increases the docking probability for both enantiomers with a slight increase in the separation factor. From Table 5.1, the additional ring favors M1 docking for both enantiomers and, from Table 5.3, replacement of a hydrogen with a phenyl group leads to interactions with the phenanthryl group that are much stronger for M1 docking. Figs. 5.6(a)-(d) present snapshots for M1 and M2 docking of PEDO and DPED. It is clear from the figure that, for both enantiomers, the added ring interacts well with the phenanthryl group for M1 docking, but for M2 it is not as well positioned. This is in direct contrast to STYO and STBO, where the ring reduced the docking probability of the R enantiomer. These two comparisons show that the impact of an added functional group (aromatic ring in this case) depends primarily on the location of the group upon docking.

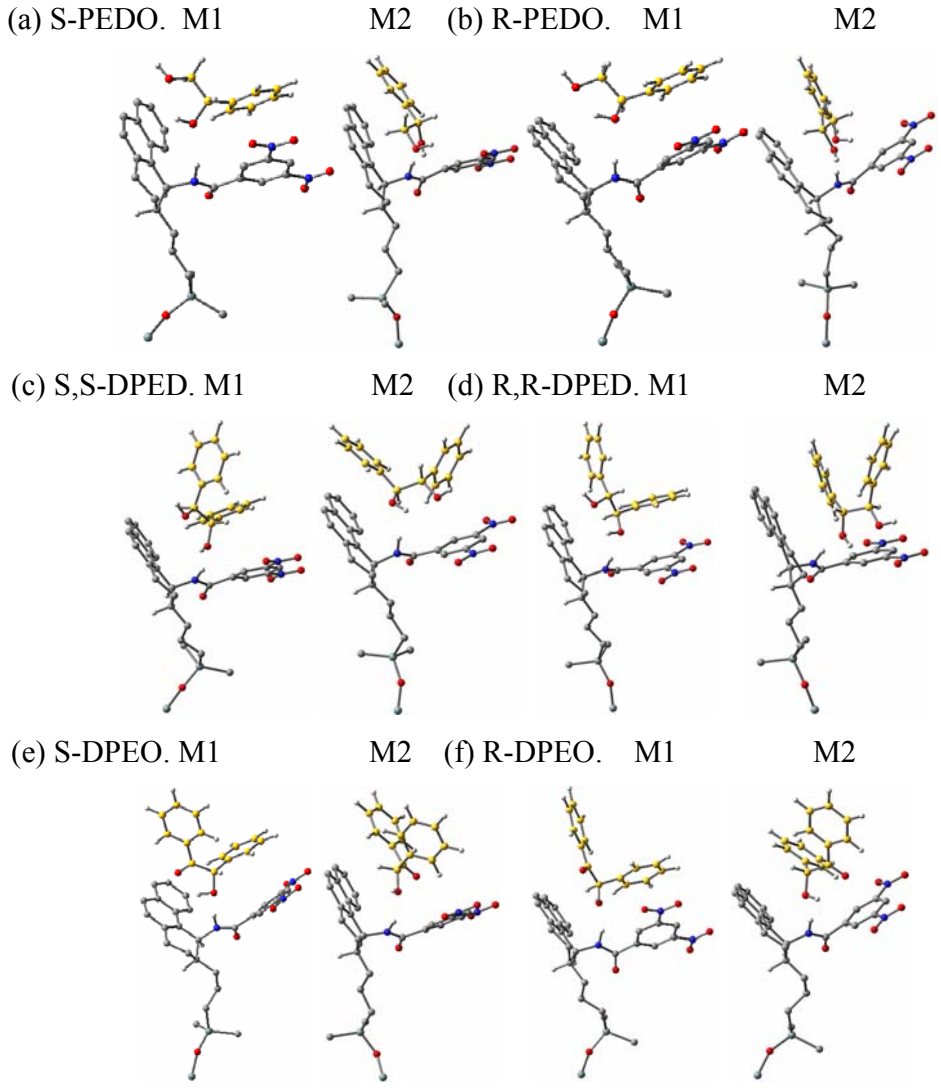


Figure 5.6: Snapshots of M1 and M2 docking of PEDO, DPED, and DPEO. *n*-hexane, end-caps, other selectors, etc. are omitted from the snapshots for clarity.

Molecular shape can also influence analyte docking. DPED and DPEO are very similar, but the replacement of an alcohol group in DPED by a carbonyl changes the molecular shape and reduces conformational freedom. The docking probabilities in Table 5.1 reveal that the impact of the bent shape is strongest for S-DPEO, and this enantiomer docks much less frequently than its analog, S,S-DPED. In fact, docking is so severely impacted that S-DPEO is predicted to elute first whereas S,S-DPED elutes last. A comparison of docking probabilities in Table 5.1 shows that the bent shape of S-DPEO discourages M1 docking and Fig 5.6(e) shows that M1 docking

forces the carbonyl oxygen into close proximity with the phenanthryl group leading to steric hindrance. DPED and DPEO include two aromatic and two H-bonding groups. There are, as a result, distinct docking modes for each analyte. Our simulations reveal that virtually all docked DPED complexes have π-π stacks formed by the analyte ring nearest the H-bonding alcohol group (see Fig. 5.6). Evidently, steric factors prevent the alternate docking arrangement. For DPEO, we find that H-bonding occurs predominantly at the alcohol oxygen.

Table 5.3 lists the total energy for docked analyte-selector pairs. This total includes the intermolecular docking energy (Int E) and the selector and the analyte energies. The latter show the conformational costs due to docking and are always positive since they are evaluated relative to the lowest energy achievable for the molecule. From Table 5.3, the energetic costs of docking differ significantly between the enantiomers of a given analyte, between the docking mechanisms, and between the analytes. It is important to note that the contribution from the solvent, other enantiomers, and from the remaining interfacial atoms are not included in the total. As a result, the total energies in Table 5.3 only provide a *rough* indication of the overall energy changes associated with docking. Nonetheless, for each analyte, the total energies are lowest for the most retained enantiomer (R-PEOL and S-PEAM for example) and correctly predict that M1 docking is preferred by these enantiomers.

The docking probabilities in Table 5.1 are extracted from snapshot analysis and they are appropriate thermodynamic averages that include all energetic and entropic contributions. It is relevant to explore the extent to which docked-pair energies can be used to predict the docking mode and separation factors. In particular, *ab initio* and force field approaches to selectivity are typically based on studies of single analyte-selector pairs [242]. Consider the use of the total energies in Table 5.3 to predict separation factors:

$$\alpha = e^{(\Delta G_A - \Delta G_B)/RT} \approx e^{(\Delta H_A - \Delta H_B)/RT} \qquad 5.1$$

where entropic contributions are assumed to be negligible. We find that substituting total pair energies into Eq. 5.1 leads to a significant overestimation of the separation factors. Thus the most

probable docking arrangements and elution orders can generally be predicted from pair energies but selectivity factors are not well predicted using energies alone.

5.3.7 Analyte Docking Locations and Conformational Changes in the Selector

It has been generally assumed [55,57] that the more retained enantiomer docks inside of the cleft while the less retained enantiomer docks outside of the cleft. In this regard, "inside-of-cleft" refers to the spatial region defined by the phenanthryl group, the amide linkage, and the DNP group. "Outside-of-cleft" encompasses all other areas around the selector. The docking probabilities in Table 5.1 and the docking energies in Table 5.3 appear to contradict this view: both enantiomers interact well with the cleft-defining elements of the Whelk-O1 selector. In this section, we explore the preferred docking locations of each analyte for M1 docking events. The Whelk-O1 selector is partially flexible and may accommodate the analyte by changing its shape. As a result, it is important to characterize the analyte location in terms of the position of the DNP ring in the selector. Therefore, two angles are employed: the torsional angle about the C(27)-C(29) [see Fig. 2.6(b) for atom numbering] bond and the angle between the vector normal to the DNP ring and the center-to-center vector joining the DNP ring and the analyte ring. The former angle determines the orientation of the DNP ring relative to the remainder of the selector while the latter determines the relative position of the analyte and selector rings. Fig. 5.7 illustrates the scheme used to identify analyte docking location from the snapshots. The locations are identified as inside of the cleft, on the side of the cleft, or under the cleft.

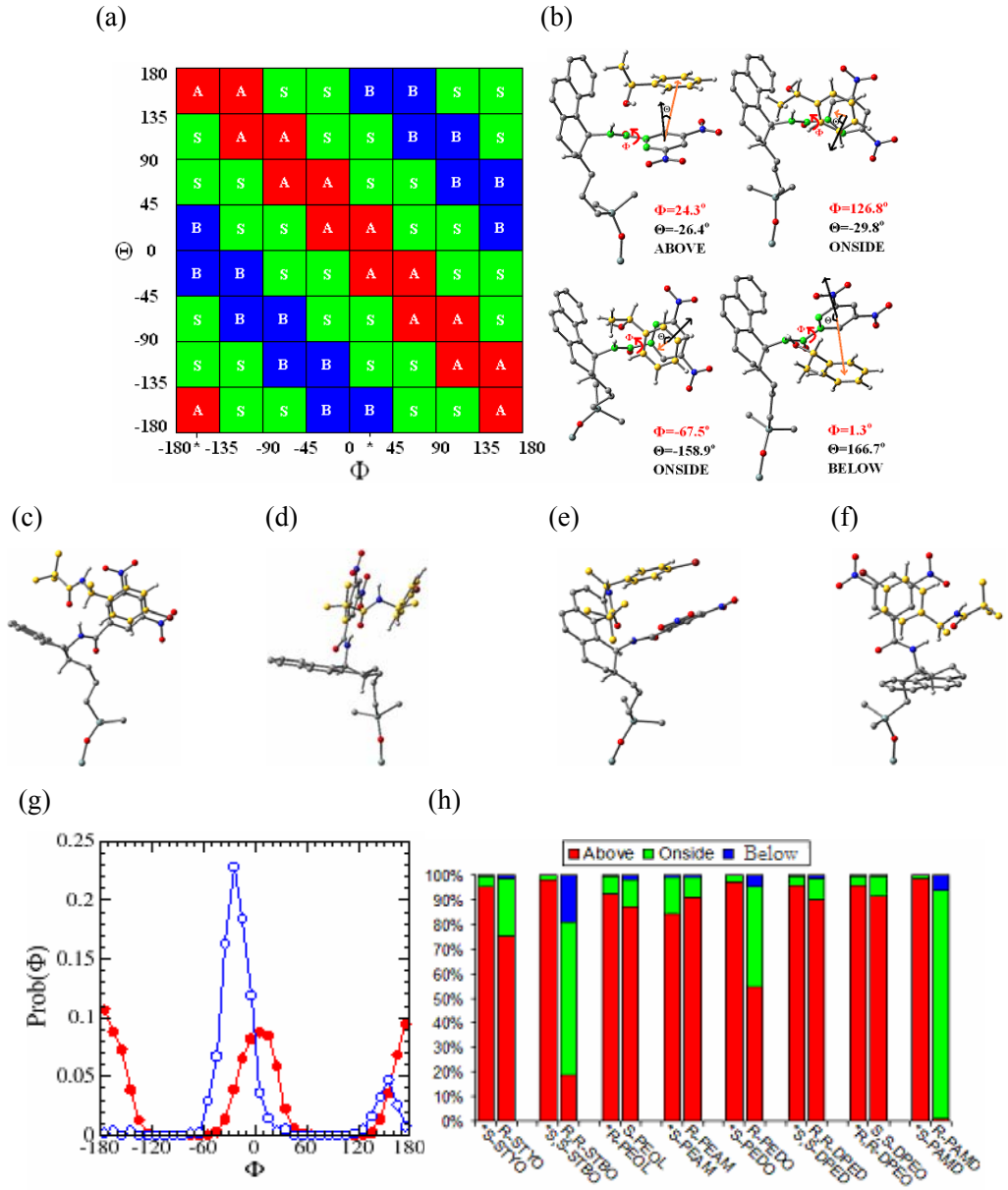


Figure 5.7: Analyte location for M1 docking events. Three locations are identified (a): "S" for on the side of the cleft (Green), "A" for above the cleft (Red), and "B" for below the cleft (Blue). Two angles are used to identify the location: the N(25)-C(27)-C(29)-C(30) dihedral (Φ), and the angle (Θ) between the vector normal to the DNP ring and the vector joining the DNP ring center to the analyte ring center. Snapshots of docked S-PEOL are shown in (b) to illustrate the criteria. Two views of R-PAMD, which docks ONSIDE, are shown in (c) and (d). Two views of S-PAMD, which docks ABOVE, are shown in (e) and (f). The probability distribution for Φ when S-PAMD (red) and R-PAMD (blue) are docked is shown in (g). Analyte location probabilities for eight analytes are given in (h).

Koscho *et al.*[55] have co-crystallized a Whelk-O1 derivative with the enantiomers of PAMD. Both homochiral and heterochiral crystal structures were consistent with the M1 docking mechanism but, in the homochiral complex, the analyte docked inside of the cleft, while in the heterochiral complex the analyte docked outside of the cleft. For the latter, a large deviation from planarity was noticed for two dihedral angles and these indicate a significant reorientation of the DNP ring. With the classification scheme in Fig. 5.7, the heterochiral complex would be identified as on-the-side. Snapshots of M1 docked PAMD are shown in Figs. 5.7(c)-(f). The docked structures from the simulations are consistent with the crystal structures: S-PAMD docks inside of the cleft while R-PAMD docks on the side of the cleft.

Docking location probabilities are compared in Fig. 5.7(h). 98.6% of the S-PAMD dock inside of the cleft, while 93.1% of R-PAMD dock on the side of the cleft. Under-the-cleft docking is uncommon for all analytes since the presence of the end-caps leaves very little room. Most analytes prefer inside-of-cleft docking for both enantiomers. This contradicts the assumption that only one enantiomer "fits" into the cleft region. For these analytes, selectivity follows from the quality of the fit, as reflected by the energies in Table 5.3. Side-of-cleft docking is observed for R-PAMD, R-STYO, R,R-STBO, and R-PEDO. When an analyte is docked on the side-of-cleft, the DNP ring in the selector rotates to form a π-π stack. Fig. 5.7(g) shows the probability distribution for the C(27)-C(29) torsional angle when PAMD is docked according to the M1 mechanism. The distribution for S-PAMD is consistent with inside-of-cleft binding while the angular distribution for R-PAMD is shifted and reflects the significant reorientation of the DNP required for on-the-side docking. Consistent with this, the Sel E in Table 5.3 are notably higher for analytes that show side-of-cleft M1 docking. With a π-π stack formed from the side, the analyte ring is farther from the phenanthryl group than is typical for inside-of-cleft docking, and the energy-lowering secondary interactions are reduced. The Int E between the phenanthryl group and the analyte in Column 6 of Table 5.3 confirm this, since they are consistently smaller for

side-of-cleft docking enantiomers. Nonetheless, on-the-side docking may proceed with a docking energy that is still quite favorable.

5.4 Conclusion

The Whelk-O1 chiral stationary phase forms the basis for this MD simulation study of selectivity. Ten analytes have been selected with the objective of identifying the factors that contribute to selectivity. We have explored a number of factors: the importance of H-bonding; the H-bonding ability of the analyte; π-π stacking and substituent effects; edge-to-face ring interactions; CH-π interactions; competition between docking modes; the energetic balance at the heart of the recognition process; steric and conformational costs in the analyte and in the selector; and, finally, the analyte position relative to the Whelk-O1 selector.

A detailed analysis of the interaction energy between the analyte and the selector shows that important secondary interactions, such as edge-to-face ring-ring and CH-π interactions, occur upon analyte docking. These interactions occur simultaneously with the H-bonding and π-π stacking that are essential to the docking process. Both enantiomers are found to benefit from these interactions but they differ in the energy lowering achieved. This observation that both enantiomers interact strongly and favourably with the selector confirms that recognition on Whelk-O1 cannot be viewed as a competition between one enantiomer that "fits" into the cleft and the other which does not. In fact, both "fit" but one can achieve lower interaction energy with the selector.

An analysis of analyte-selector complexes identifies four distinct docking arrangements: The most frequent docking arrangement, M1, consists of an H-bond with the amide H and a π-π stack with the DNP; other docking arrangements include H-bonds with the amide oxygen (M3), H-bonds with the nitro oxygens (M4), or π-π stacking with the phenanthryl group (M2). A comparison of overall docking frequencies for both enantiomers leads to predicted selectivity factors in good agreement with experiment. The M1 docking mechanism is consistent with

proposed selection mechanisms [77,96] for Whelk-O1 but the other three mechanisms have not been invoked in these mechanisms. In fact only M2 docking has been examined previously (Chapter 4). Overall, we find that M1 docking offers the most energy benefit since stabilizing interactions with the amide linkage, the dinitrophenyl, and the phenanthryl groups are significant. However, for some analytes, torsional strain on the analyte discourages M1 docking and leads to other modes of interaction. M2 docking, where the analyte forms an H-bond with the amide H and a π-π stack with DNP, is favoured by several analytes including R,R-STBO and R-PEDO. M4 is more probable for analytes that include a hydrogen bond donor, such as PEAM, while M3 docking is uncommon.

As expected, simulations predict negligible selectivity for analytes that lack either an H-bond acceptor or an aromatic ring. These analytes are found to interact favourably with the selector but they adopt a broad range of positions relative to the selector. Selectivity can also be lost when both enantiomers dock to equal extents. PEAM is a good example of an analyte that is not well discriminated by Whelk-O1 despite significant docking. Here, both enantiomers exhibit all four docking modes to some extent. A competition between H-bond donating and H-bond accepting character, as exemplified by PEAM, also reduces selectivity.

The Whelk-O1 selector retains its cleft-like shape in the fluid and, in general, both enantiomers tend to dock in the cleft. The selector is not completely rigid, however, and the energetic costs of reorienting the dinitrophenyl ring are fairly minimal. This leads to alternative docking locations and, with a twist about the C(27)-C(29) [see Fig. 2.6(b) for numbering] bond, analyte can approach the cleft from the side. An H-bond with the amide hydrogen and a π-π stack with the DNP can then form. This “on the side” docking is particularly noteworthy for R-PAMD and R,R-STBO.

Chapter 6
Rational Optimization of the Whelk-O1 Chiral Stationary Phase

6.1 Introduction

CSP design and optimization [35,41,58,61,74,99] has been an important area in chiral chromatography research, and it is now getting a stronger impetus from the drug industry. Specifically, cost is a major concern [24] and it is therefore essential to develop high loading capacity, high throughput, environmentally friendly, targeted CSPs to be used in chiral separations [24].

One way to improve chiral separations is to modify known CSPs for better enantioselectivity. Numerous examples have shown that it is possible to improve CSPs in this way. As noted in Chapter 1, Pirkle and Welch [40] improved the enantioselectivity of Whelk-O1 by shortening the hydrocarbon tether that joins the selector to the underlying silica bead. Mandl *et al*. [243] showed that the introduction of a carbamate group as a substituent on the quinine moiety broadened the scope of applicability of a quinine-derived CSP. Francotte [24] prepared a cellulose tris(3-chloro-4-methylphenyl-carbamate) based CSP and efficiently separated the LTD4 antagonist iralukast which is only poorly resolved on cellulose tris(3,5-dimethyl-phenyl-carbamate) based CSP. These examples of CSP design and optimization, however, depend on empirical reasoning and involve a lengthy trial-and-error process. Faster and more effective design and screening methods are therefore desirable. As with computer-aided drug design [244], computational chemistry provides a viable option.

The rational optimization of a CSP involves two important steps: First, the elucidation of the chiral separation mechanism of the original CSP; Second, the design, optimization and evaluation of new CSPs. Computational chemistry can be a useful tool in both steps. First, as discussed in Chapter 1, computational chemistry methods have been applied on various CSPs to

achieve an atomistic understanding of the chiral separation mechanism. An atomic level understanding of the enantioselective mechanism provides the possibility of subtle modification of the CSP to achieve better enantioselectivity. Computer experiments also enable *in silico* building and evaluation of modified CSPs and this may accelerate the CSP design and optimization process. As an early example, Däppen *et al*. [99] applied molecular mechanics to optimize the (R)-2,2-dimethyl-N-(1-(1-(4-amino)-naphthalenyl)-ethyl)-propanamide based CSP. They first searched for selector-analyte complexes with favorable docking enthalpies. Based on the 3D-structures of these complexes, an amine substituent on the naphthyl group of the original selector was introduced to allow an extra hydrogen bond for the most favorable enantiomer. The differential enthalpy of the new selector-analyte pair was thus increased and they predicted that the modified CSP would be superior to the original one. However, the modified CSP has not been synthesized. To our knowledge, no *in silico* optimized CSPs have been synthesized and tested to date. One reason is that there have been few examples of computations that are able to reproduce experimental separation factors. Our positive results in Chapters 4 and 5 on using MD simulations to predict separation factors and to study the selection mechanism of the Whelk-O1 CSP, makes it natural to proceed further and attempt to rationally optimize the Whelk-O1 CSP. Optimization of the Whelk-O1 CSP may increase the resolving power for targeted analytes. In addition, modified CSPs may be selective for new classes of analytes. We focus here on the impact of relatively minor structural modifications designed to enhance selectivity of target analytes.

In Chapter 5, Simulations of the docking of ten analytes on Whelk-O1 showed that the more retained enantiomers have one dominant docking arrangement: these analytes dock inside of the cleft, forming a hydrogen bond with the amide hydrogen and a π-π stack with the dinitrophenyl group. The less retained enantiomers are found to adopt multiple docking arrangements. One common alternative arrangement, M2 docking, involves a π-π stack with the phenanthryl group instead of with the dinitrophenyl group. Another common alternate

arrangement is side-of-cleft docking where the DNP ring twists to allow the enantiomers to form the hydrogen bond and the π-π stack from the side of the cleft. Intuitively, enantioselectivity can be improved by either facilitating the docking of the more retained analyte or restricting the docking of the less retained analyte.

The modified CSP selectors are shown in Fig. 6.1. The first modified selector, yielding CSP1, consists of saturating the top-most ring of the phenanthryl group. This modification is expected to reduce the probability for M2 docking, where ring-ring interactions form with the phenanthryl group. Docking on CSP1 is examined for styrene oxide and stilbene oxide. An analysis of docking modes for these two analytes in Chapter 4 on Whelk-O1 showed that the less retained enantiomers preferred M2 docking. Thus, the less retained enantiomers should be more strongly impacted by this modification. The second modified stationary phase, CSP2, is designed to restrict side-of-cleft docking by constraining the flexibility of the amide linkage. Simulations in Chapter 5 revealed that the least-retained enantiomers of stilbene oxide and 1-phenylethane-1,2-diol have side-of-cleft docking. We examine the docking of these two analytes on CSP2 where we expect to improve selectivity by inhibiting the docking of the least retained enantiomer. Unlike the first two modifications, CSP3 is designed to introduce a motif that had been demonstrated to be effective in chiral discrimination. Specifically, the amide linkage of Whelk-O1 selector has been reversed in CSP3. Thus, we get an N-3,5-dinitrophenyl amide from the original 3,5-dinitrobenzamide group. N-phenyl amide is widely used in cellulose tris(3,5-dimethyl-phenylcarbamate) and amylose tris(3,5-dimethyl-phenylcarbamate) based CSPs, the two most successful polysaccharide CSPs in both analytical and preparative applications. These two selectors both have an N-phenyl amide group. Studies have shown that this group plays an important role in the chiral recognition mechanism of polysaccharide CSPs through hydrogen bonding and ring-ring interactions [41,46-48]. A similar, though not identical, motif is also present in the ULMO CSP [245] which is known to be better than Whelk-O1 at separating certain chiral alcohols. Cellulose tris(3,5-dimethyl-phenylcarbamate) and amylose tris(3,5-dimethyl-

phenylcarbamate) based CSPs are also superior to Whelk-O1 for the separation of many alcohols [234]. With these observations in mind, we have tested CSP3 for 1-phenylethanol, a small chiral alcohol. For comparison, we have also examined the selectivity of CSP3 for styrene oxide.

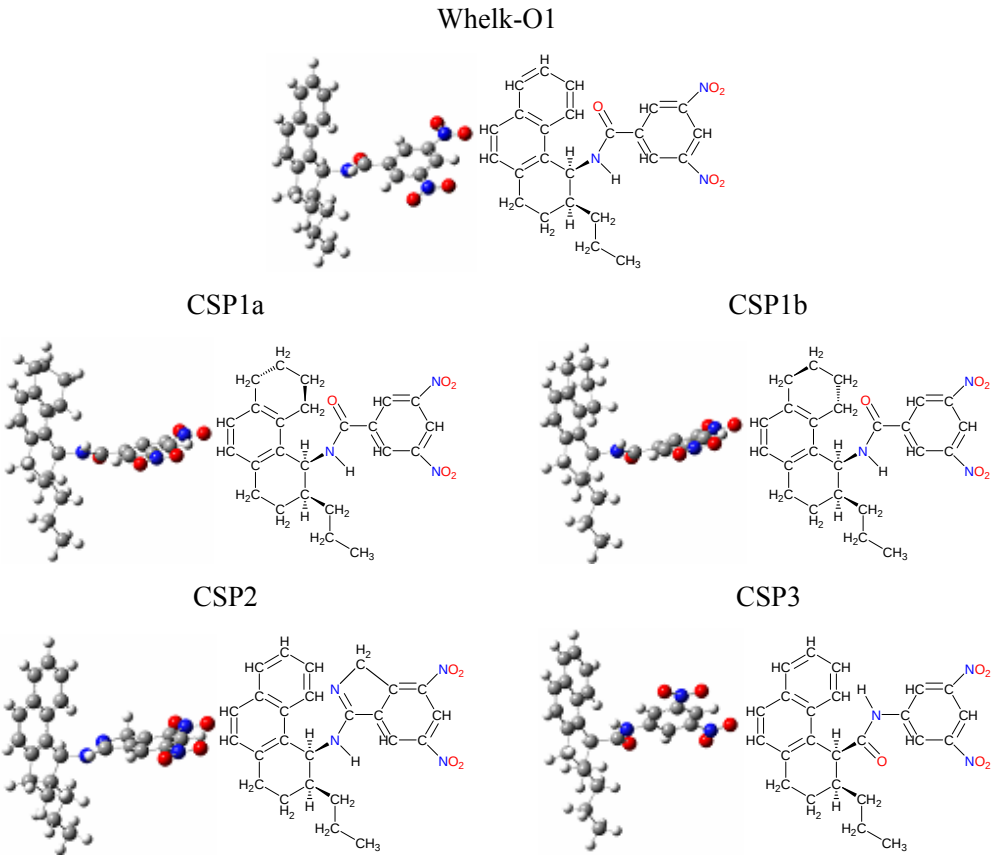


Figure 6.1: Lowest energy 3D structures, obtained from DFT calculations, and chemical structures of the selectors of the Whelk-O1 CSP and the modified CSPs. For CSP1, the two nearly-degenerate conformations of the saturated ring are shown.

The remainder of this chapter is organized as follows. Theoretical details are described in Section 6.2. This includes a brief discussion of the molecular models for the selectors and the analytes, a discussion of the chiral interface, and simulation details. The evaluation of the three modified CSPs is presented and discussed in Section 6.3. Each modified CSP is considered in turn for the targeted analytes. Conclusions are presented in Section 6.4.

6.2 Theoretical Details

As discussed in Chapter 2, the model chiral interface used in the simulations is representative of the pore environment of the Whelk-O1 CSP. The interface includes 16 chiral selectors, 48 trimethylsilyl end-caps, and 64 silanol groups corresponding to the range of experimental coverages [40,188,189]. A stationary layer of 128 Si atoms, to which end-caps, silanol groups, and selectors are attached, is included in the simulations to represent the underlying substrate. The simulation cell, similar to the ones in Chapters 2-5 includes two chiral interfaces, with solvent and racemate in between. Empty space is also included above and below the two surfaces.

Consistent to Chapter 5, an *n*-hexane solvent has been chosen for the simulations by appeal to its ubiquitous use as the major solvent in normal phase chromatography [77]. The TraPPE-UA [211] model for *n*-hexane has been selected in accordance with solvation and selectivity studies for Whelk-O1 (Chapters 2-5) . Within the simulations, the fluid between the surfaces includes 450-470 *n*-hexane molecules. We have verified that the hexane density in the center of the simulation cell is roughly consistent with the experimental hexane density [214] at 273K.

Extensive molecular dynamics simulations, as described in Chapter 5, have been performed to study analyte resolution on the CSPs. Each simulation includes 16 analytes: 8 of each enantiomer. For each analyte and CSP combination, between twenty and forty MD simulations have been performed: this corresponds to a total simulation time of 23-47 ns. The lower limit is appropriate for analytes that do not interact strongly with the CSPs while the upper limit applies to strongly interacting analytes. The first 100,000 time steps are used for equilibration. Snapshots are then saved every 2,000 iterations. The racemate, solvent, and selector distributions are assessed every 20 iterations to provide atomic distributions at the interface. All simulations are performed with the MDMC program [205].

Conformational change in the analyte and the selector can occur upon docking. It is, as a result, essential that intramolecular flexibility be well represented within the simulations. A

suitable semi-flexible model for the Whelk-O1 selector has been developed and discussed in Chapter 2. Potentials representing the end-caps and silanol groups have also been discussed in the same chapter. Full details for the models of the modified CSPs are provided in Appendix A. Briefly, an extensive series of B3LYP/6-31G* calculations have been performed to identify major conformers and to assess the energy changes required for bond stretching, angle bending, torsions, and out-of-plane bending (improper torsion). Based on these results, potential parameters are fitted to empirical forms, as shown in the Supplemental Materials.

The minimum energy structures for the four analytes of interest are provided in Fig. 5.1. Semi-flexible intramolecular models have been previously developed in Chapters 4 and 5 for styrene oxide (STYO), stilbene oxide (STBO), 1-phenylethane-1,2-diol (PEDO), and 1-phenylethanol (PEOL). Full details are provided in Appendix A.

6.3 Results and Discussions

6.3.1 Overview

The probabilities for "docked" analytes were counted from the simulation snapshots and, based on these probabilities, the separation factors were calculated according to Eq. 4.1 as discussed in Chapters 4 and 5. This approach has been able to predict the correct elution orders and reasonable separation factors for ten diverse analytes on the Whelk-O1 CSP. We apply it here for the modified CSPs. The first five columns of Table 6.1 summarize the docking counts. Specifically, for each snapshot, the analytes are examined to see whether they form zero, one, two, or more interactions with Whelk-O1 selectors. For this analysis, an interaction consists of a hydrogen bond or a π-π stack as assessed by the geometric criteria specified above. Typically, few analytes will simultaneously form multiple π-π stacks or H-bonds, and interactions with multiple selectors are uncommon, so that the docking probability is nearly equal to the probability for two interactions.

Table 6.1: Interaction summary for the modified CSPs and Whelk-O1 extracted from MD snapshot analysis. The first four columns give the percentage of analytes that have 0-3 interactions with the CSP. A "docked" interaction occurs when an H-bond and a π-π stack are present with a selector simultaneously. The central column compares the calculated and experimental (in brackets) separation factors. The last five columns provide a breakdown of the docking probability according to the docking mode, M1-M4. M1 docking is further subdivided according to the analyte location: "INSIDE" for inside-of-cleft, and "SIDE" for side-of-cleft.

		Interaction Frequency					α^{calc} (α^{exp})	**Docking Arrangement**				
		0	**1**	**2**	**3**	**Docked**		**M1**		**M2**	**M3**	**M4**
								Inside	**Side**			
Whelk												
STYO	**S**	63.6	26.3	10.1	<0.1	**9.9**	1.3 (1.37[77])	9.3	0.5	0.1	N/A	N/A
	R	66.8	25.4	7.8	<0.1	**7.7**		2.8	0.9	3.9	N/A	N/A
STBO	**S,S**	62.5	23.5	13.0	1.0	**9.6**	1.8 (2.00[77])	8.5	0.2	0.1	N/A	N/A
	R,R	65.6	24.7	9.1	0.5	**5.7**		0.2	0.8	4.5	N/A	N/A
PEOL	**S**	63.8	26.5	9.0	0.7	**7.0**	1.2 (1.00[234],1.68[235])	2.9	0.4	1.0	0.7	1.3
	R	63.7	25.8	9.9	0.7	**8.3**		4.2	0.4	0.1	0.8	2.2
PEDO	**S**	65.7	23.0	10.0	1.4	**5.5**	1.1 (1.08[77])	3.8	0.1	0.3	0.6	0.5
	R	64.0	25.0	9.3	1.7	**5.0**		0.7	0.5	2.2	0.7	0.5
CSP1a												
STYO	**S**	64.3	19.4	16.3	0	**16.2**	1.7	16.0		0.1	N/A	N/A
	R	70.5	19.5	10.0	0	**10.0**		7.5		2.5	N/A	N/A
STBO	**S,S**	77.6	13.4	8.6	0.4	**8.5**	5.7	8.0	0.1	<0.1	N/A	N/A
	R,R	82.4	15.7	1.9	<0.1	**1.6**		0.2	<0.1	1.3	N/A	N/A
CSP1b												
STYO	**S**	76.1	16.4	7.5	0	**7.4**	1.7	7.0	0.2	0.1	N/A	N/A
	R	73.8	21.7	4.5	0	**4.4**		2.7	0.2	0.1	N/A	N/A
STBO	**S,S**	73.7	16.2	10.0	1.1	**7.8**	14.0	6.7	0.1	0.1	N/A	N/A
	R,R	81.8	17.3	0.9	<0.1	**0.6**		0.2	<0.1	0.4	N/A	N/A
CSP2												
STBO	**S,S**	75.3	19.6	4.9	0.3	**1.5**	2.7	1.0	0.1	0.1	N/A	N/A
	R,R	75.9	17.2	6.7	0.2	**3.9**		0.1	<0.1	3.7	N/A	N/A
PEDO	**S**	82.4	14.4	3.1	0.2	**1.2**	2.2	0.8	<0.1	0.1	<0.1	0.2
	R	78.5	15.6	5.0	0.9	**2.6**		0.4	<0.1	1.3	<0.1	0.2
CSP3												
PEOL	**S**	68.4	25.4	6.0	0.2	**3.5**	1.7	0.4	<0.1	0.1	0.9	2.2
	R	67.7	24.7	7.4	0.2	**5.7**		0.9	<0.1	<0.1	0.5	4.5
STYO	**S**	76.5	21.9	1.7	0	**1.6**	1.6	1.5	0.2	<0.1	N/A	N/A
	R	76.8	22.1	1.1	0	**1.0**		0.7	0.2	0.2	N/A	N/A

The Whelk-O1 selector, and the modified selectors, consist of two aromatic regions: the dinitrophenyl group and the phenanthryl moiety. The selectors also have three potential hydrogen bonding sites: the amide hydrogen, the amide oxygen (nitrogen for CSP3), and the nitro oxygens. In principle, there are six possible combinations for an analyte to be regarded as "docked". However, two of these are only rarely observed. The probabilities for the remaining four docking arrangements, M1-M4, are given in Table 6.1. From simulation studies on Whelk-O1 in Chapters 4 and 5, the preferred docking arrangement for the more retained enantiomer involves a hydrogen bond with the amide hydrogen and a π-π stack with the dinitrophenyl group. In this chapter, M1 docking refers to this mechanism. The M1 docking mechanism is consistent with proposed, experiment-based, three-point interaction mechanisms [77,96] for Whelk-O1. For some less retained analytes, we have found that torsional strain discourages M1 docking and leads to other modes of interaction. M2 docking, where the analyte forms a hydrogen bond with the amide hydrogen and a π-π interaction with the phenanthryl moiety, is favoured by several less retained analytes including R-styrene oxide and (R,R)-stilbene oxide. Docking through hydrogen bonding with the amide oxygen (M3 docking) or nitro oxygen (M4 docking) is observed when the analyte has a hydrogen bond donor.

M1 docking typically occurs with the analyte inside of the cleft region. However, for some analytes, the dinitrophenyl ring can twist to allow analyte to approach from the side. In Table 6.1, the M1 docking probabilities are subdivided according to analyte location. Fig. 6.1 schematically illustrates the M1-M4 docking modes, along with inside-of-cleft and side-of-cleft M1 docking.

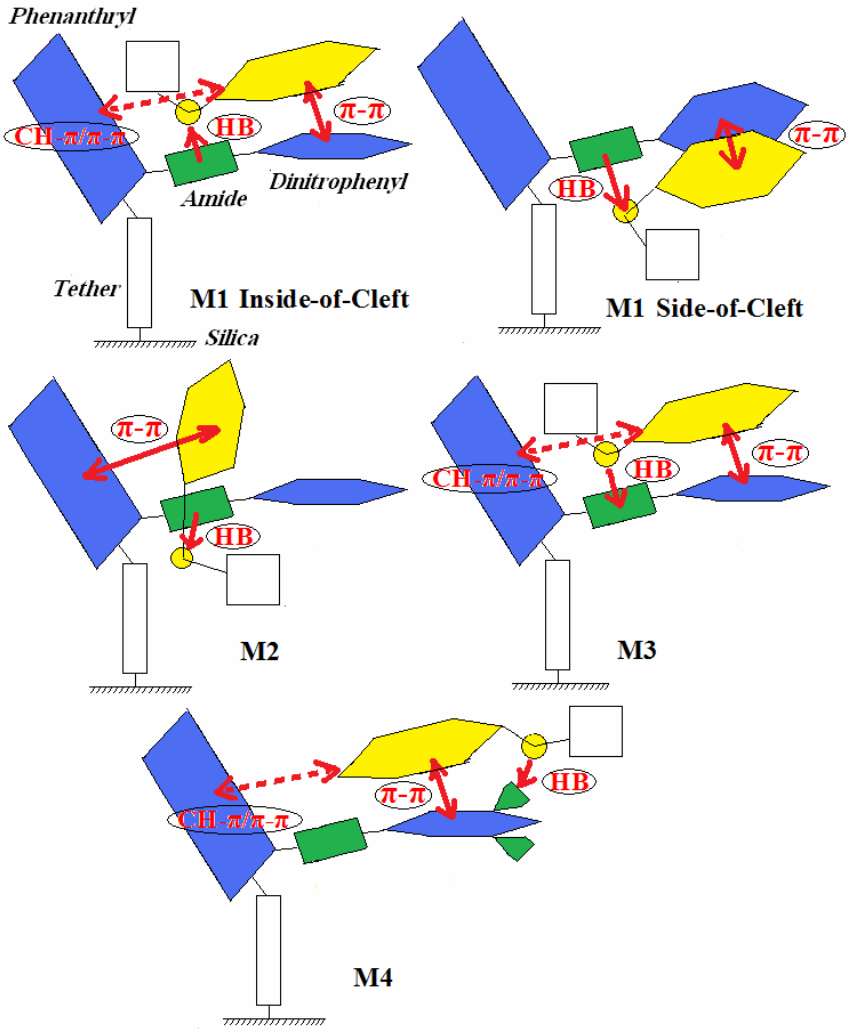


Figure 6.2: Illustration of docking arrangements. The analyte is composed of an aromatic ring (yellow hexagon), a hydrogen bonding site (yellow circle), and other structural elements(white square). The selector is composed of a tether (white rectangle) and two aromatic regions - phenanthryl group (blue rectangle) and a dinitrophenyl group (green corns for the nitro groups and blue hexagon for the phenyl group) - connected by an amide bridge (green rectangle). The four docking arrangements, M1-M4, on Whelk-O1 are shown. In addition, M1 docking is subdivided according to analyte location. Specifically, inside-of-cleft and side-of-cleft docking is shown above. Red arrows identify hydrogen bonding and π-π stacking interactions. Red dotted arrowd identify secondary CH-π or edge-to-face π-π interaction.

Changes to the Whelk-O1 selector potentially alter the surface atomic distributions. Fig. 6.3 compares the surface distributions for selected atoms after the Whelk-O1 CSP is modified to yield CSP1, CSP2, and CSP3. In all cases, the modification encourages selector positions where the DNP ring lies roughly parallel to the surface and the phenanthryl group points toward the

bulk. For Whelk-O1 in *n*-hexane, the selector also adopts an alternate arrangement whereby the phenanthryl is roughly parallel to the surface and the DNP points towards the bulk. All three modified CSPs de-emphasize this latter arrangement at the surface. The impact of this configurational shift is likely minimal since the cleft-like structure is always maintained and oriented towards the bulk. Nonetheless, the different presentations may alter the probability for outside-of-cleft docking which could have a small impact on selectivity.

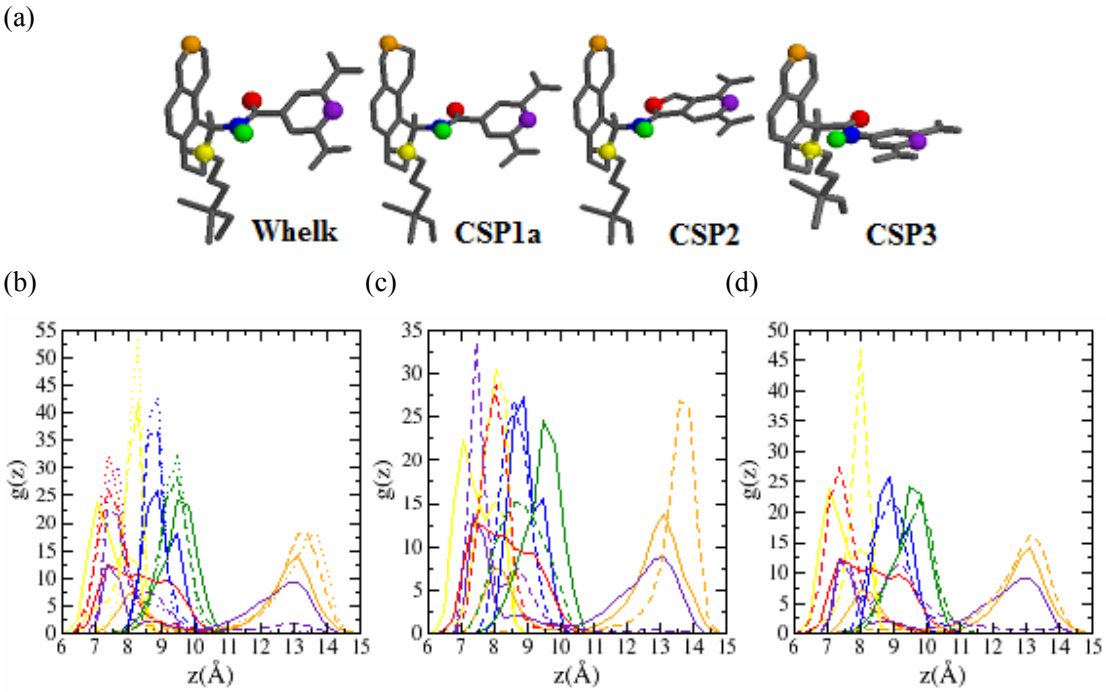


Figure 6.3: Surface atomic distributions for selected atoms before and after selector modification. The atoms being compared are shown in (a). The atom colors correspond to the curve colors, and the Whelk-O1 results are represented by solid lines while the modified CSP distributions are given by dashed lines. In (b), atomic distributions for Whelk-O1, CSP1a (dotted), and CSP1b (dashed) are compared. A comparison of CSP2 and Whelk-O1 is given in (d) while CSP3 and Whelk-O1 are compared in (d).

6.3.2 CSP1

CSP1 is designed to hamper the docking of the least retained enantiomer. Specifically, the CSP1 selector is identical to Whelk-O1 except that the topmost phenanthryl ring is saturated (Fig. 6.1). This modification should reduce the probability for M2 docking, since this docking mode

involves a π-π interaction with the phenanthryl moiety. The latter includes two aromatic rings and one might assume a reduction of 50%. However, we have found (Chapters 4 and 5) that in all the M2 docking arrangements, the top ring of the phenanthryl moiety is the major ring responsible for π-π stacking in M2 docking events. Take (R)-Styrene oxide for example: 90% of M2 docking events have a π-π stack with the topmost ring compared to 71% for the other ring. Note that analyte can simultaneously stack with both rings due to their proximity, although we count this as a single interaction in Table 6.1. Thus, saturation of the top ring should effectively reduce M2 docking.

Saturation of the lower phenanthryl ring, or to both rings, would seem to provide a viable alternative to CSP1. These alternate CSPs would impact M2 docking, but would also significantly alter other docking modes since the shape of the Whelk-O1 cleft would be impacted. In particular, the overall shape and rigidity of the phenanthryl moiety would be altered. CSP1 preserves the roughly perpendicular arrangements of the DNP and phenanthryl groups but two conformational minima are obtained due to flexibility in the saturated topmost ring. The energy difference between the two minima is 0.3kJ/mol, which is well within the error of B3LYP/6-311++G** DFT calculations. Studies on conformations and torsional barriers for cyclohexene [246] indicate that interconversion requires 22 kJ/mol. Therefore, we have proceeded by examining the individual conformers, now identified as CSP1a and CSP1b and shown in Fig. 6.1.

The differences between the two conformers extend beyond steric concerns in the cleft region. Specifically, CSP1a places a ring hydrogen close (within 2.6 Å) to the amide nitrogen while CSP1b has two ring hydrogens directed towards the amide nitrogen. The proximity of these hydrogens is sufficient to alter the charge density within the amide linkage. Indirectly, the aromatic regions of the selector are also affected. With these observations in mind, independent models for CSP1a and CSP1b have been derived: stretches, bends, torsions, and atomic charges are distinct for the conformers. In most cases, the parameters are very similar. One important

exception is the atomic charges within the amide group. Here, both the amide H and the amide O bear more positive (less negative for O) charge for CSP1a (Appendix A).

Styrene oxide (STYO) and stilbene oxide (STBO) are chosen to evaluate CSP1. On Whelk-O1, the less retained enantiomers of these analytes prefer M2 docking, as shown in Table 6.1. Docking characteristics for CSP1a and CSP1b are also given in Table 6.1 and, as expected, the saturation of the topmost ring reduces M2 docking probability. In agreement with this, Table 6.2 indicates a reduction in π-π stacking probability with the phenanthryl group. Simulations predict that CSP1 is more selective than Whelk-O1: For STYO, the selectivity factor has increased from 1.3 to 1.7 but for STBO the increase is much more substantial.

Consider the docking of STYO on CSP1a and CSP1b. The overall probability for the former is significantly higher than for the latter (fifth column in Table 6.1). This follows, in part, due to the small size of the analyte and the increased charge on the amide hydrogen for CSP1a, which encourages H-bonding to the amide hydrogen. An increase in H-bonding is evident from the probabilities for STYO on CSP1a in Table 6.2. Despite the difference in docking probability between the two conformers, both CSP1a and CSP1b have similar docking modes: S-STYO and R-STYO prefer M1 docking, and M2 docking has sharply decreased due to the saturation of the top-most ring.

For STBO, the conformers CSP1a and CSP1b have a similar impact: Both conformers have reduced docking probabilities relative to Whelk-O1 and this is particularly true for R,R-STBO; M1 docking is strongly preferred by S,S-STBO and, despite a sharp reduction in probability, M2 docking remains the most probable docking mode for R,R-STBO.

Table 6.2: Hydrogen bonding and π-π stacking statistics for the Whelk-O1 and the modified CSPs.

Analyte		H-bond interaction probability			π-π stacking interaction probability	
		Amide hydrogen	Amide oxygen	Nitro oxygen	π-π stack with Phenanthryl	π-π stack with dinitrophenyl
Whelk						
STYO	**S**	14.6	N/A	N/A	3.2	28.8
	R	11.2	N/A	N/A	8.6	21.2
STBO	**S,S**	11.9	N/A	N/A	9.3	30.8
	R,R	6.6	N/A	N/A	17.5	19.4
PEOL	**S**	9.0	2.7	9.5	5.9	19.5
	R	10.4	3.3	10.5	3.6	19.9
PEDO	**S**	13.2	3.7	15.1	2.6	12.1
	R	11.9	5.0	15.9	5.4	9.7
CSP1a						
STYO	**S**	22.7	N/A	N/A	2.1	27.2
	R	18.8	N/A	N/A	4.4	16.3
STBO	**S,S**	11.2	N/A	N/A	2.9	18.0
	R,R	5.2	N/A	N/A	5.3	9.0
CSP1b						
STYO	**S**	11.0	N/A	N/A	2.6	17.7
	R	8.5	N/A	N/A	4.0	18.2
STBO	**S,S**	9.6	N/A	N/A	9.5	20.5
	R,R	2.1	N/A	N/A	5.9	11.1
CSP2						
STBO	**S,S**	5.4	N/A	N/A	11.8	12.9
	R,R	4.3	N/A	N/A	18.0	8.8
PEDO	**S**	7.7	1.7	5.2	2.8	3.5
	R	10.9	1.2	6.6	6.3	3.0
CSP3						
PEOL	**S**	1.0	3.6	13.9	5.5	13.4
	R	2.0	2.0	15.3	3.4	17.1
STYO	**S**	2.7	N/A	N/A	4.0	18.5
	R	2.0	N/A	N/A	5.2	17.1

Clearly, both analytes have responded differently to the two conformers that result from saturation of the top-most ring. For the smaller analyte, STYO, the conformation-induced differences within the amide linkage have increased H-bonding probability for CSP1a. This is not the case for stilbene oxide, where the size of the analyte introduces steric constraints. Fig. 6.4 shows radial distributions between the analyte ring centers and the phenanthryl ring centers of Whelk-O1, CSP1a, and CSP1b selectors. In all panels, the shoulder at around 4.1Å in the ring-ring distribution is mostly due to π-π stacking between the analyte ring and the selector ring. For STBO, Figs. 6.4(d) and (e) show that the saturation of the topmost ring reduces the probability for finding the analyte ring at these probabilities, as expected from the reduction in M2 docking events. This is also the case for STYO on CSP1b, but interestingly CSP1a shows only a minimal reduction although the shoulder has shifted to larger separations. Taken together with the fact that the π-π stacking probability has decreased for CSP1a, relative to Whelk-O1, these results indicate that R-STYO still has an enhanced probability for placing its ring near P2, despite a decrease in stacking interactions with the phenanthryl group and an increase in M1 docking events.

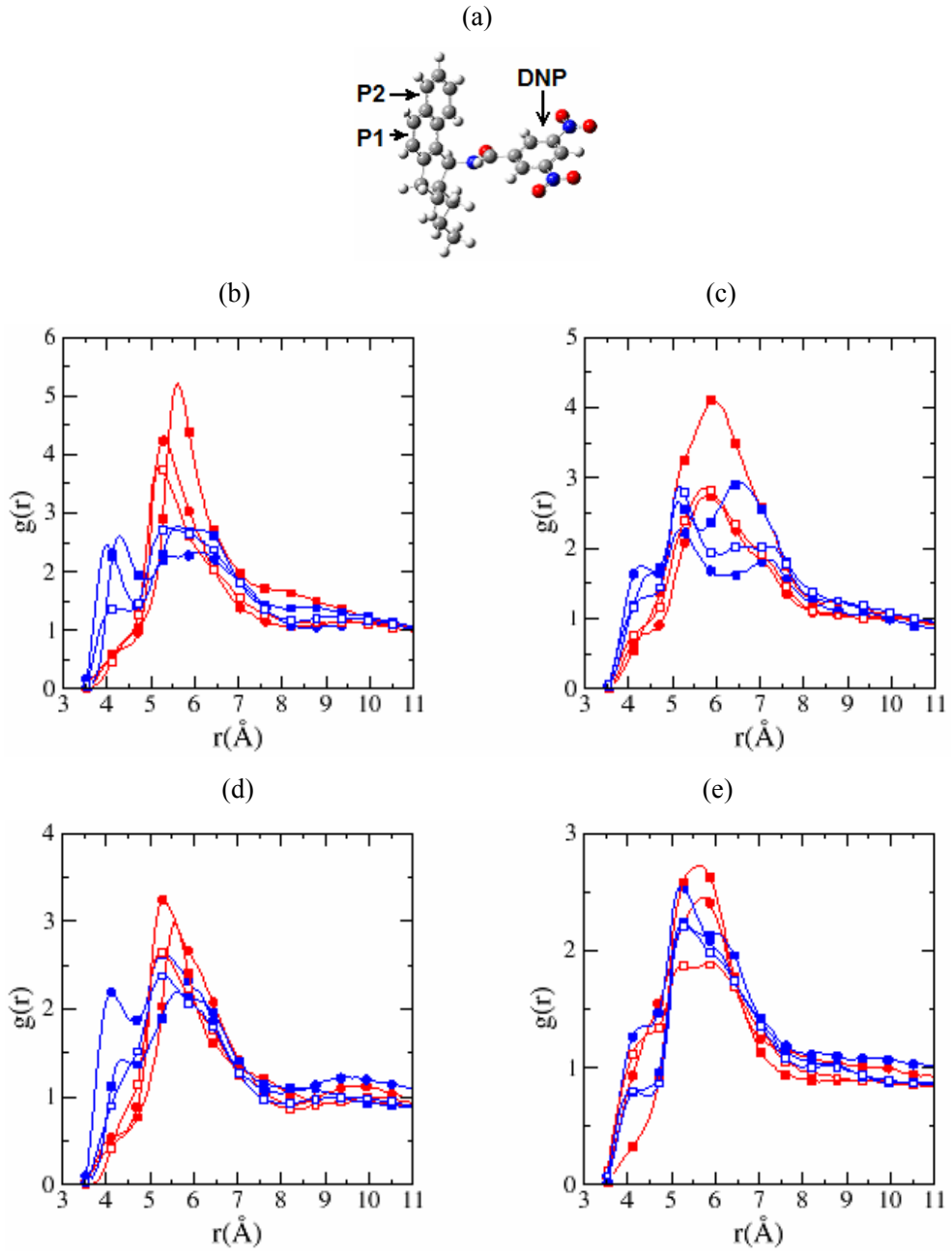


Figure 6.4: A comparison of radial distributions for CSP1a, CSP1b, and Whelk-O1 for (R/S)-styrene oxide and (R,R)/(S,S)-stilbene oxide docking. The aromatic regions of Whelk-O1 are identified by P1, P2, and DNP in (a). Radial distributions between the aromatic ring of S-STYO(red) or R-STYO(blue), and the P2 ring of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b(open squares) are shown in (b). Similarly, distributions between the aromatic ring of S-STYO(red) or R-STYO(blue), and the P1 ring of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b(open squares) are shown in (c). Radial distributions for S,S-STBO(red) and R,R-STBO(blue) are shown in panels (d) and (e). Distributions between the ring centers and P2 of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b (open squares) are shown in (d). Distributions between the ring centers and P1 of Whelk-O1 (filled circles), CSP1a (filled squares), and CSP1b (open squares) are shown in (e).

6.3.3 CSP2

CSP2 is modified to reduce the side-of-cleft docking of analytes. Experiments [55,57] and simulations in Chapter 5 have shown that side-of-cleft M1 docking events are probable for some less retained analytes. By limiting this docking arrangement, the separation factor can potentially be increased. For side-of-cleft M1 docking, the Whelk selector twists its dinitrophenyl ring (Chapter 5 and Ref. [55]) to accommodate the analytes, while for typical, inside-of-cleft M1 docking the dinitrophenyl ring remains around the equilibrium position. Thus, the expectation is that a restriction in the flexibility of the dinitrophenyl ring will preferentially disfavour side-of-cleft M1 docking while having a smaller impact on inside-of-cleft docking. To restrict DNP ring rotation, the dinitro-benzamide portion is modified to be a dinitro-1H-isoindole group (Fig. 6.1, CSP2). The energetic costs for torsion responsible for the dinitrophenyl flexibility [N(H)-C(N) bond for CSP2 and C(O)-C(C,C) bond for Whelk] are compared, before and after the modification, in Fig. 6.5. For Whelk-O1, rotation of the DNP ring was relatively facile but CSP2 offers only a restricted range of motion.

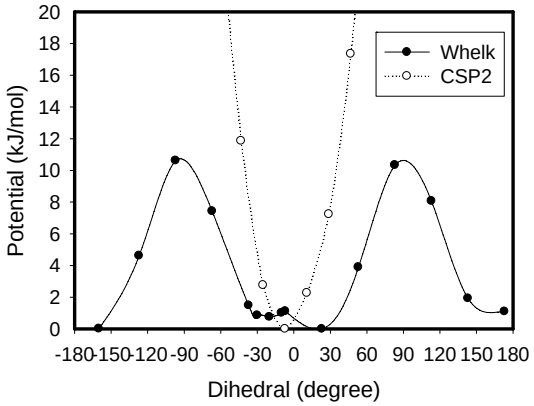


Figure 6.5: Torsional flexibility comparison between the dinitrophenyl ring of the Whelk-O1 selector and the dinitroisoindole ring in the CSP2 selector. The solid line with filled circles shows the torsional potential of N(H)-C(O)-C-C of the Whelk while the dotted line with open circles identify the torsional potential between C-N(H)-C-N of CSP2.

Stilbene oxide (STBO) and 1-phenylethane-1,2-diol (PEDO) are selected to evaluate the effect of the modification. The less retained enantiomers of these analytes often adopt side-of-

cleft M1 docking modes on Whelk-O1. Consider the interaction statistics in Table 6.1. First, the modified CSP has fewer interactions with analytes, as seen by the increase in the probability for finding analytes that are not interacting with a selector. As a result, the probability for finding a docked analyte has dropped following the CSP modification. The enantiomers are not impacted equally by the modification, as expected, and selectivity has increased: From 1.8 to 2.7 for STBO and from 1.1 to 2.2 for PEDO.

The docking characteristics have changed markedly due to the restricted orientations of the DNP ring. First, the probability for side-of-cleft docking events has dropped sharply for all enantiomers. This was the objective of the modification. However, M1 inside-the-cleft docking has also been impacted, as has the probability for M3 and M4 docking. These results indicate that all docking modes involving a π-π stack with the DNP ring rely, to some extent, on the adaptability of Whelk-O1 to position the DNP ring. Without this flexibility, these docking modes are less probable and docking is less likely overall. In addition, CSP2 may introduce extra steric and electrostatic repulsion that discourage ring-ring interactions. M2 docking, where a π-π stack forms with the phenanthryl group, rather than the DNP ring, is only slightly affected by the modification. Thus, CSP2 ultimately favours M2 docking. This is the reason for the inversion in retention order resulting from the modified CSP: S,S-STBO elutes last on Whelk-O1 but first on CSP2; S-PEDO elutes last on Whelk-O1 and first on CSP2.

6.3.4 CSP3

CSP3 is modified from Whelk by reversing the amide linkage. The resulting N-3,5-dinitrophenyl amide motif is similar to the N-3,5-dimethylphenyl amide group present in the polysaccharide derived CSPs, including cellulose/amylose tris(3,5-dimethylphenyl-carbamate). In comparison to Whelk-O1, the modification alters the spatial arrangement of the hydrogen bonding and π-π stacking sites. As shown in Fig. 6.2(c), CSP3 has a more localized carbonyl group, with the O closer to the surface, but the nitrogen and hydrogen atom placement, relative to the underlying

surface, is very similar to Whelk-O1. The latter result is crucial for selectivity since the placement of the amide hydrogen in the cleft region of Whelk-O1 is pivotal for docking. These results suggest that the largest impact of the modification will be due to a change in lateral distance between the amide hydrogen and the aromatic rings of Whelk-O1.

At the outset, the impact on docking of the modification yielding CSP3 is less predictable than for CSP1 and CSP2. A chiral alcohol, 1-phenyl-ethanol (PEOL) which is better separated by the cellulose-(3,5-dimethylphenyl-carbamate) based CSP than by Whelk-O1 [234], is picked to evaluate CSP3. Styrene oxide (STYO), which resolves well on Whelk-O1, is also tested on CSP3.

From Table 6.1, the interactions with the selector have dropped somewhat due to the modification. This is particularly true for STYO where roughly 11% more analytes are, on average, not forming any interactions with selectors. The impact of the modification is less severe for PEOL, and only 4% fewer analytes interact with selectors. In contrast to CSP2, where elution orders were reversed, CSP3 does not alter the elution order but it does increase selectivity: For STYO, the selection factor increases from 1.3 to 1.6, while for PEOL α increases from 1.2 to 1.7.

The docking statistics in Table 6.1 reveal important changes in the docking modes upon inversion of the amide bridge. For both analytes, the probability for M1 docking, both inside-of-cleft and side-of-cleft, has decreased sharply. Probability for M2 docking has also dropped. PEOL may also dock according to M3 and M4. The former is uncommon on Whelk-O1 and CSP3, but the latter becomes dominant for CSP3. These results indicate that repositioning of the amide hydrogen has negatively impacted all docking modes that involve H-bonding to this hydrogen. Table 6.2 provides hydrogen bonding statistics and, as expected, the probability for H-bonding to the amide hydrogen has dropped sharply. PEOL may also interact via M4, which does not involve the amide hydrogen, and the simulations predict that this alternate docking mode

becomes preferred. The shift towards M4 is reflected in the increase in probability for H-bonding to a nitro oxygen (Table 6.2) upon amide inversion.

6.4 Conclusions

Three CSPs have been derived from the Whelk-O1 chiral stationary phase. Their design was based on the elucidated atomic-level enantioselective mechanism for Whelk-O1 by previous molecular dynamics studies (Chapters 4 and 5) and empirical reasoning. The first and second CSPs are designed to reduce the docking of the least retained enantiomer, via hindrance of specified docking modes. The third CSP reverses the amide bridge to introduce a motif similar to one present in carbamate-derived polysaccharide CSPs. All three modified CSPs have increased selectivity, relative to Whelk-O1, for the analytes under consideration. However, for CSP2, this increase is accompanied by a reversed elution order.

Ab initio calculations are applied to set up semi-flexible models that can correctly represent the intramolecular flexibility of these selectors. For the three modified CSPs, molecular dynamics simulations are applied to evaluate their selectivity. Simulations reveal that the influence of the modifications is complex and extends beyond the originally envisioned impact on docking modes. The subtle interplay between selector characteristics and chiral selectivity is evident from our results.

This study shows that *in silico* design and testing of novel CSPs is possible as long as care is taken to correctly model the selector flexibility and the model interface. Although one could argue that ultimately the favored CSP has to be prepared and tested, computer-based CSP design, in analogy to computer-aided drug design, circumvents the slow and expensive screening process and provides an option for fast virtual discovery of CSPs.

Chapter 7
Conclusions and Outlook

The enantioselective mechanism of the Whelk-O1 chiral stationary phase has been the main subject of this thesis. *ab initio* calculations have been applied to develop semi-flexible molecular models for the Whelk selector and a series of analytes. Molecular dynamics simulations have been used to study the solvation, enantio-selectivity and optimization of the Whelk-O1 chiral stationary phase.

DFT electronic structural calculations have been carried out to explore the structural minima and to develop a semi-flexible model for the (3R,4S)-Whelk-O1 selector, as discussed in Chapter 2. We identified the equilibrium structure which shows that the phenanthryl group and dinitrophenyl group, connected by an amide linkage, form a cleft-like shape. A semi-flexible model was developed based on an extensive series of B3LYP/6-31G* constrained optimizations. Whelk-O1 selectors, together with end-caps and silanol groups, are then attached to a stationary layer of silicon atoms to form the model Whelk-O1 chiral stationary phase. A simulation cell, consisting of two surfaces with explicit solvent and analyte in between as well as empty space beyond, is then designed to mimic the environment of a pore of silica.

Molecular dynamics simulations have been carried out in Chapter 3 to study the solvation of the Whelk-O1 chiral stationary phase in three distinct solvents: *n*-hexane/2-propanol, water/methanol and a supercritical CO_2/methanol. We explored the conformations of the solvated selectors and found that the majority of them are "open": their clefts point to the bulk. We also noticed significant solvent partitioning that generates a polarity gradient between the selective region and the bulk which may encourage analyte proximity to the selectors. We also observed extensive hydrogen bonding activity between solvent and selector. Solvent molecules hydrogen bond to the amide hydrogen of the selector mostly from the cleft region, indicating that analytes have to replace them for docking. However, only about 15%-20% of the amide hydrogens are

hydrogen bonded under commonly used solvent conditions. In comparison, about twice as many amide oxygens, which are not directly involved in the three-point interaction model [77,96], are hydrogen bonded. The solvent approaches these oxygens from the side and therefore may not impede analyte docking.

Extensive molecular dynamics simulations of analyte docking were applied to study the enantio-selective mechanism of the Whelk-O1 chiral stationary phase in Chapters 4 and 5. In Chapter 4 we analyzed the docking of styrene oxide and stilbene oxide on the Whelk-O1 chiral stationary phase. Radial distributions and snapshot analysis both indicate abundant hydrogen bonding interactions with the amide hydrogen and π-π interactions with the aromatic groups of the selector. Edge-to-face π-π interactions and CH-π interactions are also evident, but they are more long-ranged and less specific. Three methods were used to predict the separation factors between enantiomers: considering only the hydrogen bonding interaction, considering both the hydrogen bonding interaction and the π-π stacking interaction, and considering all three interactions (hydrogen bonding, π-π stacking, and edge-to-face π-π). It was found that the use of the second criterion as "docked" for analytes gave excellent prediction on the separation factors. This criterion was further supported in Chapter 5 where the hydrogen bonding and π-π stacking interactions were found to be essential for enantioselectivity. In Chapter 5, we used this criterion and predicted the elution orders and separation factors for 10 analytes. The predictions are in good agreement with chromatographic experiments, which provides a sound basis for analyzing different contributions to enantioselectivity of the CSP.

Four distinct docking arrangements were identified. The most frequent docking arrangement consists of a hydrogen bond with the amide hydrogen and π-π stack with the dinitrophenyl group (M1). This mechanism is consistent with the proposed three-point interaction model [77,96]. It usually offers the most energy benefit for analyte docking and is the dominant docking arrangement for the most retained enantiomers. However, for some analytes, steric hindrance with the amide or phenanthryl group and intramolecular strain on the analyte

discourage such a docking arrangement. Other docking arrangements, which involve hydrogen bonding with oxygens and/or π-π stack with phenanthryl group could be frequent for the least retained enantiomers. A notable example is the docking of (R,R)-stilbene oxide: to dock according to M1, (R,R)-stilbene oxide has to make unfavourable conformational changes to avoid steric hindrance between its second phenyl ring and the selector. Thus, M1 is not favoured by (R,R)-stilbene oxide.

One view [55,57] of Whelk-O1 selectivity proposes that the most retained enantiomer docks inside of the cleft while the least retained enantiomer docks outside of the cleft. We found in Chapter 5 that the selector had to twist its dinitrophenyl group to accomplish an outside-of-cleft docking with an analyte. Thus, outside-of-cleft docking is energetically unfavourable but is sometimes observed for the least retained enantiomers. However, the inside/outside competition was only true for certain analytes, such as the n-(1-(4-bromophenyl)ethyl)pivalamide, but not applicable for most of the analytes examined.

In summary, we found that the hydrogen bonding with amide hydrogen and π-π stacking with the dinitrophenyl group are essential for the enantioselectivity of the Whelk-O1 CSP. A third point interaction is needed to complete the three-point interaction model. Our study shows that the third point could come from multiple sources. That is, after having formed the two primary interactions, the "rest of the molecule" is relatively constrained due to the specification of the hydrogen bond and π-π interaction and both stabilizing interactions, such as edge-to-face π-π or CH-π, and destabilizing factors, such as steric hindrance and unfavourable conformational changes involved in the analyte or the selector, contribute to the "third" interaction point and enantioselectivity. A destabilizing "third" point of interaction can often be identified for the least retained enantiomer which could lead to frequent alternative docking arrangements including outside-of-cleft docking and docking arrangements involving alternative hydrogen bonds and/or π-π stacks.

Based on the atomistic enantioselective mechanism, we rationally optimize the Whelk-O1 CSP, and perform *in silico* testing of the modified CSPs. Several CSPs were found to be improved for the separation of selected analytes, which is encouraging. It is worth noting that "tailored" CSP for high-efficiency large-scale separation of a chosen analyte is one of the focuses in current chiral separation science and technology. Our results indicate a bright future for molecular dynamics simulations to be applied in this area.

We also noticed several challenges, especially for large analytes. First, sampling efficiency needs to be improved to shorten the process of extensive molecular dynamics simulations for converged results. In this regard, modern and advanced techniques, such as replica exchange [247] molecular dynamics simulations, could be applied. Second, the limitations due to the usage of classical force fields need to be reduced. For intermolecular potentials, although the force fields we applied provide reasonable potentials based on *ab initio* energetic evaluations, it would be beneficial to use a force field that could better account for the interactions involved in the analyte docking process such as the π-π interaction and CH-π interaction. Unfortunately, *ab initio* studies on π-π and CH-π interactions have proved very challenging [236,241]. For intramolecular motion, we developed intramolecular potentials based on series of *ab initio* calculations and subtracted the intramolecular nonbonding interactions from torsional potentials. Both efforts enabled us to obtain a reasonable description of the molecular flexibility, which we believe is essential for the success of our models. However, for relatively large molecules, such as the n-(1-(4-bromophenyl)ethyl)pivalamide which has three highly correlated torsions, a more sophisticated torsional potential could be beneficial, although we still obtained reasonable results from our simulations.

Despite of all the challenges, our results indicate broad applications for molecular simulation in the area of chiral chromatography. The methods we developed could be easily apply to other brush type chiral stationary phase which usually have similar or smaller molecular size than the Whelk-O1 selector and have similar types of analyte-selector interactions.

Bibliography

[1] V.A. Davankov, Pure and Applied Chemistry **69** (1997) 1469.

[2] [Anon], Chirality **4** (1992) 338.

[3] K.M. Rentsch, Journal of Biochemical and Biophysical Methods **54** (2002) 1.

[4] D.J. Cordato, L.E. Mather, G.K. Herkes, Journal of Clinical Neuroscience **10** (2003) 649.

[5] H. Caner, E. Groner, L. Levy, I. Agranat, Drug Discovery Today **9** (2004) 105.

[6] A.J. Hutt, S.C. Tan, Drugs **52** (1996) 1.

[7] A.G. Rauws, K. Groen, Chirality **6** (1994) 72.

[8] A.M. Thayer, Chemical & Engineering News **August(6)** (2007) 11.

[9] W.S. Knowles, M.J. Sabacky, Chemical Communications (1968) 1445.

[10] R. Noyori, T. Ohkuma, M. Kitamura, H. Takaya, N. Sayo, H. Kumobayashi, S. Akutagawa, Journal of the American Chemical Society **109** (1987) 5856.

[11] T. Katsuki, K.B. Sharpless, Journal of the American Chemical Society **102** (1980) 5974.

[12] G. Subramanian, Chiral Separation Techniques: A Practical Approach, Wiley-VCH, Weinheim, 2001.

[13] L. Pasteur, C. R. Hebd. Acad. Sci. **26** (1848) 535.

[14] R.M. Secor, Chemical Reviews **63** (1963) 297.

[15] D.F. Reinhold, Fireston.Ra, W.A. Gaines, J.M. Chemerda, Sletzing.M, Journal of Organic Chemistry **33** (1968) 1209.

[16] Y.H. Wei, K. Kannappan, G.W. Flynn, M.B. Zimmt, Journal of the American Chemical Society **126** (2004) 5318.

[17] E. Fogassy, M. Nogradi, D. Kozma, G. Egri, E. Palovics, V. Kiss, Organic & Biomolecular Chemistry **4** (2006) 3011.

[18] K.S. Kim, J.S. Sack, J.S. Tokarski, L.G. Qian, S.T. Chao, L. Leith, Y.F. Kelly, R.N. Misra, J.T. Hunt, S.D. Kimball, W.G. Humphreys, B.S. Wautlet, J.G. Mulheron, K.R. Webster, Journal of Medicinal Chemistry **43** (2000) 4126.

[19] M.P. Grella, R. Danso-Danquah, M.K. Safo, G.S. Joshi, J. Kister, M. Marden, S.J. Hoffman, D.J. Abraham, Journal of Medicinal Chemistry **43** (2000) 4726.

[20] I. Ali, K. Kumerer, H.Y. Aboul-Enein, Chromatographia **63** (2006) 295.

[21] M. Juza, M. Mazzotti, M. Morbidelli, Trends in Biotechnology **18** (2000) 108.

[22] M. Schulte, J. Strube, Journal of Chromatography A **906** (2001) 399.

[23] T.J. Ward, Analytical Chemistry **78** (2006) 3947.

[24] E.R. Francotte, Journal of Chromatography A **906** (2001) 379.

[25] H.Y. Aboul-Enein, I. Ali, Chiral Separations by Liquid Chromatography and Related Technologies, Marcel Dekker, Inc., New York, 2003.

[26] J. Debowski, D. Sybilska, J. Jurczak, Journal of Chromatography **237** (1982) 303.

[27] S.M. Han, Biomedical Chromatography **11** (1997) 259.

[28] W. Lindner, J.N. Lepage, G. Davies, D.E. Seitz, B.L. Karger, Journal of Chromatography **185** (1979) 323.

[29] S. Allenmark, B. Bomgren, H. Boren, Journal of Chromatography **237** (1982) 473.

[30] A. Berthod, Analytical Chemistry **78** (2006) 2093.

[31] L.H. Easson, E. Stedman, Biochemical Journal **27** (1933) 1257.

[32] A.G. Ogston, Nature **162** (1948) 963.

[33] W.H. Pirkle, T.C. Pochapsky, Chemical Reviews **89** (1989) 347.

[34] W. Lindner, Mikrochimica Acta **2** (1991) 113.

[35] C.J. Welch, Journal of Chromatography A **666** (1994) 3.

[36] V.A. Davankov, Chirality **9** (1997) 99.

[37] S. Topiol, M. Sabio, Journal of the American Chemical Society **111** (1989) 4109.

[38] A.D. Mesecar, D.E. Koshland, Nature **403** (2000) 614.

[39] V. Sundaresan, R. Abrol, Chirality **17** (2005) S30.

[40] W.H. Pirkle, C.J. Welch, Journal of Liquid Chromatography **15** (1992) 1947.

[41] Y. Okamoto, Y. Kaida, Journal of Chromatography A **666** (1994) 403.

[42] Y. Okamoto, M. Kawashima, K. Hatada, Journal of the American Chemical Society **106** (1984) 5357.

[43] Y. Okamoto, M. Kawashima, K. Hatada, Journal of Chromatography **363** (1986) 173.

[44] Y. Okamoto, R. Aburatani, T. Fukumoto, K. Hatada, Chemistry Letters (1987) 1857.

[45] R.H. Gerhard Hesse, in, 1976, p. 996.

[46] K.B. Lipkowitz, Journal of Chromatography A **694** (1995) 15.

[47] H.Y. Aboul-Enein, I. Ali, Journal of Pharmaceutical and Biomedical Analysis **27** (2002) 441.

[48] H.Y. Aboul-Enein, I. Ali, Chirality **14** (2002) 47.

[49] J.A. Hamilton, L. Chen, Journal of the American Chemical Society **110** (1988) 5833.

[50] J.I. Seeman, H.V. Secor, D.W. Armstrong, K.D. Timmons, T.J. Ward, Analytical Chemistry **60** (1988) 2120.

[51] A.T.-K.M.P.-P. Stavroula Piperaki, Chirality **7** (1995) 257.

[52] K. Fujimura, S. Suzuki, K. Hayashi, S. Masuda, Analytical Chemistry **62** (1990) 2198.

[53] F. Mikes, G. Boshart, E. Gilav, Journal of Chromatography **122** (1976) 205.

[54] C. Altomare, S. Cellamare, A. Carotti, M.L. Barreca, A. Chimirri, A.M. Monforte, F. Gasparrini, C. Villani, M. Cirilli, F. Mazza, Chirality **8** (1996) 556.

[55] M.E. Koscho, P.L. Spence, W.H. Pirkle, Tetrahedron-Asymmetry **16** (2005) 3147.

[56] W.H. Pirkle, T.C. Pochapsky, Journal of the American Chemical Society **109** (1987) 5975.

[57] M.E. Koscho, W.H. Pirkle, Tetrahedron-Asymmetry **16** (2005) 3345.

[58] W.H. Pirkle, M.H. Hyun, B. Bank, Journal of Chromatography **316** (1984) 585.

[59] I.W. Wainer, M.C. Alembik, Journal of Chromatography **367** (1986) 59.

[60] K.K. Stewart, R.F. Doherty, Proceedings of the National Academy of Sciences of the United States of America **70** (1973) 2850.

[61] D.W. Armstrong, Y.B. Tang, S.S. Chen, Y.W. Zhou, C. Bagwill, J.R. Chen, Analytical Chemistry **66** (1994) 1473.

[62] H.Y. Aboul-Enein, I. Ali, Chromatographia **52** (2000) 679.

[63] N.M. Maier, P. Franco, W. Lindner, Journal of Chromatography A **906** (2001) 3.

[64] E. Blasius, W. Adrian, K.P. Janzen, G. Klautke, Journal of Chromatography **96** (1974) 89.

[65] L.R. Sousa, D.H. Hoffman, L. Kaplan, D.J. Cram, Journal of the American Chemical Society **96** (1974) 7100.

[66] V.A. Davankov, S.V. Rogozhin, Journal of Chromatography **60** (1971) 280.

[67] A. Kurganov, Journal of Chromatography A **906** (2001) 51.

[68] Y. Machida, H. Nishi, K. Nakamura, Chirality **11** (1999) 173.

[69] M.H. Hyun, J.S. Jin, W.J. Lee, Journal of Chromatography A **822** (1998) 155.

[70] N. Oi, M. Nagase, T. Doi, Journal of Chromatography **257** (1983) 111.

[71] W.H. Pirkle, D.W. House, J.M. Finn, Journal of Chromatography **192** (1980) 143.

[72] W.H. Pirkle, T.C. Pochapsky, Journal of the American Chemical Society **108** (1986) 352.

[73] W.H. Pirkle, T.J. Sowin, Journal of Chromatography **396** (1987) 83.

[74] W.H. Pirkle, C.J. Welch, B. Lamm, Journal of Organic Chemistry **57** (1992) 3854.

[75] W.H. Decamp, Chirality **1** (1989) 2.

[76] J. Bojarski, Journal of Liquid Chromatography **12** (1989) 2685.

[77] W.H. Pirkle, C.J. Welch, Tetrahedron-Asymmetry **5** (1994) 777.

[78] W.H. Pirkle, L.J. Brice, S. Caccamese, G. Principato, S. Failla, Journal of Chromatography A **721** (1996) 241.

[79] W.H. Pirkle, L.J. Brice, T.S. Widlanski, J. Roestamadji, Tetrahedron-Asymmetry **7** (1996) 2173.

[80] W.H. Pirkle, M.E. Koscho, Z.P. Wu, Journal of Chromatography A **726** (1996) 91.

[81] W.H. Pirkle, K.Z. Gan, Journal of Chromatography A **790** (1997) 65.

[82] W.H. Pirkle, W. Lee, C.J. Welch, Enantiomer **2** (1997) 423.

[83] C.J. Welch, T. Szczerba, S.R. Perrin, Journal of Chromatography A **758** (1997) 93.

[84] A. Magora, S. Abu-Lafi, S. Levin, Journal of Chromatography A **866** (2000) 183.

[85] B.H. Shao, X.Z. Xu, J.D. Lu, X.J. Cai, L. Zou, X.Y. Fu, Acta Chimica Sinica **61** (2003) 1635.

[86] X.J. Cai, X.Z. Xu, C.X. Pan, H.M. He, J.J. Chen, Chinese Journal of Analytical Chemistry **33** (2005) 653.

[87] H.M. He, X.J. Cai, X.Z. Xu, C.X. Pan, X.J. Zhang, J.J. Chen, Chinese Journal of Analytical Chemistry **33** (2005) 861.

[88] C.X. Pan, Q.Z. Wu, B.C. Shen, X.J. Zhang, D.T. Zhang, X.Z. Xu, Chinese Journal of Analytical Chemistry **34** (2006) 159.

[89] J. Dungelova, J. Lehotay, J. Cizmarik, D.W. Armstrong, Journal of Liquid Chromatography & Related Technologies **26** (2003) 2331.

[90] W.H. Pirkle, L.J. Brice, G.J. Terfloth, Journal of Chromatography A **753** (1996) 109.

[91] P. Oswald, K. Desmet, P. Sandra, J. Krupcik, P. Majek, D.W. Armstrong, Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences **779** (2002) 283.

[92] C.M. Kraml, D.H. Zhou, N. Byrne, O. McConnell, Journal of Chromatography A **1100** (2005) 108.

[93] A.M. Blum, K.G. Lynam, E.C. Nicolas, Chirality **6** (1994) 302.

[94] B.C. Hamper, D.R. Dukesherer, K. Moedritzer, Journal of Chromatography A **666** (1994) 479.

[95] G. Yanik, F. Mannerino, J. Kocergin, Lc Gc North America (2005) 34.

[96] A. Del Rio, J.M. Hayes, M. Stein, P. Piras, C. Roussel, Chirality **16** (2004) S1.

[97] W.H. Pirkle, C.J. Welch, Journal of Chromatography A **683** (1994) 347.

[98] K.B. Lipkowitz, Journal of Chromatography A **906** (2001) 417.

[99] R. Dappen, H.R. Karfunkel, F.J.J. Leusen, Journal of Computational Chemistry **11** (1990) 181.

[100] K.B. Lipkowitz, D.A. Demeter, R. Zegarra, R. Larter, T. Darden, Journal of the American Chemical Society **110** (1988) 3446.

[101] K.B. Lipkowitz, B. Baker, R. Larter, Journal of the American Chemical Society **111** (1989) 7750.

[102] N.M. Maier, S. Schefzick, G.M. Lombardo, M. Feliz, K. Rissanen, W. Lindner, K.B. Lipkowitz, Journal of the American Chemical Society **124** (2002) 8611.

[103] A. Ragusa, J.M. Hayes, M.E. Light, J.D. Kilburn, Chemistry-a European Journal **13** (2007) 2717.

[104] J.M. Hayes, M. Stein, J. Weiser, Journal of Physical Chemistry A **108** (2004) 3572.

[105] I. Kolossvary, Journal of the American Chemical Society **119** (1997) 10233.

[106] K.B. Lipkowitz, B. Baker, Analytical Chemistry **62** (1990) 770.

[107] K.B. Lipkowitz, Journal of Chromatography A **666** (1994) 493.

[108] K.B. Lipkowitz, G. Pearl, B. Coner, M.A. Peterson, Journal of the American Chemical Society **119** (1997) 600.

[109] K.B. Lipkowitz, Accounts of Chemical Research **33** (2000) 555.

[110] K. Mihlbachler, M.A. De Jesus, K. Kaczmarski, M.J. Sepaniak, A. Seidel-Morgenstern, G. Guiochon, Journal of Chromatography A **1113** (2006) 148.

[111] R.B. Kasat, C.Y. Chin, K.T. Thomson, E.I. Franses, N.H.L. Wang, Adsorption-Journal of the International Adsorption Society **12** (2006) 405.

[112] C. Temporini, E. Calleri, G. Fracchiolla, G. Carbonara, F. Loiodice, A. Lavecchia, P. Tortorella, G. Brusotti, G. Massolini, Journal of Pharmaceutical and Biomedical Analysis **45** (2007) 211.

[113] U. Norinder, J. Hermansson, Chirality **3** (1991) 422.

[114] R. Kaliszan, T.A.G. Noctor, I.W. Wainer, Chromatographia **33** (1992) 546.

[115] V. Andrisano, R. Gotti, M. Recanatini, A. Cavalli, L. Varoli, C. Bertucci, Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences **768** (2002) 137.

[116] T.D. Booth, I.W. Wainer, Journal of Chromatography A **737** (1996) 157.

[117] Z. Chilmonczyk, H. Ksycinska, M. Lisowaka-Kuzmicz, A. Jonczyk, M. Mazgajska, M. Jaronczyk, A. Strzelczyk, H.Y. Aboul-Enein, Analytica Chimica Acta **536** (2005) 7.

[118] W.M.F. Fabian, W. Stampfer, M. Mazur, G. Uray, Chirality **15** (2003) 271.

[119] A. Del Rio, P. Piras, C. Roussel, Chirality **18** (2006) 498.

[120] S.J. Klatte, T.L. Beck, Journal of Physical Chemistry **97** (1993) 5727.

[121] J.T. Slusher, R.D. Mountain, Journal of Physical Chemistry B **103** (1999) 1354.

[122] L. Zhang, L. Sun, J.I. Siepmann, M.R. Schure, Journal of Chromatography A **1079** (2005) 127.

[123] C. Hellriegel, U. Skogsberg, K. Albert, M. Lammerhofer, N.M. Maier, W. Lindner, Journal of the American Chemical Society **126** (2004) 3809.

[124] I.N. Levine, Quantum chemistry, Prentice Hall, Upper Saddle River, NJ, 2000.

[125] R.J. Bartlett, Annual Review of Physical Chemistry **32** (1981) 359.

[126] R.J. Bartlett, Journal of Physical Chemistry **93** (1989) 1697.

[127] P. Hohenberg, W. Kohn, Physical Review B **136** (1964) B864.

[128] W. Kohn, L.J. Sham, Physical Review **140** (1965) 1133.

[129] A.D. Becke, Physical Review A **38** (1988) 3098.

[130] C.T. Lee, W.T. Yang, R.G. Parr, Physical Review B **37** (1988) 785.

[131] R.G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules, Oxford University Press, 1989.

[132] J.P. Perdew, Y. Wang, Physical Review B **45** (1992) 13244.

[133] A.D. Becke, Journal of Chemical Physics **98** (1993) 5648.

[134] Ditchfie.R, W.J. Hehre, J.A. Pople, Journal of Chemical Physics **54** (1971) 724.

[135] M.J.T. Frisch, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.;, in, Gaussian Inc., Wallingford CT, 2004.

[136] U.C. Singh, P.A. Kollman, Journal of Computational Chemistry **5** (1984) 129.

[137] C.M. Breneman, K.B. Wiberg, Journal of Computational Chemistry **11** (1990) 361.

[138] D. Frenkel, B. Smit, Understanding Molecular Simulation: From Algorithms to Applications, Academic Press, San Diego, 2002.

[139] M.P. Allen, D.J. Tildesley, Computer simulation of liquids, Clarendon Press, Oxford, 1989.

[140] P.M. Morse, Physical Review **34** (1929) 57.

[141] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Journal of Computational Physics **23** (1977) 327.

[142] H.C. Andersen, Journal of Computational Physics **52** (1983) 24.

[143] W.L. Jorgensen, D.S. Maxwell, J. TiradoRives, Journal of the American Chemical Society **118** (1996) 11225.

[144] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, Journal of Physical Chemistry B **102** (1998) 3586.

[145] J.P. Ryckaert, A. Bellemans, Faraday Discussions (1978) 95.

[146] C.J. Cramer, Essentials of Computational Chemistry: Theories and Models Wiley, 2004.

[147] H.A. Lorentz, Annalen der physik **248** (1881) 127.

[148] D. Berthelot, Comptes rendus hebdomadaires des séances de l'Académie des sciences **126** (1898).

[149] S.W. Rick, S.J. Stuart, B.J. Berne, Journal of Chemical Physics **101** (1994) 6141.

[150] G. Lamoureux, A.D. MacKerell, B. Roux, Journal of Chemical Physics **119** (2003) 5185.

[151] P.P. Ewald, Annalen Der Physik **64** (1921) 253.

[152] I.C. Yeh, M.L. Berkowitz, Journal of Chemical Physics **111** (1999) 3155.

[153] D.M. York, T.A. Darden, L.G. Pedersen, Journal of Chemical Physics **99** (1993) 8345.

[154] A.W. Appel, Siam Journal on Scientific and Statistical Computing **6** (1985) 85.

[155] J. Barnes, P. Hut, Nature **324** (1986) 446.

[156] L. Greengard, V. Rokhlin, Journal of Computational Physics **73** (1987) 325.

[157] R.W. Hockney, J.W. Eastwood, Computer Simulations Using Particles, McGraw-Hill, New York, 1981.

[158] L. Verlet, Physical Review **159** (1967) 98.

[159] R.W. Hockney, Methods in computational physics (1970) 136.

[160] W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, Journal of Chemical Physics **76** (1982) 637.

[161] D. Beeman, Journal of Computational Physics **20** (1976) 130.

[162] W.G. Hoover, Physical Review A **31** (1985) 1695.

[163] S. Nosé, Journal of Chemical Physics **81** (1984) 511.

[164] C.W. Gear, Numerical initial value problems in ordinary differential equations, Prentice-Hall, Englewood Cliffs, 1971.

[165] M. Tuckerman, B.J. Berne, G.J. Martyna, Journal of Chemical Physics **97** (1992) 1990.

[166] D.J. Evans, G.P. MOrriss, Statistical Mechanics of Non-Equilibrium Liquids, Academic Press, London, 1990.

[167] H.C. Andersen, Journal of Chemical Physics **72** (1980) 2384.

[168] A. Rahman, Stilling.Fh, Journal of Chemical Physics **55** (1971) 3336.

[169] D.J. Evans, S. Murad, Molecular Physics **34** (1977) 327.

[170] D. Fincham, Molecular Simulation **8** (1992) 165.

[171] G.R. Kneller, K. Hinsen, Physical Review E **50** (1994) 1559.

[172] T.R. Forester, W. Smith, Journal of Computational Chemistry **19** (1998) 102.

[173] X.W. Wu, S.S. Sung, Journal of Computational Chemistry **19** (1998) 1555.

[174] D.J. Evans, Molecular Physics **34** (1977) 317.

[175] G.J. Martyna, D.J. Tobias, M.L. Klein, Journal of Chemical Physics **101** (1994) 4177.

[176] C.D. Taylor, M. Neurock, Current Opinion in Solid State & Materials Science **9** (2005) 49.

[177] M. Cacciatore, M. Rutigliano, International Journal of Quantum Chemistry **106** (2006) 631.

[178] T. Iwasaki, Jsme International Journal Series B-Fluids and Thermal Engineering **47** (2004) 470.

[179] M.S.G. Razul, E.V. Tam, M.E. Lam, P. Linden, P.G. Kusalik, Molecular Physics **103** (2005) 1929.

[180] J. Hautman, M.L. Klein, Journal of Chemical Physics **91** (1989) 4994.

[181] S.Y. Jiang, Molecular Physics **100** (2002) 2261.

[182] F. Ganazzoli, G. Raffaini, Physical Chemistry Chemical Physics **7** (2005) 3651.

[183] V.P. Raut, M.A. Agashe, S.J. Stuart, R.A. Latour, Langmuir **21** (2005) 1629.

[184] K.S. Siddiqui, R. Cavicchioli, Annual Review of Biochemistry **75** (2006) 403.

[185] S. Nita, N.M. Cann, J.H. Horton, Journal of Physical Chemistry B **108** (2004) 3512.

[186] S. Nita, J.H. Horton, N.M. Cann, Journal of Physical Chemistry B **110** (2006) 9511.

[187] A. Berthod, Journal of Chromatography **549** (1991) 1.

[188] T. Szczerba, personal communication, 2004.

[189] J. Nawrocki, Journal of Chromatography A **779** (1997) 29.

[190] W.F. Vangunsteren, H.J.C. Berendsen, Molecular Physics **34** (1977) 1311.

[191] J.G. Gay, B.J. Berne, Journal of Chemical Physics **74** (1981) 3316.

[192] G. LaPenna, D. Catalano, C.A. Veracini, Journal of Chemical Physics **105** (1996) 7097.

[193] I. Cacelli, G. Cinacchi, C. Geloni, G. Prampolini, A. Tani, Molecular Crystals and Liquid Crystals **395** (2003) 171.

[194] T. Clark, J. Chandrasekhar, G.W. Spitznagel, P.V. Schleyer, Journal of Computational Chemistry **4** (1983) 294.

[195] M.J. Frisch, J.A. Pople, J.S. Binkley, Journal of Chemical Physics **80** (1984) 3265.

[196] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, Journal of the American Chemical Society **117** (1995) 5179.

[197] J.M. Stubbs, J.J. Potoff, J.I. Siepmann, Journal of Physical Chemistry B **108** (2004) 17596.

[198] P.E.M. Lopes, V. Murashov, M. Tazi, E. Demchuk, A.D. MacKerell, Journal of Physical Chemistry B **110** (2006) 2782.

[199] W.L. Jorgensen, J.D. Madura, C.J. Swenson, Journal of the American Chemical Society **106** (1984) 6638.

[200] E04USF, NAG Fortran Library, Mark 21, Numerical Algorithms Group, Oxford, UK, 2005.

[201] W.L. Jorgensen, J. Tiradorives, Journal of the American Chemical Society **110** (1988) 1657.

[202] M.L.P. Price, D. Ostrovsky, W.L. Jorgensen, Journal of Computational Chemistry **22** (2001) 1340.

[203] J.M. Briggs, T. Matsui, W.L. Jorgensen, Journal of Computational Chemistry **11** (1990) 958.

[204] G.J. Martyna, M.E. Tuckerman, D.J. Tobias, M.L. Klein, Molecular Physics **87** (1996) 1117.

[205] N.M. Cann, E. Cressman, B. Das, D. J., R. Ghenea, Y. Huh, S. Nita, I. Paci, S. Wang, C. Zhao, unpublished.

[206] J. Vatamanu, N.M. Cann, Journal of Chemical Physics **121** (2004) 6922.

[207] Y. Huh, N.M. Cann, Journal of Chemical Physics **121** (2004) 10299.

[208] W. Cropp, E. Lusk, A. Skjellum, Using MPI: Portable Parallel Programming with the Message-Passing Interface, The MIT Press, Cambridge, Massachusetts, 1999.

[209] C. Wolf, L. Pranatharthiharan, E.C. Volpe, Journal of Organic Chemistry **68** (2003) 3287.

[210] D.H. Liu, P. Wang, W.F. Zhou, X. Gu, Z.S. Chen, Z.Q. Zhou, Analytica Chimica Acta **555** (2006) 210.

[211] B. Chen, J.J. Potoff, J.I. Siepmann, Journal of Physical Chemistry B **105** (2001) 3093.

[212] M. Levitt, M. Hirshberg, R. Sharon, K.E. Laidig, V. Daggett, Journal of Physical Chemistry B **101** (1997) 5051.

[213] J.G. Harris, K.H. Yung, Journal of Physical Chemistry **99** (1995) 12021.

[214] B.D. Smith, R. Srivastava, Thermodynamic data for pure compounds, Elsevier, New York, 1986.

[215] D.L. Goldfarb, D.P. Fernandez, H.R. Corti, Fluid Phase Equilibria **160** (1999) 1011.

[216] E.J.W. Wensink, A.C. Hoffmann, P.J. van Maaren, D. van der Spoel, Journal of Chemical Physics **119** (2003) 7308.

[217] M. Levitt, R. Sharon, Proceedings of the National Academy of Sciences of the United States of America **85** (1988) 7557.

[218] A. Luzar, D. Chandler, Physical Review Letters **76** (1996) 928.

[219] K.W. Phinney, Analytical and Bioanalytical Chemistry **382** (2005) 639.

[220] C.A. Eckert, B.L. Knutson, P.G. Debenedetti, Nature **383** (1996) 313.

[221] M. Dangelo, G. Onori, A. Santucci, Journal of Chemical Physics **100** (1994) 3107.

[222] S. Dixit, J. Crain, W.C.K. Poon, J.L. Finney, A.K. Soper, Nature **416** (2002) 829.

[223] S.E. Schaus, B.D. Brandes, J.F. Larrow, M. Tokunaga, K.B. Hansen, A.E. Gould, M.E. Furrow, E.N. Jacobsen, Journal of the American Chemical Society **124** (2002) 1307.

[224] A. Del Rio, P. Piras, C. Roussel, Chirality **17** (2005) S74.

[225] M.H. Hyun, J.J. Ryoo, W.H. Pirkle, Journal of Chromatography A **886** (2000) 47.

[226] G.E. Job, A. Shvets, W.H. Pirkle, S. Kuwahara, M. Kosaka, Y. Kasai, H. Taji, K. Fujita, M. Watanabe, N. Harada, Journal of Chromatography A **1055** (2004) 41.

[227] U.D. Neue, HPLC Columns: Theory, Technology, and Practice, Wiley-VCH, New York, 1997.

[228] R. Berardi, G. Cainelli, P. Galletti, D. Giacomini, A. Gualandi, L. Muccioli, C. Zannoni, Journal of the American Chemical Society **127** (2005) 10699.

[229] J. Dungelova, J. Lehotay, J. Krupcik, J. Cizmarik, D.W. Armstrong, J Sep Sci **27** (2004) 983.

[230] A. Luzar, D. Chandler, Nature **379** (1996) 55.

[231] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, Journal of the American Chemical Society **124** (2002) 104.

[232] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, A. Fujii, Journal of Physical Chemistry A **110** (2006) 10163.

[233] C. Wolf, W.H. Pirkle, Tetrahedron **58** (2002) 3597.

[234] J.M. Huang, P. Zhang, H. Chen, T.Y. Li, Analytical Chemistry **77** (2005) 3301.

[235] C. Moiteiro, N. Fonseca, M.J.M. Curto, R. Tavares, A.M. Lobo, P. Ribeiro-Claro, V. Felix, M.G.B. Drew, Tetrahedron-Asymmetry **17** (2006) 3248.

[236] M.O. Sinnokrot, C.D. Sherrill, Journal of Physical Chemistry A **110** (2006) 10656.

[237] C. Czerwenka, M.M. Zhang, H. Kahlig, N.M. Maier, K.B. Lipkowitz, W. Lindner, Journal of Organic Chemistry **68** (2003) 8315.

[238] W.H. Pirkle, K.Z. Gan, L.J. Brice, Tetrahedron-Asymmetry **7** (1996) 2813.

[239] T. Steiner, Angewandte Chemie-International Edition **41** (2002) 48.

[240] M.O. Sinnokrot, C.D. Sherrill, Journal of the American Chemical Society **126** (2004) 7690.

[241] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, Journal of the American Chemical Society **122** (2000) 3746.

[242] H. Dodziuk, Journal of Molecular Structure **614** (2002) 33.

[243] A. Mandl, L. Nicoletti, M. Lammerhofer, W. Lindner, Journal of Chromatography A **858** (1999) 1.

[244] W.L. Jorgensen, Science **303** (2004) 1813.

[245] N.M. Maier, G. Uray, O.P. Kleidernigg, W. Lindner, Chirality **6** (1994) 116.

[246] F.A.L. Anet, M.Z. Haq, Journal of the American Chemical Society **87** (1965) 3147.

[247] Y. Sugita, Y. Okamoto, Chemical Physics Letters **314** (1999) 141.

Appendix A
Interaction Potentials for Chiral Selectors and Analytes

Semi-flexible models have been developed for ten analytes and four modified selectors. The analyte and selector structures shown in this appendix include shaded regions. These shaded regions identify the rigid portions of the molecules. Within the simulations, these rings are treated as rigid units (RUs) and intramolecular motion within the group is not allowed. That is, the atoms within the group move together without changing their relative separations.

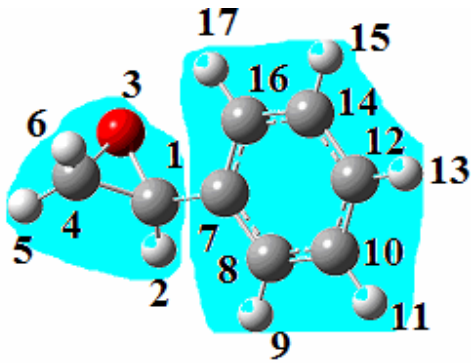
The interaction potential for each analyte is extracted from a B3LYP/6-311++G** geometry optimization, from a series of B3LYP/6-31G* scans to explore the energy required for intramolecular motion, and from the OPLS force field [143]. All electronic structure calculations are performed using the Gaussian 03 program [135].

The intramolecular potential, which dictates the energetic costs for molecular conformations, consists of four parts: bond stretching, angle bending, improper torsion for out-of-plane motion, and dihedral torsion respectively. The equations for these potentials are listed in Chapter 2. Harmonic potentials are used for bond stretching (Eq. 2.21), angle bending (Eq. 2.22) and improper torsion (Eq. 2.26) and a modified Rickaert-Bellemans [145] potential is used for torsion (Eq. 2.25).

For both the analytes and the modified selectors, atomic positions and CHELPG charges are extracted from the B3LYP/++G** geometry optimization. However, for the selectors, the configuration of the tether from Si(1) to C(6) is obtained from geometry optimization of $SiH_3OSi(CH_3)_2CH_2CH_3$. For nonbonding potential, the OPLS [143] parameters for Lennard-Jones potential (Eq. 2.28) and CHELPG charges from the *ab initio* global minimum are applied for the analytes. For the selectors, parameters for silicon is not available in OPLS and CHARMM [198] parameters are used instead.

Parameters for analytes and modified selectors developed in this thesis are provided in this appendix. For chiral analytes, only the parameters for one enantiomer are provided. For the other enantiomer, while most of the parameters stay the same, atomic positions are obtained through flipping the signs of all three coordinates, torsional potentials are obtained through flipping the signs of the phase shifts, and the equilibrium improper torsions are obtained through flipping the signs of that of the provided enantiomers.

Styrene Oxide



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.155	-1.544	0.061	0.739
2	H	0.150	-2.463	-0.197	0.216
3	O	-0.287	-1.498	-0.407	2.093
4	C	0.012	-1.662	0.996	1.875
5	H	0.091	-2.635	1.398	2.144
6	H	0.068	-0.803	1.609	2.138
7	C	0.271	-0.302	-0.052	-0.078
8	C	-0.257	-0.376	0.090	-1.468
9	H	0.106	-1.339	0.261	-1.940
10	C	-0.019	0.773	0.013	-2.250
11	H	0.081	0.702	0.127	-3.326
12	C	-0.126	2.011	-0.217	-1.651
13	H	0.091	2.906	-0.283	-2.258
14	C	-0.087	2.089	-0.374	-0.267

15	H	0.093	3.047	-0.567	0.202
16	C	-0.145	0.940	-0.293	0.515
17	H	0.111	0.994	-0.435	1.588

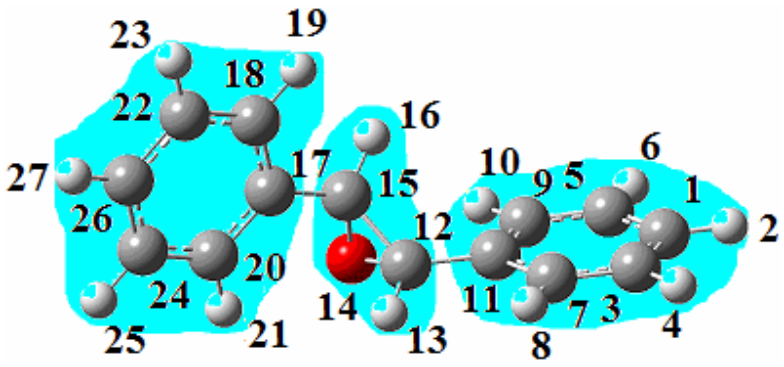
Stretch	k (kJ· mol^{-1} · nm^{-2})
1,7	1.16*10^{5}

Angle	k (kJ· mol^{-1} · rad^{-2})
2,1,7	244.9
3,1,7	385.9
4,1,7	305.4
1,7,8	369.0
1,7,16	491.7

Improper Torsion	k (kJ· mol^{-1} · rad^{-2})
7,8,16,1	261.4

Torsion	c0, c1, c2, c3, c4, c5, c6 (kJ/mol)	φ1, φ2, φ3, φ4, φ5, φ6
16,7,1,4	7.68, 0.19, 8.83, -1.52, 6.61, 1.13, 3.28	2.09, 2.54, 0.09, 0.05, 0.03, 1.00

Stilbene Oxide



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.139	-4.622	-0.558	-0.168
2	H	0.095	-5.660	-0.818	-0.347

3	C	-0.020	-3.887	-1.239	0.800
4	H	0.079	-4.351	-2.031	1.379
5	C	-0.065	-4.020	0.465	-0.899
6	H	0.087	-4.591	1.006	-1.645
7	C	-0.223	-2.556	-0.900	1.033
8	H	0.096	-1.990	-1.427	1.794
9	C	-0.155	-2.689	0.803	-0.670
10	H	0.105	-2.223	1.611	-1.221
11	C	0.232	-1.944	0.117	0.294
12	C	-0.061	-0.511	0.443	0.540
13	H	0.107	-0.147	0.217	1.541
14	O	-0.277	0.000	1.666	-0.000
15	C	-0.061	0.511	0.443	-0.540
16	H	0.107	0.147	0.217	-1.541
17	C	0.232	1.944	0.117	-0.294
18	C	-0.154	2.689	0.803	0.6701
19	H	0.105	2.223	1.611	1.221
20	C	-0.223	2.556	-0.900	-1.033
21	H	0.096	1.990	-1.427	-1.794
22	C	-0.064	4.020	0.465	0.899
23	H	0.087	4.591	1.006	1.645
24	C	-0.020	3.887	-1.239	-0.800
25	H	0.079	4.351	-2.031	-1.379
26	C	-0.138	4.622	-0.558	0.168
27	H	0.095	5.660	-0.818	0.347

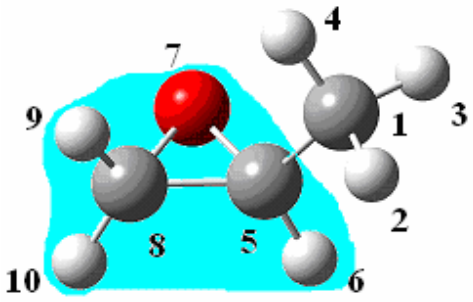
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
11,12/17,15	$1.16*10^5$

Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
7,11,12/18,17,15	428.7
9,11,12/20,17,15	437.3
11,12,13/17,15,16	267.3
11,12,14/17,15,14	372.4
11,12,15/17,15,12	287.3

Improper Torsion	k (kJ· mol^{-1}· rad^{-2})
11,7,9,12/17,18,20,15	261.1

Torsion	c0, c1, c2, c3, c4, c5, c6 (kJ/mol)	φ1, φ2, φ3, φ4, φ5, φ6
15,12,11,7/ 12,15,17,20	7.12, 0.74, 8.53, 1.07, 8.83, 0.85, 3.24	2.11, 0, 0.13, 0.63, 0, 1.37

2-Methyloxirane



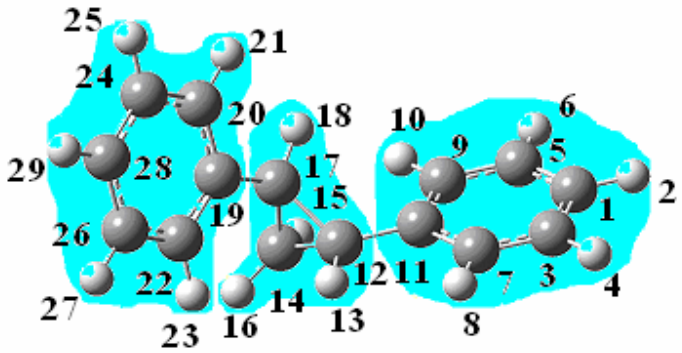
ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.143	1.511	0.098	-0.149
2	H	0.026	2.081	0.900	0.332
3	H	0.075	2.078	-0.831	-0.042
4	H	0.065	1.421	0.324	-1.214
5	C	0.112	0.153	-0.035	0.486
6	H	0.077	0.154	-0.253	1.554
7	O	-0.322	-0.830	-0.790	-0.240
8	C	-0.122	-1.043	0.617	-0.061
9	H	0.126	-0.956	1.215	-0.966
10	H	0.106	-1.869	0.880	0.596

Stretch	k (kJ· mol^{-1}· nm^{-2})
1,5	1.44*10^{5}

Angle	k (kJ· mol^{-1} · rad^{-2})
1,5,6	252.1
1,5,7	302.6
1,5,8	292.6

Torsion	c0, c1, c2, c3, c4, c5, c6 (kJ/mol)	φ1, φ2, φ3, φ4, φ5, φ6
2,1,5,7	5.95, -15.56, -10.97, 21.09, 9.21, -0.59, 6.29	4.44, 6.07, 4.45, 5.77, 1.57, 3.42

1,2-Diphenylcyclopropane



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.142	4.521	-0.155	-1.166
2	H	0.088	5.523	-0.326	-1.148
3	C	-0.015	3.818	0.993	-1.810
4	H	0.074	4.271	1.722	0.065
5	C	-0.028	3.924	-1.082	0.354
6	H	0.070	4.461	-1.979	-0.661
7	C	-0.233	2.531	1.208	-0.949
8	H	0.099	1.992	2.104	0.553
9	C	-0.240	2.638	-0.863	1.219
10	H	0.141	2.198	-1.598	0.199
11	C	0.257	1.918	0.286	0.685
12	C	-0.161	0.535	0.551	0.407
13	H	0.099	0.155	1.528	1.982
14	C	-0.197	0.000	0.000	2.571

15	H	0.088	0.649	-0.637	2.571
16	H	0.088	0.649	0.637	0.685
17	C	-0.161	0.535	-0.551	0.407
18	H	0.099	0.155	-1.528	0.199
19	C	0.257	1.918	-0.286	-0.661
20	C	-0.233	2.531	-1.208	-0.949
21	H	0.099	1.992	-2.104	0.553
22	C	-0.240	2.638	0.863	1.219
23	H	0.141	2.198	1.598	-1.148
24	C	-0.015	3.818	-0.992	-1.810
25	H	0.074	4.271	-1.722	0.065
26	C	-0.028	3.924	1.082	0.354
27	H	0.070	4.461	1.979	-0.789
28	C	-0.142	4.521	0.155	-1.166
29	H	0.088	5.523	0.326	-1.166

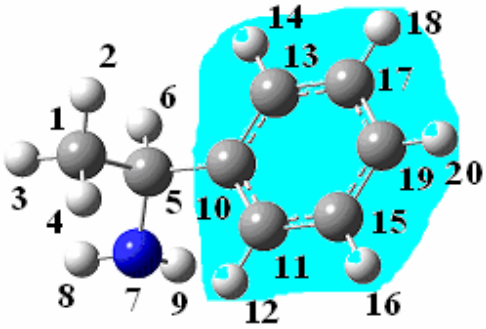
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
12,11/17,19	$1.15*10^5$

Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
7,11,12/20,19,17	486.4
9,11,12/22,19,17	453.6
11,12,13/20,17,18	252.6
11,12,14/20,17,14	261.9
11,12,17/20,17,12	256.0

Improper Torsion	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
11,7,9,12/19,20,22,17	275.9

Torsion	**c0, c1, c2, c3, c4, c5, c6**	**φ1, φ2, φ3, φ4, φ5, φ6**
7,11,12,14/ 20,19,17,15	4.13, 0.26, -7.64, -0.57, 7.21, 0.46, 4.33	2.35, 0.00, 0.27, 0.81, 0.30, 6.28

1-phenylethanamine



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.142	2.273	-0.535	1.161
2	H	0.003	1.827	-1.391	1.674
3	H	0.054	3.350	-0.712	1.079
4	H	0.037	2.116	0.358	1.771
5	C	0.357	1.647	-0.348	-0.229
6	H	-0.044	1.814	-1.277	-0.796
7	N	-0.898	2.279	0.814	-0.876
8	H	0.338	3.285	0.693	-0.922
9	H	0.296	1.940	0.928	-1.827
10	C	0.117	0.141	-0.147	-0.138
11	C	-0.072	-0.403	1.099	0.194
12	H	0.081	0.263	1.941	0.343
13	C	-0.205	-0.726	-1.222	-0.349
14	H	0.101	-0.318	-2.192	-0.617
15	C	-0.126	-1.782	1.260	0.317
16	H	0.094	-2.189	2.232	0.572
17	C	-0.061	-2.106	-1.065	-0.224
18	H	0.080	-2.764	-1.910	-0.395
19	C	-0.093	-2.638	0.179	0.110
20	H	0.084	-3.711	0.306	0.203

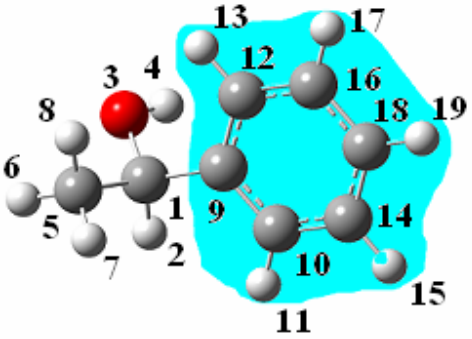
Stretch	k (kJ· mol^{-1}· nm^{-2})
5,10	1.33*10^{5}

Angle	k (kJ· mol^{-1} · rad^{-2})
5,10,11	438.4
5,10,13	442.5

Improper Torsion	k (kJ· mol^{-1} · rad^{-2})
10,13,11,5	272.4

Torsion	c0, c1, c2, c3, c4, c5, c6	φ1, φ2, φ3, φ4, φ5, φ6
7,5,10,11	18.90, -7.04, -8.29, 2.19, -21.65, 1.79, 11.25	0.20, 1.20, 0.40, 0.31, 1.63, 0.42
8,7,5,10	11.97, -26.80, -21.12, -25.41, 12.03, -9.02, 9.70	4.10, 0.79, 1.12, 0.65, 0.30, 1.46
2,1,5,10	1.19, -25.45, 8.75, 31.09, 7.85, 0.91, 5.61	6.24, 4.41, 0.04, 0.00, 0.38, 5.45

1-phenylethanol



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	0.316	-1.648	-0.263	-0.311
2	H	-0.004	-1.844	-0.999	-1.103
3	O	-0.622	-2.203	1.000	-0.697
4	H	0.340	-3.163	0.932	-0.670
5	C	-0.097	-2.282	-0.763	0.993
6	H	0.054	-3.363	-0.892	0.874
7	H	-0.004	-1.860	-1.728	1.284
8	H	0.040	-2.101	-0.048	1.800
9	C	0.063	-0.146	-0.101	-0.183
10	C	-0.174	0.690	-1.204	-0.382
11	H	0.091	0.262	-2.158	-0.675

12	C	-0.120	0.425	1.124	0.173
13	H	0.104	-0.217	1.986	0.305
14	C	-0.055	2.069	-1.091	-0.218
15	H	0.082	2.704	-1.956	-0.380
16	C	-0.094	1.805	1.239	0.333
17	H	0.091	2.236	2.197	0.603
18	C	-0.099	2.631	0.133	0.141
19	H	0.087	3.705	0.225	0.264

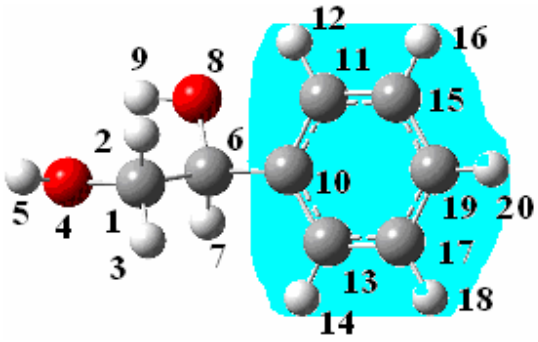
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
1,9	$1.34*10^5$

Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
1,9,10	429.3
1,9,12	431.5

Improper Torsion	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
9,10,12,1	263.1

Torsion	**c0, c1, c2, c3, c4, c5, c6**	**φ1, φ2, φ3, φ4, φ5, φ6**
3,1,9,12	13.89, -1.23, -11.27, 7.25, -20.95, -9.08, 14.67	6.28, 0.40, 1.45, 2.45, 1.44, 2.50
3,1,5,6	1.43, -22.33, 17.25, 29.26, -10.93, -1.11, 7.06	6.25, 0.58, 0.04, 3.73, 1.43, 2.14
4,3,1,5	11.35, -8.02, -14.28, 8.79, 5.51, -2.86, -2.87	6.28, 0.12, 6.28, 3.42, 0.81, 1.72

1-phenylethane-1,2-diol



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	0.327	-1.966	-0.553	0.713
2	H	-0.053	-1.739	0.078	1.580
3	H	-0.048	-1.658	-1.583	0.927
4	O	-0.704	-3.361	-0.492	0.389
5	H	0.430	-3.882	-0.619	1.187
6	C	0.266	-1.211	-0.022	-0.508
7	H	0.034	-1.430	-0.692	-1.352
8	O	-0.667	-1.658	1.293	-0.815
9	H	0.413	-2.622	1.270	-0.772
10	C	0.080	0.282	-0.013	-0.261
11	C	-0.077	0.927	1.131	0.216
12	H	0.103	0.353	2.038	0.358
13	C	-0.205	1.031	-1.175	-0.475
14	H	0.104	0.544	-2.064	-0.865
15	C	-0.114	2.296	1.109	0.480
16	H	0.089	2.787	2.005	0.843
17	C	-0.051	2.398	-1.199	-0.205
18	H	0.079	2.966	-2.106	-0.380
19	C	-0.090	3.034	-0.055	0.274
20	H	0.084	4.099	-0.069	0.478

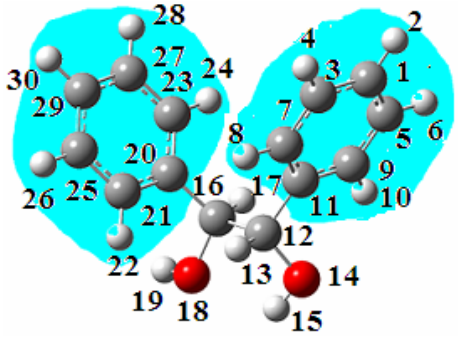
Stretch	k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)
6,10	$1.38*10^{5}$

Angle	k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)
6,10,11	430.8
6,10,13	427.6

Improper Torsion	k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)
10,11,13,6	256.0

Torsion	c0, c1, c2, c3, c4, c5, c6	φ1, φ2, φ3, φ4, φ5, φ6
8,6,10,11	19.24, 1.26, -8.27, 2.03, -29.90, -1.53, 22.90	0.00, 0.07, 1.34, 2.85, 1.54, 5.92
4,1,6,8	37.77, -13.08, -70.68, 18.05, 46.48, -7.19, -22.68	5.77, 0.13, 0.15, 3.30, 0.89, 1.75
9,8,6,1	8.50, -13.70, 41.59, -12.81, -36.60, -11.34, -29.24	3.93, 1.24, 0.73, 4.03, 1.22, 1.76
5,4,1,6	5.08, -13.71, -10.26, -18.49, 9.17, 11.80, 12.48	3.23, 0.06, 0.62, 3.65, 2.39, 6.08

1,2-diphenylethane-1,2-diol



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.113	3.214	-1.939	0.022
2	H	0.090	3.864	-2.797	0.159
3	C	-0.046	2.210	-1.969	-0.943
4	H	0.088	2.074	-2.849	-1.562
5	C	-0.068	3.388	-0.796	0.801
6	H	0.085	4.177	-0.761	1.544
7	C	-0.201	1.381	-0.863	-1.123
8	H	0.078	0.605	-0.890	-1.880
9	C	-0.159	2.558	0.308	0.622
10	H	0.111	2.708	1.207	1.208
11	C	0.181	1.542	0.282	-0.339
12	C	0.099	0.620	1.470	-0.514
13	H	0.086	0.185	1.444	-1.522
14	O	-0.608	1.372	2.666	-0.342
15	H	0.377	0.729	3.383	-0.255
16	C	0.099	-0.558	1.464	0.495
17	H	0.086	-0.127	1.440	1.503

18	O	-0.608	-1.208	2.732	0.282
19	H	0.377	-1.919	2.831	0.924
20	C	0.181	-1.515	0.304	0.324
21	C	-0.159	-2.435	0.287	-0.732
22	H	0.111	-2.477	1.128	-1.415
23	C	-0.201	-1.488	-0.774	1.213
24	H	0.078	-0.779	-0.773	2.034
25	C	-0.068	-3.305	-0.788	-0.895
26	H	0.085	-4.015	-0.788	-1.715
27	C	-0.046	-2.356	-1.853	1.049
28	H	0.088	-2.322	-2.682	1.746
29	C	-0.113	-3.266	-1.863	-0.006
30	H	0.090	-3.945	-2.699	-0.132

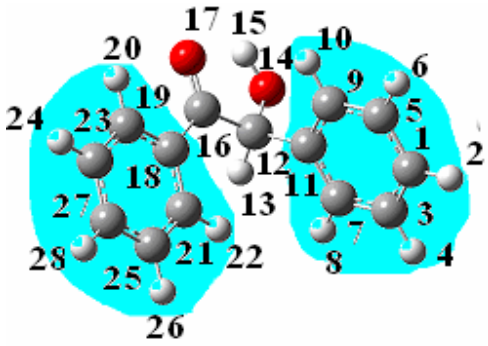
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
11,12/16,20	$1.36*10^5$

Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
12,11,7/16,20,23	443.5
12,11,9/16,20,21	445.8

Improper Torsion	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
11,7,9,12/20,23,21,16	263.3

Torsion	**c0, c1, c2, c3, c4, c5, c6**	**φ1, φ2, φ3, φ4, φ5, φ6**
14,12,11,7/ 23,20,16,18	31.74, 10.40, -30.52, 9.57, -20.03, -8.88, 9.11	5.91, 3.30, 3.28, 2.08, 5.46, 2.24
14,12,16,18	6.40, -11.54, -65.54, -20.35, 64.21, -5.56, 48.86	4.74, 1.47, 1.12, 4.90, 0.29, 4.09
15,14,12,16/ 19,18,16,12	15.32, -10.77, 35.90, 15.92, -37.97, -7.00, -23.47	3.77, 1.12, 3.80, 3.94, 3.53, 4.75

2-hydroxyl-1,2-diphenylethanone



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	-0.080	3.395	-1.650	0.785
2	H	0.084	4.086	-2.377	1.198
3	C	-0.116	3.088	-1.669	-0.573
4	H	0.094	3.545	-2.405	-1.225
5	C	-0.051	2.814	-0.688	1.612
6	H	0.081	3.053	-0.666	2.670
7	C	-0.106	2.201	-0.731	-1.100
8	H	0.0950	1.977	-0.740	-2.162
9	C	-0.198	1.929	0.246	1.084
10	H	0.119	1.490	0.997	1.733
11	C	0.121	1.609	0.230	-0.278
12	C	0.315	0.678	1.282	-0.873
13	H	0.041	0.424	0.983	-1.896
14	O	-0.691	1.307	2.546	-0.929
15	H	0.400	0.833	3.109	-0.294
16	C	0.340	-0.631	1.408	-0.063
17	O	-0.448	-0.804	2.439	0.567
18	C	-0.043	-1.643	0.316	-0.053
19	C	0.009	-2.776	0.474	0.761
20	H	0.075	-2.863	1.374	1.356
21	C	-0.135	-1.520	-0.842	-0.833
22	H	0.040	-0.654	-0.990	-1.464
23	C	-0.111	-3.759	-0.506	0.797

24	H	0.096	-4.629	-0.376	1.430
25	C	-0.029	-2.508	-1.822	-0.798
26	H	0.085	-2.403	-2.714	-1.404
27	C	-0.077	-3.627	-1.656	0.017
28	H	0.090	-4.394	-2.422	0.044

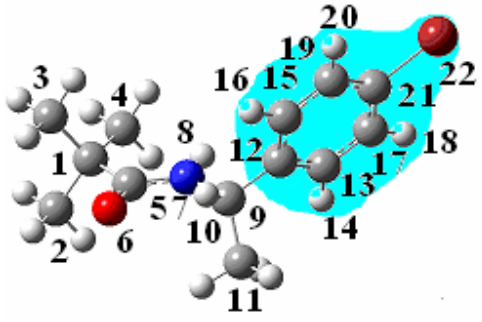
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
11,12	$1.40*10^5$
16,18	$1.49*10^5$

Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
12,11,7	500.1
12,11,9	353.8
16,18,19	498.1
16,18,21	486.3

Improper Torsion	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
11,7,9,12	231.2
18,16,19,21	268.4
16,12,17,18	328.0

Torsion	**c0, c1, c2, c3, c4, c5, c6**	**φ1, φ2, φ3, φ4, φ5, φ6**
7,11,12,14	17.70, -0.49, -14.56, 1.23, -30.30, -0.71, 16.08	2.55, 1.70, 2.11, 3.57, 2.18, 3.75
19,18,16,17	48.58, -3.41, -34.89, -2.11, -17.39, -0.41, -10.70	3.16, 0.35, 4.18, 2.35, 1.27, 3.00
14,12,16,17	49.11, 12.19, -55.39, 33.91, -32.38, 19.43, 18.92	1.55, 0.81, 4.95, 5.65, 1.61, 1.02
15,14,12,16	9.92, -8.40, -9.62, -8.38, 5.25, -1.94, -2.25	4.30, 0.21, 0.98, 3.79, 1.36, 3.59

N-(1-(4-bromophenyl)ethyl)pivalamide



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	C	0.325	-0.025	-0.044	-0.111
2	CH3	-0.077	-0.985	-0.205	1.087
3	CH3	-0.082	1.435	-0.089	0.387
4	CH3	-0.120	-0.260	-1.188	-1.113
5	C	0.603	-0.262	1.371	-0.692
6	O	-0.581	-0.121	2.366	0.008
7	N	-0.643	-0.631	1.467	-2.001
8	H	0.271	-0.765	0.631	-2.546
9	C	0.358	-0.909	2.764	-2.641
10	H	0.033	-1.334	3.394	-1.857
11	CH3	-0.070	0.382	3.436	-3.127
12	C	0.063	-1.961	2.570	-3.723
13	C	-0.124	-1.636	2.091	-4.996
14	H	0.114	-0.604	1.879	-5.253
15	C	-0.204	-3.304	2.844	-3.444
16	H	0.125	-3.580	3.216	-2.464
17	C	-0.080	-2.617	1.887	-5.965
18	H	0.084	-2.350	1.519	-6.948
19	C	-0.018	-4.298	2.652	-4.401
20	H	0.075	-5.332	2.874	-4.171
21	C	0.049	-3.943	2.171	-5.657
22	Br	-0.101	-5.300	1.901	-6.985

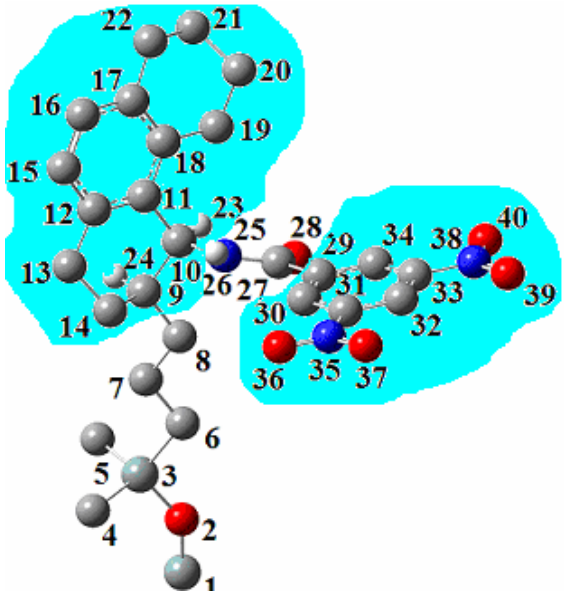
Stretch	k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)
9,12	$1.12*10^{5}$

Angle	k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)
9,12,13	420.4
9,12,15	466.8

Improper Torsion	k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)
5,1,6,7	226.8
7,5,8,9	28.5
12,9,13,15	200.9

Torsion	c0, c1, c2, c3, c4, c5, c6	φ1, φ2, φ3, φ4, φ5, φ6
7,9,12,13	17.85, -3.53, -33.7, 1.77, 27.11, 0.87, -14.32	0.16, 0.24, 0.23, 0.27, 3.45, 1.96
5,7,9,12	33.14, 9.08, -50.33, 14.98, 32.44, -13.60, -28.88	0.39, 0.30, 0.28, 0.21, 0.42, 1.82
1,5,7,9	32.71, 20.01, -400.20, -39.70, 325.18, 47.52, 271.94	0.60, 0.59, 1.19, 0.33, 1.41, 1.14
2,1,5,7	3.37, 0.47, -6.23, 2.21, 16.41, -3.75, -11.06	0.00, 1.07, 1.02, 1.08, 0.03, 4.21

CSP1a



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	Si	0.0	0.000	0.000	0.000
2	O	0.0	1.044	-0.379	-1.191
3	Si	0.0	2.320	0.193	-2.094
4	CH_3	0.0	1.804	1.755	-3.004
5	CH_3	0.0	3.755	0.573	-0.944
6	CH_2	-0.069	2.752	-1.186	-3.292
7	CH_2	0.099	3.813	-0.839	-4.337
8	CH_2	0.012	4.268	-2.062	-5.141
9	C	-0.085	5.082	-1.650	-6.242
10	C	0.387	5.365	-2.733	-7.307
11	C	-0.116	6.158	-2.203	-8.501
12	C	0.074	6.975	-1.074	-8.345
13	CH_2	0.005	7.017	-0.262	-7.062

14	CH_2	0.257	6.397	-0.963	-5.854
15	CH	-0.188	7.735	-0.635	-9.436
16	CH	-0.156	7.669	-1.288	-10.652
17	C	0.031	6.839	-2.402	-10.835
18	C	-0.136	6.080	-2.871	-9.752
19	CH2	-0.005	5.136	-4.051	-9.933
20	CH2	0.089	5.383	-4.866	-11.209
21	CH2	0.101	5.571	-3.947	-12.415
22	CH2	0.015	6.804	-3.068	-12.199
23	H	0.009	4.407	-3.136	-7.639
24	H	-0.018	4.447	-0.906	-6.747
25	N	-0.496	6.104	-3.891	-6.751
26	H	0.210	7.103	-3.852	-6.886
27	C	0.560	5.598	-5.111	-6.437
28	O	-0.362	4.430	-5.453	-6.555
29	C	0.051	6.621	-6.085	-5.889
30	CH	-0.039	6.429	-7.464	-5.802
31	C	0.112	7.450	-8.264	-5.300
32	CH	-0.021	8.664	-7.747	-4.877
33	C	0.079	8.829	-6.371	-4.983
34	CH	-0.119	7.843	-5.536	-5.488
35	N	0.662	7.233	-9.731	-5.220
36	O	-0.415	6.156	-10.160	-5.611
37	O	-0.396	8.141	-10.411	-4.768
38	N	0.674	10.120	-5.779	-4.541
39	O	-0.405	10.974	-6.547	-4.123
40	O	-0.402	10.237	-4.566	-4.624

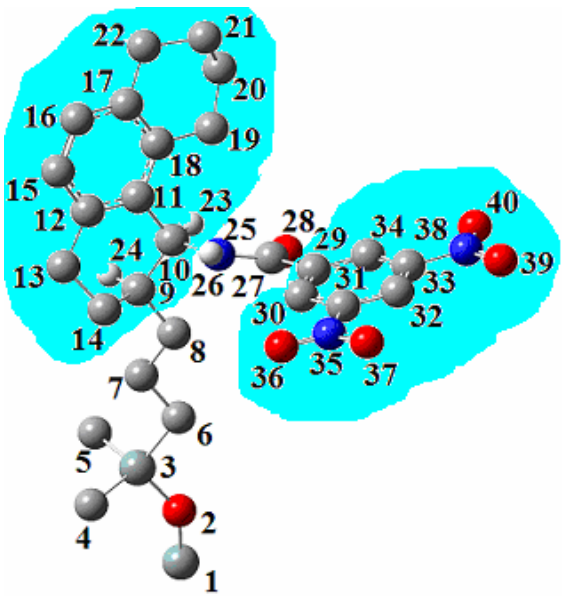
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
8,9	$1.18*10^5$
10,25	$1.16*10^5$
27,29	$1.24*10^5$

Angle	k (kJ· mol^{-1} · rad^{-2})
8,9,10	421.3
8,9,14	363.1
8,9,24	262.5
9,10,25	297.3
11,10,25	394.4
23,10,25	297.3
27,29,30	425.2
27,29,34	453.8

Improper Torsion	k (kJ· mol^{-1} · rad^{-2})
25,10,26,27	50.3
27,28,29,25	349.5
29, 27,34,30	283.7

Torsion	c0, c1, c2, c3, c4, c5, c6	φ1, φ2, φ3, φ4, φ5, φ6
7,8,9,10	0.00, -20.55, 29.38, -14.05, -11.54, -15.67, 14.09	3.71, 0.51, 1.12, 3.78, 0.84, 2.06
9,10,25,27	22.96, -49.56, -45.38, -37.94, 83.05, -26.11, -46.99	4.03, 1.02, 1.24, 4.08, 0.36, 4.18
25,27,29,30	2.74, -1.43, -3.99, 1.27, 16.64, -1.58, 2.97, 3.16	0.22, 1.50, 4.68, 1.28, 3.29

CSP1b



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	Si	0.0	0.000	0.000	0.000
2	O	0.0	-0.035	-1.625	0.115
3	Si	0.0	-0.588	-2.895	1.037
4	CH_3	0.0	0.091	-2.757	2.784
5	CH_3	0.0	-2.465	-2.841	1.081
6	CH_2	-0.077	0.030	-4.448	0.182
7	CH_2	0.099	-0.230	-5.757	0.929
8	CH_2	0.031	0.106	-6.992	0.086
9	C	-0.089	-0.157	-8.182	0.834
10	C	0.303	0.587	-9.380	0.201
11	C	-0.126	0.617	-10.617	1.098
12	C	0.111	-0.347	-10.766	2.106
13	CH_2	-0.012	-1.371	-9.694	2.427
14	CH_2	0.066	-1.530	-8.638	1.335
15	CH	-0.089	-0.339	-11.932	2.881
16	CH	-0.045	0.605	-12.918	2.668
17	C	0.079	1.587	-12.779	1.680
18	C	-0.136	1.597	-11.624	0.884
19	CH2	0.136	2.636	-11.472	-0.216
20	CH2	-0.051	3.835	-12.418	-0.085
21	CH2	0.077	3.369	-13.845	0.195
22	CH2	-0.026	2.619	-13.880	1.526
23	H	-0.018	1.603	-9.054	-0.023
24	H	0.030	0.444	-7.900	1.711
25	N	-0.379	-0.009	-9.765	-1.097
26	H	0.182	-0.765	-10.430	-1.036
27	C	0.504	0.496	-9.517	-2.333
28	O	-0.532	1.557	-8.954	-2.559
29	C	0.060	-0.353	-9.991	-3.491
30	CH	-0.035	-1.707	-10.311	-3.373
31	C	0.102	-2.420	-10.684	-4.507
32	CH	0.002	-1.840	-10.757	-5.764
33	C	0.064	-0.494	-10.426	-5.851

34	CH	0.044	0.254	-10.036	-4.750
35	N	0.669	-3.860	-11.014	-4.367
36	O	-0.414	-4.348	-10.932	-3.247
37	O	-0.399	-4.463	-11.345	-5.376
38	N	0.681	0.163	-10.489	-7.183
39	O	-0.407	-0.530	-10.814	-8.136
40	O	-0.403	1.353	-10.216	-7.236

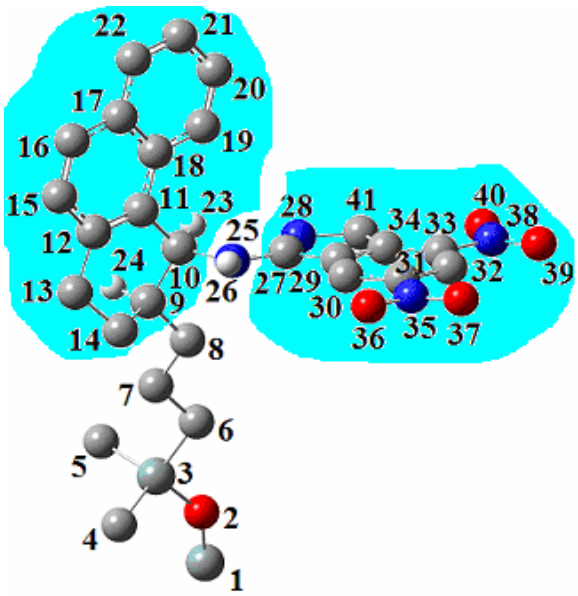
Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
8,9	$1.18*10^5$
10,25	$1.24*10^5$
27,29	$1.24*10^5$

Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
8,9,10	422.1
8,9,14	352.2
8,9,24	257.1
9,10,25	420.0
11,10,25	394.3
23,10,25	300.0
27,29,30	416.8
27,29,34	446.2

Improper Torsion	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
25,10,26,27	37.3
27,28,29,25	334.6
29,27,34,30	264.3

Torsion	**c0, c1, c2, c3, c4, c5, c6**	**φ1, φ2, φ3, φ4, φ5, φ6**
7,8,9,10	12.80, -21.23, -12.25, -18.65, 28.38, -10.52, -14.50	3.58, 0.53, 0.97, 3.47, 0.85, 3.51
9,10,25,27	16.36, -7.75, 79.87, 37.84, -48.33, -37.78, -50.36	4.83, 0.23, 0.29, 3.87, 0.45, 3.07
25,27,29,30	31.66, -2.20, -29.46, -2.57, -22.40, 1.39, -1.71	3.62, 0.41, 0.82, 5.50, 0.73, 1.75

CSP2



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	Si	0.0	0.000	0.000	0.000
2	O	0.0	-0.260	-1.602	0.140
3	Si	0.0	0.467	-3.031	0.571
4	CH_3	0.0	0.996	-2.945	2.363
5	CH_3	0.0	-0.741	-4.439	0.345
6	CH_2	-0.072	1.967	-3.317	-0.508
7	CH_2	0.047	-0.092	-5.789	0.665
8	CH_2	0.125	-1.055	-6.969	0.478
9	C	-0.217	-0.428	-8.348	0.748
10	C	0.635	-1.293	-9.501	0.190
11	C	-0.341	-0.705	-10.877	0.487
12	C	0.174	0.183	-11.052	1.535
13	CH_2	-0.049	0.683	-9.900	2.385
14	CH_2	0.096	-0.123	-8.611	2.226
15	CH	-0.081	0.695	-12.348	1.817
16	CH	-0.074	0.331	-13.437	1.075
17	C	0.180	-0.571	-13.300	-0.011
18	C	0.175	-1.100	-12.005	-0.314
19	CH	-0.071	-1.996	-11.899	-1.413
20	CH	-0.004	-2.344	-13.002	-2.160

21	CH	0.038	-1.820	-14.276	-1.853
22	CH	-0.123	-0.950	-14.417	-0.798
23	H	0.024	0.519	-8.398	0.193
24	H	-0.082	-1.373	-9.359	-0.887
25	N	-0.734	-2.680	-9.451	0.702
26	H	0.342	-2.867	-9.988	1.534
27	C	0.483	-3.752	-9.101	-0.056
28	N	-0.616	-3.698	-8.646	-1.266
29	C	0.072	-5.141	-9.195	0.456
30	CH	-0.039	-5.658	-9.614	1.674
31	C	0.036	-7.043	-9.573	1.827
32	CH	0.041	-7.898	-9.135	0.824
33	C	-0.068	-7.340	-8.719	-0.381
34	C	-0.007	-5.963	-8.740	-0.585
35	N	0.719	-7.630	-10.013	3.116
36	O	-0.443	-6.855	-10.391	3.984
37	O	-0.404	-8.846	-9.971	3.229
38	N	0.790	-8.254	-8.251	-1.445
39	O	-0.432	-9.453	-8.258	-1.212
40	O	-0.457	-7.742	-7.885	-2.496
41	CH2	0.336	-5.053	-8.367	-1.727

Stretch	**k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)**
8,9	$1.17*10^5$
10,25	$1.15*10^5$
27,29	$2.03*10^5$

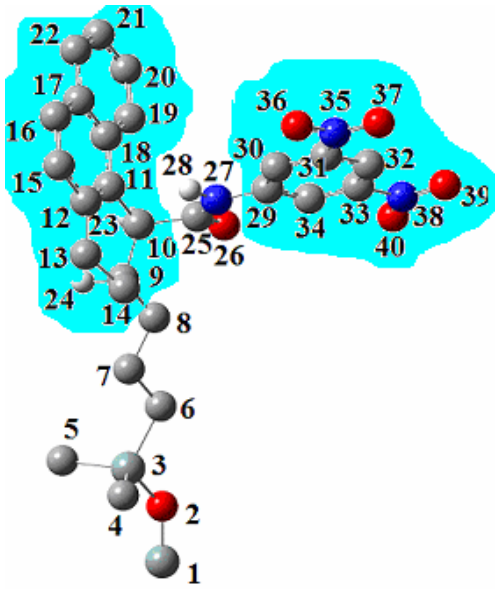
Angle	**k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)**
8,9,10	413.4
8,9,14	351.7
8,9,24	262.2
9,10,25	413.6
11,10,25	399.7
23,10,25	298.5
25,27,28	538.4

25,27,29	495.9

Improper Torsion	k (kJ· mol^{-1} · rad^{-2})
25,10,26,27	34.2
27,25,29,28	288.3

Torsion	c0, c1, c2, c3, c4, c5, c6	φ1, φ2, φ3, φ4, φ5, φ6
7,8,9,10	13.71, -21.03, -15.76, 15.44, 38.97, -12.89, -19.44	2.67, 0.48, 2.16, 3.59, 0.85, 3.61
9,10,25,27	37.05, -7.66, 46.22, 6.18, -55.27, 5.67, -35.90	3.51, 1.78, 1.20, 4.46, 2.26, 2.01
10,25,27,28	14.13, 61.58, 46.78, 52.67, 153.55, -55.03, -92.30	3.23, 3.09, 1.25, 1.56, 1.22, 1.61

CSP3



ATOM NO.	ATOM/ GROUP	q (|e|)	ATOMIC POSITION x(Å), y (Å), z (Å)		
1	Si	0.000	0.000	0.000	0.000
2	O	0.000	1.044	-0.379	-1.1917
3	Si	0.000	2.320	0.193	-2.094
4	CH_3	0.000	2.752	-1.186	-3.292
5	CH_3	0.000	1.804	1.755	-3.004
6	CH_2	-0.056	3.755	0.573	-0.944
7	CH_2	0.077	3.813	-0.839	-4.337

8	CH_2	0.012	4.268	-2.062	-5.141
9	C	0.013	5.082	-1.650	-6.242
10	C	0.177	5.365	-2.733	-7.307
11	C	-0.218	6.158	-2.203	-8.501
12	C	0.103	6.975	-1.074	-8.345
13	CH_2	0.046	7.017	-0.262	-7.062
14	CH_2	-0.014	6.397	-0.963	-5.854
15	CH	-0.054	7.735	-0.635	-9.436
16	CH	-0.083	7.669	-1.288	-10.652
17	C	0.179	6.839	-2.402	-10.835
18	C	0.099	6.080	-2.871	-9.752
19	CH	-0.086	5.136	-4.051	-9.933
20	CH	0.026	5.383	-4.866	-11.20
21	CH	0.030	5.571	-3.947	-12.41
22	CH	-0.102	6.804	-3.068	-12.19
23	H	-0.020	4.407	-3.136	-7.639
24	H	-0.014	4.447	-0.906	-6.747
25	C	0.659	6.104	-3.891	-6.751
26	O	-0.572	7.103	-3.852	-6.886
27	N	-0.595	5.598	-5.111	-6.437
28	H	0.242	4.430	-5.453	-6.555
29	C	0.440	6.621	-6.085	-5.889
30	CH	-0.044	6.429	-7.464	-5.802
31	C	0.072	7.450	-8.264	-5.300
32	CH	-0.022	8.664	-7.747	-4.877
33	C	0.086	8.829	-6.371	-4.983
34	CH	-0.110	7.843	-5.536	-5.488
35	N	0.706	7.233	-9.731	-5.220
36	O	-0.418	6.156	-10.160	-5.611
37	O	-0.420	8.141	-10.411	-4.768
38	N	0.707	10.120	-5.779	-4.541
39	O	-0.414	10.974	-6.547	-4.123
40	O	-0.431	10.237	-4.566	-4.624

Stretch	k ($kJ \cdot mol^{-1} \cdot nm^{-2}$)
8,9	$1.25*10^5$
10,25	$1.54*10^5$
27,29	$1.40*10^5$

Angle	k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)
8,9,10	398.6
8,9,14	356.9
8,9,24	262.4
9,10,25	386.9
11,10,25	396.6
27,29,30	521.1
27,29,34	541.8

Improper Torsion	k ($kJ \cdot mol^{-1} \cdot rad^{-2}$)
25,10,26,27	317.2
27,28,29,25	57.6
29,27,34,30	311.9

Torsion	c0, c1, c2, c3, c4, c5, c6	φ1, φ2, φ3, φ4, φ5, φ6
7,8,9,10	24.32, -14.38, -14.04, -5.76, -32.56, -30.81, 26.99	3.78, 4.64, 3.59, 3.11, 0.86, 3.21
9,10,25,27	4.06, 7.64, 24.48, -3.40, -7.22, -4.04, -10.04	5.56, 6.25, 4.51, 6.12, 1.01, 3.18
25,27,29,30	80.04, -3.13, -65.92, -4.72, -44.60, 5.18, -37.40	3.27, 0.68, 1.52, 2.80, 1.41, 2.17