

**ENANTIO-SELECTIVE MECHANISM OF THE POLY-PROLINE
CHIRAL STATIONARY PHASE:
A MOLECULAR DYNAMICS STUDY**

By

Mohammad Ashtari

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

(January, 2013)

Copyright © Mohammad Ashtari, 2013

Abstract

Poly-proline-based chiral stationary phases are relatively new stationary phases and have shown to be competitive to other commercially available chiral stationary phases for high performance liquid chromatography (HPLC). The conformational studies, solvation properties and enantio-selective mechanism of this chiral stationary phase are the main focus of this thesis. Semi-flexible models are developed based on an extensive series of *ab initio* calculations for proline selectors from di- to hexa-proline and a series of six chiral analytes. Then molecular dynamics simulations are performed to study the solvation, conformational preferences at the interface, and the selectivity.

The solvation and conformational preferences of poly-proline selectors at the interface are examined in a normal phase *n*-hexane/-2propanol and a reverse phase water/methanol solvent. We noticed a significant difference between conformational preferences of poly-proline chains in these solvents indicating the effect of solvent polarity and hydrogen bonding on the relative stabilities of poly-proline conformers. Solvent partitioning occurs at the interface and this creates a polarity gradient between the stationary phase and the bulk that encourages analyte docking at the interface. Hydrogen bonding to the poly-proline selectors is shown to be a function of solvent composition and poly-proline conformation at the interface.

The selectivity of the poly-proline chains was studied by molecular dynamics simulations of chiral analytes docking into the interface. The selectivity factors for a set of enantiomers were predicted successfully. Enantio-resolution has been shown to mostly happen with hydrogen bonding to poly-proline carbonyl oxygens located close to the interface. Steric interactions and conformational flexibility of the analytes are also contributing factors for enantio-resolution.

Acknowledgments

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. Nicholas Mosey, and Gang Wu, for being on my supervisory committee and for many helpful suggestions. I thank Dr. Marcel Nooijen for helpful discussions. I would also like to thank the group members in Dr. Natalie Cann's lab: Drs. Chunfeng Zhao, Shihao Wang, Rodica Pecheanu, and Christopher Daub, and also Kyle Hall, for many interesting discussions, shared experiences and friendship.

I thank my wife, Parisa Akhshi, for all the love, scientific discussions and support she has offered me in the past years. She has always been supporting and believing in me.

Financial support from NSERC and Queen's University, and the use of computer resources of compute Canada, particularly HPCVL, SHARCNET, CLUMEQ 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.

Mohammad Ashtari

January, 2013

Table of Contents

Abstract.....	ii
Acknowledgments.....	iii
Statement of Originality.....	iv
List of Abbreviations and Symbols.....	xxii
Chapter 1 Introduction.....	1
1.1 Chirality.....	1
1.2 Chiral High Performance Liquid Chromatography.....	2
1.3 Theoretical Studies on Chiral Stationary Phases.....	8
1.4 Poly-proline: The Biological Importance.....	10
1.5 Proline Stationary Phases.....	11
1.6 Theoretical Studies of Chiral Interfaces.....	13
1.7 Research Direction.....	15
Chapter 2 Methods and Models.....	17
2.1 Introduction.....	17
2.2 <i>Ab initio</i> Methods.....	17
2.2.1 <i>Ab initio</i> Calculations.....	17
2.2.2 Basis Functions and Basis Sets.....	19
2.2.3 Density Functional Theory.....	21
2.2.4 DFT Functionals.....	22
2.2.5 Second order Moller-Plesset Perturbation Theory.....	24
2.2.6 Coupled Cluster Methods.....	25
2.2.7 Solvation method.....	28
2.3 Molecular Dynamics Simulations.....	28
2.3.1 Introduction to Molecular Dynamics Simulations.....	28
2.3.2 Force Field.....	30
2.3.3 Periodic Boundary Conditions.....	35
2.3.4 Truncation of the Lennard-Jones Potential.....	37
2.3.5 Ewald Summations.....	39
2.3.6 Integrating the Equations of Motion.....	41

2.3.7	Nosé-Hoover Thermostat.....	47
2.3.8	Quaternions for Rotation of Rigid Moieties	48
2.4	Method Improvement.....	50
2.4.1	Equilibrating the Surface	50
2.4.2	Impact of Electrostatics.....	59
Chapter 3	Conformational Studies of Poly-proline Chains.....	62
3.1	Introduction.....	62
3.2	Computational Methods.....	67
3.3	Results and Discussions	68
3.3.1	Stability Studies in the Gas Phase and in Continuum Solvents	68
3.3.2	Backbone Flexibility	89
3.3.3	Puckering Studies.....	94
3.3.4	Other Stable Conformers	98
3.4	Conclusion	99
Chapter 4	Solvation Studies of Di-proline Chiral Stationary Phases.....	101
4.1	Introduction.....	101
4.1.1	The Di-proline Selectors	101
4.1.2	Force Field Development.....	103
4.1.3	The Chiral Interface	109
4.1.4	The Assessment of Interfacial Structure	111
4.2	Simulation Details.....	112
4.3	Results.....	115
4.3.1	The structure of the TMA-(Pro) ₂ -N(CH ₃)-tether and BOC-(Pro) ₂ -N(CH ₃)-tether Chiral Interface	115
4.3.2	Solvation and Hydrogen Bonding at the TMA-(Pro) ₂ -N(CH ₃)-tether Chiral Interface	121
4.3.3	Solvation and Hydrogen Bonding at the BOC-(Pro) ₂ -N(CH ₃)-tether Chiral Interface.....	127
4.4	Conclusions.....	132
Chapter 5	Solvation Studies of Tri-, Tetra-, Penta-, and Hexa-proline Chiral Stationary Phases.....	134
5.1	Introduction.....	134
5.1.1	Force Field Development.....	136

5.2 The Chiral Interface	139
5.3 Simulation Details.....	140
5.4 Simulation Results	142
5.4.1 Interfacial Poly-proline	142
5.4.2 Solvation and Hydrogen Bonding.....	152
5.4.3 Terminal Group Effect.....	160
5.5 Conclusions.....	161
Chapter 6 Docking Studies of Poly-proline Chiral Stationary Phases.....	164
6.1 Introduction.....	164
6.2 Theoretical Details.....	166
6.2.1 Potentials.....	166
6.2.2 The Simulation Cell.....	167
6.2.3 MD Simulations	167
6.3 Results and Discussions.....	169
6.3.1 Distribution at the Interface	169
6.3.2 The Selectivity	183
6.3.3 The Docking Mechanism.....	197
6.4 Conclusions.....	207
Chapter 7 Conclusions and Outlook	209
Bibliography	213
Appendix A Interaction Potentials for Chiral Selectors and Analytes.....	226

List of Tables

Table 3-1. Dihedral angles, relative energies and free energies of TMA-(Pro)-N(CH ₃) ₂ and BOC-(Pro)-N(CH ₃) ₂ . For gas phase calculations and solvation studies, CCSD//mPW2PLYP-D and CCSD//mPW2PLYP-D/COSMO are used, respectively. All energies (for gas phase calculations) and free energies (for solvent phase calculations) are in kJ/mol, and all angles are in degree. 6-311G(d,p) basis set is used for all calculations.	69
Table 3-2. Dihedral angles, relative energies and free energies of TMA-(Pro) ₂ -N(CH ₃) ₂ and BOC-(Pro) ₂ -N(CH ₃) ₂ . For gas phase and solvation studies, LPNO-CCSD//mPW2PLYP-D and LPNO-CCSD//mPW2PLYP-D/COSMO methods are used, respectively. All energies (for gas phase calculations) and free energies (for solvent phase calculations) are in kJ/mol, and all angles are in degree. 6-311G(d,p) basis set is used for all calculations.	70
Table 3-3. Conformers TMA-(Pro) ₂ -N(CH ₃) ₂ and BOC-(Pro) ₂ -N(CH ₃) ₂ . The first column identifies the conformation about the amide bonds. Relative energies, in kJ/mol, are reported relative to the lowest energy conformer for each method. The energies in brackets, for CCSD and LPNO-CCSD, are relative to the overall minimum energy structure. Results are provided for: 1: B3LYP/6-311G(d,p); 2: B97-D/6-311G(d,p); 3: mPW2PLYP-D/cc-pVTZ; 4: MP2/6-311G(d,p); 5: CCSD/6-311G(d,p)//B3LYP/6-311G(d,p); 6: CCSD/6-311G(d,p)//B97-D/6-311G(d,p); 7: LPNO-CCSD/6-311G(d,p)//B3LYP/6-311G(d,p); 8: LPNO-CCSD/6-311G(d,p)//B97-D/6-311G(d,p) calculations.	71
Table 3-4. The conformer energies for TMA-(Pro) ₃ -N(CH ₃) ₂ . All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets provide the LPNO-CCSD energy relative to the overall lowest energy conformer. The lowest energy structure, for each conformer, is identified with a shaded cell.	74
Table 3-5. The conformer energies for BOC-(Pro) ₃ -N(CH ₃) ₂ . All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets are the LPNO-CCSD energy relative to the overall lowest energy conformer. The lowest energy structure, for each conformer, is identified with a shaded cell.	77

Table 3-6. Relative stabilities of TMA-(Pro) ₃ -N(CH ₃) ₂ in continuum solvents. The LPNO-CCSD//mPW2PLYP-D/COSMO/6-311G(d,p) methods is used for all calculations. All free energies are in kJ/mol.	78
Table 3-7. The conformer energies for TMA-(Pro) ₄ -N(CH ₃) ₂ . All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets provide the LPNO energy relative to the overall lowest energy conformer (TTCT structure obtained from mPW2PLYP-D optimization). The lowest energy structure, for each conformer, is identified with a shaded cell.	80
Table 3-8. Relative stabilities of TMA-(Pro) ₄ -N(CH ₃) ₂ in continuum solvents. LPNO-CCSD//mPW2PLYP-D/COSMO/6-311G(d,p) methods is used for all calculations. All free energies are in kJ/mol.	82
Table 3-9. The conformer energies for TMA-(Pro) ₅ -N(CH ₃) ₂ . All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets provide the LPNO energy relative to the overall lowest energy conformer (TTCCT structure obtained from RI-MP2 optimization). The lowest energy structure, for each conformer, is identified with a shaded cell.	83
Table 3-10. Relative stabilities of TMA-(Pro) ₅ -N(CH ₃) ₂ in continuum solvents. LPNO-CCSD//mPW2PLYP-D/COSMO/6-311G(d,p) methods is used for calculations. All free energies are in kJ/mol.	85
Table 3-11. The conformer energies for TMA-(Pro) ₆ -N(CH ₃) ₂ . All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method.	86
Table 4-1. Conformational preferences of TMA-(Pro) ₂ -N(CH ₃)-tether and BOC-(Pro) ₂ -N(CH ₃)-tether at the interface. The percentage of selectors adopting the specific conformer is given....	118
Table 4-2. Hydrogen bonding statistics for the TMA-(Pro) ₂ -N(CH ₃)-tether and BOC-(Pro) ₂ -N(CH ₃)-tether interfaces with <i>n</i> -hexane/2-propanol and water/methanol. The probability of one, or more, H-bonds as a function of conformation are given. All probabilities are shown in percent.	123
Table 4-3. Hydrogen bonding statistics for the TMA-(Pro) ₂ -N(CH ₃)-tether and BOC-(Pro) ₂ -N(CH ₃)-tether interfaces with <i>n</i> -hexane/2-propanol and water/methanol. The H-bonding location and solvent identity, for water/methanol, are given. The probability of a hydrogen bond at the	

specified carbon, in the given conformer, is presented In water/methanol, the total probability is given with the probability for methanol given in brackets. All probabilities are shown in percent.

.....	123
Table 5-1. Average conformer populations at the interface. Only conformers with relative population of more than 5% are listed.	144
Table 5-2. Hydrogen bonding statistics for carbonyl oxygens at the interface. The last column gives the percentage of selectors with an H-bond at the specified oxygen. In the water/methanol, the contribution from methanol is listed in brackets.....	156
Table 5-3. Hydrogen bonding statistics for poly-proline selectors. The last column gives the percentage of selectors with number of H-bonds. Only probabilities above 1% are listed.	158
Table 6-1. Hydrogen bonding statistics and selectivity factors for analyte AnC-CH ₃ . Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.69% of O(8) oxygens of TMA-di-proline in <i>n</i> -hexane/2-propanol hydrogen bond to the R enantiomer while 0.28% of them hydrogen bond to the S enantiomer.. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].....	186
Table 6-2. Hydrogen bonding statistics and selectivity factors for analyte AnS-CH ₃ . Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.81% of O(8) oxygens of TMA-di-proline in <i>n</i> -hexane/2-propanol hydrogen bond to the R enantiomer while 0.57% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].....	188
Table 6-3. Hydrogen bonding statistics and selectivity factors for analyte Ph-CF ₃ . Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.7% of O(8) oxygens of TMA-di-proline in <i>n</i> -hexane/2-propanol hydrogen bond to the R enantiomer while 0.7% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].....	190
Table 6-4. Hydrogen bonding statistics and selectivity factors for analyte Ph-CH ₃ . Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.66% of O(8) oxygens of TMA-di-proline in <i>n</i> -hexane/2-propanol hydrogen bond to the R enantiomer while 0.66% of	

them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47]..... 191

Table 6-5. Hydrogen bonding statistics and selectivity factors for analyte Ph-C₂H₅. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.42% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.44% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47]..... 192

Table 6-6. Hydrogen bonding statistics and selectivity factors for analyte AnC-CF₃. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.46% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.2% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47]..... 194

List of Figures

Figure 1-1. A typical representation of the separation in chromatography where two enantiomers shown by yellow and red colors are separated.....	4
Figure 1-2. A schematic representation of silica pores. The figure is adopted from ref [22].	4
Figure 1-3. Examples of common chiral stationary phases. a) polysaccharide based, b) whelk-O1, c) crown based, d) protein based, e) β -cyclo-dextrin, f) Vancomycin CSPs.	6
Figure 1-4. Proline monomer (a) and decamer (b). Note that in a poly-proline, the backbone nitrogen does not have any hydrogen atom and the backbone is unable to participate in hydrogen bonding as a donor.....	11
Figure 2-1. Illustration of intra-molecular motions.....	31
Figure 2-2. A typical stretching potential between two atoms i and j.....	33
Figure 2-3. LJ potential as a function of the inter-atomic distance between two atoms i and j. ...	35
Figure 2-4. A schematic representation of a three dimensional periodic boundary conditions. Picture is taken from: http://isaacs.sourceforge.net/phys/pbc.html . [Last Accessed at 2012].	37
Figure 2-5. A schematic representation of the cut-off and the truncated-shifted LJ potential.....	38
Figure 2-6. Illustration of the point charge treatment in the Ewald method.	41
Figure 2-7. Representation of Euler angles that connect the space-fixed frame to the body-fixed frame. Picture is taken from at: http://hepweb.ucsd.edu/ph110b/110b_notes/node31.html [Last Accessed at 2012].	49
Figure 2-8. The simulation cell, replicated in three dimensions but shown in 2D.....	51
Figure 2-9. Representation of selectors used for method development. A) TMA-(Pro) ₂ -N(CH ₃)-tether and C ₁₈ H ₃₇	54
Figure 2-10. The impact of equilibration method on the surface distribution of backbone carbons in TMA-(Pro) ₂ -N(CH ₃)-tether. The simulations begin from all <i>trans</i> selector conformation with a 150 <i>ps</i> equilibration period that includes annealing at 600 K (long dashed) or of torsion barrier reduction (short dashed). The distributions obtained without annealing or barrier reduction are shown by solid lines.....	56
Figure 2-11. The impact of equilibration method on the surface distribution of C(19), the topmost carbon in surface-bound C ₁₈ H ₃₇ Si. The simulations begin from all <i>trans</i> conformation and a 150 <i>ps</i> equilibration period that includes 90 <i>ps</i> of annealing at 600 K (long dashed) or of torsion	

barrier reduction (short dashed). The distribution obtained without measures to expedite the selector conformational equilibration is shown by a solid line..... 58

Figure 2-12. The impact of the treatment of electrostatics on the surface-solvent structure distributions for the TMA-(Pro)₂-N(CH₃)-tether interface with *n*-hexane/2-propanol. Surface distributions for C(7), C(9), and C(11) are shown in (a), (b), and (c), respectively. Radial distributions between the alcohol H and the carbonyl oxygens, O(8), O(10), and O(12) are shown in (d), (e), and (f), respectively. The Ewald correction appropriate for infinite plates (solid line) and for rectangular prism (long dashes) are shown. The short-dashed line gives distributions for a spherical cutoff. 61

Figure 3-1. a) A proline monomer with labeling of the ring atoms. b) TMA-(Pro)₂-N(CH₃)₂ with the backbone and endo-cyclic dihedral angles. The endo-cyclic dihedral angle is shown by χ which is defined between N, C ^{α} , C ^{β} , and C ^{γ} . Here the all trans conformer of TMA terminated diproline is shown. c) The PPII conformer of deca-proline. d) The PPI conformer of deca-proline. 63

Figure 3-2. Stable conformers from B3LYP/6-311G(d,p) calculations of TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂..... 72

Figure 3-3. The lowest energy conformers of TMA-(Pro)₃-N(CH₃)₂. The structures are obtained from mPW2PLYP-D/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations. 75

Figure 3-4. The lowest energy conformers of BOC-(Pro)₃-N(CH₃)₂. The structures are obtained from mPW2PLYP-D/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations. 78

Figure 3-5. The lowest energy conformers of TMA-(Pro)₄-N(CH₃)₂. The structures are obtained from mPW2PLYP-D/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations. 81

Figure 3-6. The lowest energy conformers of TMA-(Pro)₅-N(CH₃)₂. The structures are obtained from RI-MP2/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations. 84

Figure 3-7. The lowest energy conformers of TMA-(Pro)₆-N(CH₃)₂. The structures and relative energies are obtained from RI-MP2/6-311G(d,p) geometry optimizations. 87

Figure 3-8. The relative conformer energy in the gas phase as a function of the *cis* amide contribution for TMA-(Pro)_n-N(CH₃)₂ with n=3-6. For each conformer, only the most stable is chosen. For instance, there are 4 conformers in tetra-proline that have one *cis* amide linkage (CTTT,TCTT, TTCT, and TTTC). The relative energy is shown for TTCT, the conformer with the lowest energy. 88

Figure 3-9. Results of scanning backbone dihedral angles using the mPW2PLYP-D/6-311G(d,p) method and basis set starting from TT conformer of the TMA-(Pro)₂-N(CH₃)₂ for a) ω_2 , b) ψ_1 , and c) ϕ_2 90

Figure 3-10. The coupling of backbone dihedrals with the ω_2 angle for TMA-(Pro)₂-N(CH₃)₂. Angles are obtained using relaxed scans at the mPW2PLYP-D/6-311G(d,p). The x-axis is the ω_2 angle and the y-axis is the other coupled dihedral angles. Black circles correspond to ψ_1 , red triangles correspond to ψ_2 , yellow diamonds correspond to ϕ_2 , and green squares correspond to ϕ_1 91

Figure 3-11. Counter plot representation of the potential energy surface with respect to the ω_2 and ψ_1 angles for TMA-(Pro)₂-N(CH₃)₂ using mPW2PLYP-D/6-311G(d,p). The white circles and black triangles correspond to two transition pathways. The lowest energy is around $\omega_2=180^\circ$ and $\psi_1=125^\circ$ 92

Figure 3-12. The energy barrier for *cis-trans* isomerization for TMA-(Pro)₂-N(CH₃)₂ using mPW2PLYP-D/COSMO/6-311G(d,p) method and basis. The black line corresponds to the gas phase, the red line corresponds to water, the green line corresponds to chloroform and the blue line corresponds to hexane..... 93

Figure 3-13. The endocyclic torsion χ energy barrier for both rings of the TMA-(Pro)₂-N(CH₃)₂ using mPW2PLYP-D/COSMO/6-311G(d,p) in (a) gas phase for CC, (b) water for the CC conformer, (c) chloroform for the CC conformer, (d) hexane for the CC conformer, (e) gas phase for the TT conformer, (f) water for the TT conformer, (g) chloroform for the TT conformer, (h) hexane for the TT conformer. The black lines with filled circles show first ring and the red line with filled up triangles shows the second ring energy barriers. 96

Figure 3-14. The relative energy barrier for rotating around ψ_2 (black, filled circle) and ψ_3 (red, filled triangle) backbone dihedral angles for di- and penta-proline chains, respectively, using mPW2PLYP-D/6-311G(d,p). 99

Figure 4-1. The two di-proline selectors, TMA-(Pro)₂-N(CH₃)-tether (left) and BOC-(Pro)₂-N(CH₃)-tether (right). The atom numbering shown above is used throughout this thesis. A single number identifies each methyl and methylene group since a united-atom representation is adopted for these groups..... 102

Figure 4-2. Torsional barriers for rotation about the inter-ring amide torsion [C(17)-N(21)-C(9)-C(23)] are shown for TMA-(Pro)₂-N(CH₃)₂. Black circles identify the initial B3LYP/6-311G(d,p) energies. Yellow circles identify the CCSD-corrected torsional potential where the B3LYP energies have been scaled to yield the CCSD *cis-trans* energy difference. Blue circles include corrections for LJ and electrostatic interactions between atoms separated by more-than-three bonds. The final torsional potential, obtained by least squares fitting, is shown by a solid line. 106

Figure 4-3. Force field assessment for TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂ are shown in (a) and (b) respectively. Each open symbol represents a final conformer and its energy, obtained from a single-molecule molecular dynamics simulation that begins with a random molecular structure that is optimized to a minimum energy conformation by gradual cooling. Black circles, green triangles, red triangles, and yellow squares correspond to TT, CT, TC, and CC structures, respectively. The energy of the structure, as predicted from the molecular model, is plotted against the sum of squares, calculated relative to the optimized minimum energy structure obtained for B3LYP/6-311G(d,p) calculations. The stars identify the relative energies (calculated with CCSD//B3LYP/6-311G(d,p)) and sum of squares evaluated for the four B3LYP/6-311G(d,p) minimum energy conformers..... 108

Figure 4-4. Illustrative snapshots of the BOC-(Pro)₂-N(CH₃)-tether interface is shown (a) top view of the surface in the presence of a *n*-hexane/2-propanol, (b) in the absence of solvent, (c) in the presence of a water/methanol and (d) in the presence of a *n*-hexane/2-propanol. Corresponding snapshots of the TMA-(Pro)₂-N(CH₃)-tether interface are shown in (e), (f), (g) and (h). The underlying Si layer is omitted for clarity. Carbon, nitrogen, hydrogen, oxygen, and silicon atoms are shown in grey, blue, white, red, and yellow, respectively. The surfaces include di-proline selectors, trimethylsilyl end-caps, silanol groups, and underlying Si atoms. 110

Figure 4-5. Surface distributions of carbonyl carbons (see Figure 4-1 for atom numbering) for the TMA-(Pro)₂-N(CH₃)-tether (left) and the BOC-(Pro)₂-N(CH₃)-tether (right) selector. Results are

shown in the absence of solvent (blue, short dashes), in *n*-hexane/2-propanol (black, solid), and in water/methanol (red, long dashes). 116

Figure 4-6. The distribution of carbonyl orientations at interface. Results are given for (a) *n*-hexane/2-propanol at the TMA-(Pro)₂-N(CH₃)-tether interface, (b) TMA-(Pro)₂-N(CH₃)-tether interface in water/methanol, (c) BOC-(Pro)₂-N(CH₃)-tether interface in *n*-hexane/2-propanol, and (d) BOC-(Pro)₂-N(CH₃)-tether interface in water/methanol. The percentage of carbonyls with the given angle relative to the Si layer is shown: An angle of zero degrees corresponds to the carbonyl pointing directly at the underlying Si layer. Distributions for O(8), O(10), and O(12) are represented by (black, solid), (red, long dashed), and (blue, short dashed) lines, respectively. ... 120

Figure 4-7. Solvent distributions above the TMA-(Pro)₂-N(CH₃)-tether and the BOC-(Pro)₂-N(CH₃)-tether interfaces. Results are shown for (a) hexane/2-propanol at the TMA-(Pro)₂-N(CH₃)-tether interface, (b) water/methanol at the TMA-(Pro)₂-N(CH₃)-tether interface, (c) *n*-hexane/2-propanol at the BOC-(Pro)₂-N(CH₃)-tether interface, and (d) water/methanol at the BOC-(Pro)₂-N(CH₃)-tether interface. The surface distribution for a central carbon of *n*-hexane (red, medium dashes), the O in 2-propanol (black, solid), the O in water (blue, long dashes), and the O in methanol (green, short dashes) are shown. 122

Figure 4-8. The 2D distribution between the hydrogen-bonding H of 2-propanol and the carbonyl oxygens of TMA-(Pro)₂-N(CH₃)-tether for the *n*-hexane/2-propanol solvated interface. Results are given for O(8), O(10), and O(12) in (a), (b), and (c) respectively. Contour plots are also provided to highlight the locations of the features..... 126

Figure 4-9. The 2D distribution between the hydrogen-bonding H of methanol (a-c) and of water (d-f) and the carbonyl oxygens of TMA-(Pro)₂-N(CH₃)-tether for the water/methanol solvated interface. Results are given for O(8) in (a) and (d), O(10) in (b) and (e), and O(12) in (c) and (f). Contour plots are also provided. 129

Figure 4-10. The 2D distribution between the hydrogen-bonding H of 2-propanol and the carbonyl oxygens of BOC-(Pro)₂-N(CH₃)-tether for the *n*-hexane/2-propanol solvated interface. Results are given for O(8), O(10), and O(12) in (a), (b), and (c) respectively. Contour plots are also provided to highlight the locations of the features. 130

Figure 4-11. The 2D distribution between the hydrogen-bonding H of methanol (a-c) and of water (d-f) and the carbonyl oxygens of BOC-(Pro)₂-N(CH₃)-tether for the water/methanol

solvated interface. Results are given for O(8) in (a) and (d), O(10) in (b) and (e), and O(12) in (c) and (f). Contour plots are also provided.	131
Figure 5-1. Poly-proline selectors: (a) TMA-(Pro) ₃ -N(CH ₃)-tether; (b) BOC-(Pro) ₃ -N(CH ₃)-tether; (c) TMA-(Pro) ₄ -N(CH ₃)-tether; (d) TMA-(Pro) ₅ -N(CH ₃)-tether; and (e) TMA-(Pro) ₆ -N(CH ₃)-tether. The atom numbering shown above is used in this thesis. We represent methyl and methylene groups with a single number since a united-atom representation is used for these groups.....	135
Figure 5-2. A comparison between the force field for TMA-(Pro) ₅ -N(CH ₃) ₂ and the results of LPNO-CCSD calculations. The relative conformer energies and structure (as assessed by a sum of squares over atomic positions, relative the lowest energy conformer) are shown. The filled stars identify the results from <i>ab initio</i> calculations. Each circle represents the lowest-energy structure from a single-molecule molecular dynamics simulation. Different conformers are indicated by color coding. The proximity between a sequence of circles and the corresponding star (yellow for CTCTT for example) indicates the level of agreement between the force field and the LPNO-CCSD results.	137
Figure 5-3. Assessment of the intramolecular potential for a) TMA-tri-proline, b) BOC-tri-proline, c) TMA-tetra-proline, d) TMA-penta-proline and e) TMA-hexa-proline. The open black circles show the relative energy of the force field conformers while the open red circles show the corresponding relative LPNO-CCSD (RI/MP2) energies. Conformers having relative energies of more than 20 kJ/mol are not shown.	139
Figure 5-4. A snapshot of the TMA-(Pro) ₅ -N(CH ₃)-tether interface with a water/methanol solvent. Silicon, carbon, oxygen, nitrogen and hydrogen atoms are shown in yellow, light blue, red, dark blue, and white, respectively. Water molecules are in blue and methanol molecules are yellow. The surfaces include silanol groups, trimethylsilyl end-caps, and the underlying silicon atoms.....	140
Figure 5-5. Surface distributions of carbonyl carbons for TMA-tri-proline-tether (left) and BOC-tri-proline-tether (right). Results are shown for <i>n</i> -hexane/2-propanol (solid black line) and water/methanol (dashed red line) solvents.....	146
Figure 5-6. Surface distribution of carbonyl carbons. Results are shown for <i>n</i> -hexane/2-propanol (solid lines) and water/methanol (dashed lines). Black, blue, and red lines correspond to TMA-	

(Pro) ₄ -N(CH ₃)-tether, TMA-(Pro) ₅ -N(CH ₃)-tether, and TMA-(Pro) ₆ -N(CH ₃)-tether, respectively.	147
Figure 5-7. An illustration of the convergence of surface distributions. The surface distribution of a) C(13) in the triproline selector and b) C(19) for the hexaproline selector in <i>n</i> -hexane/2-propanol. Results for 10 independent simulations are shown. The thick black line is the average surface distribution and error bars are provided.	149
Figure 5-8. The distribution of the two topmost carbonyl groups at the interface. The probability of having a carbonyl group with a given angle relative to the underlying silicon layer is presented: an angle of 0° corresponds to a carbonyl group pointing directly towards the surface while an angle of 180° corresponds to a C=O group that points into the bulk. Results for TMA-(Pro) _n -N(CH ₃)-tether are shown for <i>n</i> -hexane/2-propanol (black line) and water/methanol (red line). Filled circles in panels (a) and (b) identify the corresponding distributions for BOC-(Pro) ₃ -N(CH ₃)-tether.....	150
Figure 5-9. Top-down and side-view snapshots of the TMA-(Pro) _n -N(CH ₃)-tether selective surface. All selector atoms are shown in blue, except for the carbonyl oxygens shown in red. The underlying Si layer, the end-caps and the silanol groups are represented in yellow. The top two rows show the selective surface in the <i>n</i> -hexane/2-propanol. The corresponding snapshots for water/methanol are provided in the bottom two rows.....	151
Figure 5-10. Solvent distribution above the interface for TMA-(Pro) _n -N(CH ₃)-tether in <i>n</i> -hexane/2-propanol (left hand side) and water/methanol (right hand side). The results for the tri-proline, tetra-proline, penta-proline and hexa-proline are in the first, second, third, and fourth row, respectively, with the <i>n</i> -hexane and methanol distributions in red, and the 2-propanol and water distributions in black. The solvent distribution above the BOC-terminated tri-proline interface is given in the first row, with distributions identified by filled circles.	153
Figure 5-11. 2D distributions for the TMA-(Pro) ₅ -N(CH ₃)-tether interface. Distributions between potential H-bonding pairs: Each row corresponds to a backbone carbonyl oxygen and each column corresponds to a solvent H. The first column shows the distribution for the alcohol H in <i>n</i> -hexane/2-propanol, the second and third correspond to H in water and the alcohol H, respectively, in water/methanol. The distance perpendicular to the surface is <i>z</i> and a positive value means that the solvent atom is further from the Si layer. <i>r</i> is the component of the O···H separation parallel to the surface.....	155

Figure 6-1. The B3LYP/6-311G(d,p) optimized structures of the analytes. “AnC” indicates an attachment to a central carbon in anthracyl, “AnS” indicates an attachment to a side carbon in anthracyl, and Ph represents a phenyl ring. Numbers in bracket indicate the analyte numbering used by the Li group [47]...... 165

Figure 6-2. A snapshot of a typical simulation cell for TMA terminated hexa-proline in *n*-hexane/2-propanol with AnC-CH₃. The R and S enantiomers are shown in blue and red, respectively. Yellow, red, ocean-blue, blue and white correspond to silicon, oxygen, carbon, nitrogen and hydrogen atoms..... 168

Figure 6-3. Surface distributions for backbone carbonyl oxygens. The poly-proline selectors are identical in each panel. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The color identification for backbone oxygens is shown on top right panel..... 170

Figure 6-4. Surface distribution for the hydroxyl hydrogen of AnC-CH₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel. 172

Figure 6-5. Surface distribution for the hydroxyl hydrogen of AnS-CH₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel. 174

Figure 6-6. Surface distribution for the hydroxyl hydrogen of Ph-CF₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel. 176

Figure 6-7. Surface distribution for the hydroxyl hydrogen of Ph-CH₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel. 178

Figure 6-8. Surface distribution for the hydroxyl hydrogen of Ph-C₂H₅. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel. 180

Figure 6-9. Surface distribution for the hydroxyl hydrogen of AnC-CF₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel. 182

Figure 6-10. Predicted (red stars and error bars) and experimental (black circles) selectivity factors for di- and hexa-proline CSPs with six analytes. All selectivity factors are in *n*-hexane/2-

propanol solvent. No experimental error is reported for the experimental selectivity factors. The selectors and analytes are shown on the top and bottom of the graph.....	195
Figure 6-11. Uncertainty calculations for selectivity factors. The black solid line with filled circles is for AnC-CH ₃ , the black dashed line with filled squares is for AnS-CH ₃ , the red solid line with filled circles is for Ph-CF ₃ , the red dashed line with filled squares is for Ph-CH ₃ , the blue solid line with filled circles is for Ph-C ₂ H ₅ , the blue dashed line with filled squares is for AnC-CF ₃ . The selector is identified within each panel and the right and left hand panels correspond to water/methanol and <i>n</i> -hexane/2-propanol solvents.....	196
Figure 6-12. Side and top view for the docking of both enantiomers of AnC-CH ₃ on TMA-hexa-proline CSP. The top left picture shows the side view when the S analyte hydrogen bonds to O(20). The top right view clearly shows why the analyte (pink color) has steric constraints and can only “slide” into the surface. The lower left picture is showing the docking of the R enantiomer. The lower right graph shows the torsional distribution about the bond joining the aromatic group to the chiral carbon, for enantiomers that hydrogen bond to O(20). The equilibrium angles for the S and R enantiomers are 120 and 240°, respectively.....	200
Figure 6-13. Torsional angle distributions for AnS-CH ₃ (top left), Ph-CF ₃ (top right), Ph-CH ₃ (bottom left), and Ph-C ₂ H ₅ (bottom right) at the events of hydrogen bonding with O(8). For AnS-CH ₃ , the equilibrium angles for S and R enantiomers are 140 and 220°, respectively. For Ph-CF ₃ , the equilibrium angles for S and R enantiomers are 35 and 315°, respectively. For Ph-CH ₃ , the equilibrium angles for S and R enantiomers are 35 and 315°, respectively. For Ph-C ₂ H ₅ , the equilibrium angles for S and R enantiomers are 155 and 205°, respectively.	202
Figure 6-14. The relative angle between the plane of aromatic group and the Si layer. For AnC-CH ₃ (black line) and Ph-CF ₃ (red line) that hydrogen bond to O(20) and O(8), respectively, on hexa-proline and in <i>n</i> -hexane/2-propanol. An angle of 90° shows that the aromatic group is perpendicular to the Si layer and the angle of either 0 or 180° means that the aromatic group is perpendicular.	203
Figure 6-15. Double hydrogen bonding event for AnC-CH ₃ with the tri-proline selector in <i>n</i> -hexane/2-propanol.....	205
Figure 6-16. Hydrogen bonding of the R enantiomer of AnS-CH ₃ with the hexa-proline selector at water/methanol.....	206

List of Abbreviations and Symbols

2D	2-Dimensional
3D	3-Dimensional
AnC-CF₃	1-(9-anthryl)-2,2,2-trifluoroethanol
AnC-CH₃	α -methyl-9-anthracenemethanol
AnS-CH₃	1-anthracen-2-yl-ethanol
BOC	t-butyl carbamate
CC	Coupled cluster
CCD	Coupled cluster doubles
CCSD	Coupled cluster singles and doubles
CCSDT	Coupled cluster singles, doubles, and triples
CGTF	Contracted gaussian-type functions
CHARMM	Chemistry at Harvard molecular mechanics
CHELPG	Charges from electrostatic potentials using a grid based method
CI	Configuration interaction
CSP	Chiral stationary phase
COSMO	Conductor like solvation model
DFT	Density functional theory
Fmoc	Fluorenylmethoxycarbonyl
GGA	Generalized gradient approximation
H-bond	Hydrogen bond
HF	Hartree-Fock

HPLC	High performance liquid chromatography
LJ	Lennard-Jones
LDA	Local density approximation
LPNO	Local pair natural orbital
LSDA	Local-spin-density approximation
MC	Monte Carlo
MD	Molecular dynamics
MP	Moller-Plesset
MP2	Second order Moller-Plesset
NMR	Nuclear magnetic resonance
OPLS	Optimized potential for liquid simulations
Ph-C₂H₅	1-phenyl-1-propanol
Ph-CF₃	α -(trifluoromethyl)benzyl alcohol
Ph-CH₃	1-phenylethanol
PNO	Pair natural orbitals
QSPR	Quantitative structure property relationships
RI-MP2	Resolution of identity second order Moller-Plesset
RU	Rigid unit
SCF	Self-consistent field
TMA	Trimethylacetyl

Chapter 1

Introduction

1.1 Chirality

A chiral molecule is a molecule which is not superimposable on its mirror image. Such a molecule does not possess any plane of symmetry, center of inversion and any rotation-reflection axes [1]. Two chiral molecules that are mirror images of each other are called enantiomers. Enantiomers of a given molecule have the exact same physio-chemical properties except that they interact distinctly with polarized light and with other chiral molecules [2].

Many building blocks of life are chiral. Sugars, nearly all amino acids, proteins and nucleic acids are chiral. Our body, in nature, is chiral and it is expected that chirality plays an important role in biological processes occurring in the body. Enantiomers interact differently when they enter the body. About 80% of current drugs are chiral and 75% of them are used as racemates [3]. Two enantiomers of a given chiral compound might have different impacts on the body [4-7]. R and S enantiomers of ketamine and labetalol [8], for example, have different pharmacological activities. As a result, there is a high demand for producing enantiomerically pure drugs. The production of such drugs is strongly encouraged in many countries [2,9].

To produce pure enantiomers of a drug, two approaches can be taken. Developing new synthetic methods that solely produce one of the enantiomers [10-12] and developing means to efficiently separate two enantiomers of a drug. The development of synthetic pathways is difficult and time consuming. As a result, the second approach, separation of enantiomers, is usually taken in the early stages of drug development [13].

Separation methods have been widely used and can be categorized into three categories: direct, indirect and spontaneous resolution. In a direct resolution, enantiomers are resolved based on their distinct affinity to form transient diastereomeric complexes with another chiral molecule, often named the chiral selector. Chromatography, with the use of a chiral stationary phase (CSP) [14], is an example of this category. Here, enantiomers are introduced into the chromatographic column, which is packed with the stationary phase. Chiral selectors on the stationary phase interact with the enantiomers that are carried through by means of a mobile phase (that could be either polar or apolar). If the chiral selector can interact more strongly with one of the enantiomers, this enantiomer is retained in the column longer than the other. The stronger-interacting enantiomer elutes last. In the indirect resolution method, enantiomers react with another pure enantiomer and form a set of diastereomers [1]. Since diastereomers have distinct physical and chemical properties, they can be separated by appropriate methods. Then a chemical process is needed to extract each enantiomer. In industry, resolution of bases and acids via diastereomeric salt formation followed by crystallization is widely used [15,16]. Spontaneous resolution is based on the fact that some enantiomers separate upon crystallization. This method was first used by Louis Pasteur [17]. Unfortunately, this is a rare event and cannot be used for separation of many enantiomers.

1.2 Chiral High Performance Liquid Chromatography

In general, two methods are used in liquid chromatography. The first one is based on mobile phase additives, while in the second the chiral selector is bound to the stationary phase and interacts with the analyte. Chiral mobile phase additives can be used to form transient diastereomeric complexes with the chiral analytes and then these diastereomers can be separated on an achiral stationary phase [18-21]. The additive chiral compounds range from β -cyclodextrins

to proteins. Although this method has been shown to be useful, it suffers from several drawbacks. It wastes chiral additives and the initial enantiomers need to be obtained from the diastereomers. On the other hand, introducing chirality to the stationary phase is more common and does not have the problems noted above. Here, the enantiomers form transient diastereomers with the chiral selectors that are bound to the stationary phase and will be separated based on the free energy of formation of these diastereomers. In this method, both enantiomers enter the chromatographic column (see Figure 1-1). The mobile phase which could be a polar or apolar solvent elutes the enantiomers throughout the column. Based on the difference of enantiomers interactions with the selectors in the pores of the stationary phase, enantiomers are separated. The time in which each enantiomer is eluted is recorded using a detector. Often, an inert analyte is also added to the column that does not interact with the stationary phase. Such analyte is eluted first and the retention time of this analyte is shown by t_M . The retention time of the enantiomers could be identified as well: t_R and t_S , in which R and S stand for the absolute configuration of the analytes. The selectivity factor can then be calculated experimentally using the below equation:

$$\alpha = \frac{t_R - t_M}{t_S - t_M}, \quad 1-1$$

where it is assumed that the R enantiomer is retained so that $t_R > t_S$. Figure 1-2 shows the pores of the silica in which the enantioselectivity happens and the location in which the simulation is performed.

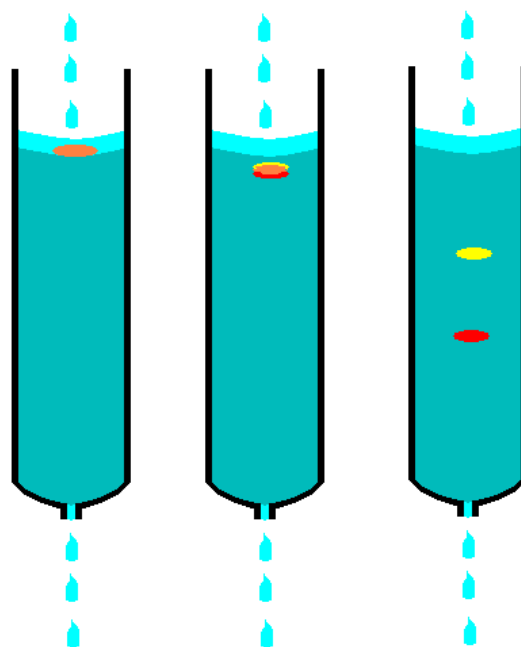


Figure 1-1. A typical representation of the separation in chromatography where two enantiomers shown by yellow and red colors are separated.

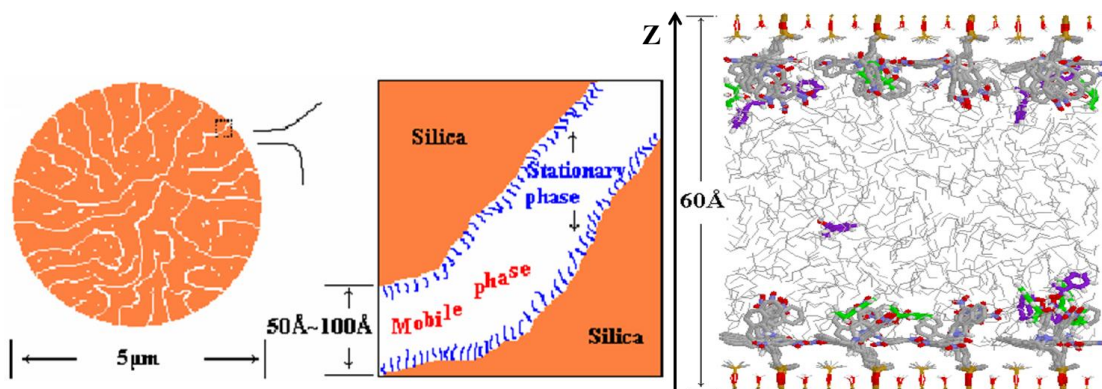


Figure 1-2. A schematic representation of silica pores. The figure is adopted from ref [22].

The selection mechanism includes several types of interactions such as hydrogen bonding, π - π interactions, dipole-dipole interactions, steric interactions and more [22]. The enantiomer that interacts more strongly with the CSP will stay longer in the column while the enantiomer with the

weaker interaction will be in the mobile phase more frequently and elutes first. CSPs can be classified into several types [23] including polysaccharide-based CSPs, brush-type CSPs, protein-based CSPs, macrocyclic glycopeptide antibiotic-based CSPs and cyclodextrin-based CSPs. Figure 1-3 shows examples of the selective moieties for each of the selector types. Figure 1-3(a) shows a polysaccharide based CSP. Examples of this type are cellulose and amylose based CSPs [24-27]. The enantio-separation mechanism is thought to be due to the presence of several chiral centers close to hydrogen bonding sites and also the steric effect from the inclusion of enantiomers into the helical chiral grooves [28]. Other factors such as π - π stacking can also play a role in enantio-separation. The whelk-O1 selector, as an example of a brush-type stationary phase, is shown in Figure 1-3(b). The whelk-O1 selector consists of one electron deficient and one electron rich aromatic group along with two chiral centers close to hydrogen bonding donor and acceptor sites [23,29-32]. For whelk-O1, the important interactions are hydrogen bonding, π - π stack, CH- π interactions, and steric interactions [22].

Figure 1-3(c) shows a crown-ether based CSP. These CSPs work based on the selective inclusion of enantiomers into the cavity [33,34] but they are limited in application since they only resolve molecules with specific functional groups. Protein-based CSPs are logical choices since proteins are responsible for chiral recognition in the body (see Figure 1-3(d)). Usually such amino-acid chains contain several hydrogen bonding donor and acceptor sites and the proposed mechanism involves the contribution of multiple sites.

In achiral chromatography analytes having different sizes or functional groups interact differently with the stationary phase. Here, the two enantiomers behave exactly the same in an achiral environment. Although the introduction of a chiral stationary phase leads to distinct

interactions of the enantiomers, the mechanism and the underlying interactions responsible for chiral resolutions can only be inferred from chromatographic data.

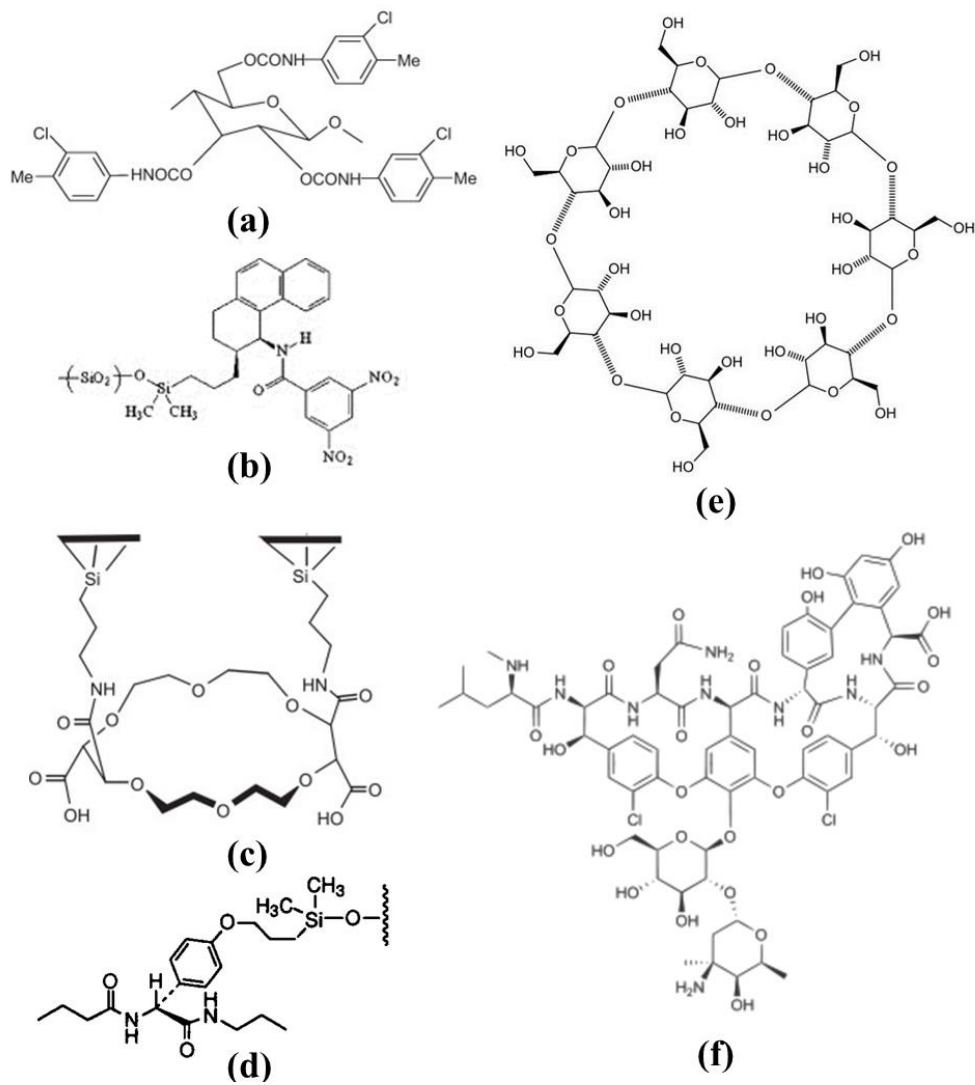
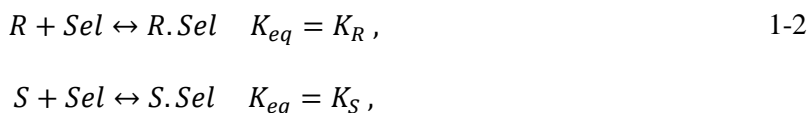


Figure 1-3. Examples of common chiral stationary phases. a) polysaccharide based, b) whelk-O1, c) crown based, d) protein based, e) β -cyclo-dextrin, f) Vancomycin CSPs.

The number of interactions responsible for chiral discrimination is a subject of controversy. Some believe that only two interactions are necessary [35], some believe in the three-point attachment model [36-40]. Four interaction models are also proposed [41,42]. These interactions could be any type of known interactions such as hydrogen bonding, dipole-dipole, π - π stack, and more. It is also argued that the direction the enantiomer approaches the stationary phase could be considered as an interaction [43]. In actual fact, enantio-recognition follows the fact that two diastereomeric complexes are distinct, whether specific interactions can be identified or not.

1.3 Theoretical Studies on Chiral Stationary Phases

Here we briefly discuss the principles of resolution on stationary phases based on molecular dynamics (MD) simulations. The resolution of enantiomers depends on two equilibria between the *R* and the *S* enantiomer with the selector:



where *R* and *S* refer to absolute configurations of the enantiomers. The equilibrium constant of each equilibrium is denoted by K_R and K_S . The selectivity factor, α , is defined as the ratio of the larger equilibrium constant to the smaller one so that it is always larger than one. Assuming that the *S* enantiomer is eluted first, we get:

$$\alpha = \frac{K_R}{K_S}. \quad 1-3$$

Also each equilibrium has an associated free energy change ΔG :

$$\Delta G = -R_{gas}T \ln K_{eq}, \quad 1-4$$

where R_{gas} and T are the ideal gas constant and temperature, respectively. Note that the above equation is held in an NPT ensemble. Combining this relationship with Eq. 1-3 we get:

$$\Delta\Delta G = \Delta G_R - \Delta G_S = -R_{gas}T \ln \alpha . \quad 1-5$$

As a result, the selectivity factor is determined by the difference in free energy change for the formation of two transient diastereomeric complexes. To calculate $\Delta\Delta G$ from Eq. 1-5 one needs to calculate the free energy of both diastereomeric complexes because the free energies of single enantiomers in an achiral solvent are equal. This reduces the number of required free energy calculations. The success of this method depends on knowing the exact mechanism of interaction between each enantiomer and the selector. Although methods have been developed to guess the interactions responsible for resolution [44], the potential energy surfaces need to be thoroughly scanned to find the best docking arrangements. This is computationally expensive in practice and one needs to limit the degrees of freedom. Usually some guess of the docking arrangement is used but this is not necessarily unbiased. Moreover, in such methods, the flexibility of the selectors and the enantiomers are not included and only the most stable conformers are chosen. It is possible that these molecules adopt other conformations when docking. One approach to solve this problem is to use Monte-Carlo (MC) searches or molecular dynamics simulations [45,46]. Here, some constraints are applied to fix the relative position of the enantiomer and the selector and the other degrees of freedom are allowed to relax. Finally, these methods ignore the effects of solvent and the surface environment. The chromatographic interfaces are usually covered with the selectors, and free and end-capped silanol groups. In fact, surface crowding limits enantiomer docking. For instance, Huang [47] showed the role of linker length that connects the CSP to the surface on the separation of 53 chiral enantiomers on di-proline stationary phases. Solvent also can interact strongly with the selector. It can block the

enantiomer docking or alter the potential energy surface while in most cases solvent is simply ignored or it is treated implicitly [48]. Despite these limitations, methods based on calculating the free energy difference have been successful [34,44,46,49-53].

Another approach is to apply force field-based molecular dynamics methods or Monte-Carlo simulation to sample the configuration space for transient diastereomeric complexes [46,54,55]. These methods sample many configurations for the enantiomer-selector interaction and have more success in finding the preferred modes of docking. Here, force fields are used to model the molecular flexibility and interaction with other molecules. Often the analyte and selector are positioned close to each other and then the trajectory of the system is analyzed to find the best docking arrangements. The analysis of the trajectory using the snapshots and distribution functions provide some insight into docking arrangements.

Another approach to predict α is based on structure-function relationships [56-62]. Quantitative structure property relationships (QSPR) are based on developing quantitatively defined relationships between the properties of interest and some observed known properties of the system, such as the presence of specific functional groups. In this method it is assumed that knowledge about structural properties and energy features of the molecules is enough to predict molecular properties. So, by employing statistical methods, the structure and energy of a family of related molecule are fitted to observed selectivity factors and retention times. The developed model is then used to predict the properties of new molecules. QSPR methods are purely statistical: They do not explain why or how a process happens at the atomic level. On the other hand, they are fast and efficient. Several excellent reviews have discussed the application of QSPR methods to chromatography [62-67].

1.4 Poly-proline: The Biological Importance

Proline is unique among amino acids since the backbone nitrogen is a secondary amide group that makes it unable to form hydrogen bond as a hydrogen bonding donor site (see Figure 1-4). Unlike other amino acids, the α -carbon is part of a pyrrolidine ring. The lack of hydrogen bond donor sites makes it impossible for poly-proline to adopt a α helix or a β sheet. It also makes it harder to be identified in protein structures by hydrogen bonding patterns, as it is common for other structures [68,69]. Proline residues are found at the end of α helices and β sheets, where turns occur to make a loop, and in non-repetitive structures [70]. The helical structure of poly-proline was unknown for several years and often confused with irregular or random coil conformations [71]. Poly-proline helical conformation plays an important role in biological systems: signal transduction, immune response, transcription, and cell motility [69-71]. The pyrrolidine ring dihedral angle is fixed at around -60° in contrast to other amino acids [72]. Proline coexists in both *cis* and *trans* forms, unlike other amino acids that almost exclusively adopt a *trans* conformer. *Cis-trans* isomerization of proline residues is responsible for folding and slow denaturation of small proteins [73-76]. This transition also provides a switch for inter-conversion of the open and the closed states of neurotransmitter-gated ion channels [77].

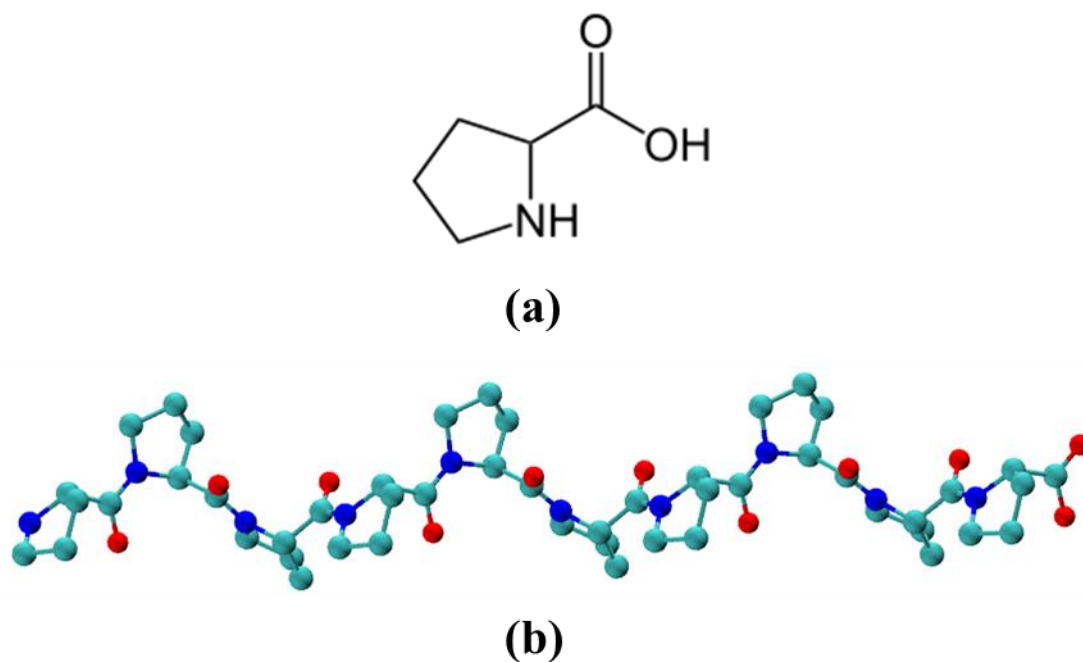


Figure 1-4. Proline monomer (a) and decamer (b). Note that in a poly-proline, the backbone nitrogen does not have any hydrogen atom and the backbone is unable to participate in hydrogen bonding as a donor.

Poly-proline adopts two major helical conformers which are known as PPI and PPII. In PPI, all amide linkages adopt a *cis* dihedral angle; as a result this conformer is compact. In PPII, however, the amide linkages are *trans* and the structure is extended. The PPII conformer is common in biological systems [68,69]. Even without proline residues, short chains of amino acids, which are unable to adopt α -helices, can adopt a PPII conformation.

1.5 Proline Stationary Phases

The high demand to develop new stationary phases arises from the fact that there are a number of chiral analytes that cannot be resolved or resolved well on available CSPs [78]. Due to the diversity of chiral analytes, there is no stationary phase which can resolve all chiral analytes. In

fact, only a few CSPs have demonstrated broad chiral selectivity. Daicel columns [24,79], the Chirobiotic columns [80,81], and the whelk O-1/2 [82] columns are among these CSPs. As a result of ongoing research on developing new CSPs, proline stationary phases as new brush-type stationary phases were introduced by Welch and Millet [82-84] and further developed by several research groups [81,83-89]. Welch used one proline unit in conjunction with an electron rich terminal group and obtained surprisingly high selectivity factors for several di-nitro-benzyl-amino acid derivatives. Ramiandrasoa *et al.* [83] used proline based CSPs with benzyloxycarbonyl and *tert*-butoxycarbonyl as terminal groups to resolve N-3,5-dinitrobenzoyl amino acid esters. They studied the effect of terminal group and solvent polarity on enantio-resolution and found that the nature of the terminal group and the solvent polarity change the selectivity factor.

In recent years, several other experimental studies have been performed on proline-based stationary phases to improve the resolution properties of these CSPs [90-92] by increasing the number of proline residues, and by changing the tether, solvent, and the terminal group. Proline CSPs, although newly developed, have been shown to be competitive with other commercially available stationary phases [47,78,93-95]. The first report on broad selectivity of poly-proline CSPs was in 2005 by Huang *et al.* [78]. They used mono- to tetra-proline selectors along with Fluorenylmethyloxycarbonyl chloride (Fmoc) as the terminal group to resolve 53 chiral analytes. Interestingly, they found that enantio-separation only occurs for longer chains and mono-proline did not resolve any enantiomer. Later, they replaced the Fmoc with several other terminal groups for di-proline selectors and found that the best resolution was achieved with trimethylacetyl (TMA) [94]. To further improve the efficiency of enantio-resolution, they optimized the peptide length and linker to the surface [47]. They studied di- to deca-proline stationary phases and also changed the linker by adding a non-selective hydrogen bonding site. Solvent composition was

also varied for resolution of 53 enantiomers. They observed that, in general, the selectivity factor and the number of resolved analytes increases with chain length. Specifically, the number of resolved analytes is reduced from di- to hexa- and then increases to deca-proline. Non-specific hydrogen bonding at the surface has been shown to reduce the enantio-separation. Also, changing solvent polarity impacts the enantio-separation. For instance, while the deca-proline CSP resolves 40 out of 53 analytes with an average selectivity factor of 1.32 in *n*-hexane/2-propanol, this selector resolves 47 analytes with an average selectivity factor of 1.49 in *n*-hexane/2-propanol/CH₂Cl₂.

It is believed that the PPII conformer can offer more hydrogen bonding possibilities since it is much more extended compared to the PPI conformer. From a chromatographic point of view, these two different backbone conformations can lead to distinct interactions with the enantiomers in different solvent. So, simply by changing the mobile phase, the proline chain could adopt a new backbone conformation which might be able to resolve racemates.

Collectively, proline based CSPs are new, competitive CSPs. They are stable in most of commonly used mobile phases, and they are commercially inexpensive [47,78,93-95]. They offer considerable flexibility in terms of the proline chain length and solvent composition.

1.6 Theoretical Studies of Chiral Interfaces

Chiral recognition occurs in nature and is needed in several industries such as drug discovery, food additives, and fragrances [96,97]. To understand the chiral recognition process, molecular simulations have been used extensively [98]. Molecular simulations of interfaces present many challenges. They require an atomistic description of the underlying complex interface. Despite this, several studies have been performed on the chromatographic interfaces. These studies aim to understand the solvation, conformational preferences, and resolution mechanism at the interface.

Molecular modeling starts with two steps: building or selecting molecular models that can accurately describe the interactions between the atoms; and building a model system that resembles the real system. To this end, two simulation approaches are usually used: Monte-Carlo simulations [99] and molecular dynamics simulation [100]. In the first method a collection of replicas are generated randomly and analyzed. In the latter, such a replica is the trajectory of the system in time obtained from Newtonian equations of motion.

MC methods are employed to study achiral interfaces extensively. Rafferty *et al.* [101] used configurational-bias MC simulations, in the Gibbs ensemble, to investigate the retention time and successfully predicted the retention times of analytes. Rafferty *et al.* [102] studied the mobile phase effect in reversed phase chromatography using a C₁₈ stationary phase. They examined the role of partitioning and adsorption on the retention of alkane solutes and showed that changing the solvent from acetonitrile/water to methanol/water does not affect the resolution mechanism. Szabelski and Sholl performed MC simulations to differentiate the adsorption of enantiomers on a chiral stationary phase [103,104].

MD studies are employed to study enantio-resolution on chiral stationary phases. Lipkowitz *et al.* identified chiral recognition regions by finding the Boltzmann-weighted intermolecular energies between analytes and β -cyclodextrin that are obtained from molecular dynamics trajectories [53,105]. Jiang *et al.* [106] used MD to study chiral separation of racemic phenylglycines in thermolysin crystals. They observed that the D-phenylglycine is retained and interacts more strongly. The chiral discrimination is assumed to result from several chiral centers of thermolysin. Kasat *et al.* studied the effects of solute composition on enantio-resolution in cellulose using MD simulations, and the elution orders predicted were consistent with HPLC results [107,108]. Sebestik and Bour used MD to study the chiral resolution of CHFCIBr on

enantiomerically pure stationary phase composed of CHFCIBr and water as solvent. Cann *et al.* studied the enantio-resolution and solvation of whelk-O1 stationary phase in the presence of water/methanol, *n*-hexane/2-propanol, and supercritical CO₂/methanol [32,109-111]. They successfully predicted the selectivity factor of several chiral analytes including styrene oxide, stilbene oxide, 2-methyloxirane, 1,2-diphenylcyclopropane, 1-phenylethanamine, 1-phenylethanol, 1-phenylethane-1,2-diol, 1,2-diphenylethane-1,2-diol, 2-hydroxyl-1,2-diphenylethanone, and N-(1-(4-bromophenyl)ethyl)pivalamide. They also proposed a rational optimization of the whelk-O1 stationary phase. Nita and Cann [112,113] studied the solvation of (R)-N-(3,5-dinitrobenzoyl)phenylglycine- and (R)-N-(3,5-dinitrobenzoyl) leucine-derived chiral stationary phases and found distinct interfaces despite the structural similarity of leucine and phenylglycine.

1.7 Research Direction

An in-depth understanding of the chiral recognition process is an important step in discovering and modifying chiral stationary phases and molecular dynamics simulations have shown strength in this field [109-113]. Usually, experimental methods such as NMR and X-ray [98] are limited in defining the mechanism of separation and they suffer from the fact that the measurement conditions are not the same as conditions in which the separation occurs.

With the rapid development in computational resources, more efficient computational techniques can now be used. Molecular dynamics simulations over the past decade have become more practical for systems consisting of 100,000 atoms or even more [114]. Here, we aim to use molecular dynamics simulations to study the interfacial structure, solvation, and selectivity of proline based stationary phases. This is the first time that a molecular modeling approach has been applied to proline based stationary phases.

We have chosen to focus on proline-based stationary phases for several reasons. First, they are a new type of stationary phase. Second, they are competitive to other commercially available stationary phases. From a chromatographic point of view, they are stable and they work in both normal and reversed phase methods. Experimental studies on proline based CSPs relate the selectivity to several factors such as the chain length, terminal group and solvent type, but the underlying mechanism of action is yet unclear.

Here, we study the solvation properties of proline selectors with 2-6 proline units in *n*-hexane/2-propanol and water/methanol and in the presence of interfacial groups such as silanols and trimethylsilyl endcaps. The role of the terminal group is also studied for di- and tri-proline chains where TMA and t-butyl carbamate (BOC) are studied. Selectivity studies are performed for di-, tri-, and hexa-proline CSPs with six analytes that are experimentally studied [47,78,93-95] on these selectors. Specifically, we study the role of chain length, solvent and terminal group in solvation and selectivity of poly-proline CSPs at the interface.

We begin with force fields that are developed from, and tested against, *ab initio* calculations. Molecular dynamics simulations are then performed to study the solvation and selectivity. Chapter 2 includes a review of the methods and models employed in this research. Chapter 3 discusses the conformational preferences of the poly-proline chains. Chapter 4 and 5 report solvation studies of poly-proline selectors. The selectivity of proline selectors is analyzed in Chapter 6. A conclusion of our study and an outlook for future work is provided in Chapter 7.

Chapter 2

Methods and Models

2.1 Introduction

In this chapter, the theoretical methods employed in this research are discussed. *Ab initio* calculations form the basis for understanding the relative stability of conformers and for developing the force fields. They are discussed in Section 2.2. The basics of the molecular dynamics simulations are addressed in Section 2.3. Some of the methodological improvements that are performed to specifically improve the accuracy of the simulations are discussed in Section 2.4.

2.2 *Ab initio* Methods

2.2.1 *Ab initio* Calculations

To obtain details of conformational balance, flexibility and structural properties of the polyproline selectors and chiral analytes, quantum mechanical calculations are employed. According to quantum mechanical postulates, any physical property can be extracted from the wave function of the molecule. Such molecular wave functions can be obtained by solving the Schrödinger equation:

$$\hat{H}\Psi(\vec{r}_e, \vec{r}_n) = E\Psi(\vec{r}_e, \vec{r}_n), \quad 2-1$$

where \hat{H} is the system Hamiltonian operator and E is the energy eigenvalue of the system. The wave function $\Psi(\vec{r}_e, \vec{r}_n)$ depends on the coordinates of electrons \vec{r}_e and the nuclei \vec{r}_n . The quantum mechanical Hamiltonian consists of kinetic and potential energy operators and is defined as follows:

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_n \frac{Z_n}{r_{in}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} - \frac{1}{2} \sum_n \frac{\nabla_n^2}{M_n} + \sum_n \sum_{m>n} \frac{Z_n Z_m}{r_{nm}}, \quad 2-2$$

where i and j sums go over electrons, n and m sums go over all nuclei, Z is the nuclear charge, and M is the nuclear mass. Note that the equation is written in atomic units for simplicity. The first term is the kinetic energy operator of all electrons, the second term is the attraction potential operator between the electrons and the nuclei in the molecule, the third term is the electron-electron repulsion operator, the fourth term is the kinetic energy operator of all nuclei, and the fifth term is the nuclear-nuclear repulsion operator. To reduce the complexity of this equation, the Born-Oppenheimer approximation can be used where the electronic and nuclear motions are treated separately. It is assumed that since nuclei are much heavier and move much more slowly than electrons, the last two terms can be studied together after the electronic degrees of freedom are addressed. The electronic Hamiltonian is:

$$\hat{H}_e = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_n \frac{Z_n}{r_{in}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}}. \quad 2-3$$

This Hamiltonian is then used to find the electronic wave function and energy eigenvalues by solving the $\hat{H}_e \Psi(\vec{r}_e) = E_e \Psi(\vec{r}_e)$ equation. Unfortunately, with a few exceptions, Eq. 2-3 cannot be solved exactly. So, several approximate methods are employed to solve this equation, such as the variational and the perturbation approaches [115]. In the variational approach, a trial wave function is chosen and the total energy is then estimated based on this trial wave function. The wave function is not exact and the variational theorem guarantees that the corresponding energy is always higher than, or equal to, the exact energy. The success of this method is based on systematically approaching the exact wave function.

The Hartree-Fock (HF) method [115] uses the variational principle and self-consistent field (SCF) theory to approximately solve the Schrödinger equation. The trial wave function is guessed and a Slater determinant is constructed from spatial-spin orbitals to satisfy the parity constraints. This trial wave function is then optimized to minimize the expectation energy. The spatial part of the trial wave function consists of linear combinations of basis functions. In practice, the type and the number of basis functions affect the final trial wave function and the accuracy of the result. The HF method is simple and fast, but it has some disadvantages. The use of a single Slater determinant does not provide the flexibility needed for a multi-electron wave function. Moreover, this method only considers the average effective potential of the electrons and ignores instantaneous electronic correlations. To include electron correlation, various methods have been developed such as Moller-Plesset perturbation theory (MP), coupled cluster (CC) approaches, and configuration interaction (CI) methods. Although these methods are more accurate than HF, they are expensive and this limits their applicability. Another approach is based on using electron density instead of the wave function to capture the electron correlation.

2.2.2 Basis Functions and Basis Sets

Linear combinations of basis functions are used to represent the molecular orbitals in *ab initio* calculations:

$$\varphi_i = \sum_r c_{ri} X_r, \quad 2-4$$

where c_{ri} is the coefficient of each basis function X_r in the φ_i molecular orbital. Gaussian type basis functions are most widely used because their integrals can be solved analytically so that numerical integration (which is expensive) is not necessary. A typical Gaussian function, in Cartesian coordinates, is defined as:

$$g_{ijk} = Nx_b^i y_b^j z_b^k \exp(-\alpha r_b^2), \quad 2-5$$

where i, j, k are any non-negative integers, N is the normalization constant, and α is a parameter that controls the width of the Gaussian peak. The Gaussian functions are centered on nucleus b . Depending on the sum $i + j + k$, a Gaussian basis set is an s-, p-, or d-type basis set. Currently, contracted Gaussian type functions (CGTF) are usually used to build the molecular orbitals.

$$X_r = \sum_l c_l g_{ijk}, \quad 2-6$$

where c_l is a normalized Cartesian coordinate coefficient.

Several types of basis sets exist. Some are called minimal since each core and valence atomic orbital is represented by only one CGTF. Double zeta basis sets have two CGTF for each core and valence atomic orbital. Split-valence basis sets use one CGTF for core atomic orbitals and two or more for valence shells. The well-known Pople basis sets are split-valence basis sets denoted generally by X-YZG. X shows the number of primitive Gaussian functions to build one CGTF for the inner shell, Y and Z show that valence orbitals have two basis functions each. The first one has Y primitive Gaussian functions and the second one has Z.

To account for the distribution of atomic orbitals in a molecular environment, polarizable functions of higher angular momentum can be added. For instance 6-31G(d) includes 6 primitive Gaussian functions for core orbitals, 3 and 1 CGTFs for valence orbitals and 6 d-type Gaussian functions added to all atoms except hydrogen. For molecules that have higher electron density far from the nuclei, adding diffusive terms is also necessary. A diffuse orbital simply decays more slowly with distance. For instance 6-31+G is formed from 6-31G by adding four diffusive functions to each non-hydrogen atom.

In this thesis, we used the 6-311G(d,p) basis set which is a split-valence triple-zeta basis set. This basis set has 6 primitive Gaussian functions for core orbitals, and three contracted Gaussians for each valence orbital along with 6d- and 3p-type Gaussian functions for non-hydrogen and hydrogen atoms, respectively.

2.2.3 Density Functional Theory

Density functional theory (DFT) aims for the molecular electron density rather than the wave function. The application of DFT to molecular systems gained momentum with the Hohenberg and Kohn existence principle and variational theorem [116]. Briefly, these theorems state that for a non-degenerate molecular ground state, the wave function, and as a result, all other properties, are uniquely determined by the ground state electron probability density. Similar to the variational principle, a good probability density minimizes the energy if the exact functional relating energy to density is known. Kohn and Sham [117] showed a method to find the ground state density. They introduced a reference system containing N non-interacting electrons and each feels the same external potential. This potential makes the ground state electron probability equal to the exact density. The total energy can be written as:

$$E(\rho) = \int \rho(\vec{r})v(\vec{r})d\vec{r} + \bar{T}_s(\rho) + \frac{1}{2} \iint \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 + \Delta T(\rho) + \Delta \bar{V}_{ee}(\rho). \quad 2-7$$

The first term is the electron-nuclei attraction, the second term is the kinetic energy, and third term is the classic repulsion energy between electrons. The last two terms are the difference between the kinetic and potential energy between the reference system and the real system. The first and third terms can be trivially calculated [115]. The second term is equal to:

$$\bar{T}_s(\rho) = -\frac{1}{2} \sum_i \langle \theta_i^{KS}(1) | \nabla_1^2 | \theta_i^{KS}(1) \rangle, \quad 2-8$$

where the θ_i^{KS} is the spatial part of the Kohn-Sham orbital that can be obtained from the equation below for the one-electron non-interacting reference system:

$$\left[-\frac{1}{2}\nabla_i^2 + v_s(\vec{r}_i) \right] \theta_i^{KS} = \varepsilon_i^{KS} \theta_i^{KS}, \quad 2-9$$

where ε_i^{KS} is the Kohn-Sham orbital energy and $v_s(\vec{r}_i)$ is the external potential in the reference system. The last two terms in the Eq. 2-7 are not known. A new term, the exchange-correlation energy functional, is introduced that contains both of them.

$$E_{xc}(\rho) = \Delta T(\rho) + \Delta \bar{V}_{ee}(\rho). \quad 2-10$$

By combining the last four equations, we have:

$$\left[-\sum_n \frac{Z_n}{r_{1n}} - \frac{1}{2}\nabla_1^2 + \int \frac{\rho(\vec{r}_2)}{r_{12}} d\vec{r}_2 + v_{xc}(1) \right] \theta_i^{KS}(1) = \varepsilon_i^{KS} \theta_i^{KS}(1). \quad 2-11$$

The operator in the left hand side of the Eq. 2-11 is the Kohn-Sham operator. In this equation:

$$v_{xc}(\vec{r}) = \frac{\delta E_{xc}[\rho(\vec{r})]}{\delta \rho(\vec{r})}, \quad 2-12$$

where δ shows the functional derivative of the exchange correlation energy. The term $v_{xc}(\vec{r})$ is the exchange-correlation potential. At present, the form of the exact exchange-correlation potential is unknown. In practice, approximate functionals are used. In the next Section, some of the functionals used in this thesis are introduced.

2.2.4 DFT Functionals

Various approximations are introduced to find a proper exchange-correlation functional ($v_{xc}(\vec{r})$) [118-122]. In fact, the accuracy of the DFT calculations depends on the accuracy of $v_{xc}(\vec{r})$. The exchange-correlation functional includes all contributions to the energy that are not included in

the DFT theory such as electron correlation, exchange energy, self-interaction energy and kinetic energy correction. Unfortunately, the exact functionals for exchange and correlation are not known except for the free electron gas. However, several approximations exist. The most widely used approximation is the local-density approximation (LDA), where the functional depends only on the density where the functional is evaluated. Including the electron spin results in local-spin-density approximation (LSDA) method. Generalized gradient approximation (GGA) is another local approximation that takes advantage of the gradient of the density as well. The meta-GGA approach uses the second derivative of the density. Hyper-GGA functionals include hybrid functionals and offer good performance.

The B3LYP hybrid functional is one of the most widely used exchange-correlation functionals [118,122]. This functional combines the gradient-corrected HF exchange energy, local-spin-density approximation functional [120], B88 exchange functional [118], and Lee-Yang-Parr correlation functional [119]. Note that this functional and several other DFT functionals use empirical parameters and in fact cannot be strictly called *ab initio* methods. Throughout this thesis, we also have used other functionals such as B97-D [123], and mPW2PLYP-D [124-129]. As will be discussed, the B3LYP functional is missing long range dispersion contributions and for dispersion dominant interactions, it is not accurate enough.

The B97-D functional is a new GGA functional introduced by Grimme in 2006 that includes a post-DFT calculation dispersion correction term $\frac{C_6}{R^6}$. The total energy is given by [123]:

$$E_{DFT-D} = E_{KS-DFT} + E_{dis} , \quad 2-13$$

Here, the dispersion energy is simply added to the Kohn-Sham energy. The dispersion is calculated from:

$$E_{dis} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{damp}(R_{ij}), \quad 2-14$$

where N_{at} is the number of atoms in the system, C_6^{ij} is the dispersion term between atom i and j , R_{ij} is the inter-atomic distance, $f_{damp}(R_{ij})$ is a damping function to avoid singularity for small R_{ij} , and s_6 is a global scaling factor. This functional has performed well in predicting properties of non-covalent bound complexes and thermochemical properties of transition metals [123].

The mPW2PLYP-D functional was introduced by Grimme and Schwab in 2006 and is a hybrid-GGA functional [124]. Here, the calculated Kohn-Sham orbitals based on the hybrid-GGA method are used as input for a standard second-order Moller–Plesset type perturbation theory correction. This method has shown to be good for capturing weak interactions and gives the lowest mean absolute deviation over the whole G3/05 set reported for DFT functionals [123,124].

2.2.5 Second order Moller-Plesset Perturbation Theory

In 1934, Moller and Plesset [115] introduced a perturbation method that works based on the Hartree-Fock wave function as the unperturbed wave function. The first applications of this theory were in 1975-1981 with works of Pople and Bartlett [130] where second order corrections to the energy were calculated. In this method, the perturbed Hamiltonian, \hat{H}' , is the difference between the molecular Hamiltonian \hat{H} and \hat{H}^0 , which is the sum of one electron Fock operators:

$$\hat{H}' = \hat{H} - \hat{H}^0. \quad 2-15$$

To the first order, the energy correction is simply equal to the Hartree-Fock energy. To improve the method, the second energy correction is added:

$$E_0^{(2)} = \sum_{s \neq 0} \frac{|\langle \psi_s^{(0)} | \hat{H}' | \Phi_0 \rangle|^2}{E_0^{(0)} - E_s^{(0)}}, \quad 2-16$$

where $E_0^{(0)}$ is the ground state energy of the unperturbed system, Φ_0 is the ground state unperturbed wave function, and $\psi_s^{(0)}$ is unperturbed wave function for all states except 0. Higher MP calculations are also possible by calculating the third, fourth, ... energy corrections to the energy of the ground state. Second order Moller-Plesset (MP2) theory captures most of the electron correlation and is a relatively fast method [130].

For short poly-proline chains, we have used MP2 calculations while for longer chains, where MP2 is not feasible, resolution of identity MP2 (RI-MP2) is employed [131]. This method is an efficient implementation of orbital-optimized MP2 within the resolution of identity approximation [131]. In this method product of orbitals are expanded in an auxiliary basis set [132,133]. This method speeds up the calculations dramatically while the associated error is very small [131].

2.2.6 Coupled Cluster Methods

The coupled cluster (CC) method was introduced by Cizek [134] in 1966. The main equation of this method is:

$$\psi = \exp(\hat{T}) \Phi_0, \quad 2-17$$

where ψ is the exact ground-state wave function of the molecule, Φ_0 is the normalized ground state Hartree-Fock wave function, and \hat{T} is the cluster operator. The $\exp(\hat{T})$ is defined as below:

$$\exp(\hat{T}) = 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{\hat{T}^k}{k!}, \quad 2-18$$

And the cluster operator is:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_n. \quad 2-19$$

Here, n is the number of electrons. Each of the \hat{T} operators corresponds to simultaneous excitation of n electrons. The one-particle excitation operator, \hat{T}_1 , and two-particle excitation operators, \hat{T}_2 , are:

$$\hat{T}_1 = \sum_{a=n+1}^{\infty} \sum_{i=1}^n t_i^a \Phi_{ij}^a, \quad 2-20$$

$$\hat{T}_2 = \sum_{b=a+1}^{\infty} \sum_{a=n+1}^{\infty} \sum_{j=i+1}^n \sum_{i=1}^{n-1} t_{ij}^{ab} \Phi_{ij}^{ab},$$

where Φ_i^a is the singly excited Slater determinant in which the occupied spin-orbital number i is replaced with a virtual spin-orbital a . The \hat{T}_1 operator converts the Slater determinant of Φ_0 into linear combination of all possible singly excited Slater determinants, Φ_{ij}^a . Here Φ_{ij}^{ab} corresponds to a Slater determinant in which two occupied spin-orbitals (denoted by i and j) are replaced with two virtual spin-orbitals (denoted by a and b). By including higher excitation operators, all possible excited states are taken into account. So, the ψ wave function is now expressed as a linear combinations of Slater determinants. Mixing of such excited states allows electrons to be kept away from each other which in turn produces electron correlation. The goal then is to calculate the $t_i^a, t_{ij}^{ab}, t_{ijk}^{abc}, \dots$ coefficients. Along with truncation of the basis set, not all of the cluster operators are used. The most important contribution in cluster operator come from the \hat{T}_2 operator. So:

$$\psi_{CCD} = \exp(\hat{T}_2) \Phi_0 . \quad 2-21$$

This approximate method is called the coupled cluster doubles (CCD) method. The CCD wave function only includes determinants with double, and some quadruple, hexuple, and ... substitutions. Note that, for example, the quadruple substitutions are not treated fully in CCD method since the corresponding Slater determinates are produced by two operations of \hat{T}_2^2 and thus the coefficients are not determined independently. By assuming that the double excitations are the most important contribution and by ignoring other cluster operators:

$$E_{CCD} = \langle \Phi_0 | \hat{H} | \exp(\hat{T}_2) \Phi_0 \rangle . \quad 2-22$$

To improve the accuracy, other cluster operators can be added too. For instance adding the single excitation operator gives coupled cluster singles and doubles (CCSD) method. Including the triple excitation operator also gives coupled cluster singles, doubles, and triples (CCSDT) method. CCSDT is considered the “gold standard” of *ab initio* calculations and is very accurate but computationally demanding.

During this thesis, we have used CCSD single point calculations to capture the energies of mono- and di-proline selector conformations. For longer chains, though, the CCSD method was not feasible. In such cases, the local pair natural orbital (LPNO) [135,136] method is used as an approximation to full CCSD. The LPNO-CCSD method works based on the idea of local correlation domains. Here, localized internal orbitals and projected atomic orbitals are used to span the virtual space. This method tries to only choose orbitals that contribute most to the total wave function correlation. It is assumed that only electron pairs that are close have important contributions into total correlation. So, the wave function is compacted by using local orbitals. Such local orbitals are obtained by pair natural orbitals (PNOs). They are a highly compact set of

orbitals that are different for each electron pair. The accuracy of this method is controlled by two adjustable parameters: $T_{CutPairs}$ and T_{CutPNO} . The first term determines the threshold used to ignore or include electron correlation energy. The second term shows where the PNO expansion is truncated. The advantage is that this method can be linearly scaled in terms of the CPU, memory and disk usage. This method captures most of the total electron correlation [135,136].

2.2.7 Solvation method

To calculate the effects of the solvent on the conformation and stability of the conformers, solvation studies are also performed. Here, it is assumed that each atom is surrounded by a sphere which is 1.2 times larger than its van der Waals radius [115]. The resulting overlapping 3D shape of the connected atoms is called cavity. The charge of the solute is assumed to be distributed on the surface of the cavity. In conductor like solvation model (COSMO), these charges are calculated using a condition suitable for a solvent medium. Here solvent is treated as an electrical conductor rather than a dielectric medium. The initial charges are then scaled by the solvent dielectric constant. Then numerical methods are used to find the solvent-solute interaction potential energy term.

2.3 Molecular Dynamics Simulations

2.3.1 Introduction to Molecular Dynamics Simulations

Quantum mechanical calculations have developed rapidly in recent years with the advent of new computers and methods to accelerate the calculations. However, these methods are still expensive and are limited to systems with a relatively small number of atoms. On the other hand, classical molecular dynamics methods are applicable to much larger systems and are widely used. Computer simulations were first reported by Metropolis [137] in 1952 and have been developed

extensively since then. They are useful tools to verify a new theory since they can be tested against experimental results. In fact, computer simulations are “experiments on computers” [137].

Molecular dynamics simulations are often employed to study the dynamics and equilibrium properties of systems. Briefly, the molecular or atomic interactions are calculated at a given time. Then the system moves to a new position based on the forces experienced. On the new configuration, the forces are again calculated and are used to advance the atoms, again. This process happens for a long time, enough to explore the full configuration space. Properties of interest are calculated during the time that the system evolves.

For a typical system of N atoms, the classical Hamiltonian can be written as a sum of the kinetic and potential terms

$$H(\vec{p}^N, \vec{q}^N) = \sum_i \frac{\vec{p}_i^2}{2m_i} + U(\vec{q}^N), \quad 2-23$$

where \vec{p}_i , \vec{q}_i and m_i , are the linear momentum, position and mass of atom i , respectively. Here, the summation goes over all the atoms in the system. The first term is the kinetic energy and the second term is the potential energy. The Newtonian equations of motion for such a system are defined as:

$$\begin{aligned} \frac{d\vec{q}_i}{dt} &= \frac{\vec{p}_i}{m_i}, \\ \frac{d\vec{p}_i}{dt} &= -\frac{\partial U(\vec{q}^N)}{\partial \vec{q}_i} = \vec{f}_i, \end{aligned} \quad 2-24$$

where \vec{f}_i is the total force on atom i , \vec{q}_i is the position of atom i , \vec{p}_i is the linear momentum of atom i , dt is the time step, and $U(\vec{q}^N)$ is the potential energy of the system. During the simulation, these equations are solved at each time step.

2.3.2 Force Field

The force fields describe the atomic interactions in the system. They are used to calculate the overall force on each atom that is the result of contributions from other atoms. These forces, \vec{f}_i , are used to advance the positions of the atoms during the simulation (Eq. 2-24).

Most potentials can be divided into two categories: intra-molecular and inter-molecular potentials. The intra-molecular potential reflects the energetic costs associated with the molecule's flexibility. For instance, when an angle increases or a twist about a dihedral occurs intra-molecular potentials can be used to calculate the energy cost for these changes. The non-bonded potential is defined between atoms of different molecules or atoms within a molecule that are far from each other. We first begin with the intra-molecular potentials. Most force fields include stretching, bending, proper and improper torsion potentials. Figure 1-3 shows each of these degrees of freedom. The stretching potential defines the energy cost for changing the distance between two bonded atoms. The bending potential calculates the energy costs for changing the angle between three consecutive atoms. The proper torsion potential is for four consecutively connected atoms and the improper torsion applies to out-of-plane motion.

The stretching potential defines the energy cost associated with pushing two atoms towards each other or pulling them apart. To accurately model this energy-distance relation, a Morse potential model can be used. Figure 2-2 shows a typical stretching potential between two atoms obtained by this potential.

$$U^{stretch}(r) = D_{eq} [1 - \exp\{-a(r - r_{eq})\}]^2, \quad 2-25$$

where r_{eq} is the equilibrium bond length, r is the distance between the two atoms, D_{eq} is the well depth and a is a constant defining the width of the potential well. This simple model is flexible enough to produce energy-distance behavior that is close to the experiment. However, this

potential is not often used for computational purposes. The reason is that this term should be calculated many times during each time step and much more during the whole simulation. Typically calculation of exponential terms is computationally more demanding than other simple mathematical functions. For this reason, instead, a harmonic potential is often used:

$$U^{stretch}(r) = k_r(r - r_{eq})^2, \quad 2-26$$

where k_r is the force constant. This potential is acceptable only near the minimum, at bond lengths close to r_{eq} . In practice a harmonic potential is satisfactory since the energy penalty to

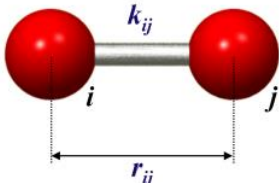
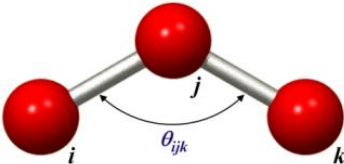
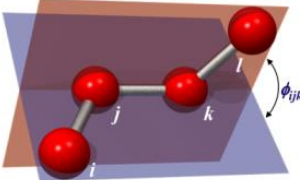
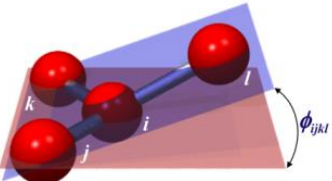
	stretching
	bending
	proper torsion
	improper torsion

Figure 2-1. Illustration of intra-molecular motions.

reduce or increase the bond length is relatively high. Indeed, the bond length is practically fixed because of the steepness of the stretching potential at room temperature. It is expected that a harmonic potential is a good representation for simulations at typical temperatures.

The bending potential is also defined by a harmonic potential:

$$U^{bend}(\theta) = k_{\theta}(\theta - \theta_{eq})^2, \quad 2-27$$

In the above equation, θ is the angle, θ_{eq} is the equilibrium value of the angle and k_{θ} is the associated force constant.

The torsional behavior, however, is more complex and needs a more flexible model. At typical temperatures, all torsional angles may be sampled. We have used the Ryckaert-Bellemans [138] model:

$$U^{torsion} = \sum_t \sum_{i=0}^6 c_{it} \left(\cos(\varphi_t + \varphi_t^{i;0}) \right)^i, \quad 2-28$$

where $\varphi_t^{i;0}$ is a phase shift, φ_t is the torsional angle, and c_{it} is a linear coefficient. We include the total of seven terms so that model is flexible enough to closely mimic the torsional energy predicted from *ab initio* calculations.

Out-of-plane motion is also modeled by a harmonic potential. Consider three atoms j , k , and l that are connected to a central atom i . The potential for improper torsion (ω) is defined by the angle between the vectors normal to ijk and jkl planes:

$$U^{improper-torsion}(\omega) = k_{\omega}(\omega - \omega_{eq})^2, \quad 2-29$$

where ω_{eq} , and k_{ω} are the improper torsion equilibrium angle, and the force constant, respectively. Improper torsional potential is often employed to conserve chirality or the planarity in a molecule [139].

It is worth noting that for enantiomers most of the potentials terms stay the same. By calculating the related force constants for one enantiomer, one does not need to re-do this for the other enantiomer. The only differences come from proper and improper torsional potentials. Considering the fact that the enantiomers are mirror images of each other, the vectors used to define the proper and improper torsions can be reversed to produce the related potentials for the other enantiomer. This is equivalent to changing the sign of the phase shifts (Eq. 2-28) and equilibrium improper torsional angle (Eq. 2-29).

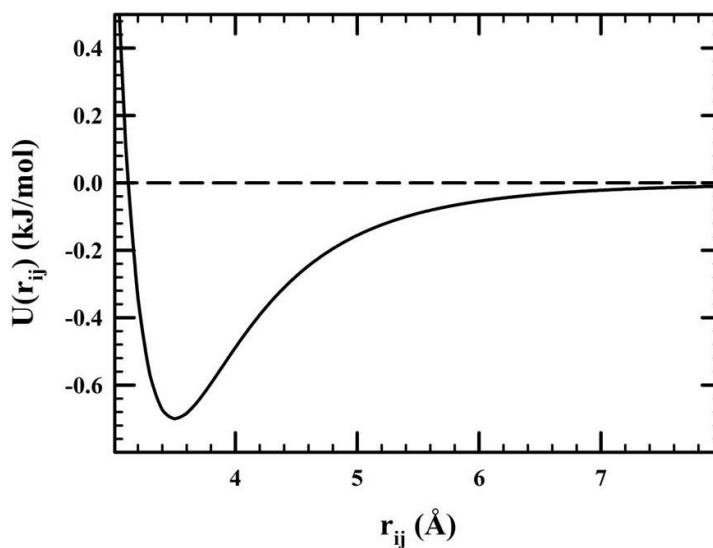


Figure 2-2. A typical stretching potential between two atoms i and j .

Inter-molecular potentials consist of short range repulsion and dispersion and also electrostatic interactions. A Lennard-Jones (LJ) potential is introduced for the first two and is defined as:

$$U^{LJ}(r_{ij}) = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad 2-30$$

where r_{ij} is the distance between atoms i and j . The definitions of the energy (ε_{ij}) and length (σ_{ij}) parameters can be understood from the plot of this potential against the inter-atomic distance in Figure 2-3. The σ_{ij} is the distance parameter, defining the distance where the LJ potential is zero. The ε_{ij} is the well depth of the potential curve and shows the strength of the interaction between atoms i and j . Each atom has distinct values for σ and ε . When calculating the LJ potential between pair of atoms, these atomic terms are mixed for two atoms according to the Lorentz-Berthelot combining rules [140,141]:

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}, \varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2}. \quad 2-31$$

The first term in the LJ potential corresponds to the very steep repulsion when atoms are closer than the equilibrium bond length. The second term corresponds to the dispersion between the two atoms. The LJ potential is a short range potential since it is quite small when atoms are more than, say 10 Å apart.

During the simulation, the LJ potential is calculated between any pairs of atoms that do not share any intra-molecular potential. So, for a system of N atoms, this calculation should be done about $N(N-1)$ times at each time step.

A Coulomb potential is used for the electrostatic interactions between ions or partially charged atoms:

$$U^{ELE}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}, \quad 2-32$$

where q_i is the charge on atom i , r_{ij} is the inter-atomic distance and ϵ_0 is the dielectric constant. In most of force fields, the charges are constant. However, in recent years, polarizable force fields with fluctuating charges or atomic dipoles are also developed [142,143].

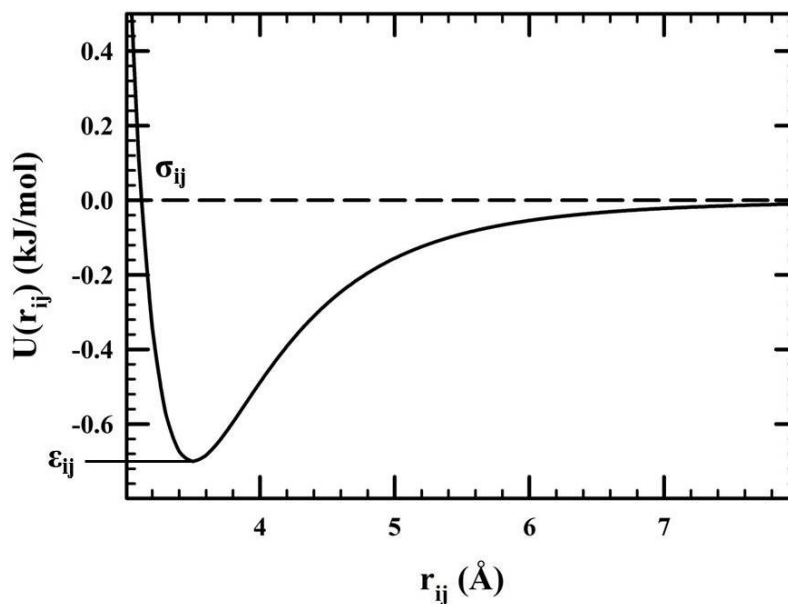


Figure 2-3. LJ potential as a function of the inter-atomic distance between two atoms i and j .

2.3.3 Periodic Boundary Conditions

In practice, all computer simulations are limited in the number of atoms they can handle. Although computing facilities are extensive nowadays, they are still far away from simulating moles of atoms. As a result, we are limited to systems with a small number of atoms compared to real systems. This causes an artificial surface effect since the ratio of atoms at the surface of the system to the number of atoms in the bulk is much larger than the same ratio in a real system: There is always a non-negligible number of atoms that do not interact with any atom from, say,

above and do not “feel” similar to bulk atoms. To overcome this problem, periodic boundary conditions are applied [144]. Here, the simulation cell is replicated in all three directions. Figure 2-4 shows a representation of such conditions. Consider the central cell. During the simulation, when an atom moves outside of its box, its mirror atoms in all replicated cells move the same way. This causes an image atom to enter the cell from the other side, removing the surface effect. For a system with N atoms, each atom interacts with all other atoms in other replicated cells. So, the total potential energy of the system, assuming a cubic cell, is:

$$U = \frac{1}{2} \sum_{\vec{n}} \sum_{i=1}^N \sum_{j=1}^N U(|\vec{r}_{ij} + \vec{n}L|), \quad 2-33$$

where i and j are two atoms in the central cell, L is the side length of the box, and \vec{n} is defined as $\vec{n} = (n_x, n_y, n_z)$ with three n_i integers so that the summation over \vec{n} goes over all periodic boxes. Note that the term $i = j$ for the central cell is not calculated in the summation since the particle does not interact with itself, but it does with its images. At the moment, periodic boundary conditions seem to be impractical since the summation is infinite and it allows the calculations to run forever. However, in the next section we will see how this summation can be dealt with.

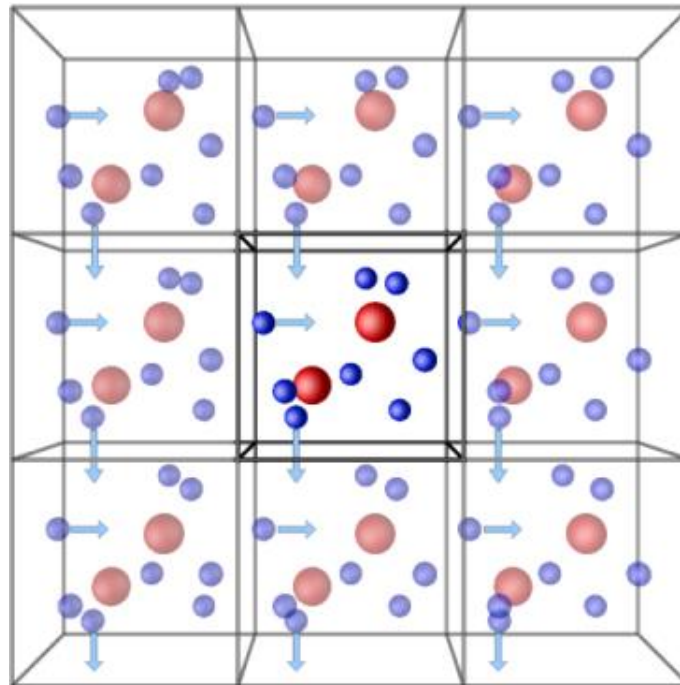


Figure 2-4. A schematic representation of a three dimensional periodic boundary conditions. Picture is taken from: <http://isaacs.sourceforge.net/phys/pbc.html>. [Last Accessed at 2012].

2.3.4 Truncation of the Lennard-Jones Potential

As mentioned, the LJ potential decays rapidly with distance. Physically, it means that when the atoms are far from each other, they only interact weakly and these interactions can be ignored without loss of accuracy. So, for the LJ potential it is reasonable to define a cut-off distance, r_c , and ignore atomic interactions when they are further than this distant. The LJ potential term is then truncated and shifted:

$$\begin{cases} U^{tru-shi}(r) = U^{LJ}(r) - U^{LJ}(r_c) & , \text{when } r \leq r_c \\ U^{tru-shi}(r) = 0 & , \text{when } r > r_c \end{cases} \quad 2-34$$

Figure 2-5 shows the definition of the cut-off potential.

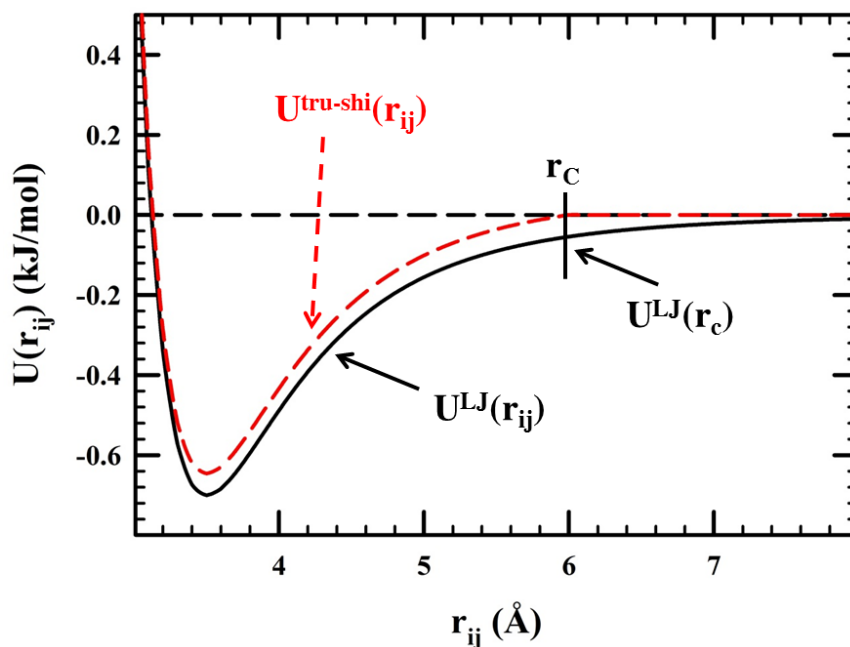


Figure 2-5. A schematic representation of the cut-off and the truncated-shifted LJ potential.

Note that the LJ potential needs to be both truncated and shifted, otherwise the potential term will be a discontinuous function of the distance and the force, defined from the negative of the first derivative of the potential, would be infinite at r_c . In practice, the cut-off distance should be chosen with care. It should be large enough that missing interactions have a small contribution to the potential. Also, assuming a cubic cell, it should be less than half of the cell diameter so that an atom does not interact with its image from other cells. When the cell is not cubic, the cut-off is chosen to be less than half the smallest side length of the cell. Since the chosen cut off distance is around 10-12 Å in our simulations, the discontinuity of the force is not supposed to have considerable effects on the results.

2.3.5 Ewald Summations

The application of periodic boundary conditions is consistent with short ranged interactions: the LJ potential can simply be truncated and shifted. However, the electrostatic potential does not decay as fast as the LJ potential and using the same cut-off potentials can result in large errors [145]. Some simulations do use truncation for electrostatics, nonetheless, but we employ Ewald summations [137,146]. For a system of N atoms in a cubic simulation cell, the total electrostatic potential is:

$$U^{ELE}(r_{ij}) = \frac{1}{2} \sum_{i=1}^N q_i \varphi(r_i), \quad 2-35$$

where the term $\varphi(r_i)$ is the electrostatic field at position where the charge i is located and is defined as:

$$\varphi(r_i) = \sum_{\vec{n}} \sum_{j=1}^N \frac{q_j}{|\vec{r}_{ij} + \vec{n}L|}, \quad 2-36$$

Note that the term $4\pi\epsilon_0$ is omitted. The summation in the above equation is conditionally convergent [137]: We need to use another approach to treat electrostatic potentials.

Ewald summation is a simple method to treat the electrostatic potentials [146,147]. Here we assume that each point charge q_i is surrounded by a charge cloud of equal and opposite sign. This cloud can adopt any shape, in general, but a Gaussian shape is usually chosen for mathematical simplicity. The cloud charge screens and cancels the charge on each particle.

$$\varrho(r) = -q_i \left(\frac{\alpha}{\pi}\right)^{\frac{3}{2}} \exp(-\alpha r^2), \quad 2-37$$

where α is a parameter determining the width of the screening cloud. For each of the charges in the simulation cell, one screening charge is assigned that exactly cancels the point charge (see

Figure 2-6). Also, for each screening charge, another compensating charge is assigned that cancels it. So point charges, screening charge cloud, and compensating charge cloud all contribute to the total potential where the point charge is located. Now, the electrostatic field for the middle box in Figure 2-6 is short range and can be calculated by a summation over all screening charges in all cells. The field produced by the compensating charge, right box in Figure 2-6, is calculated in Fourier space since the field is a smooth function and is periodic. Note that the compensating cloud for point charge i is also included. In the end, the interaction between a point charge and its screening charge must be removed. The mathematical derivation of the Ewald summation is not given here; only the final expression is presented in Eq. 2-38. For a complete derivation see reference [137].

$$\begin{aligned} \varphi(r_i) = & \sum_{\vec{n}} \sum_{j=1}^N \frac{q_j}{|\vec{r}_{ij} + \vec{n}L|} \operatorname{erfc}(\sqrt{\alpha}|\vec{r}_{ij} + \vec{n}L|) \\ & + \frac{1}{V} \sum_{\vec{k} \neq 0} \sum_{j=1}^N \frac{4\pi q_j}{k^2} \exp(i\vec{k} \cdot \vec{r}_{ij}) \exp\left(-\frac{k^2}{4\alpha}\right) + 2(-q_i) \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}}, \end{aligned} \quad 2-38$$

where V is the volume of the central simulation cell, and \vec{k} is the lattice vector in Fourier space and is equal to $\left(\frac{2\pi}{L}\right)\vec{l}$ and $\vec{l} = (l_x, l_y, l_z)$. The function $\operatorname{erfc}(x)$ is the complementary error function [137]. Each term in the above equation corresponds to a box in Figure 2-6. The α parameter needs to be adjusted since the convergence of the above equation depends on the value of α . A large α causes the first term to converge quickly, but makes the second term converge slowly. The value of α is usually chosen such that the first term can be truncated within the central cell (the sum over \vec{n} can be eliminated).

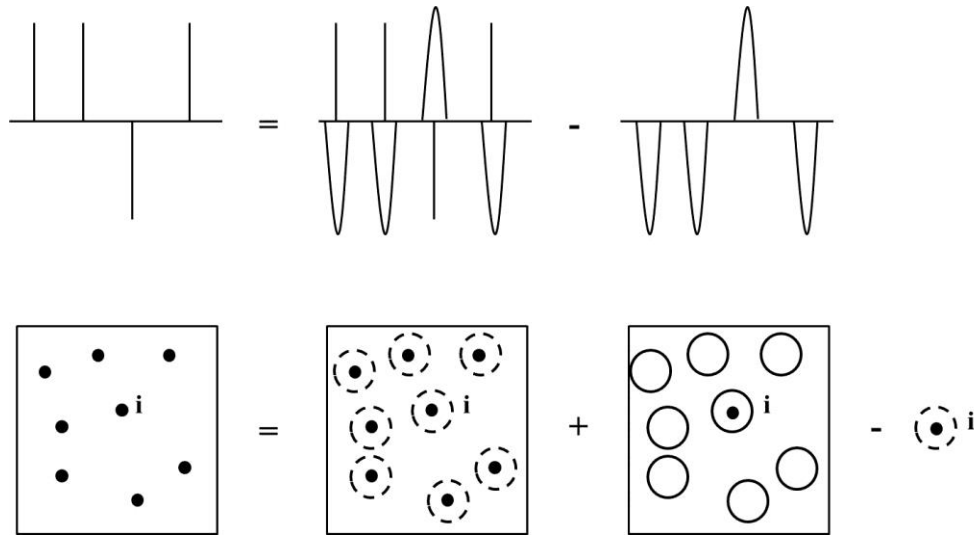


Figure 2-6. Illustration of the point charge treatment in the Ewald method.

2.3.6 Integrating the Equations of Motion

Once a force field is defined, all forces can be calculated on all atoms and then used to advance the positions of atoms. To move each atom, Newton's equations of motion must be solved as discussed in Section 2.3.1. Several algorithms are commonly used to advance the system in time. Perhaps, the simplest is the normal Verlet algorithm [137,148]. If, the positions are expanded using Taylor's expansion around time t and using time steps of Δt ,

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{\vec{f}(t)}{m} \frac{\Delta t^2}{2} + \mathcal{O}(\Delta t^3) + \dots, \quad 2-39$$

$$\vec{r}(t - \Delta t) = \vec{r}(t) - \vec{v}(t)\Delta t + \frac{\vec{f}(t)}{m} \frac{\Delta t^2}{2} + \mathcal{O}(\Delta t^3) + \dots,$$

where $\vec{r}(t + \Delta t)$ and $\vec{r}(t - \Delta t)$ are new and old positions, respectively, $\vec{r}(t)$ is the current position, $\vec{v}(t)$ is the current velocity, and $\vec{f}(t)$ is the current force. By adding these two equations we get the normal Verlet equation:

$$\vec{r}(t + \Delta t) = 2\vec{r}(t) - \vec{r}(t - \Delta t) + \frac{\vec{f}(t)}{m} \Delta t^2 + \mathcal{O}(\Delta t^4) + \dots \quad 2-40$$

So, the position at some time in the future can be found based on the current position, the old position and the current force acting on the atom. The errors associated with this algorithm are in the order of Δt^4 . By choosing a small time step, this error can be reduced down to an acceptable level. In this algorithm, the velocity is not necessarily calculated to advance the positions. To calculate the kinetic energy, however, it can be independently calculated using the positions. In our simulations, we use the velocity version of the Verlet algorithm along with the Leap-Frog algorithm [149]. In the Velocity Verlet algorithm, velocities are explicitly used to advance the position. The other advantage of this algorithm is that it is consistent with the application of RATTLE [150] and Nosé-Hoover thermostats. The equations of motion for Velocity Verlet are

$$\begin{aligned} \vec{r}(t + \Delta t) &= \vec{r}(t) + \vec{v}(t)\Delta t + \frac{\vec{f}(t)}{m} \frac{\Delta t^2}{2}, \\ \vec{v}(t + \Delta t) &= \vec{v}(t) + \frac{\Delta t}{2m} [\vec{f}(t) + \vec{f}(t + \Delta t)]. \end{aligned} \quad 2-41$$

The velocities and positions are advanced in a two-step process. In the first step, the positions are advanced fully and velocities are advanced a half time step:

$$\vec{v}\left(t + \frac{\Delta t}{2}\right) = \vec{v}(t) + \frac{\Delta t}{2m} \vec{f}(t). \quad 2-42$$

Then, the forces at the new positions, $\vec{f}(t + \Delta t)$, are calculated and the velocity is advanced another half time step.

$$\vec{v}(t + \Delta t) = \vec{v}\left(t + \frac{\Delta t}{2}\right) + \frac{\Delta t}{2m} \vec{f}(t + \Delta t). \quad 2-43$$

The Velocity Verlet algorithm is time reversible, area preserving and has good energy conservation. On the whole, Verlet algorithms might not have good short-term energy conservation, but in the long run, they conserve the energy reasonably well.

Several other algorithms are equivalent to the Verlet algorithm. For instance, Leap-Frog algorithm calculates velocities at half time steps and uses them to find the new positions:

$$\vec{v}\left(t - \frac{\Delta t}{2}\right) = \frac{\vec{r}(t) - \vec{r}(t - \Delta t)}{\Delta t}, \quad 2-44$$

$$\vec{v}\left(t + \frac{\Delta t}{2}\right) = \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t}.$$

From the second equation, the new positions are:

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}\left(t + \frac{\Delta t}{2}\right) \Delta t. \quad 2-45$$

For the update of the velocities, Verlet algorithm could be used:

$$\vec{v}\left(t + \frac{\Delta t}{2}\right) = \vec{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\vec{f}(t)}{m} \Delta t. \quad 2-46$$

Leap-Frog algorithm produces the same trajectory as Verlet algorithm. One may note that the velocities are calculated at half time steps, so, kinetic and potential energies are not calculated at the same time.

Clearly, small time steps will reduce the error associated with integrating the equations of motion, but the simulations advance very slowly in time. In contrast, large time steps decreases the accuracy but the simulation will finish in a reasonable time. To capture properties of interest, the time step of the simulation must be appreciably smaller than the period for the fastest motion of the system. For molecular systems, this motion is intra-molecular bond stretching. So, the time step must be considerably smaller than the vibrational period for the fastest stretch in the

molecule. For a classical molecular dynamics simulation, since no bond breaks and no bonds form between atoms, and changes in bond lengths are small, bond stretching often has little impact. Bond lengths can be frozen during the simulation using algorithms such as SHAKE [151] proposed by Ryckaert *et al.* or RATTLE [150] proposed by Andersen. In the Ryckaert approach an extra force is added to each atom which is part of the stretching potential that needs to be frozen.

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{f}_i + \vec{g}_i, \quad 2-47$$

where \vec{f}_i is the stretching force and the \vec{g}_i is the constraint force to keep the bond length the same.

By using the Verlet algorithm:

$$\vec{r}(t + \Delta t) = \vec{r}(t + \Delta t) + \frac{\Delta t^2}{m} \vec{g}_i, \quad 2-48$$

where the $\vec{r}(t + \Delta t)$ is the position obtained from normal application of the Verlet algorithm without any constraint. Since the constraint force should be directed along the bond and should also follow Newton's law, then:

$$\vec{g}_i = \sum_j \lambda_{ij} \vec{r}_{ij}, \quad 2-49$$

where $\lambda_{ij} = \lambda_{ji}$, and i represents all atoms having stretching potential with the atom j . λ_{ij} are undetermined time-dependent Lagrange multipliers. By combining the last two equations we will have:

$$\begin{aligned} \vec{r}_{ij}(t + \Delta t) = & \vec{r}_{ij}(t + \Delta t) - \sum_{h \neq j} \frac{\Delta t^2}{m_h} \lambda_{hi} \vec{r}_{hi}(t) + \frac{\Delta t^2 (m_i + m_j)}{m_i m_j} \lambda_{ij} \vec{r}_{ij}(t) \\ & - \sum_{k \neq i} \frac{\Delta t^2}{m_h} \lambda_{jk} \vec{r}_{jk}(t), \end{aligned} \quad 2-50$$

where h goes over all atoms except atom j which is bonded to atom i . Similarly, k runs over all atoms except i which is bonded to atom j . By taking the square module of both sides of the above equation, we will have:

$$|\vec{r}_{ij}(t + \Delta t)|^2 = |\vec{r}_{ij}(t)|^2 = d_{12}^2, \quad 2-51$$

where d_{12} is the equilibrium bond length. The result is a series of quadratic equation for λ_{ij} that are solved by iteration. The answers of λ_{ij} are then used to calculate the \vec{g}_i .

In the RATTLE algorithm, the constraints are added to the Velocity Verlet algorithm:

$$\begin{aligned} \vec{r}(t + \Delta t) &= \vec{r}(t) + \vec{v}(t + \Delta t) + \frac{\vec{f}(t)}{2m} \Delta t^2, \\ \vec{v}(t + \Delta t) &= \vec{v}(t) + \frac{[\vec{f}(t) + \vec{f}(t + \Delta t)]}{2m} \Delta t. \end{aligned} \quad 2-52$$

This algorithm allows calculating the positions and velocities at time $t + \Delta t$ by using the position and velocity at time t without any need to know the position and velocity at previous times. In the Rattle algorithm two independent constraints are satisfied at the same time for position and velocity. Rattle has two major advantages over the SHAKE algorithm. First, it has more precision and it is easier to be modified to be compatible with constant temperature or constant pressure MD. In Rattle algorithm, we have:

$$\sigma_{12}\{\vec{r}(t)\} = |\vec{r}_1(t) - \vec{r}_2(t)|^2 - d_{12}^2, \quad 2-53$$

where d_{12} is the fixed distance between atoms 1 and 2 and $\sigma_{12}\{\vec{r}(t)\}$ is the restraint $\sigma_{12}\{\vec{r}(t)\} = 0$. The time derivative of this constraint equation gives the constraints for velocities.

$$\left| \frac{d\vec{r}_1(t)}{dt} - \frac{d\vec{r}_2(t)}{dt} \right| \cdot |\vec{r}_1(t) - \vec{r}_2(t)| = 0. \quad 2-54$$

The constraint equations are

$$m_1 \frac{d^2 \vec{r}_1(t)}{dt^2} = F_1 + G_1, \quad 2-55$$

where F and G are the real forces on the atom 1 and the constraint force on atom 1, respectively.

This constraint force is:

$$G_1 = - \sum_j \lambda_{1j}(t) \nabla_1 \sigma_{1j}, \quad 2-56$$

the sum goes over all atoms that are connected to atom 1. The λ_{1j} is a time-dependent Lagrange multiplier. In the Rattle algorithm, λ_{1j} is approximated. The Eqs. 2-52 change to the below equations.

$$\begin{aligned} \vec{r}(t + \Delta t) &= \vec{r}(t) + \vec{r}'(t + \Delta t) + \frac{[\vec{f}(t) - 2 \sum_j \lambda_{1j}(t) \vec{r}_{1j}(t)]}{2m_1} \Delta t^2, \quad 2-57 \\ \vec{v}(t + \Delta t) &= \vec{v}(t) + \\ &\frac{[\vec{f}(t) - 2 \sum_j \lambda'_{1j}(t) \vec{r}_{1j}(t) \times \vec{f}(t + \Delta t) - 2 \sum_j \lambda'_{1j}(t + \Delta t) \vec{r}_{1j}(t + \Delta t)]}{2m_1} \Delta t. \end{aligned}$$

Note that two distinct time-dependent Lagrange multipliers are used for position and velocity. In this thesis, we used the RATTLE algorithm. In general, constrained dynamics provide a way to fix atomic distances using a simple and fast algorithm. The example shown above was concerned about constraining a bond, although similar approaches can be taken to freeze any distance.

The Rattle algorithm, although used extensively by the molecular simulation community, does not represent the exact correct dynamics. It has been shown that the dynamics of the Rattle or SHAKE has a slight deviation from the exact dynamics. Patriciu et al. [152], for example, have shown that equilibrium statistics and conformational transitions by constrained methods are not

exact. They have proposed a correction term which could be added to the potential. Although, their method is interesting, it is supposed to have very minor effects on the results.

2.3.7 Nosé-Hoover Thermostat

In the canonical NVT ensemble, the temperature is constant. Experimentally, the system can be kept in contact with a thermal reservoir. In simulations, several methods are available to yield canonical properties such as velocity rescaling, Nosé-Hoover thermostats, Berendsen thermostats, and Langevin dynamics [137,144]. In this work, extended dynamics methods are used to keep the temperature at the desired value. Here, one additional degrees of freedom is added along with an associated momentum [153]. Constant temperature dynamics for a simulation with fixed time step was introduced by Hoover [154] who expanded upon earlier work by Nosé [152]. The extended Hamiltonian is:

$$H_{NH} = \sum_i^N \frac{\vec{p}_i^2}{2m_i s^2} + U(\vec{r}^N) + \frac{p_s^2}{2Q} + g \frac{\ln s}{\beta}, \quad 2-58$$

where s is the additional degree of freedom, Q is the associated mass, p_s is the momentum for this degree of freedom, and $\beta = (k_B T)^{-1}$. The parameter g is equal to the number of degrees of freedom in the system plus one. Using this Hamiltonian, the ensemble average of the whole extended system is equal to the NVT ensemble average

$$\langle A(\vec{p}/s, \vec{r}) \rangle_{NH} = \langle A(\vec{p}, \vec{r}) \rangle_{NVT}, \quad 2-59$$

where A is any physical property. The equations of motion for a system are then:

$$\begin{aligned}\frac{d\vec{r}_i}{dt} &= \frac{\vec{p}_i}{m_i}, \\ \frac{d\vec{p}_i}{dt} &= -\frac{\partial U(\vec{r}^N)}{\partial \vec{r}_i} - \zeta \vec{p}_i, \\ \frac{d\zeta}{dt} &= \frac{\sum_i \frac{p_i^2}{m_i} - \frac{g}{\beta}}{Q}, \\ \frac{d \ln s}{dt} &= \zeta,\end{aligned}$$

where ζ is a thermodynamic friction coefficient. The last equation is only solved to check the conservation of the Nosé-Hoover Hamiltonian. These equations need to be solved by iterations since ζ and p_i are coupled.

2.3.8 Quaternions for Rotation of Rigid Moieties

Often it is needed to treat molecules, or part of molecules, as rigid. A rigid body translates and rotates according to the forces and torques applied on its center of mass. Such systems need specific equations of motion [144]. Note that in rigid moieties, the relative positions of atoms within the rigid moieties do not change during the simulation. Figure 2-7 is an illustration of Euler angles ϕ , θ , and ψ that are used to relate the space-fixed frame ($z_1 y_1 x_1$) to the body-fixed frame ($z_B y_B x_B$). This rotation happens through three consecutive counter-clockwise rotations, about the Z_1 axis by ϕ , about the line of nodes (x') axis by θ , and about the Z_B axis by ψ . So, a rotation matrix could be defined to rotate the rigid moiety from space-fixed frame to the body-fixed frame.

However, equations of motions using Euler angles are problematic because of the singularity in the first derivative of the trigonometric functions. An alternative approach,

introduced by Evans [155], uses quaternions as generalized coordinates. Quaternions are a set of four values that are normalized:

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1. \quad 2-61$$

These quaternions are related to the Euler angles by [144]:

$$q_0 = \cos \frac{\theta}{2} \cos \frac{\phi + \psi}{2}, q_1 = \sin \frac{\theta}{2} \cos \frac{\phi - \psi}{2}, \quad 2-62$$

$$q_2 = \sin \frac{\theta}{2} \sin \frac{\phi - \psi}{2}, q_3 = \cos \frac{\theta}{2} \sin \frac{\phi + \psi}{2}.$$

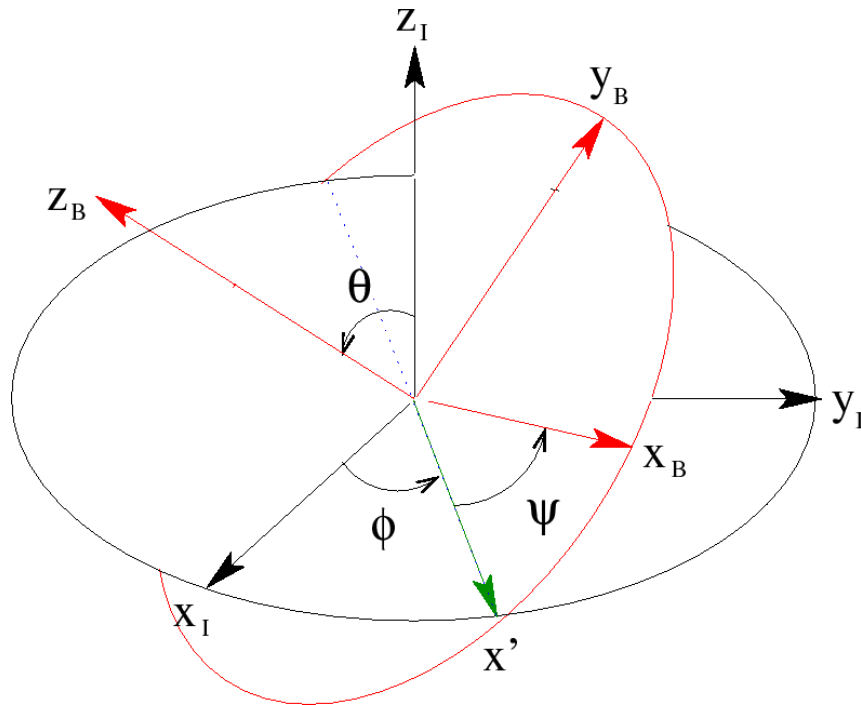


Figure 2-7. Representation of Euler angles that connect the space-fixed frame to the body-fixed frame. Picture is taken from at: http://hepweb.ucsd.edu/ph110b/110b_notes/node31.html [Last Accessed at 2012].

In terms of quaternions, the rotation matrix is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_0q_1) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad 2-63$$

This way the quaternion derivatives are:

$$\begin{pmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \begin{pmatrix} 0 \\ \omega_x^b \\ \omega_y^b \\ \omega_z^b \end{pmatrix}, \quad 2-64$$

where ω^b is the body-fixed angular velocity. A Nosé-Hoover thermostat is also necessary to keep the rotational temperature [153,154] consistent with the system temperature. It is possible to use the same thermostat for both translational and rotational degrees of freedom, but there is always a chance that the total temperature is conserved by compensation between these translational and rotational temperatures. We define separate independent thermostats for translations and rotations.

2.4 Method Improvement

In this section, we briefly review the methodological changes specifically performed to ensure the accuracy of the simulation results. To this end, we have studied the surface equilibration and the electrostatic treatment to ensure the proper timing for dihedral torsion scaling. Our results on this Section are reported in J. Chem. Phys. 136, 114705 (2012).

2.4.1 Equilibrating the Surface

The model interface is shown in Figure 2-8. It includes two parallel surfaces with the solvent in between and empty space above or below the layers. The selectors are connected to the interface by covalent bonds. The interface also includes silanol groups and trimethylsilyl end-caps.

Full equilibration of a complex solid-liquid interface involves multiple dynamical factors, each with their own time scales. First for large, flexible selectors tethered to the surface, equilibration may be slow due to surface crowding. As well, these selectors have intramolecular degrees of freedom, torsions for example, that are inherently slow to equilibrate between conformers. Third, a multi-component solvent will partition, to some extent, at the interface. Finally, if analyte is present, then analyte must also partition between the bulk and the interface. The latter process will be particularly slow for larger analytes, where conformational changes may accompany their “docking” with the surface-bound selectors.

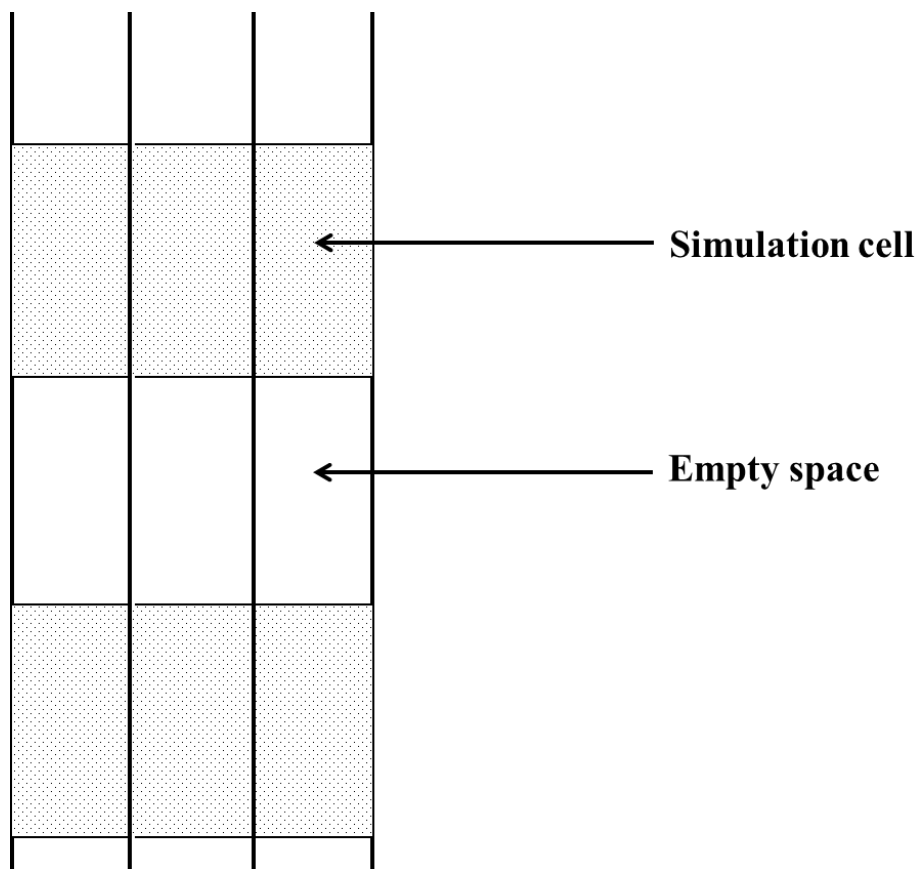


Figure 2-8. The simulation cell, replicated in three dimensions but shown in 2D.

Annealing offers one approach to equilibration. Certainly, a higher temperature increases molecular motion, expedites solvent-interface partitioning, and allows some energetic barriers to be overcome to arrive at minima which would not have been explored at the lower temperature. For interfacial systems, annealing will be helpful provided that the proper conformational balance, for the temperature of interest, is ultimately restored. This would occur for an indefinitely long simulation but equilibration of large surface bound moieties can be slow.

An alternative to annealing, or possibly in addition to annealing, we explore the targeted reduction of torsional barriers for surface-bound moieties. This reduction has been done such that the relative energy of the minima is largely unchanged – the barriers are directly targeted. In this way, as the original barrier is gradually returned, the correct conformational balance should remain. Consider the application of barrier reduction to chiral selectors. Many selectors have conformations that, despite being energetically close in energy, have large barriers to inter-conversion. More precisely, amide bonds are common in chiral stationary phases and the barrier to inter-conversion between *cis* and *trans* amides is over 50 kJ/mol. Exceedingly high temperatures would be required to equilibrate via annealing. Consider di-proline selectors which include three amide bonds (see Figure 2-9): The first amide is between C(13)-C(11)-N(27)-C(23), the second amide is between C(23)-C(9)-N(21)-C(17), and C(17)-C(7)-N(5)-C(4). We have accelerated the equilibration of each amide bond in di-proline by introducing a time-dependent scaling function that temporarily lowers the barrier to amide *cis-trans* inter-conversion without altering the relative energies of the conformers. Specifically, the amide torsions (Eq. 2-28) are scaled by the below multiplicative factor:

$$S(\varphi_t, t) = \frac{(t - t_s) + (t_f - t) \cos^2 \varphi_t}{(t_f - t_s)}. \quad 2-65$$

Alternate scaling functions, involving $\cos(\varphi_t/2)$ instead of $\cos(\varphi_t)$ for example, can be defined and may be beneficial for certain torsional barriers. For longer chains, the adjacent coupled torsions are also scaled. As will be discussed in Chapter 3, for TMA terminated diproline, the backbone dihedrals have shown to be highly coupled. It means when a torsion is changing, the next adjacent torsion is altered to reduce the molecule instability. For longer chains, we have also scaled the coupled torsions to allow for easier equilibration of the interface. The scaling function above is both time and angle dependent. For *cis* and *trans* conformers (with angles close to either 0 or 180°), the scaling function is one, at all times, but for in-between torsional angles joining these two conformers, the scaling function is a small number at times near t_s and approaches unity at t_f . At other times, the scaling function is simply set to one. Thus, the barrier is effectively eliminated at the start time, t_s , and is returned to its original height by the final time, t_f . The start and final times were 3 ps and 60 ps, respectively. A further 90 ps of equilibration time followed the “scaling in” of the torsional barrier. It is important to note that the final amide conformational balance reflects the presence of solvent and surface elements such as end-caps and silanol groups. Figure 2-9 shows the representation of different CSPs used for studying the effect of scaling.

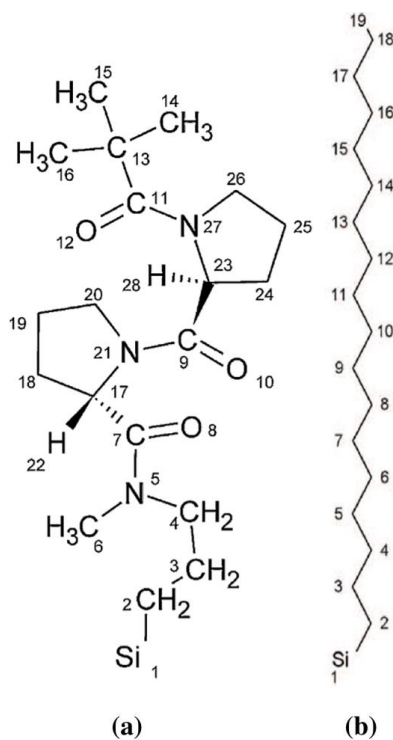


Figure 2-9. Representation of selectors used for method development. A) TMA-(Pro)₂-N(CH₃)-tether and C₁₈H₃₇.

The assessment of interfacial structure begins with surface distribution functions. These distributions are obtained from

$$g(z) = \frac{n^{actual}(z)}{n^{ideal}(z)}, \quad 2-66$$

where $n^{actual}(z)$ is the number of atoms found at a distance between z and $z + \Delta z$ from the underlying Si layer, and $n^{ideal}(z)$ is the number expected for an unstructured, random distribution. In this way, $g(z)$ identifies areas of concentration and depletion for each atom above the surface. For surface distributions, we have a total of N_s surface atoms and $n^{ideal}(z) = N_s/100$ where the distance from a surface to the center of the cell has been divided into one hundred segments ($L_z/\Delta z = 100$). The surface $g(z)$ distributions integrate to N_s . Since the

simulation cell includes two identical surfaces, the two surface distributions are averaged. The $g(z)$ is calculated for all of atoms of the selector. A typical simulation cell is shown on Figure 1-2. Note that the z direction is defined as the distance between two parallel surfaces.

Figure 2-10 shows the surface distribution of backbone carbons for the TMA-(Pro)₂-N(CH₃)-tether in *n*-hexane/2-propanol with different approaches to equilibrate the selector conformation. All simulations begin with each selector in the global energy minimum structure, where all amides are *trans*. In the first simulation, equilibration proceeds at 298 K and without altering the torsional barrier. In the second simulation, equilibration begins at a temperature of 600 K for 60 *ps*, then gradually decreases to 298 K over the next 60 *ps*. Finally, in the third simulation, equilibration proceeds at 298 K but with torsional scaling implemented for all three amides.

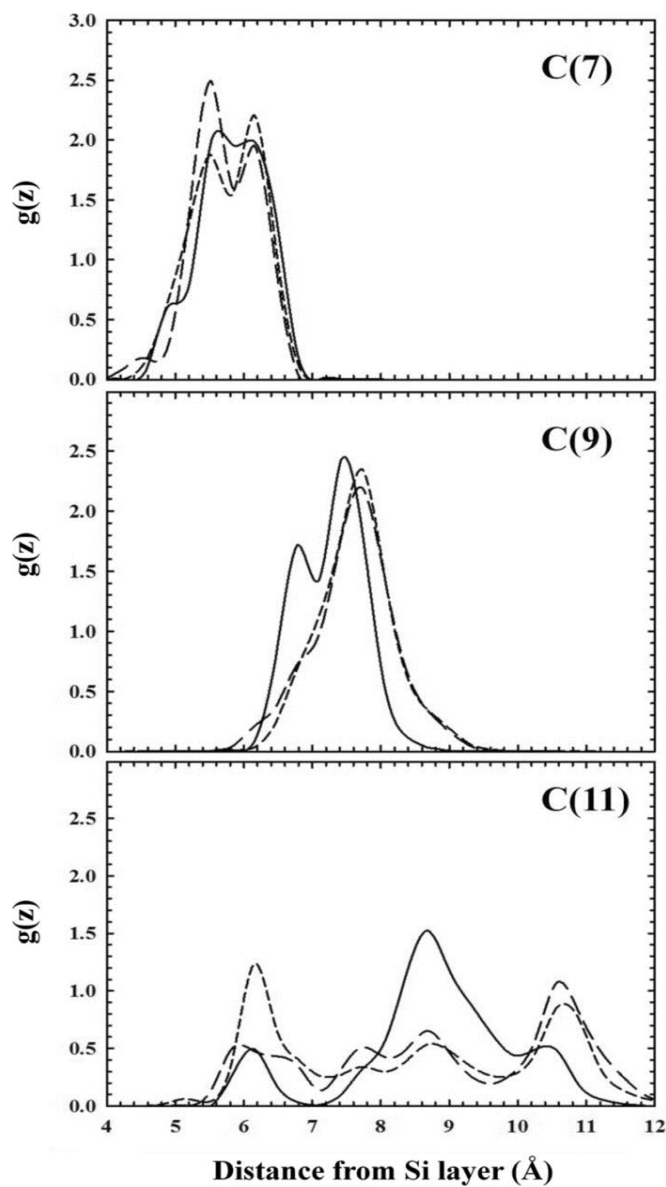


Figure 2-10. The impact of equilibration method on the surface distribution of backbone carbons in TMA-(Pro)₂-N(CH₃)-tether. The simulations begin from all *trans* selector conformation with a 150 ps equilibration period that includes annealing at 600 K (long dashed) or of torsion barrier reduction (short dashed). The distributions obtained without annealing or barrier reduction are shown by solid lines.

In all three cases, a 150 ps equilibration period is followed by a 0.6 ns collection period where surface distributions are assessed. For the di-proline interface, where the amide torsions have a high energy barrier and interactions between selectors is limited due to their separation at the surface, all selectors remain in an all *trans* arrangement if equilibration proceeds without annealing or barrier reduction. When either of the latter two methods is undertaken, the surface distributions for the backbone carbons change dramatically indicative of a conformational shift. Both approaches yield very similar distributions for C(7) and C(9). However, for C(11), the top-most carbonyl carbon in the selector, the surface distributions are in qualitative agreement but differ somewhat in their predictions at smaller distances. Specifically, the barrier-scaling approach leads to a higher probability of finding C(11) at around 6 Å from the underlying Si layer. Analysis of the simulations shows a 72% chance of finding the selectors in the all *trans* conformer when annealing is performed during equilibration, 54% when torsion scaling is performed instead, and 89% when no special measures are taken to expedite equilibration. *Cis* and *trans* conformer are defined for the first, C(13)-C(11)-N(27)-C(23) and second amide, C(23)-C(9)-N(21)-C(17), torsions, of the TMA terminated di-proline (see Figure 2-9). Measures taken to expedite the equilibration of chiral selectors are generally directed towards generating an appropriate population of selector conformers. This focus on individual selectors is justified since the selectors are large, typically do not display [156] specific selector-selector interactions such as H-bonds, and are separated on the surface by silanol groups and end-caps.

An entirely different situation arises for self-assembled materials, such as surface bound hydrocarbons, where the molecules are closely packed and display a high degree of molecule-to-molecule correlation. Following Siepmann *et al.* [101,102], we have examined the surface distribution of C₁₈H₃₇ but, in our case, we explore the impact of annealing and torsion-scaling on

the surface structure. Figure 2-11 shows the surface distribution for the top-most carbon in $C_{18}H_{37}$ at 298K and in water/methanol 70/30 v/v solvent. The torsional barriers are small for $C_{18}H_{37}Si$ and, for this reason, torsion-scaling has a negligible impact on the distribution of C(19) above the surface. Annealing at 600 K, in contrast, has greatly broadened the distribution. This occurs since the higher temperature promotes disorder leading to chain entanglement. Disentanglement is slow and does not occur on the time scale of the simulation.

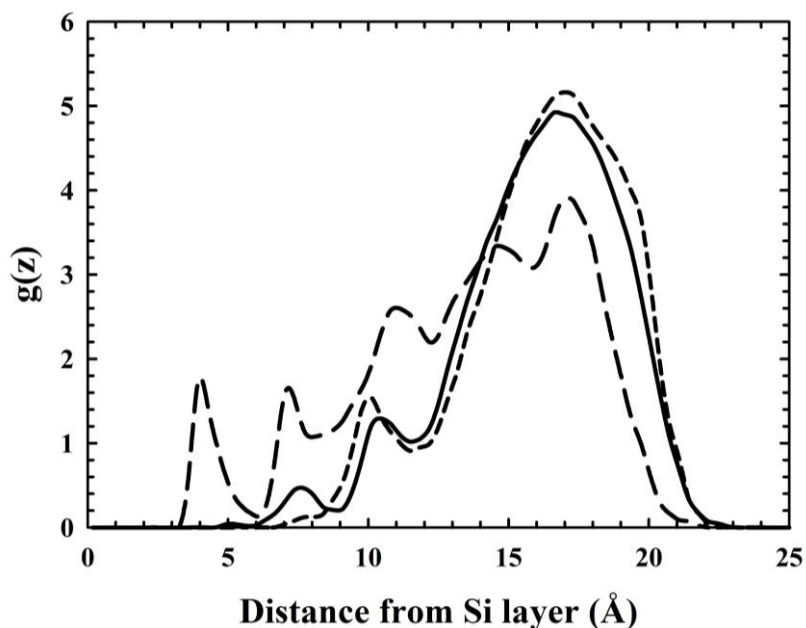


Figure 2-11. The impact of equilibration method on the surface distribution of C(19), the topmost carbon in surface-bound $C_{18}H_{37}Si$. The simulations begin from all *trans* conformation and a 150 *ps* equilibration period that includes 90 *ps* of annealing at 600 K (long dashed) or of torsion barrier reduction (short dashed). The distribution obtained without measures to expedite the selector conformational equilibration is shown by a solid line.

Regardless of the approach chosen to equilibrate the interfaces, it is important to repeat the simulations multiple times in order to examine a range of surfaces. For example, within the

simulations of the di-proline interface, the conformational balance for individual simulations varied from 48% all *trans* to 67% all *trans* conformers. The use of annealing for surface equilibration should be considered on a case-by-case basis. In many instances, barrier reduction will be a preferred option.

2.4.2 Impact of Electrostatics

3D Ewald summations include empty space beyond the interfacial system, as shown in Figure 2-8. We also include a correction to account for the interaction with the environment [147,158,159] that depends on the dielectric of the surrounding medium and the shape of the simulation cell. The latter has been chosen to be consistent with infinite plates. That is, although the simulation cell is replicated in three dimensions, the evaluation of the electrostatic sum is consistent with replication in two dimensions. Berkowitz *et al.* [147,158,159] performed a detailed comparison of 2D and 3D Ewald sums to validate this approach.

We explore the impact of the treatment of electrostatics on the interfacial structure. To this end, we increase the amount of empty space in the cell, examine the use of a spherical cutoff for electrostatics, and implement a new shape dependent correction that is representative of neighboring pores in a silica bead. The dielectric of the surrounding medium is assumed to be unity but, given that the dielectric for various forms of silica is between 2 and 4 [160], this approximation is acceptable.

In all cases, the simulation cell is a rectangular prism with volume V and side lengths L_x , L_y , L_z . Following Smith *et al.* [161], the shape-dependent correction is

$$\frac{1}{V} \sum_{i=1}^3 B_i M_i^2, \tag{2-67}$$

where M_i is the cell dipole moment in direction i and

$$B_1 = 2\pi \int_0^\infty dt \frac{d(\text{erf}(\xi_1 t))}{dt} \text{erf}(\xi_2 t) \text{erf}(\xi_3 t), \quad 2-68$$

where $\xi_i = L_i/\sqrt[3]{V}$ and $\text{erf}(\xi_1 t)$ is the error function. Equivalent expressions apply for B_2 and B_3 . We have numerically evaluated the B_i cofactors for our specific simulation cell and included the correction from Eq. 2-68 within our simulations. In this way, the correction is appropriate to fluid contained between separated surfaces.

Figure 2-12 shows the impact of electrostatic treatment for the TMA-(Pro)₂-N(CH₃)-tether interface with *n*-hexane/2-propanol. Surface distributions for backbone carbons are shown along with radial distributions between the alcohol H and carbonyl oxygens. The radial distributions are quite similar although the infinite plate correction leads to a higher probability for all three potential H-bonding pairs. The surface distributions show a stronger dependence on the treatment of electrostatics. Specifically, the three approaches are generally in agreement but they differ slightly in their conformational predictions. For instance C(7) is more likely to be found within 5.5 Å of the underlying Si layer when the Ewald sum includes a correction for the actual shape of the simulation cell.

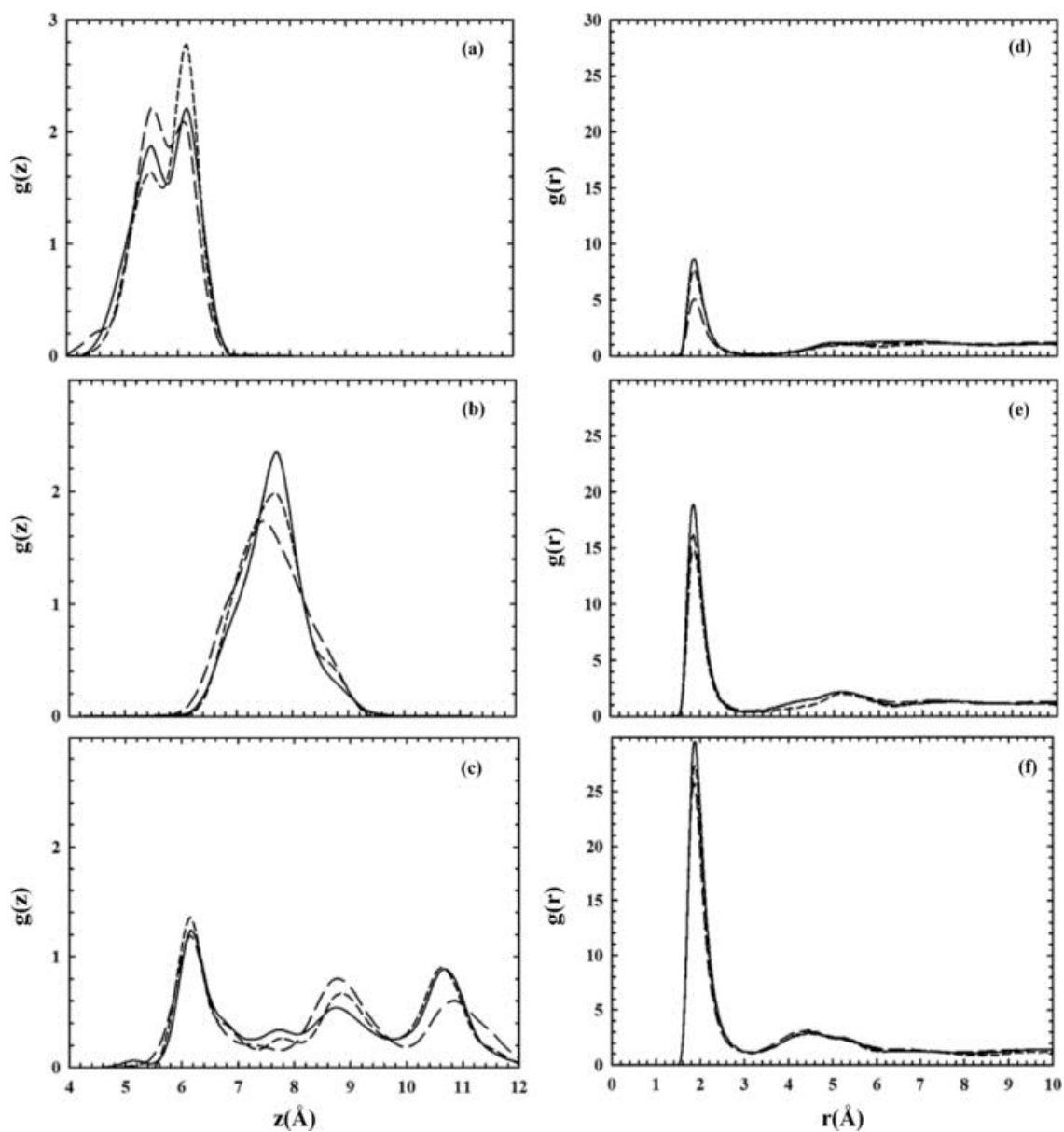


Figure 2-12. The impact of the treatment of electrostatics on the surface-solvent structure distributions for the TMA-(Pro)₂-N(CH₃)-tether interface with *n*-hexane/2-propanol. Surface distributions for C(7), C(9), and C(11) are shown in (a), (b), and (c), respectively. Radial distributions between the alcohol H and the carbonyl oxygens, O(8), O(10), and O(12) are shown in (d), (e), and (f), respectively. The Ewald correction appropriate for infinite plates (solid line) and for rectangular prism (long dashes) are shown. The short-dashed line gives distributions for a spherical cutoff.

Chapter 3

Conformational Studies of Poly-proline Chains

3.1 Introduction

As the name implies, poly-proline is made of proline residues (Figure 3-1) joined by peptide bonds. The poly-proline backbone dihedral angles are defined by ω , ϕ , and ψ as shown in Figure 3-1(b) for TMA terminated di-proline. These dihedrals are defined by the backbone atoms. For instance, ω_2 is the dihedral angle between C^α of the topmost pyrrolidine ring, C(CO), N of the second pyrrolidine ring, and C^α of the second pyrrolidine ring. The poly-proline backbone can adopt two distinct helical conformations known as PPI and PPII [162]. The PPI helix has all *cis* amide torsions (ω) in the backbone, is right handed and has 3.3 prolyl residues per turn with a pitch of 5.6 Å. On the other hand, the PPII conformer has all *trans* amide torsions in the backbone, a left handed helix containing 3 residues per turn, and a more extended backbone as reflected by a 9.4 Å pitch (see Figure 3-1(c)-(d)). Although *cis* isomers are rare in proteins because of unfavorable contact between adjacent amino acids [163], proline has a moderate probability to be found in *cis* conformer [164]. Both conformers have similar ϕ angles (around -80°) [165]. However, the ψ angle adopts an angle of 140° to 170° for PPI conformer and 110° to 150° for PPII. The backbone dihedral ψ can also adopt an angle around -30° which is less probable [164].

In this thesis, different conformers of poly-proline are defined by their amide torsions in the backbone. To be more specific, the abbreviated name of a poly-proline conformer indicates if each amide torsion is *cis* or *trans*. Here, *cis* and *trans* are denoted by “C” and “T” labels and naming starts from the first pyrrolidine ring which is closer to the terminal group. For instance,

TC means a di-proline in which the first amide torsion (ω_1) of the first ring which is closer to the terminal group adopts the *trans* conformer, and the second amide torsion (ω_2) of the second ring which is closer to the $-\text{N}(\text{CH}_3)_2$ adopts the *cis* conformer (Figure 3-1). As another example, TTCCTC means a hexa-proline in which the first and second amide torsions are both *trans*, the third and fourth amide torsions are both *cis*, the fifth amide torsion is *trans*, and the sixth amide torsion is *cis*.

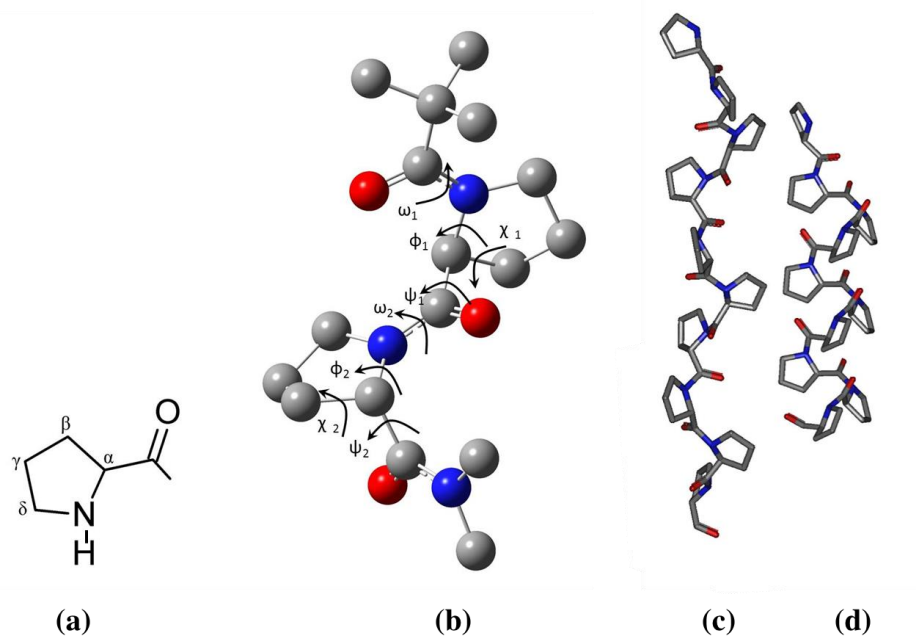


Figure 3-1. a) A proline monomer with labeling of the ring atoms. b) TMA-(Pro)₂-N(CH₃)₂ with the backbone and endo-cyclic dihedral angles. The endo-cyclic dihedral angle is shown by χ which is defined between N, C ^{α} , C ^{β} , and C ^{γ} . Here the all *trans* conformer of TMA terminated di-proline is shown. c) The PPII conformer of deca-proline. d) The PPI conformer of deca-proline.

Considerable spectroscopic research has been performed to study the conformations of poly-proline and the effect of chain length and solvent polarity on the conformational balance.

Zhang *et al.* [166] used $^1\text{H-NMR}$ for $\text{BOC-(Pro)}_n\text{-O-Benzyl}$ with $n=2-5$ in CDCl_3 and observed that the all *trans* conformer is more populated. Several other studies [167-175] showed that for short chains, less than five proline residues, poly-proline adopts a mixed *cis-trans* backbone in water. However, by reducing solvent polarity PPI becomes more populated. This effect is more pronounced for longer chains. In fact, in aliphatic alcohols, poly-proline chains are found to adopt a PPI conformer [69,169,175-182]. For longer chains, PPII becomes the dominant conformer in water. Isemura *et al.* [167] used the IR spectra where PPI and PPII show different peaks at far IR region. They found that chains of tert-amylloxycarbonyl-proline for five, six, and eight proline residues adopt the PPII conformation while four residues show a mixed *cis-trans* conformer. In another study they employed far UV spectroscopy for poly-proline in water and found that by increasing the chain length, the PPII conformer becomes more dominant. Several other studies also confirm that PPII is more stable in aqueous solvents while PPI is more stable in aliphatic alcohols [175-178]. For instance, Kakinoki *et al.* [175] studied proline oligomers with 4, 6, 13 proline residues using circular dichroism and showed that a proline chain adopts a PPII conformer in water and trifluoroethanol. The 13 residue poly-proline showed a transition to PPI in methanol and 1-propanol. Chao and Bersohn [179] used ^{13}C and $^1\text{H-NMR}$ to study $^+\text{H}_2\text{N-Pro-(Pro)}_n\text{-CO}_2^-$ with $n=1,2$ in aqueous solution. They observed that the percentage of *trans* amides increases when the chain length increases. Kurtz *et al.* [181] observed a PPII-PPI transition when the solvent changed from water to either *n*-propanol or *n*-butanol. Madison and Schellman [183] used circular dichroism and $^1\text{H-NMR}$ for N-acetyl-L-proline and found that the all *trans* conformer is more stable in less polar solvents. Deber *et al.* [169] studied a series of oligomers (t-butyloxycarbonyl)-(Pro)_n-benzyl esters, ($n=2-6$) and analyzed their NMR spectra in chloroform. They showed that nearly equal random distribution of *cis* and *trans* exists for $n=2,4$, but for $n=5,6$

PPII is dominant. Doose [184] used single-molecule spectroscopy based on Förster resonance energy transfer and showed that mix *cis-trans* short chains (1-10 proline residues) are more stable in water. Using $^1\text{H-NMR}$, Higashijima *et al.* [185] found that for N-acetyl-L-proline-N-methylamide, *cis* conformer is more stable in polar solvents. To summarize, experimental studies show that poly-proline chains adopt PPII conformer in more polar solvents although some studies show the reverse [185]. Increasing the chain length can also increase the PPII population. For shorter chains, usually mixed *cis-trans* conformers are highly populated.

To accompany experimental work, several theoretical studies have been performed to find the relative stability of PPI and PPII conformers. The theoretical studies use a variety of methods including force field based free energy calculations, Hartree-Fock calculations or density functional calculations to find the relative stability of PPI and PPII conformers [69,165,186-189]. The *cis-trans* amide barrier for $\text{HCO}-(\text{Pro})_2-\text{NH}_2$ has been examined by Kang *et al.* [165] using B3LYP/6-311++G(d,p), and found to be 89.7 kJ/mol. In contrast, the relative energy difference between the *cis* and *trans* conformer of mono-proline, obtained by CCSD(T), showed only a 3.9 kJ/mol energy difference, with *cis* being more stable, which is in accordance with experimental observations [190]. Profant *et al.* [189] studied $(\text{Pro})_n$ with $n=2,3,4,6,9,12$ using the B3LYP functional in a continuum water solvent and found that for $n=2-6$, mixed *cis-trans* conformers were dominant but increasing the chain length increased the *trans* amide dihedral population. Kang *et al.* [165] used HF/6-31+G(d) to find minima for $\text{CH}_3\text{O}-(\text{Pro})_n-\text{N}(\text{CH}_3)_2$ with $n=2-5$. For di-proline the lowest energy minima are TT, followed by CT, and CC. For $n=3$ they found that TTT is the most stable conformer in the gas phase followed by CTT. For $n=4,5$ they find that the all *trans* conformer is most stable while TCCC and TCCCC are the least stable, respectively. In contrast, Zhong and Carlson [186] explored a slightly different di-proline, $\text{CH}_3\text{CO}-(\text{Pro})_2-$

NH(CH₃), using B3LYP/6-31G(d) but they did not identify any CT or TC local minima. They only captured all *cis* and all *trans* conformers since they used optimized di-prolines to build hexa-prolines. They concluded that, in the gas phase, the all *trans* hexa-proline is more stable than the all *cis* conformer by almost 30 kJ/mol. They also analyzed conformer stability in the presence of continuum solvents and, contrary to experiment, found that the all *trans* conformers are more stable in non-polar solvents and all *cis* conformers are more stable in aqueous solvents. In both cases, the all *trans* conformation was identified as the most stable in the gas phase. RI-MP2/SVP calculations of Kuemin *et al.* [187] yield results in better overall agreement with experiment. They studied chains of 6,9, and 12 proline residues with different terminal groups using RI-MP2/SVP and found that longer chains prefer all *cis* amides in the gas phase while in the presence of explicit water molecules, the all *trans* conformer is more stable.

The different conclusions in some of these studies, and the discrepancy between some of the predictions and experiment, reflect the considerable challenge of modeling poly-prolines. Disagreement between studies originates from the effect of temperature, concentration, chain length, pH and ionic strength, terminal group and the type of adjacent amino acid [175,185, 191-201]. All of these differences contribute to inconsistency between the results. For instance, poly-proline backbone is sensitive to the type of the terminal group, for instance. As will be discussed in Section 3.3.1, adding only one oxygen atom to the terminal group can dramatically change the order of stability of conformers. Furthermore, HF, B3LYP functional and MP2 are not able to quantitatively predict the relative stability of poly-proline conformers even for mono-proline [190,202], although they work well for other amino acids such as glycine and alanine. More recent theoretical studies of poly-proline include dispersion corrections. Sherrill *et al.* [203-205] performed a comprehensive systematic study of recent density functionals that include dispersion

with CCSD(T) reference calculations, to assess their accuracy in treating non-covalent interactions. They have concluded that XYG3, B3LYP-D3, B2PLYP-D3 functionals perform the best when used with triple zeta basis sets. Kang *et al.* employed a variety of functionals to study the effect of empirical dispersion forces on the relative stability of alanine and proline dipeptide in the gas phase and continuum water [206]. Their detailed analysis reveals that hybrid functionals such as B2LYP-D and mPW2PLYP-D outperform B3LYP-D and even MP2.

Here we report extensive conformational studies on poly-proline from mono- to hexa-proline using several levels of the theory. The effect of two terminal groups and three continuum solvents are also studied. Specifically, we studied the relative stability for all possible combinations of *cis* and *trans* amide torsions (from mono- to hexa-proline), in the gas phase, in water, in chloroform and in hexane (from mono- to penta-proline).

3.2 Computational Methods

Several levels of theory are employed to better capture the right energetics of poly-proline chains. All of the structures are first optimized using B3LYP, then further optimized using B97-D [123], mPW2PLYP-D [124-130], MP2 [130], and RI-MP2 [129] methods. For all calculations, the 6-311G(d,p) basis set was used. The geometry optimizations start with a set of random conformations that cover all possible combinations of *cis* and *trans* amides. We found that the conformer structures at different levels of theory are similar: the backbone dihedral torsions for the same conformer optimized at different levels of theory agree to within 12°. Each local minimum on the potential energy surface was confirmed by the lack of any imaginary frequencies. To obtain more accurate conformer energies, single point CCSD calculations were performed on the optimized structures of mono- and di-proline. For longer chains, LPNO-CCSD/6-311G(d,p) [135,136] calculations were performed for lower cost and close efficiency to

CCSD. All B3LYP, B97-D, MP2 and CCSD calculations are performed using Gaussian 09 [207]. RI-MP2, mPW2PLYP-D and LPNO-CCSD calculations are done using ORCA [208]. For the solvation calculations, the COSMO solvation model is used [209].

As a test of basis set, for TMA terminated di-proline, the cc-pVTZ basis set was also used and compared with 6-311G(d,p). Because of very close relative energies 6-311G(d,p) is adopted for all of calculations.

To understand the effect of terminal group, for mono- to tri-proline, two different terminal groups were studied: TMA and BOC. Figure 3-1(b) shows a TMA terminated di-proline. The BOC terminal group is similar to TMA, but has one more oxygen between the topmost carbonyl group and the *tert*-carbon of the terminal group. The proline residues are connected to either TMA or BOC at the N-terminal and the C-terminal is linked to $-N(CH_3)_2$: TMA-(Pro)_n-N(CH₃)₂ with n=1-6 or BOC-(Pro)_m-N(CH₃)₂ with m=1-3.

3.3 Results and Discussions

3.3.1 Stability Studies in the Gas Phase and in Continuum Solvents

Table 3-1 summarizes the *ab initio* calculation results for mono-proline which is either terminated with TMA or BOC. For TMA-(Pro)-N(CH₃)₂, the *trans* conformer is more stable while for the BOC terminated mono-proline, *cis* conformer is preferred. The backbone torsions and the dihedral angle associated with the ring atoms (see Figure 3-1(b)) are listed in Table 3-1. Continuum solvent free energies are also listed in Table 3-1 for TMA terminated mono-proline. Water increases the energy gap between *cis* and *trans*. On the contrary, apolar solvents reduce this gap.

Table 3-1. Dihedral angles, relative energies and free energies of TMA-(Pro)-N(CH₃)₂ and BOC-(Pro)-N(CH₃)₂. For gas phase calculations and solvation studies, CCSD//mPW2PLYP-D and CCSD//mPW2PLYP-D/COSMO are used, respectively. All energies (for gas phase calculations) and free energies (for solvent phase calculations) are in kJ/mol, and all angles are in degree. 6-311G(d,p) basis set is used for all calculations.

Conformer	ω	ϕ	χ	ψ	$\Delta E/\Delta G$
TMA-(Pro)-N(CH₃)₂ (vacuum)					
T	176.4	-71.3	24.3	127.1	0.0
C	-5.4	-61.5	29.2	151.6	3.8
BOC-(Pro)-N(CH₃)₂ (vacuum)					
T	172.6	-67.7	23.1	122.0	2.9
C	-10.7	-58.0	28.1	148.3	0.0
TMA-(Pro)-N(CH₃)₂ (water)					
T	174.3	-74.5	25.1	134.5	0.0
C	-4.6	-63.7	31.2	157.8	4.7
TMA-(Pro)-N(CH₃)₂ (chloroform)					
T	177.2	-74.8	24.2	130.0	0.0
C	-2.3	-64.2	29.6	154.9	2.6
TMA-(Pro)-N(CH₃)₂ (hexane)					
T	176.3	-76.1	25.8	134.5	0.0
C	-3.4	-63.6	28.7	156.1	2.1

Table 3-2 shows the relative energies for TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂. Again, the impact of terminal group is clear where the more compact conformer (CT) is preferred when the terminal group is BOC. The overall conformational stability order is TT > TC > CT > CC for TMA terminated di-proline and CT > TT > TC > CC for BOC terminated di-proline. The value of ψ depends on the adjacent amide dihedral which could be either *cis* or *trans*. The *cis* conformation shifts ψ to around 160° while the *trans* conformation changes it to around 120°. Solvent effects are small on relative stability of conformers. In all solvents, all *trans* is the most stable conformer.

Table 3-2. Dihedral angles, relative energies and free energies of TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂. For gas phase and solvation studies, LPNO-CCSD//mPW2PLYP-D and LPNO-CCSD//mPW2PLYP-D/COSMO methods are used, respectively. All energies (for gas phase calculations) and free energies (for solvent phase calculations) are in kJ/mol, and all angles are in degree. 6-311G(d,p) basis set is used for all calculations.

Conformer	ω_1	ϕ_1	χ_1	ψ_1	ω_2	ϕ_2	χ_2	ψ_2	$\Delta E/ \Delta G$
TMA-(Pro)₂-N(CH₃)₂ (vacuum)									
TT	177.3	-67.5	30.5	134.4	-178.7	-90.8	29.6	118.8	0.0
TC	-179.7	-85.3	32.6	105.8	-5.53	-89.2	39.2	169.1	6.3
CT	-4.5	-87.9	36.6	162.3	174.1	-75.4	30.2	127.3	7.8
CC	-2.9	-88.3	37.5	167.9	-0.7	-83.4	34.5	160.9	23.2
BOC-(Pro)₂-N(CH₃)₂ (vacuum)									
TT	-178.7	-57.0	32.4	135.3	-176.8	-88.6	29.5	119.5	1.5
TC	-174.8	-97.8	33.1	101.8	-2.0	-98.8	40.5	172.7	8.7
CT	-19.90	-51.6	34.1	152.9	168.5	-58.5	31.1	138.9	0.0
CC	-23.9	-51.4	-6.5	167.7	-6.2	-77.5	36.1	159.3	17.3
TMA-(Pro)₂-N(CH₃)₂ (water)									
TT	172.7	-64.3	33.1	156.4	167.5	-66.3	34.6	161.1	0.0
TC	179.6	-108.3	32.7	98.7	-2.9	88.5	40.3	177.3	9.1
CT	-7.7	-84.8	38.5	162.1	168.6	-67.1	34.0	162.2	8.5
CC	1.6	-92.0	39.2	165.9	-2.1	-81.4	34.4	141.7	25.3
TMA-(Pro)₂-N(CH₃)₂ (chloroform)									
TT	172.5	-64.0	33.2	156.0	168.0	-61.7	29.7	144.5	0.0
TC	179.5	-107.2	33.0	98.0	-2.9	-89.0	40.6	177.8	7.8
CT	-7.3	-84.9	38.4	161.2	167.6	-66.3	34.6	160.9	9.2
CC	1.3	-91.8	39.2	165.6	-2.4	-81.7	34.4	137.9	29.2
TMA-(Pro)₂-N(CH₃)₂ (hexane)									
TT	172.2	-63.6	33.6	185.5	168.1	-61.0	30.4	140.5	0.0
TC	179.4	-106.2	33.2	97.6	-3.2	-89.2	40.9	178.5	7.1
CT	-8.9	-83.6	38.3	161.2	167.7	-60.7	31.0	142.6	9.1
CC	-0.4	-90.6	39.5	165.7	-2.9	-81.7	34.3	133.5	24.6

To carefully examine the effect of the *ab initio* method on the order of stability, several methods have been employed to optimize the geometry and to calculate the single point energy of TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂. Table 3-3 summarizes the relative energies

obtained at each level of theory. Of the methods chosen, CCSD is expected to capture more of the correlation energy and provides a reference point for assessing the other approaches. The other methods are less computationally demanding than CCSD and should remain feasible for longer chains. Dispersion is expected to be most significant for the TC and CC conformers since, from Figure 3-2, these structures are more compact. B3LYP does not account for dispersion and, for this reason, overestimates the energy gap between conformers.

Table 3-3. Conformers TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂. The first column identifies the conformation about the amide bonds. Relative energies, in kJ/mol, are reported relative to the lowest energy conformer for each method. The energies in brackets, for CCSD and LPNO-CCSD, are relative to the overall minimum energy structure. Results are provided for: 1: B3LYP/6-311G(d,p); 2: B97-D/6-311G(d,p); 3: mPW2PLYP-D/cc-pVTZ; 4: MP2/6-311G(d,p); 5: CCSD/6-311G(d,p)//B3LYP/6-311G(d,p); 6: CCSD/6-311G(d,p)//B97-D/6-311G(d,p); 7: LPNO-CCSD/6-311G(d,p)//B3LYP/6-311G(d,p); 8: LPNO-CCSD/6-311G(d,p)//B97-D/6-311G(d,p) calculations.

	ΔE							
	1	2	3	4	5	6	7	8
TMA-(Pro)₂-N(CH₃)₂								
TT	0.0	7.1	0.0	0.0	0.0 (0.0)	0.0 (2.5)	0.0 (0.0)	0.0 (7.4)
TC	17.6	0.0	8.9	6.0	6.7 (6.7)	5.9 (8.4)	4.9 (4.9)	2.9 (9.4)
CT	14.5	10.5	8.8	8.7	8.2 (8.2)	6.3 (8.9)	6.9 (6.9)	4.9 (12.1)
CC	30.8	17.3	22.7	26.2	21.7 (21.7)	20.2 (22.7)	19.7 (19.7)	17.3 (25.2)
BOC-(Pro)₂-N(CH₃)₂								
TT	0.0	5.4	2.7	2.4	1.9 (1.9)	0.3 (6.6)	0.0 (0.0)	0.0 (6.0)
TC	19.4	0.0	11.3	9.5	10.1 (10.1)	7.5 (13.8)	12.3 (12.3)	8.6 (14.6)
CT	4.2	0.9	0.0	0.0	0.0 (0.0)	0.0 (6.2)	2.1 (2.1)	1.7 (7.7)
CC	23.8	14.0	20.4	16.3	18.9 (18.9)	17.1 (23.4)	19.4 (19.4)	21.1 (27.1)

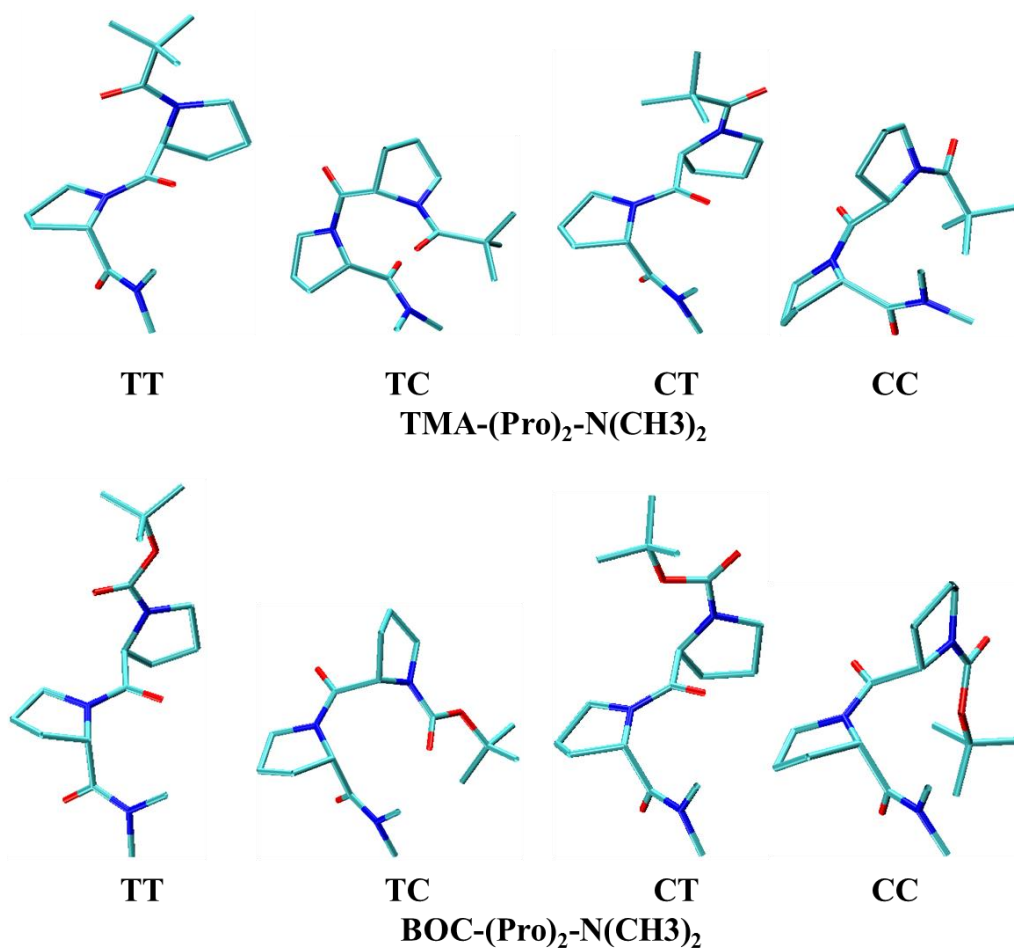


Figure 3-2. Stable conformers from B3LYP/6-311G(d,p) calculations of TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂.

In contrast to B3LYP, mPW2PLYP-D, a dispersion-corrected double-hybrid density functional, predicts the correct conformational sequence. B97-D, which adopts the generalized gradient approximation B97 functional with an added dispersive correction, overestimates the stabilization from dispersion. In fact, the lowest energy conformer is predicted to be TC, in contrast to all other approaches which indicate that TT is the most stable. Aside from B3LYP and B97-D, the methods are in reasonable agreement. For the TMA terminated di-proline, the

conformers increase in energy according to $TT < TC \approx CT < CC$. The replacement of TMA with BOC leads to $CT \approx TT < TC < CC$. Thus, the addition of an oxygen to the terminal-group has a dramatic effect: the conformers are closer in energy and the CT conformer is now predicted to be very close in energy to TT (their energies are equal within the error of the methods). The energy differences between TMA-(Pro)₂-N(CH₃)₂ conformers obtained from CCSD//mPW2PLYP-D/6-311G(d,p) calculations with respect to the most stable TT conformer are 6.3, 7.8, and 23.2 kJ/mol (Table 3-2).

Solvation free energies follow the same order as the gas phase energies with only small deviations. Table 3-2 lists the relative free energies of all conformers. Overall, the relative stabilities do not change in the presence of the continuum solvent except for water, where CT becomes more stable than TC by 0.6 kJ/mol.

Table 3-4 summarizes the relative stabilities of TMA terminated tri-proline. Our calculations indicated that the TTC conformer is the most stable conformer in the gas phase. Interestingly, for the most stable conformers, the middle ω dihedral angle stays *trans*. This makes the whole backbone to be more extended compared to *trans* terminal amide torsions. As shown in the Figure 3-3, when an inner amide linkage is *cis*, the resulting backbone structure is more compact than the corresponding *trans* conformer. For instance, the TTC conformer is much more compact compared to the TTT conformer. For the topmost amide linkage, this transition has little impact as it simply repositions the terminal group; CTT and TTT conformers are almost the same in terms of compactness. The conformational stability order is $TTC > TTT > CTC > CTT > TCT > CCT > CCC > TCC$ for TMA terminated tri-proline.

Table 3-4. The conformer energies for TMA-(Pro)₃-N(CH₃)₂. All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets provide the LPNO-CCSD energy relative to the overall lowest energy conformer. The lowest energy structure, for each conformer, is identified with a shaded cell.

Conformer /Method		TTC	TTT	CTT	TCT	CTC	CCT	CCC	TCC
ΔE	B3LYP	14.3	0.0	13.4	16.8	26.2	27.6	35.5	38.3
	B97-D	0.0	7.1	9.7	8.1	4.6	6.4	10.0	23.6
	mPW2PLYP-D	0.0	5.1	9.6	6.9	4.5	12.9	16.0	29.0
	RI-MP2	0.0	9.2	9.4	8.1	6.2	9.7	13.2	25.1
	MP2	0.0	14.3	12.1	9.6	9.7	11.8	15.3	29.6
	LPNO-CCSD //B3LYP	0.0 (3.0)	1.4 (4.4)	6.7 (9.7)	6.9 (9.9)	4.4 (7.4)	12.0 (15.0)	13.2 (16.1)	23.9 (26.8)
	LPNO-CCSD //B97-D	0.8 (14.7)	1.2 (15.1)	0.0 (13.9)	4.3 (18.1)	2.9 (16.7)	8.8 (22.7)	8.5 (22.4)	23.8 (37.6)
	LPNO-CCSD// mPW2PLYP-D	0.0 (0.0)	4.7 (4.7)	5.2 (5.2)	6.9 (6.9)	7.3 (7.3)	9.2 (9.2)	13.5 (13.5)	24.6 (24.6)
	LPNO-CCSD //RI-MP2	0.0 (1.7)	5.0 (6.7)	7.3 (9.0)	8.5 (10.2)	9.0 (10.7)	12.5 (14.2)	17.4 (19.1)	27.4 (29.1)
	LPNO-CCSD //MP2	0.0 (2.5)	8.1 (10.6)	9.0 (11.4)	12.1 (14.6)	10.0 (12.5)	15.7 (18.2)	21.0 (23.5)	29.8 (32.3)

B3LYP overestimates the energy gap between conformers of TMA-(Pro)₃-N(CH₃)₂ compared to other methods and it identifies the all *trans* TTT conformer as the most stable. The inclusion of dispersion, which is missing in B3LYP, changes the stability of the conformers such

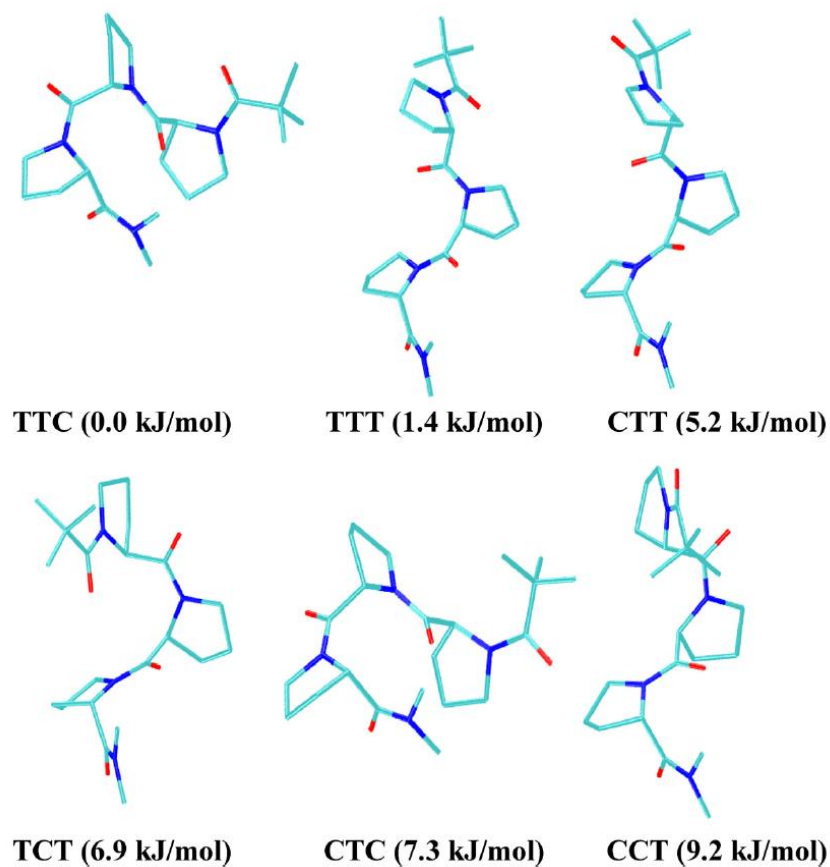


Figure 3-3. The lowest energy conformers of TMA-(Pro)₃-N(CH₃)₂. The structures are obtained from mPW2PLYP-D/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations.

that TTC is now the most stable. The LPNO-CCSD single point calculations, even for the B3LYP optimized structures, identifies TTC as the most stable conformer. The conformer structures corresponding to the lowest LPNO-CCSD energies are identified with shading. These structures and energies are retained and used for force field development discussed in the next Chapters. Overall, the TMA terminated tri-proline has five conformers within 10 kJ/mol of the global energy minimum. The six most stable structures are shown in Figure 3-3. Interestingly,

for the first three stable conformers the middle amide linkage adopts a *trans* angle which shows that a *cis* middle amide is somewhat destabilizing. The most stable conformer has a bent structure that maximizes dispersive interactions.

The results of *ab initio* calculations, for BOC-(Pro)₃-N(CH₃)₂ are summarized in Table 3-5 and the most stable conformers are shown in Figure 3-4. As expected, B3LYP overestimates the stability of the TTT conformer but the other methods identify CTC as the most stable conformer. The first three stable conformers of BOC-(Pro)₃-N(CH₃)₂ all have a *trans* amide as the central amide linkage and the two most stable conformers are bent. For di-proline, the TMA and BOC terminated di-prolines differed in the identity of the most stable conformer and in the energy gap between the conformers. We find that this persists for the tri-proline. Specifically, for the di-proline and the tri-proline, the change from TMA to BOC changes the backbone structure of the lowest energy conformer: the outer amide torsion is preferentially *cis* for BOC and *trans* for TMA. In addition, the two lowest energy conformers are closer in energy when the chain is terminated by BOC. This effect is more pronounced for the di-proline and suggests that the terminal group will have decreasing impact for the longer proline chains.

Table 3-5. The conformer energies for BOC-(Pro)₃-N(CH₃)₂. All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets are the LPNO-CCSD energy relative to the overall lowest energy conformer. The lowest energy structure, for each conformer, is identified with a shaded cell.

Conformer /Method		CTC	TTC	CTT	CCT	TTT	TCT	CCC	TCC
ΔE	B3LYP	16.4	14.3	3.4	17.6	0.0	16.0	29.0	35.3
	mPW2PLYP-D	0.0	5.8	6.9	10.0	15.7	14.1	17.0	32.6
	MP2	0.0	4.8	7.5	10.9	12.6	5.1	13.6	27.4
	LPNO-CCSD //B3LYP	0.0 (5.7)	6.0 (11.6)	2.2 (7.8)	6.8 (12.4)	7.4 (13.1)	10.1 (16.1)	13.9 (19.5)	26.0 (31.6)
	LPNO-CCSD// mPW2PLYP-D	0.0 (0.0)	3.9 (3.9)	5.6 (5.6)	6.6 (6.9)	11.6 (11.6)	13.9 (13.9)	14.3 (14.3)	28.5 (28.5)
	LPNO-CCSD //MP2	0.0 (1.1)	3.5 (4.6)	6.1 (7.2)	6.8 (7.9)	10.7 (11.8)	14.1 (15.1)	12.6 (13.7)	28.4 (28.4)

In contrast to di-proline, solvation effects change the order of stability of tri-proline conformers significantly (see Table 3-6). For all three solvents, TTT is the most stable followed by CTT which is 8-13 kJ/mol higher in free energy.

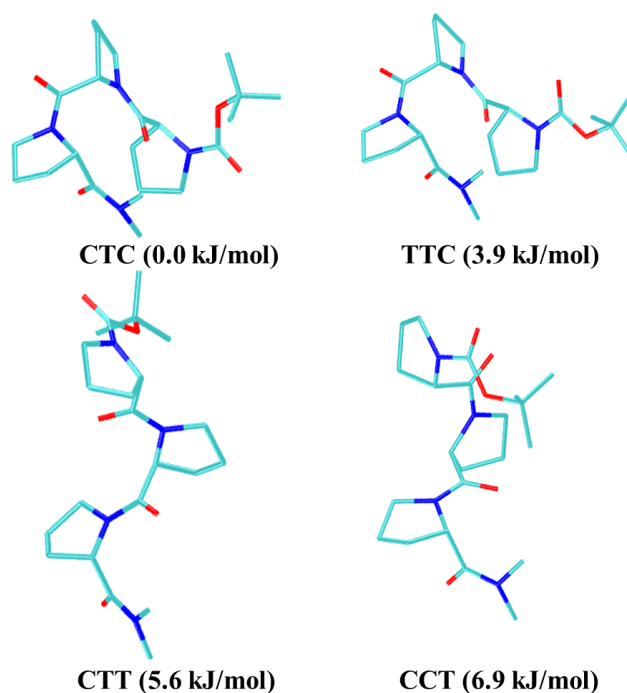


Figure 3-4. The lowest energy conformers of BOC-(Pro)₃-N(CH₃)₂. The structures are obtained from mPW2PLYP-D/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations.

Table 3-6. Relative stabilities of TMA-(Pro)₃-N(CH₃)₂ in continuum solvents. The LPNO-CCSD//mPW2PLYP-D/COSMO/6-311G(d,p) methods is used for all calculations. All free energies are in kJ/mol.

Conformer	ΔG		
	water	hexane	chloroform
TTT	0.0	0.0	0.0
CTT	7.5	9.8	12.6
TCC	19.8	24.8	38.1
CCT	26.5	26.1	35.8
TCT	29.2	24.3	25.9
CCC	29.2	29.8	35.8
TTC	31.5	23.0	22.3
CTC	41.2	35.1	35.9

Table 3-7 lists the relative energies for TMA-(Pro)₄-N(CH₃)₂ in the gas phase. TTCT is more stable than the other conformers. Tetra-proline is distinct from other proline chains in that it offers much more flexibility at room temperature. Specifically, 11 out of 16 conformers have relative energies of less than 10 kJ/mol. Since the energies are very close, it is hard to strictly assign one of them as the most stable one. Interestingly, B3LYP/6-311G(d,p) calculations only identify TTCT as a local minimum that is 12 kJ/mol higher in energy relative to TTTT. For tetra-proline, the geometry optimizations began from 16 initial structures, corresponding to the various *cis-trans* amide combinations. The most stable conformers, within 10kJ/mol of the global energy minimum are shown in Figure 3-5. LPNO-CCSD single point energies reveal that TTCT is the most stable conformer with nine other conformers within 10 kJ/mol of the global. The tetra-proline represents a transition point where the growing number of possible conformers is not yet offset by energetic contributions, such as dispersion, which will restrict the number of energetically accessible structures for longer chains. The subtlety of the balance between these competing effects is evident from the fact that B3LYP does not capture this effect at all but LPNO-CCSD and mPW2PLYP-D do. Comparing with TMA-(Pro)₃-N(CH₃)₂ and BOC-(Pro)₃-N(CH₃)₂, most of the lowest energy conformers for tetra-proline are also bent.

The stability order in solvent phase is different from the gas phase (see Table 3-8). Our calculations show that introduction of solvent increases the energy gap between the tetra-proline conformers such that only two of them are accessible at room temperature: TTTT and CTTT.

Table 3-7. The conformer energies for TMA-(Pro)₄-N(CH₃)₂. All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets provide the LPNO energy relative to the overall lowest energy conformer (TTCT structure obtained from mPW2PLYP-D optimization). The lowest energy structure, for each conformer, is identified with a shaded cell.

Conformer	ΔE					
	B3LYP	RI-MP2	mPW2PLYP-D	LPNO-CCSD //B3LYP	LPNO-CCSD //RI-MP2	LPNO-CCSD //mPW2PLYP-D
TTCT	12.2	4.1	3.5	0.0 (3.9)	0.0 (1.2)	0.0 (0.0)
CTCT	25.7	0.0	0.0	6.6 (10.5)	0.7 (0.6)	0.4 (0.4)
TCTC	27.8	3.2	3.8	4.4 (8.3)	1.9 (1.8)	1.6 (1.6)
CCCT	32.7	6.4	7.1	3.1 (6.9)	2.1 (2.5)	2.4(2.4)
CTTT	15.2	13.7	12.5	2.6 (6.4)	3.2 (3.7)	3.4 (3.4)
TTTC	14.7	5.1	4.0	0.9 (4.7)	3.9 (3.9)	3.7 (3.7)
CTTC	28.9	6.9	6.2	6.4 (10.3)	4.8 (4.8)	4.0 (4.0)
TTTT	0.0	8.4	9.6	2.0 (5.9)	4.7 (5.2)	5.1 (5.1)
CCCC	37.6	7.6	7.7	5.4 (9.3)	6.3 (5.9)	5.3 (5.3)
TCTT	16.1	11.6	11.0	6.8 (10.7)	8.1 (8.9)	8.7 (8.7)
CCTT	26.2	14.5	15.8	10.5 (14.4)	10.8 (10.4)	10.0 (10.0)
CCTC	40.8	11.7	9.8	11.6 (15.5)	10.9 (11.8)	11.4 (11.4)
TCCT	36.6	19.6	21.0	17.1 (21.0)	13.5 (13.9)	14.0 (14.0)
TCCC	42.7	24.8	22.9	18.5 (22.4)	18.6 (17.8)	17.8 (17.8)
TTCC	34.7	22.2	24.7	18.9 (22.8)	22.3 (21.3)	20.0 (20.0)
CTCC	44.0	23.5	22.8	21.6 (25.4)	21.6 (22.4)	21.0 (21.0)

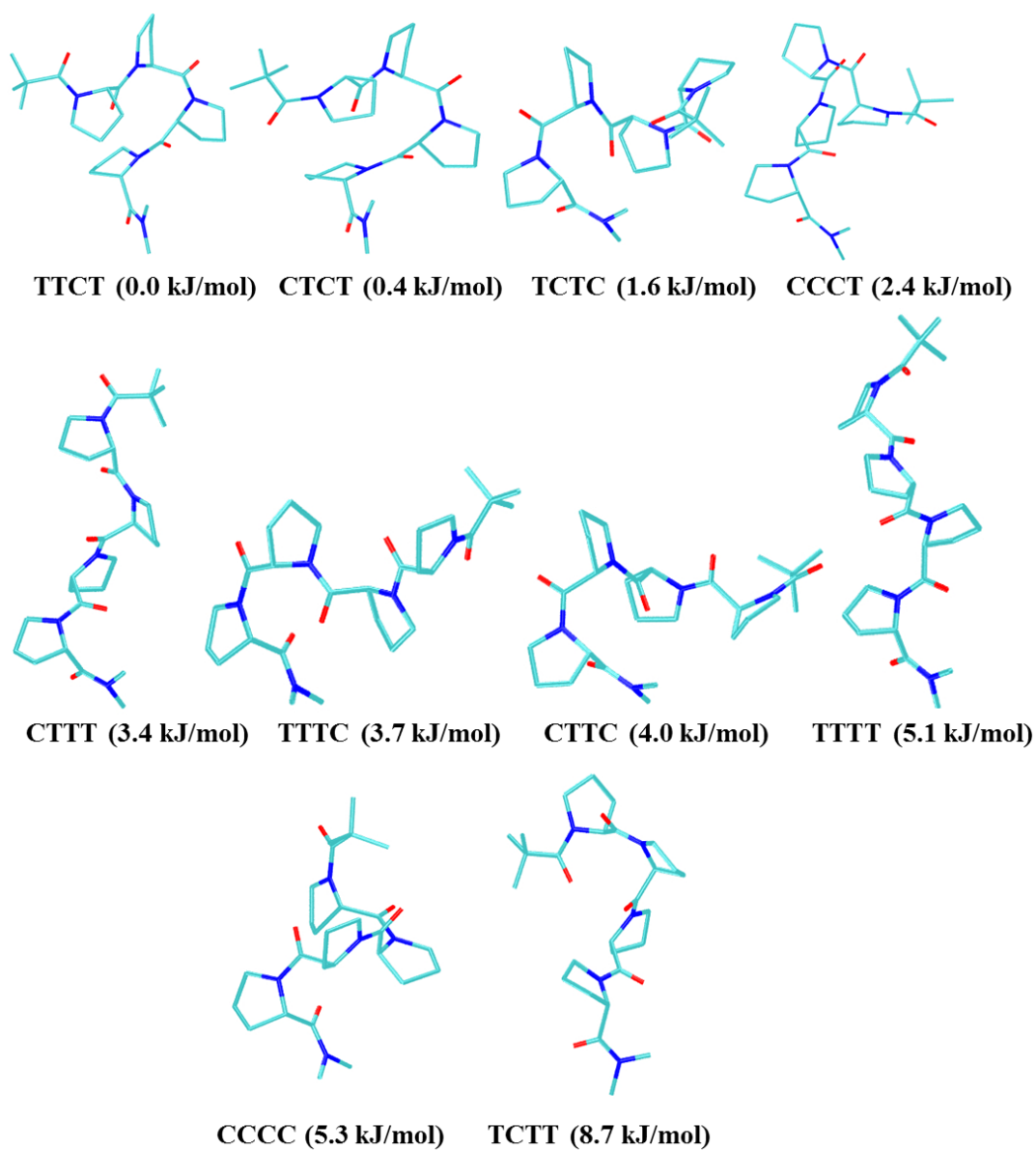


Figure 3-5. The lowest energy conformers of TMA-(Pro)₄-N(CH₃)₂. The structures are obtained from mPW2PLYP-D/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations.

Table 3-8. Relative stabilities of TMA-(Pro)₄-N(CH₃)₂ in continuum solvents. LPNO-CCSD//mPW2PLYP-D/COSMO/6-311G(d,p) methods is used for all calculations. All free energies are in kJ/mol.

Conformer	ΔG			Conformer	ΔG		
	water	hexane	chloroform		water	hexane	chloroform
TTCT	43.3	31.9	33.7	CCCC	38.5	47.7	38.4
CTCT	52.1	44.5	44.3	TCTT	20.2	27.9	17.6
TCTC	65.5	51.4	55.4	CCTT	27.4	31.7	25.2
CCCT	43.6	50.9	44.4	CCTC	62.9	61.4	58.2
CTTT	9.6	12.9	10.6	TCCT	53.2	65.2	55.4
TTTC	56.6	40.1	45.3	TCCC	28.5	46.1	32.8
CTTC	43.5	44.5	37.3	TTCC	53.3	54.4	50.1
TTTT	0.0	0.0	0.0	CTCC	40.4	61.7	54.7

For the tetra-proline, mPW2PLYP-D calculations are feasible but for the penta-proline and hexa-proline, they are highly demanding. Instead, we employ RI-MP2 geometry optimizations that have been shown to perform very well for TMA terminated tetra-proline to find the relative stability of the conformers. The RI-MP2 results are used for developing the penta- and hexa-proline force fields in Chapter 5. The results of mPW2PLYP-D calculations are also presented in Table 3-9. LPNO-CCSD single point energies are reported for tri-proline, tetra-proline and penta-proline, but these calculations are too demanding for hexa-proline.

Table 3-9 lists the relative energies for TMA terminated penta-proline, TMA-(Pro)₅-N(CH₃)₂. Clearly, mixed *cis* and *trans* conformers are more stable for penta-proline. Interestingly, in the gas phase, the *cis* conformer contribution increases when moving from short chains to longer chains. The reason is that long-range dispersion forces tend to make the structure more compact in the absence of any solvent effects [198]. This becomes more obvious when HF and B3LYP methods, that lack dispersion interactions, are not able to accurately predict the order of stability.

Table 3-9. The conformer energies for TMA-(Pro)₅-N(CH₃)₂. All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method. The numbers in brackets provide the LPNO energy relative to the overall lowest energy conformer (TTCCT structure obtained from RI-MP2 optimization). The lowest energy structure, for each conformer, is identified with a shaded cell.

Conformer	ΔE				
	B3LYP	RI-MP2	LPNO-CCSD //B3LYP	LPNO-CCSD //mPW2PLYP-D	LPNO-CCSD //RI-MP2
TTCCT	28.5	0.0	3.6 (23.9)	0	0.0 (0.0)
TTCTC	25.2	5.3	3.3 (23.7)	4.9	5.4 (5.4)
TCTCT	24.6	11.9	3.0 (23.4)	8.1	8.6 (8.6)
TTTCT	13.1	17.7	5.9 (26.2)	8.9	8.7 (8.7)
TTTTT	0.0	26.8	6.7 (27.1)	10.7	10.1 (10.1)
CCCTT	35.6	9.1	12.3 (32.7)	11.5	10.8 (10.8)
TTCTT	11.4	21.6	3.2 (23.6)	11.9	11.0 (11.0)
TTTTC	14.4	22.6	6.2 (26.6)	12.1	12.7 (12.7)
CTCTT	25.6	19.6	12.0 (32.3)	13.1	12.9 (12.9)
CCCCC	39.5	12.2	0.0 (20.4)	14.3	13.2 (13.2)
TCCTC	43.3	15.4	12.6 (33.0)	18.6	15.1 (15.1)
CTCTC	39.7	18.7	13.8 (34.1)	18.9	18.5 (18.5)
TCCCT	41.8	18.2	15.4 (35.8)	19.7	19.0 (19.0)
CCCCT	40.6	20.3	8.2 (28.5)	19.9	19.1 (19.1)
CTTCT	29.8	24.5	15.5 (35.9)	21.3	19.4 (19.4)
CTTTT	16.0	33.1	13.0 (33.3)	21.9	20.2 (20.2)
CCTTT	25.1	29.6	13.2 (33.6)	23.4	20.7 (20.7)
CCTCT	43.9	22.2	24.2 (44.6)	23.4	21.8 (21.8)
TCTTC	31.2	27.5	11.9 (32.3)	23.8	21.9 (21.9)
CTTTC	31.8	30.2	13.9 (34.2)	24.1	22.4 (22.4)
CCCTC	48.1	18.3	12.2 (32.6)	25.5	22.7 (22.7)
CTCCT	47.4	26.2	24.3 (44.7)	25.7	25.5 (25.5)
TCTTT	16.0	37.9	11.8 (32.2)	27.8	26.2 (26.2)
TCCCC	46.0	29.1	14.1 (34.5)	27.9	26.5 (26.5)
TTTCC	35.1	33.4	24.0 (44.4)	28.3	27.4 (27.4)
TCCTT	35.2	31.3	19.7 (40.0)	28.9	28.6 (28.6)
TTCCC	40.2	33.5	18.3 (38.6)	29	29.4 (29.4)
CTCCC	50.8	31.3	23.0 (43.3)	33.4	31.3 (31.3)
CCTTC	46.1	37.8	21.6 (42.0)	33.4	32.0 (32.0)
CTTCC	46.7	41.6	28.4 (48.8)	39.5	37.8 (37.8)
TCTCC	51.8	41.5	27.7 (48.1)	42.7	41.0 (41.0)
CCTCC	58.13	43.7	33.2 (53.6)	43.1	42.2 (42.2)

The most stable conformers of TMA-(Pro)₅-N(CH₃)₂ are shown in Figure 3-6. The initial 32 *cis-trans* amide structures were generated from the optimized tetra-proline geometries with the addition of one proline residue with either a *cis* or a *trans* amide.

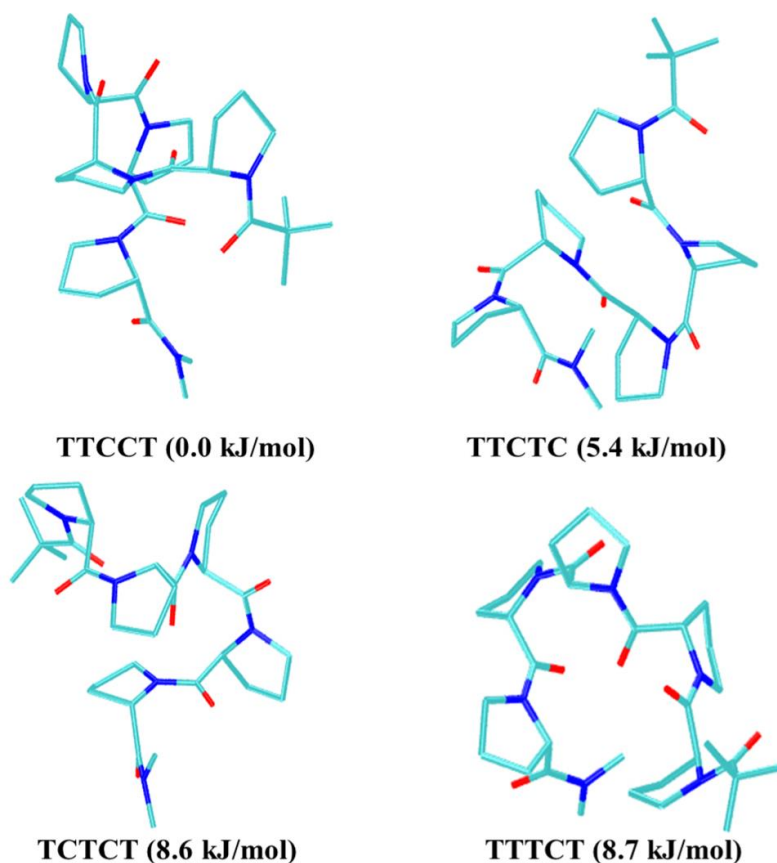


Figure 3-6. The lowest energy conformers of TMA-(Pro)₅-N(CH₃)₂. The structures are obtained from RI-MP2/6-311G(d,p) geometry optimizations and the relative energies are from LPNO-CCSD single point calculations.

As expected, the B3LYP functional predicts TTTTT as the most stable conformer. LPNO-CCSD single point calculations predict TTCCT as the most stable conformer, with three other conformers within 10 kJ/mol. The number of possible *cis-trans* amide conformers is double

the number for the tetra-proline but, for the penta-proline, only four are energetically competitive. These conformers are all bent, with a mix of *cis* and *trans* amide bonds.

For penta-proline, solvation calculations predict PPII as the most stable conformer (see Table 3-10). The next lower energy conformer is CTTTT which is 15-20 kJ/mol higher in energy. Although *cis* rich conformers become more and more preferred in the gas phase, solvation studies show that PPII is dominant in all solvents.

Table 3-10. Relative stabilities of TMA-(Pro)₅-N(CH₃)₂ in continuum solvents. LPNO-CCSD//mPW2PLYP-D/COSMO/6-311G(d,p) methods is used for calculations. All free energies are in kJ/mol.

Conformer	ΔG			Conformer	ΔG		
	water	hexane	chloro form		water	hexane	chloro form
TTCCT	93.2	74.6	68.9	CCTTT	34.5	35.6	31.6
TTCTC	98	76.1	84.4	CCTCT	75.1	74.9	71.8
TCTCT	69.3	59.1	61.6	TCTTC	72.6	57.9	66.2
TTTCT	41.2	34.4	34.5	CTTTC	50	45.8	45.2
TTTTT	0	0.0	0.0	CCCTC	70.2	77.8	71.6
CCCTT	70.5	63.5	67.1	CTCCT	66.3	77.2	71.8
TTTTC	58.2	36.8	27.5	TCTTT	20.8	25.8	24.6
CTCTT	46	45.0	40.7	TCCCC	35.7	53.6	41.0
TTCTT	44.9	36.3	37.6	TTTCC	61.1	57.9	57.8
CCCCC	46.8	54.6	46.7	TCCTT	71.1	76.7	75.6
TCCTC	67.6	71.4	66.6	TTCCC	33.4	61.4	65.0
CTCTC	79	76.8	75.2	CTCCC	62.9	79.3	68.8
TCCCT	88.9	86.4	85.3	CCTTC	60.9	61.3	57.3
CCCCT	58.4	61.3	56.5	CTTCC	71.7	72.4	62.3
CTTCT	58.7	54.5	53.6	TCTCC	92.8	92.2	93.8
CTTTT	15.7	18.9	17.5	CCTCC	68	79.1	68.4

The longest poly-proline examined here, TMA-(Pro)₆-N(CH₃)₂, has 64 *cis-trans* conformers. RI-MP2 geometry optimizations predict that the CTTCT conformer is the most stable in the gas phase followed by CCCTCT which lies only 1.5 kJ/mol higher in energy.

Table 3-11. The conformer energies for TMA-(Pro)₆-N(CH₃)₂. All calculations are performed with the 6-311G(d,p) basis set. All energies are reported in kJ/mol, relative to the lowest energy conformer for each method.

Conformer	ΔE			Conformer	ΔE		
	B3LYP	mPW2PLYP-D	RI-MP2		B3LYP	mPW2PLYP-D	RI-MP2
CTTCT	33.3	0.0	0.0	TCTTCC	53.9	35.4	34.8
CCCTCT	32.6	1.9	1.5	TCCCTT	39.1	35.4	34.9
TCTCTC	37.7	14.6	12.0	CCTTCT	45.2	35.8	35.7
TTCCTT	26.7	16.7	14.4	CTCCCC	57.3	36.1	36.2
CTCTCT	42.7	17.1	16.3	TTTTCC	35.2	36.1	36.6
TCCCCT	43.9	17.5	16.4	TCCTCT	46.7	36.2	38.0
TTCCTC	32.8	17.5	16.4	CCTCTC	50.8	38.1	38.4
CCCCCC	47.7	17.8	16.8	CCTCCC	60.7	38.3	38.5
TTTCTC	26.2	20.4	18.6	TTTCTT	12.2	38.4	38.6
TTCTCT	21.7	21.4	21.3	CTTTCT	32.8	38.8	39.0
CCCCCT	45.4	22.1	21.5	CTCCCT	53.0	39.1	39.3
CCCCTC	58.3	22.8	23.2	TCTCCC	58.0	39.4	40.5
TCTTTC	32.0	24.1	23.9	TTCCCT	40.7	41.5	43.3
CCCTTC	54.2	25.4	24.0	CCTCTT	46.5	41.6	43.3
TTCTTC	25.8	25.5	24.9	TCCTTT	36.6	41.8	43.4
TCCCTC	55.6	25.7	25.0	TTTTTC	15.0	43.7	43.6
CCCTTT	36.2	25.9	25.7	CCTTTC	46.9	43.8	44.8
TCCCCC	48.8	27.4	26.8	CTTTTC	34.8	43.9	46.1
CCCCTT	41.1	27.4	27.0	CTTCTT	18.9	44.3	46.5
TTCCCC	45.2	28.3	28.5	CTTTTT	18.6	44.7	46.5
TTTCCT	30.3	28.4	28.9	CTTCTC	44.6	44.9	47.3
TCTTCT	30.7	28.6	29.5	CCTTTT	31.5	45.6	47.4
CCCTCC	58.3	29.0	29.8	TCCTCC	66.8	45.6	48.1
TCTCTT	26.8	31.1	30.5	TCCTTC	54.6	45.7	49.3
CTCTCC	55.2	31.3	30.8	CTTCCC	54.1	46.3	49.4
CTCTTT	30.2	32.3	32.2	CCTTCC	60.2	47.9	49.7
TTCTTT	11.5	32.7	32.7	TTTTTT	0.0	48.4	54.4
TTTCCC	41.2	33.1	32.9	TCTTTT	17.9	54.4	55.8
CCTCCT	57.9	33.3	33.2	CTTTCC	51.3	56.9	56.8
TTCTCC	49.5	33.5	33.4	CTCCTC	71.1	58.1	56.8
CTCCTT	49.1	34.8	34.2	CTCTTC	48.2	58.5	59.6
TTTTCT	12.5	35.2	34.3	TCTCCT	54.9	58.6	60.1

All other conformers are at least 12 kJ/mol higher in energy. The hexa-proline energies are provided in Table 3-11 and the most stable conformers are shown in Figure 3-7. For hexa-proline LPNO-CCSD calculations were not feasible, but mPW2PLYP-D and RI-MP2 have shown to be in good agreement for shorter chains and are used here.

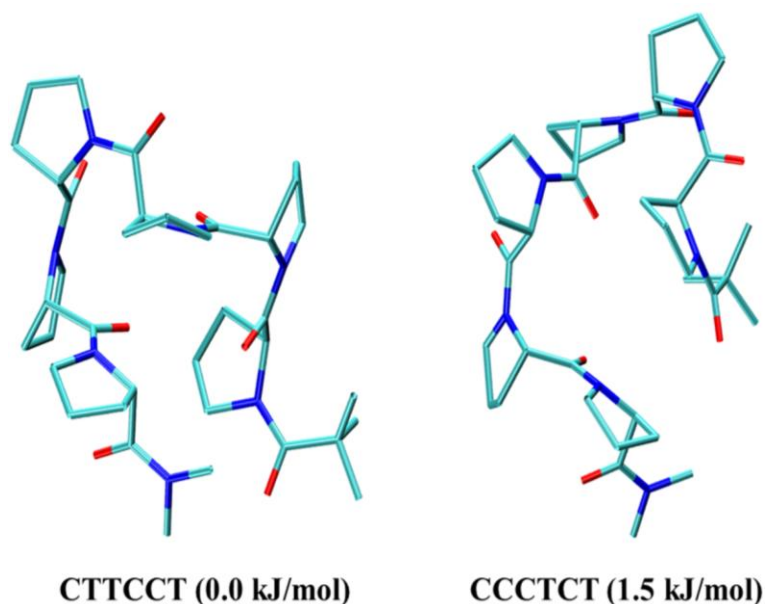


Figure 3-7. The lowest energy conformers of TMA-(Pro)₆-N(CH₃)₂. The structures and relative energies are obtained from RI-MP2/6-311G(d,p) geometry optimizations.

To summarize, the number of conformers increases from four in di-proline to sixty-four in hexa-proline. In terms of interfacial properties, only the most probable conformers are of interest. Our *ab initio* calculations show that the number of energetically-competitive conformers increases from the di-proline to the tetra-proline, where nine conformers are within 10 kJ/mol of the most stable conformer. As the number of prolines increases further, the number of energetically competitive conformers drops sharply. For the penta-proline, we find that only

three, out of thirty-two, conformers are within 10 kJ/mol of the global minimum and the hexa-proline has only two energetically competitive conformers. The lowest energy gas-phase conformers are bent, rather than helical, since a bent structure maximize the dispersive interactions between the upper and lower portions of the backbone. We also find that the fraction of *cis* amides increases with chain length. The relative conformer stability, as a function of the % *cis* amide in the backbone, is shown in Figure 3-8. Looking at the most stable conformers, the *cis* contribution increases from 33% in the tri-proline to 50% in the hexa-proline.

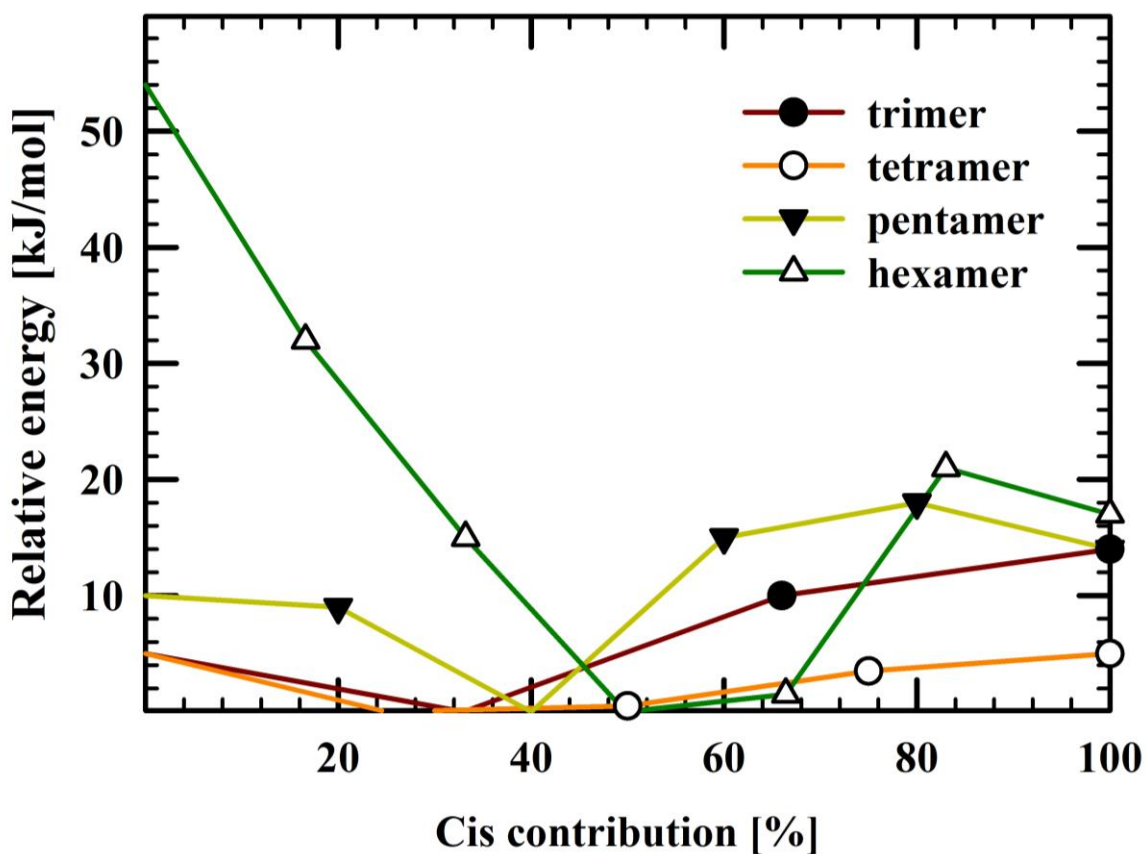


Figure 3-8. The relative conformer energy in the gas phase as a function of the *cis* amide contribution for TMA-(Pro)_n-N(CH₃)₂ with n=3-6. For each conformer, only the most stable is chosen. For instance, there are 4 conformers in tetra-proline that have one *cis* amide linkage (CTTT,TCTT, TTCT, and TTTC). The relative energy is shown for TTCT, the conformer with the lowest energy.

Continuum solvent studies have shown that the all *trans* is always the most stable conformer. For longer chains, this conformer becomes more and more stable. Obviously, this is not consistent with the observed experimental trends discussed in the Section 3.1. The most important reason is that continuum solvent models fail to capture hydrogen bonding accurately. The stability of poly-proline chains is a function of solvent hydrogen bonding capability since it lacks internal hydrogen bonding to stabilize the backbone.

3.3.2 Backbone Flexibility

The backbone torsional angle scans are performed using relaxed scans. Here, all degrees of freedom other than the torsion under consideration are allowed to relax. The scans cover 360° with steps of 10°. All scanning calculations are performed using mPW2PLYP-D/6-311G(d,p). Figure 3-9 shows the energies for rotation over each backbone dihedral starting from the TT conformer of TMA-(Pro)₂-N(CH₃)₂.

The ω_2 angle (see Figure 3-1(b)), as expected, prefers either a *cis* or *trans* angle. A high energy barrier (around 80 kJ/mol) separates the minima and is responsible for the slow *cis-trans* inter-conversion. The ϕ_2 angle exhibits a very high energy barrier for rotation (about 260 kJ/mol). This strongly limits the flexibility of the chain backbone. In fact, ϕ_2 plays the role of a de-coupler since it separates the behavior of ω angle with the torsions of the next proline. The ψ_1 angle adopts two values for *cis* and *trans* amide linkage and the energy barrier for rotation around ψ_1 is about 55 kJ/mol.

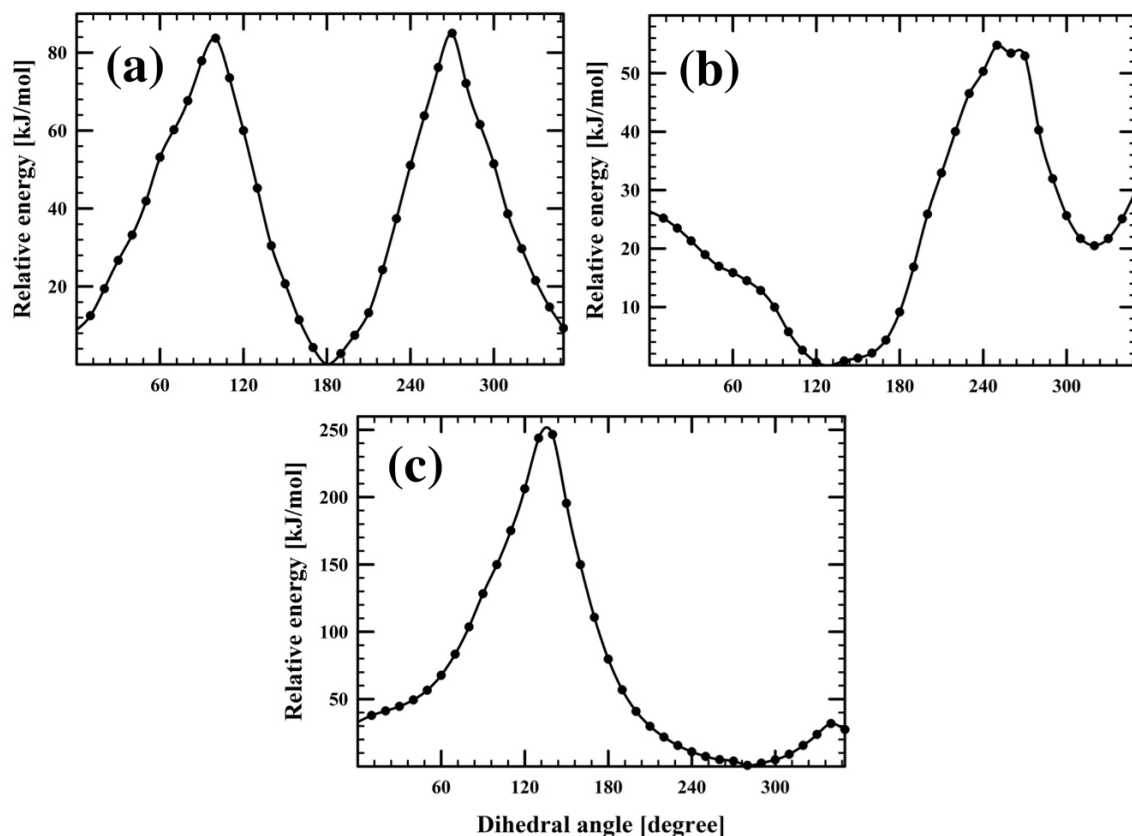


Figure 3-9. Results of scanning backbone dihedral angles using the mPW2PLYP-D/6-311G(d,p) method and basis set starting from TT conformer of the TMA-(Pro)₂-N(CH₃)₂ for a) ω_2 , b) ψ_1 , and c) ϕ_2 .

The coupling of adjacent backbone dihedrals to the ω_2 torsion is shown in Figure 3-10. The analysis of the trajectory shows that the adjacent ψ couples strongly to ω angle while the adjacent ϕ is only weakly coupled. The values of ϕ_1 and ϕ_2 do not change considerably when ω_2 is scanned over the whole range of rotation. On the contrary, the coupled ψ_1 changes by around 180° while ψ_2 changes by around 70°.

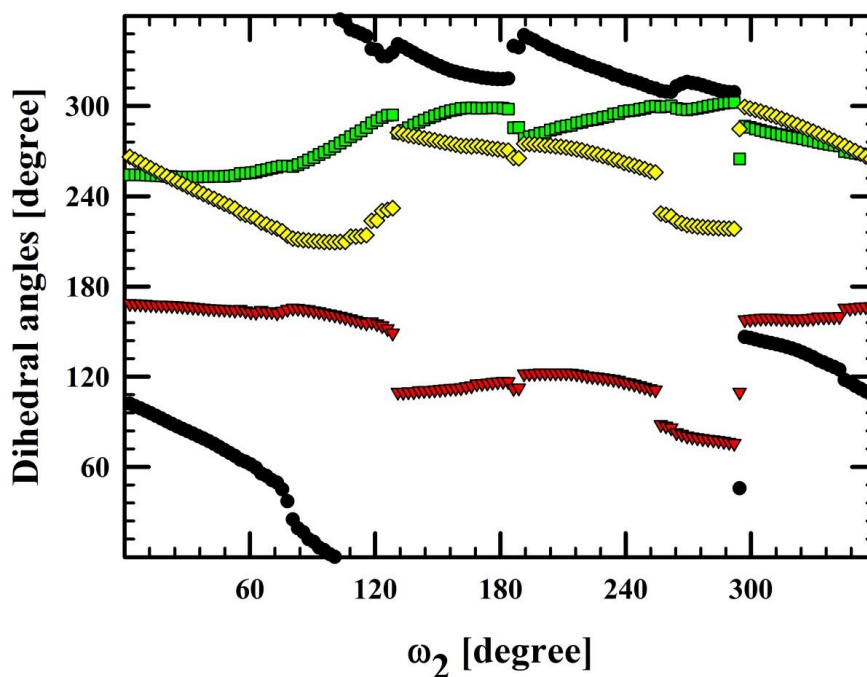


Figure 3-10. The coupling of backbone dihedrals with the ω_2 angle for TMA-(Pro)₂-N(CH₃)₂. Angles are obtained using relaxed scans at the mPW2PLYP-D/6-311G(d,p). The x-axis is the ω_2 angle and the y-axis is the other coupled dihedral angles. Black circles correspond to ψ_1 , red triangles correspond to ψ_2 , yellow diamonds correspond to ϕ_2 , and green squares correspond to ϕ_1 .

A 3D scan of the potential energy surface was also performed when both ω_2 and the coupled ϕ_1 are fixed and the rest of the molecule is relaxed at intervals of 5°. A Counter plot for ω_2 and ψ_1 is presented in Figure 3-11. The circles and triangles show the lower and higher energy pathways, respectively. The triangles represent a pathway that forms the angle of around 320° for the ψ_1 . This is in accordance with what Zhong and Carlson [185] observed using B3LYP/6-31G(d) method and basis set. They showed the presence of a new conformer, named PPIII that is necessary for *cis-trans* isomerization as the change in handedness of the backbone is necessary for such isomerization. Note that this local minimum has $\omega_2 \approx 180^\circ$, but it is similar to PPI.

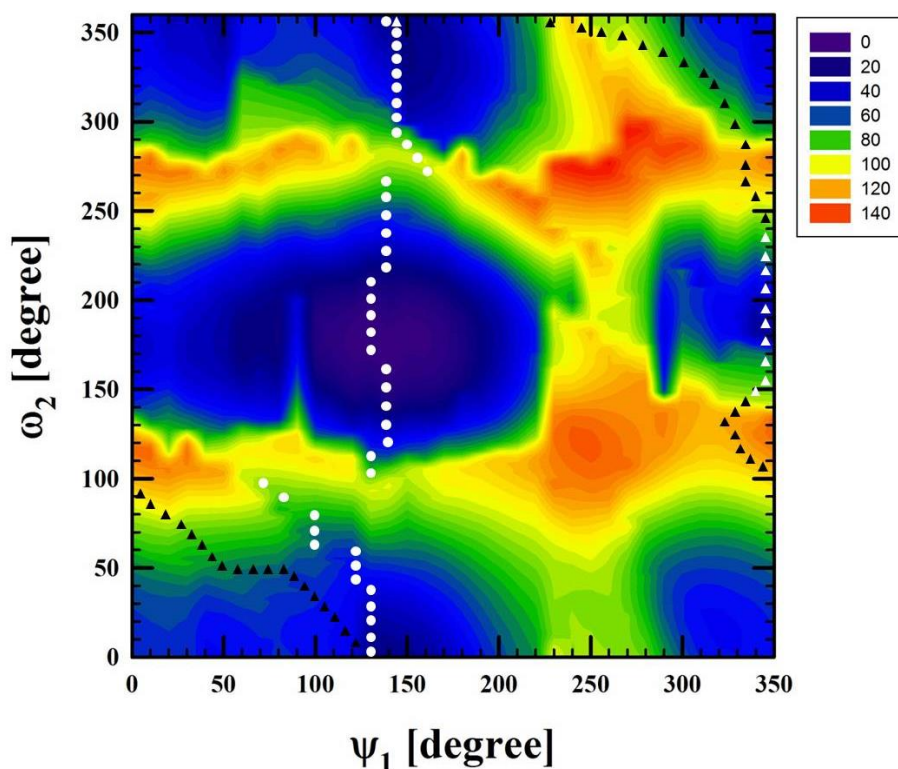


Figure 3-11. Counter plot representation of the potential energy surface with respect to the ω_2 and ψ_1 angles for TMA-(Pro)₂-N(CH₃)₂ using mPW2PLYP-D/6-311G(d,p). The white circles and black triangles correspond to two transition pathways. The lowest energy is around $\omega_2=180^\circ$ and $\psi_1=125^\circ$.

To study the solvent impact on *cis-trans* inter-conversion barrier, relaxed scans were performed using mPW2PLYP-D/COSMO/6-311G(d,p) for the 360° of rotation with steps of 10°. Figure 3-12 shows the energy barrier for the *cis-trans* isomerization of TMA-(Pro)₂-N(CH₃)₂ in the gas phase, and in continuum water, chloroform, and hexane. The curve shape does not change, however, the height of the barrier decreases in more polar solvents. The barrier for *cis-trans* conversion is around 84, 82, 81, and 74 kJ/mol for vacuum, hexane, chloroform, and water, respectively. The presence of implicit water stabilizes the transition state somewhat.

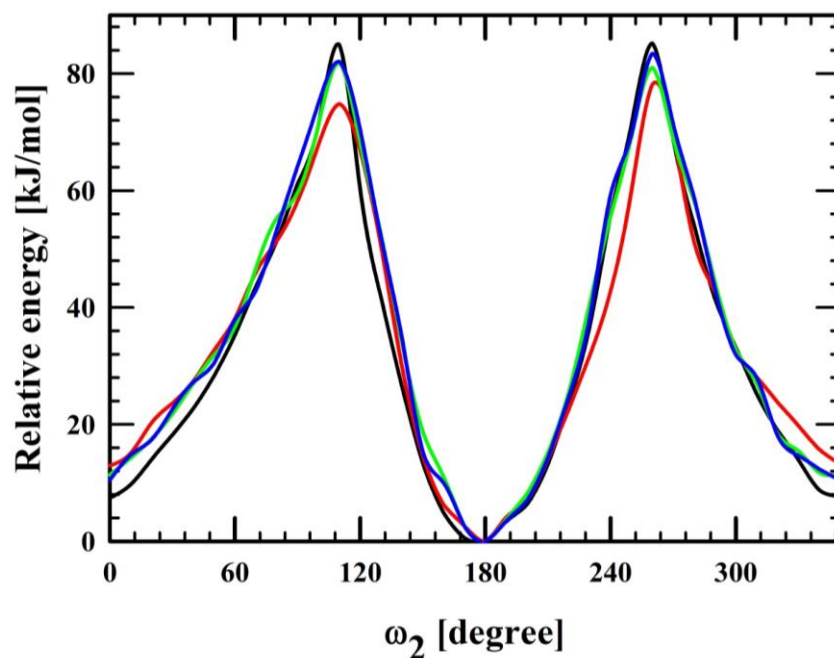


Figure 3-12. The energy barrier for *cis-trans* isomerization for TMA-(Pro)₂-N(CH₃)₂ using mPW2PLYP-D/COSMO/6-311G(d,p) method and basis. The black line corresponds to the gas phase, the red line corresponds to water, the green line corresponds to chloroform and the blue line corresponds to hexane.

Our solvation studies showed that the all *trans* conformer is always the most stable conformer in water, chloroform, and hexane and that the energy differences between all *trans* and all *cis* decreases as the solvent becomes more polar. Although it is believed that in polar solvents, all *trans* is predominant and in aliphatic alcohols all *cis* is, this also is a function of chain length. On the other hand, we should keep in mind that hydrogen bonding, as one of the most important stabilizing factors in poly-proline chains [210], is not treated explicitly in a continuum solvation model.

3.3.3 Puckering Studies

The five membered pyrrolidine ring in proline can adopt two distinct conformations which are characterized by different ring puckers. If the C^γ (see Figure 3-1) of the pyrrolidine ring and the carbonyl group lie on the same side of the plane defined by C^δ , N and C^α , it is called “down”. If they are on the opposite side, the ring pucker is described as “up”. An alternative way is to define the endocyclic angle χ defined between N- C^α - C^β - C^γ . If this angle is positive, the ring pucker is “down” and if the angle is negative it is called “up”.

Pyrrolidine ring puckering has been the subject of several theoretical and experimental studies [164,187,190,194,211]. Experimental studies show that “down” and “up” are almost equally probable, although planar rings are also observed [212]. Experimental studies show that there are two puckered conformations in water for poly-proline chains [213,214]. Theoretical studies mainly confirm very close stabilities for pucker “down” and “up” conformers (although pucker “down” is more stable in all cases) and the transition energy barrier is almost 10-12 kJ/mol [164,187,190,194,211,215]. The energy barrier for “down”-“up” inter-conversion has been studied for Ac-Pro-NHMe using B3LYP/6-311++G(d,p) by Kang [215]. They have found energy barriers of 13.0 and 12.6 kJ/mol for *trans* and *cis* conformers in water and “down” and “up” conformers had very close energies.

Although these findings are interesting, they are only about mono-proline. Here, we have performed mPW2PLYP-D/COSMO/6-311G(d,p) calculations for both rings of TMA-(Pro)₂-N(CH₃)₂ in gas, water, chloroform, and hexane and for both all *cis* and all *trans* conformers to study the energy barrier of inter-conversion between “down” and “up”. The ring pucker and peptide bond dihedral angles can be studied as decoupled degrees of freedom in a sense that a

“down” to “up” transition associated with a change in χ (see Figure 3-1) from almost 30 to -30° does not alter the backbone dihedrals [216].

Figure 3-13 shows the energy barrier for such a conversion in the gas phase, water, chloroform, and hexane for both the CC and TT conformers. Generally, the energy barrier is higher for di-proline than mono-proline. Gas phase and water energy barriers are about 25 kJ/mol and this clearly shows the higher rigidity of rings in longer proline chains. For the CC conformation, water provides a more stable transition state with the energy of 23.0 and 24.7 kJ/mol for the first and second ring compared to gas phase values that are 32.1 and 25.1 kJ/mol. These energy barriers are even higher for chloroform and hexane. Furthermore, C-terminal and N-terminal rings do not show the same flexibility: For all cases, except water, energy barrier for the rotation of the first ring is higher. The position of the transition state does not depend considerably on the solvent except for water where the transition state for the “down”-“up” interconversion in the ring nearest to TMA terminal group appears at around 10° . This value is around $0-5^\circ$ for other cases. Interestingly, the transition state could also be planar. For instance, the transition state of the first ring appears at around 0° and 10° for the gas phase and water, respectively.

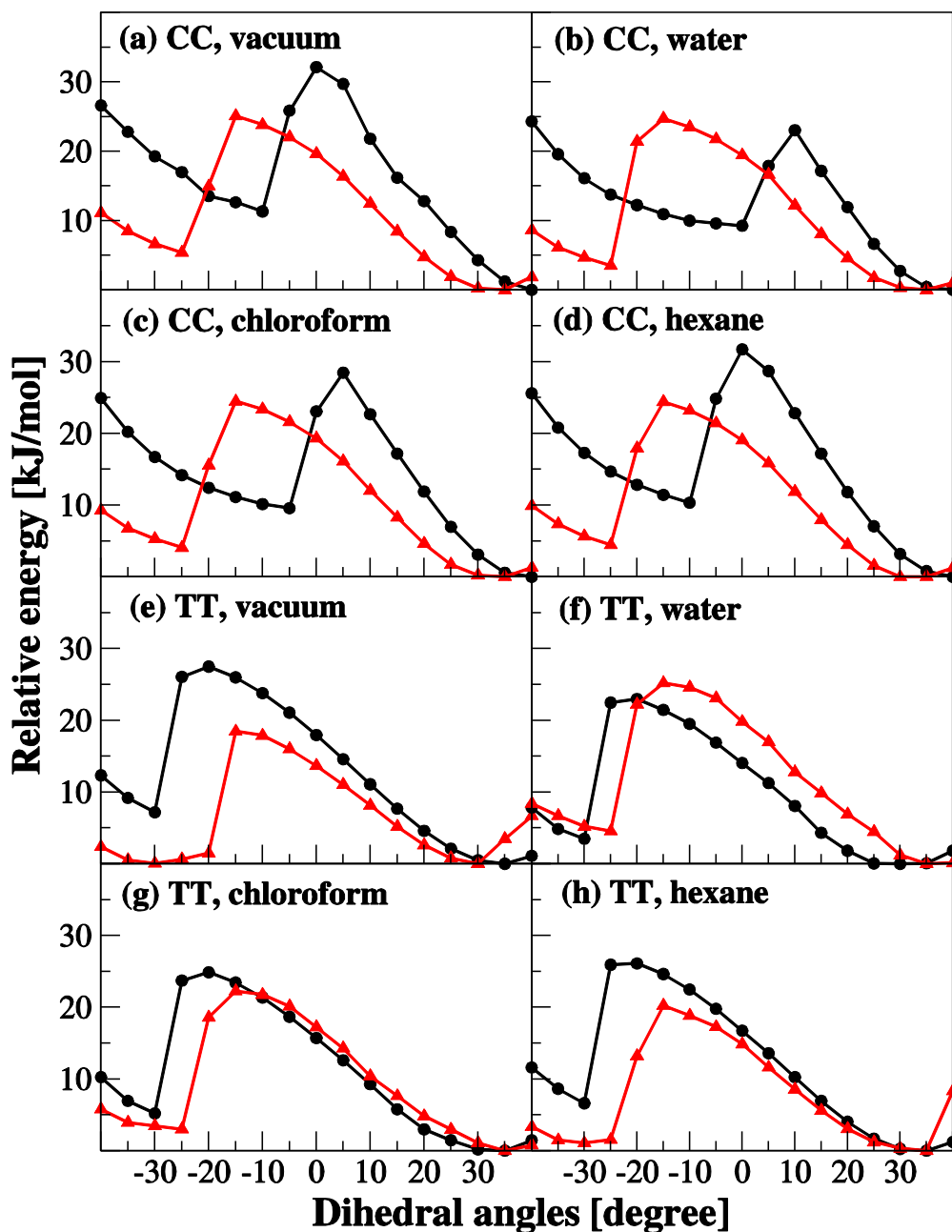


Figure 3-13. The endocyclic torsion χ energy barrier for both rings of the TMA-(Pro)₂-N(CH₃)₂ using mPW2PLYP-D/COSMO/6-311G(d,p) in (a) gas phase for CC, (b) water for the CC conformer, (c) chloroform for the CC conformer, (d) hexane for the CC conformer, (e) gas phase for the TT conformer, (f) water for the TT conformer, (g) chloroform for the TT conformer, (h) hexane for the TT conformer. The black lines with filled circles show first ring and the red line with filled up triangles shows the second ring energy barriers.

Ring pucker stability mainly depends on the ring position on the chain. For the first ring in the gas phase, water, chloroform, and hexane, the “down” conformer is more stable than “up”. In fact, the “up” conformer dihedral angle shifts towards a planar form with the angles of -10° , 0° , -5° and -10° for the gas phase, water, chloroform, and hexane, respectively. The “down” conformer of the first ring is more stable by almost 10kJ/mol than the nearly “planar” conformer in all cases. For the second ring, “up” conformer is more stable than planar conformer, but still less stable than “down” by about 4-5 kJ/mol for vacuum, water, chloroform, and hexane.

Comparison of all *cis* and all *trans* conformers is also interesting. It appears that the N-terminal ring is more influenced by the *trans* to *cis* isomerization since its flexibility differs for all *cis* and all *trans* while for the C-terminal ring it is not the case. The all *trans* energy barriers are relatively smaller than all *cis* conformer and energy gap between “down” and “up” conformers are reduced. The energy barrier, the transition state dihedral angle and the relative stability of the “down” and “up” conformers are almost the same for the gas phase, chloroform, and hexane. In water, however, the energy barrier is reduced by almost 5 kJ/mol for the first ring. Moreover, in water, “down” is more stable than “up” by only 3 kJ/mol while this number is around 6 kJ/mol for other cases. The second ring is more sensitive to the less polar solvents. For the gas phase, chloroform, and hexane, the energy barrier is lowered compared to the water. The “down” conformer is more stable by about 0.1, 5, 3, and 1 kJ/mol in the gas phase, water, chloroform, and hexane.

The structural changes that accompany a ring flip are small and, in particular, we have found that the backbone remains largely unchanged. For these reasons, our force fields are developed for the “down” configuration of the rings.

3.3.4 Other Stable Conformers

By increasing the length of the proline chain, the number of possible conformers drastically increases. To examine the possibility of other stable conformers, beyond those generated from *cis* and *trans* amides, all backbone angles are changed based on their torsional-energy behavior (Figure 3-9) using B3LYP/6-311G(d,p) and then further optimization at mPW2PLYP-D/6-311G(d,p) level of the theory, if a conformer with relative energy of less than 20 kJ/mol is located. As expected, varying the ϕ dihedral angle does not produce any new conformer and always converges at around 270°. On the other hand, ψ can adopt an angle of around 320° in addition to its more stable value at around 120°. For di-proline, this variant of TT conformer has a relative energy of about 20 kJ/mol higher than the global energy minimum. Conformer searches have been performed for all conformers of mono- to penta-proline. All ψ angles are initially changed to around 320° and then the structure is allowed to relax to examine the stability of this conformation. With one exception, the resulting local minima are found to be higher in energy by 20-40 kJ/mol in comparison with the most stable conformer for mono- to tetra-proline. The exception occurs for penta-proline where the variant of the most stable conformer, TTCCT, has a ψ_3 angle of 6.6° and a relative energy of 2.9 kJ/mol obtained with LPNO-CCSD//mPW2PLYP-D/6-311G(d,p) method. Figure 3-14 shows the rotational energy barrier for changing ψ_2 and ψ_3 (between the second and third pyrrolidine rings in penta-proline) for di- and penta-proline, respectively. This energy barrier for ψ rotation follows the same pattern for mono- to tetra-proline, but at angles around 300-360° penta-proline has a different behavior. To the best of our knowledge, such a stable conformer has not been previously observed for proline chains. This result clearly shows that for longer chains with more complexity and flexibility, other conformers may have significant contributions to the total population.

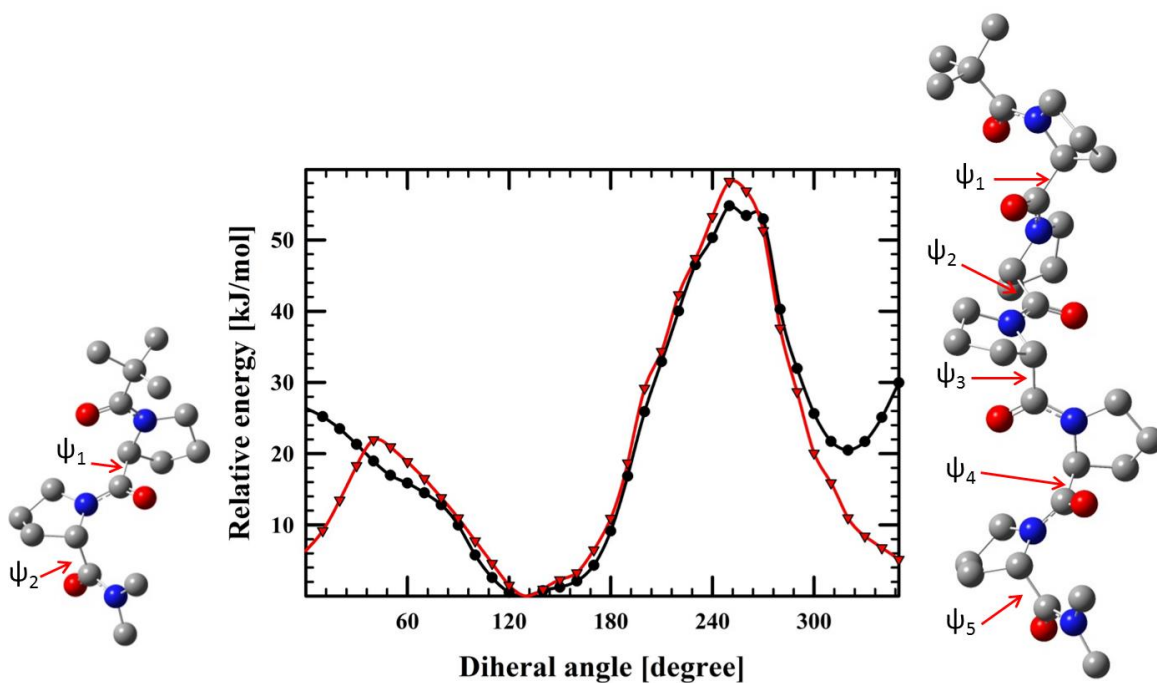


Figure 3-14. The relative energy barrier for rotating around ψ_2 (black, filled circle) and ψ_3 (red, filled triangle) backbone dihedral angles for di- and penta-proline chains, respectively, using mPW2PLYP-D/6-311G(d,p).

3.4 Conclusion

Extensive *ab initio* calculations were performed to investigate the stability and flexibility of TMA-(Pro)_n-N(CH₃)₂ with n=1-6 and BOC-(Pro)_m-N(CH₃)₂ with m=1-3 in the gas phase and in three different continuum solvents: water, hexane, and chloroform. The results show that the terminal group plays an important role in the relative stability of conformers. For instance, for mono-proline chains terminated with either TMA or BOC, the most stable conformers have *trans* and *cis* amides, respectively. For mono- and di-proline, the most stable conformer is all *trans* for TMA terminated proline chains. For longer chains, the all *cis* conformers and mixed *cis-trans* conformers become more and more stable in the gas phase.

The B3LYP functional has been shown to always find the all *trans* as the most stable conformer in the gas phase with the closest competing minima at least 10 kJ/mol higher in energy for TMA-(Pro)_n-N(CH₃)₂ with n=1-6, respectively. Dispersion is missing for B3LYP functional and inclusion of dispersion forces produces more accurate results.

Continuum models are used to find the stability of TMA-(Pro)_n-N(CH₃)₂ with n=1-5 for water, chloroform, and hexane. These studies show the domination of all *trans* amides for all solvents, especially for longer chains.

Ring pucker conformations are also studied for TMA-(Pro)₂-N(CH₃)₂ in the gas phase and in continuum water, chloroform, and hexane. The rings are shown to predominantly adopt “down” conformations. The conversion energy barrier is found to be dependent on adjacent backbone dihedrals and increases for di-proline relative to mono-proline. The relative stability of “down” and “up” conformers depends on the adjacent backbone dihedrals, the position of the proline residue in the chain, and the solvents.

Chapter 4

Solvation Studies of Di-proline Chiral Stationary Phases

4.1 Introduction

In this chapter, the structure and solvation of TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether chiral selectors are studied. These selectors are shown in Figure 4-1. The di-proline structural properties are discussed in Section 4.1.1 and then the associated force field development is discussed in Section 4.1.2. The interfacial system is presented in Section 4.1.3 along with a description of the interaction potential for the solvent-surface interface. Properties of interest, such as solvent distribution at the interface, are discussed in Section 4.1.4. Simulation details are presented in Section 4.2 and Section 4.3 presents the results. This chapter ends with a conclusion in Section 4.4.

4.1.1 The Di-proline Selectors

The two di-proline selectors of interest, TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether (see Figure 4-1 for numbering), can be roughly divided into three components: the tether joining the selector to the surface; the di-proline moiety; and the terminal group which differs slightly for the two selectors. Each tether consists of a propyl chain covalently bonded to an underlying Si surface. A simple representation of the underlying silicon surface is chosen for three reasons. First, the solvent is sterically excluded from the Si layer. In fact, even water is never within 3 Å of the surface. Second, for brush-type chiral selectors, the crucial interactions occur well above the surface. Finally, the atomic details of the underlying silica are unknown but expected to be varied. The two selectors each have three amide bonds. The amide bond joining the two proline units, C(17)-N(21)-C(9)-C(23) in Figure 4-1, is important in describing the

relative conformations of the backbone. When the amide torsional angle is near zero, the carbonyl group is *cis* to the alpha carbon C(17). Likewise an angle near 180° places the carbonyl oxygen *trans* to the alpha carbon. The second amide group joins the terminal group to the di-proline unit. Finally, an amide bond joins the tether to the di-proline and, for this amide, a *cis* or *trans* arrangement merely leads to a reorientation of the di-proline moiety relative to the surface.

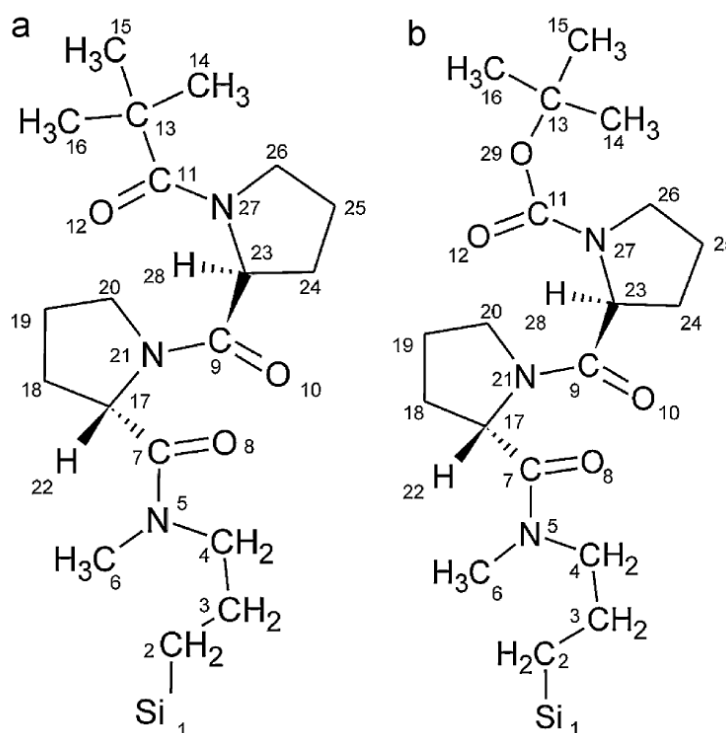


Figure 4-1. The two di-proline selectors, TMA-(Pro)₂-N(CH₃)-tether (left) and BOC-(Pro)₂-N(CH₃)-tether (right). The atom numbering shown above is used throughout this thesis. A single number identifies each methyl and methylene group since a united-atom representation is adopted for these groups.

When the amide joining the prolines is *trans*, the selector is extended but, with a *cis* amide between the two prolines, the molecular structure is more compact. A transition from

trans to *cis* on the amide joining the di-proline to the end-group has a smaller impact: the terminal group is simply repositioned. The TMA terminal group has been chosen for two reasons. First, experiments show [94] that this choice of terminal group leads to better chiral selectivity, relative to seven other terminal groups. Also, TMA is achiral and relatively small such that its interactions with solvent will be mostly due to H-bonding at the carbonyl oxygen. The BOC terminal group has been selected based on its similarity with TMA and the fact that it is an inferior chiral selector [94]. The elucidation of the reasons for this difference is one of the objectives in this work.

4.1.2 Force Field Development

Independent force fields have been developed for TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether based on *ab initio* calculations. Specifically, model development follows the procedure outlined below with the force field extracted from B3LYP/6-311G(d,p) calculations. However, the parameters are adjusted to yield an energy difference between conformers that is consistent with CCSD calculations. All force field parameters are provided in Appendix A.

The energetic cost for intramolecular changes is given by

$$U^{intra} = U^{stretch} + U^{bend} + U^{improper} + U^{torsion} + U^{NB}, \quad 4-1$$

where $U^{stretch}$ is the bond stretching potential, U^{bend} accounts for the energetic costs of angle changes, $U^{improper}$ incorporates the energetic costs for out-of-plane motions, the energy associated with twisting about bonds is captured by $U^{torsion}$, and U^{NB} is a non-bonding term between atoms that are separated by four or more bonds. Within the force field, a united-atom representation is adopted for all methyl and methylene groups. As well, the pyrrolidine rings are kept rigid.

Bond stretching is included only between the carbonyl carbons and the adjoining alpha carbons and nitrogens. From Figure 4-1, these bonds connect the pyrrolidine rings to the backbone. All other selector bonds are kept fixed using the Rattle algorithm [150]. To account for the energetic cost of changing a bond, a series of ten single point calculations is performed, with the bond length varied by up to ± 0.075 Å about the equilibrium value. The resulting energies are least squares fitted to a harmonic potential (Eq. 2-26). The force field includes potentials for all bends except those that involve three ring atoms, since the rings are rigid in our model. The bending potentials are calculated by sequentially varying each angle about the equilibrium value, in steps of 2.0° up to $\pm 10.0^\circ$. The bending potential is given in Eq. 2-27. Improper torsions are included to properly account for the energetics of distortion about a central atom. Specifically improper torsion at the alpha carbons, the carbonyl oxygens, and the ring nitrogens are included in the potential. A series of *ab initio* calculations is undertaken with the central atom moving up or down relative to the plane of the adjacent three atoms, by as much as 10.0° , in steps of 2.0° . The improper torsion potential has the form of Eq. 2-29. The torsion potential accounts for the energetic changes that accompany twisting about a bond. This form of the potential (Eq. 2-28) is flexible enough to reproduce the energy variation with torsional angle for the entire 360° range of motion. For each torsion *ab initio* energies are obtained for a series of 36 angles between 0 and 360° .

Atoms separated by four or more bonds do not share stretch, bend, improper torsion, or torsion potentials. Nonetheless, some interaction is present between distant, intramolecular atoms. It is customary to introduce “non-bonding” potentials between these atomic pairs but to reduce the potential when transitioning between the close atom regime and the distant atom regime. With these considerations in mind, the non-bonding potential has the form:

$$U^{NB} = C \sum_{\{\geq 4 \text{ bonds}\}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + D \sum_{\{\geq 4 \text{ bonds}\}} \frac{q_i q_j}{r_{ij}}, \quad 4-2$$

where the first sum is a LJ potential and the second is the electrostatic potential. The constants C and D are unity except for atoms separated by exactly four bonds, where C and D are less than one. Transferrable force fields typically transition between the two regimes when atoms are separated by three bonds. Here, due to the small differences between conformer energies, we have chosen to introduce non-bonding interactions beginning with 1-5 pairs. Aside from the two parameters, C and D, the “non-bonding” potential is calculated as though the atoms belonged to distinct molecules. The LJ parameters (see Eq. 2-30) are taken from the OPLS force field [216-219] while the atomic charges are obtained from the CHELPG algorithm [220] applied to the most stable conformer (TT for B3LYP and TC for B97-D).

The evaluation of $U^{torsion}$ is considerably more difficult than an assessment of energetic cost of bending, stretching, or improper torsion. For the latter three, least squares fitting leads directly to the parameters required for Eqs. 2-26, 2-27, and 2-29. Three factors are incorporated into our fitting process for $U^{torsion}$: the torsional angles are coupled such that a change in one angle may lead to a significant change in a nearby torsion (see Figure 3-10); the B3LYP torsion potential is adjusted to be consistent with CCSD energies; non-bonding interactions change with torsional angle and, in order to avoid double counting of non-bonding in the force field, the torsion potential must be adjusted. Figure 4-2 shows their impact on the torsional potential for the inter-ring amide torsion of TMA-(Pro)₂-N(CH₃)₂.

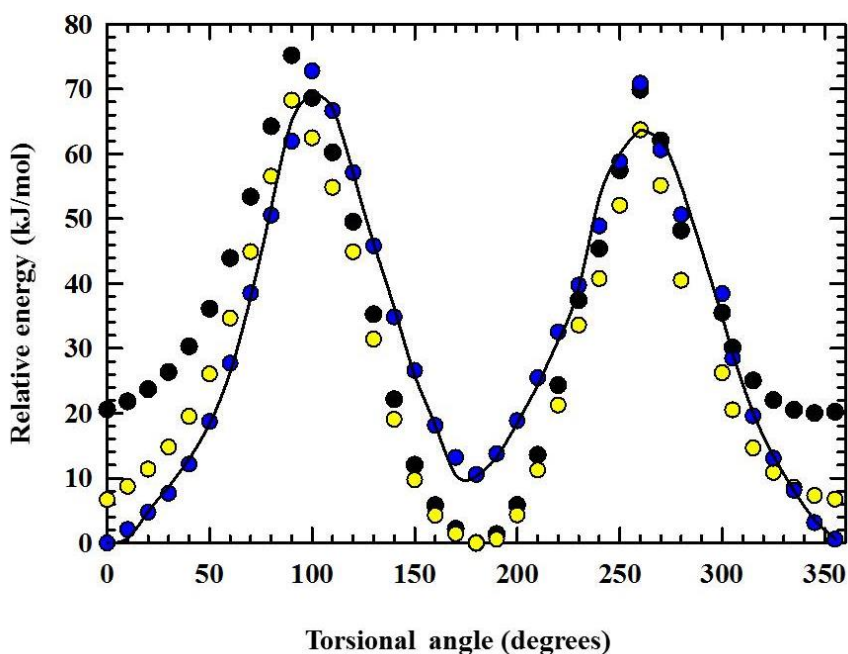


Figure 4-2. Torsional barriers for rotation about the inter-ring amide torsion [C(17)-N(21)-C(9)-C(23)] are shown for TMA-(Pro)₂-N(CH₃)₂. Black circles identify the initial B3LYP/6-311G(d,p) energies. Yellow circles identify the CCSD-corrected torsional potential where the B3LYP energies have been scaled to yield the CCSD *cis-trans* energy difference. Blue circles include corrections for LJ and electrostatic interactions between atoms separated by more-than-three bonds. The final torsional potential, obtained by least squares fitting, is shown by a solid line.

The proper conformational balance between di-proline backbone configurations (TT, CT, TC, and CC) is essential for our simulations. We parameterize our models to accurately reproduce the *ab initio* energies for bond stretching, bending, improper torsion, and torsions. However, a final parameterization step – the optimization of C and D in Eq. 4-2 - is introduced to optimize the force field. For long, flexible molecules, this final adjustment of the model is important as was shown for leucine- and phenylglycine-based stationary phases [112,113]. Figure 4-3 shows the final model assessment for TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂. Each point in Figure 4-3 corresponds to the relative energy and sum of squares obtained from a

single-molecule MD simulation. The simulation starts from a random structure and is performed as the temperature is gradually lowered from 298 to a few degrees Kelvin. The lowest-energy structure obtained in the simulation is kept and analyzed. If the force field accurately predicts the presence and energies of the four conformers, then these single-molecule simulations should converge to the *ab initio* results.

As shown in Figure 4-3, for both selectors, the force field reproduces the relative energy and structure of the global energy minimum and the lowest-energy conformer. The second lowest energy conformer is generally well represented but, for the highest energy conformer (CC) the structure is well reproduced but the energy is 5-10 kJ/mol too low. However, the force field still predicts that the CC conformer is much higher in energy than the global minimum. In the end, we obtain force fields for the di-proline chiral selectors that are structurally and energetically consistent with *ab initio* optimized structures. Importantly, our models also reproduce the energetic cost for stretching, bending, twisting about a bond, and out-of-plane motion.

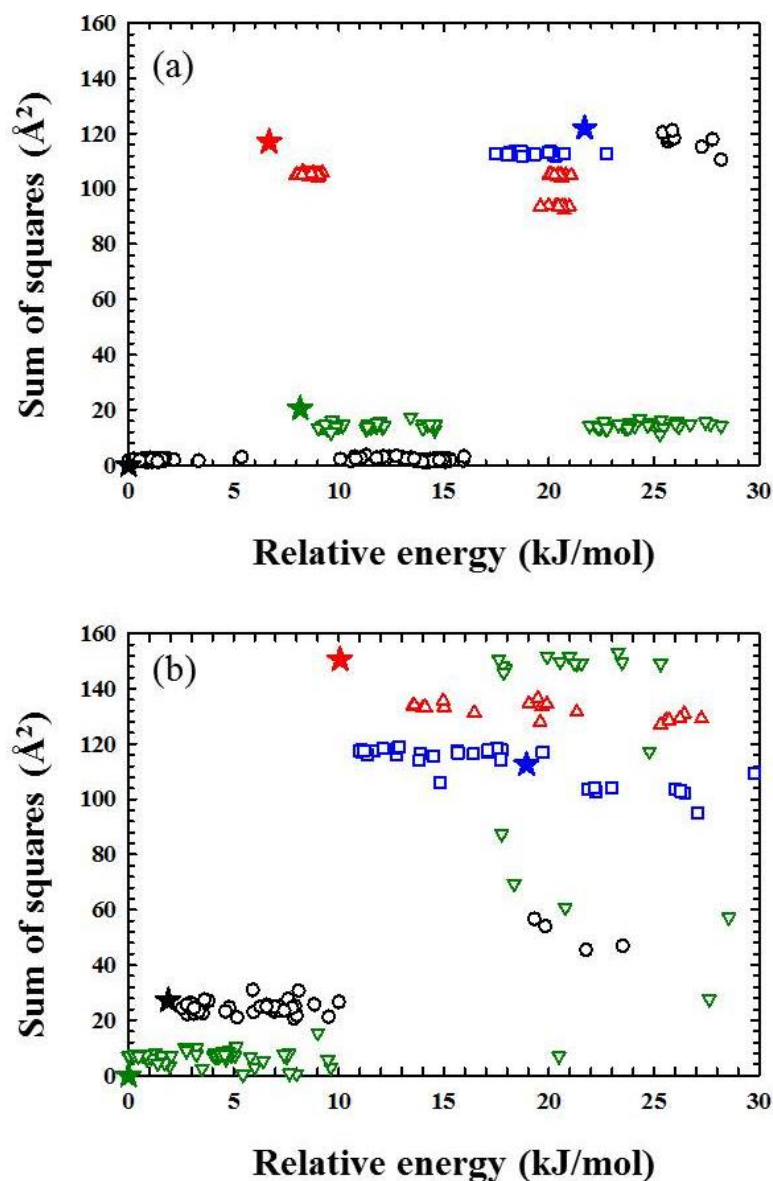


Figure 4-3. Force field assessment for TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂ are shown in (a) and (b) respectively. Each open symbol represents a final conformer and its energy, obtained from a single-molecule molecular dynamics simulation that begins with a random molecular structure that is optimized to a minimum energy conformation by gradual cooling. Black circles, green triangles, red triangles, and yellow squares correspond to TT, CT, TC, and CC structures, respectively. The energy of the structure, as predicted from the molecular model, is plotted against the sum of squares, calculated relative to the optimized minimum energy structure obtained for B3LYP/6-311G(d,p) calculations. The stars identify the relative energies (calculated with CCSD//B3LYP/6-311G(d,p)) and sum of squares evaluated for the four B3LYP/6-311G(d,p) minimum energy conformers.

4.1.3 The Chiral Interface

Figure 4-4 provides illustrative snapshots of the TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether interfaces in the absence of solvent, in water/methanol, and in *n*-hexane/2-propanol. As shown in Figure 2-8 each simulation cell has empty space beyond the two interfaces, such that roughly two thirds of the cell is empty. The empty space allows for the use of 3D periodic boundary conditions. The solvent in our simulations is effectively confined between two infinite, planar chiral surfaces. In the snapshots of Figure 4-4, the surfaces are 34.2 Å apart in water/methanol and 63.1 Å apart in *n*-hexane/2-propanol. Although the chiral selectors are regularly distributed on the surface, the 1.10 μmol/m² coverage is chosen to be consistent with experiment [30]. The surface also includes 54 trimethylsilyl end-caps, and 72 silanol groups with coverages of 3.20 and 4.26 μmol/m², respectively, and 144 silicon atoms.

Experimental surface distributions of selectors, end-caps, and silanols, are expected to be irregular but, assuming that each selector acts individually, the distribution on the surface should have minimal impact for a brush-type chiral stationary phase. The underlying, immovable layer of silicon atoms, shown in yellow in Figure 4-4, is a simple cubic lattice with atoms 3.12 Å apart.

The energy of the interfacial system is given by:

$$E^{full} = \frac{1}{2} \sum_i m_i v_i^2 + \frac{1}{2} \sum_{RU} I_{RU} \omega_{RU}^2 + U^{LJ} + U^{electro} + U^{intra}, \quad 4-3$$

where U^{intra} is the intramolecular potential. The first term in Eq. 4-3 is the translational kinetic energy while the second term is the rotational kinetic energy. The latter is calculated for the rigid units (RU) in the selectors. The translational kinetic energy includes contributions from the solvent, the end-caps, the silanol groups, and the selectors but the underlying Si layer is immovable and does not contribute. For the selectors, only the centers of

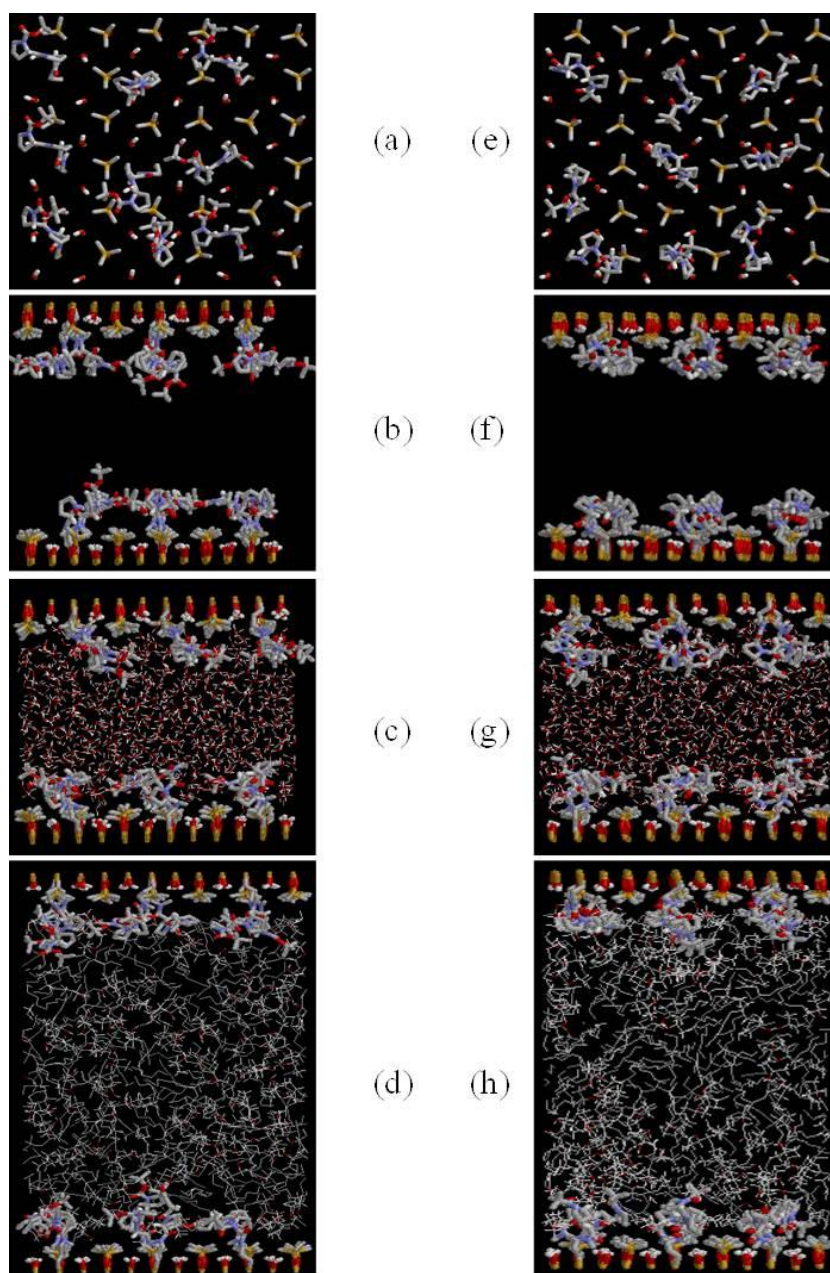


Figure 4-4. Illustrative snapshots of the BOC-(Pro)₂-N(CH₃)-tether interface is shown (a) top view of the surface in the presence of a *n*-hexane/2-propanol, (b) in the absence of solvent, (c) in the presence of a water/methanol and (d) in the presence of a *n*-hexane/2-propanol. Corresponding snapshots of the TMA-(Pro)₂-N(CH₃)-tether interface are shown in (e), (f), (g) and (h). The underlying Si layer is omitted for clarity. Carbon, nitrogen, hydrogen, oxygen, and silicon atoms are shown in grey, blue, white, red, and yellow, respectively. The surfaces include di-proline selectors, trimethylsilyl end-caps, silanol groups, and underlying Si atoms.

the pyrrolidine rings contribute to the translational kinetic energy. m_{RU} is the mass of the ring, I_{RU} and ω_{RU} are the moments of inertia and angular velocity, respectively, in the principal coordinate system of the ring. Interactions between the surface and solvent, between solvent components, and between surface-bound molecules are represented by LJ potentials (Eq. 2-30), and electrostatic potentials (Eq. 2-32). For the surface components, OPLS [217-220] parameters are chosen except for Si where OPLS parameters are not available and CHARMM values are employed instead [222].

Electrostatic interactions are also included via Ewald sums. Specifically, for our simulations the selectors, the 2-propanol, the methanol, and the water have atoms that bear partial charges. Literature values [223] have been adopted for the solvent atomic charges.

4.1.4 The Assessment of Interfacial Structure

The assessment of interfacial structure begins with surface distribution functions (Eq. 2-66). For solvent distributions, the number of solvent molecules expected at a distance z above the surface is $n^{ideal}(z) = \rho \Delta z L_x L_y$ where L_x and L_y are the surface dimensions, and ρ is the density. Detailed information of solvent location at the interface is provided by 2-dimensional (2D) cylindrical distributions $g(r,z)$. In this case, interatomic separations are divided into a component perpendicular to the surface, z , and a component parallel to the surface, r . These 2D distributions are used to study the relative positions of chosen solvent atoms around the chiral selector.

$$g(r, z) = \frac{n^{actual}(r, z)}{n^{ideal}(z)} = \frac{n^{actual}(r, z)}{2\pi r \Delta r \Delta z \rho}, \quad 4-5$$

where ρ is the number density of the solvent, and $\Delta r = \Delta z = 0.094 \text{ \AA}$ for *n*-hexane/2-propanol and $\Delta r = \Delta z = 0.107 \text{ \AA}$ for water/methanol. A positive z means the solvent atom is located above

(further from the underlying Si layer) the surface atom. 2D distributions are particularly instructive for hydrogen bonding interactions between solvent and selector.

Snapshots of the interfacial system provide direct confirmation of the selector conformations and solvent-selector interactions. The occurrence of hydrogen bonding is detected from the snapshots via the application of a geometric criterion [224,225]: The distance between H and a hydrogen bond acceptor should be less than 2.6 Å; the distance between the H-bond donor and H-bond acceptor should be less than 3.5 Å; and the donor-H-acceptor angle should be larger than 150°. We analyze hydrogen bonding probabilities for both selectors and for the amide oxygens and amide nitrogens in the backbone. The orientation of selector carbonyl groups is also analyzed from the snapshots. In this case, the angle between the C=O bond and the surface normal is averaged and results are reported.

4.2 Simulation Details

Extensive MD simulations were performed for the chiral surfaces in the presence of a vacuum, a 70/30 v/v *n*-hexane/2-propanol and a 70/30 v/v water/methanol solvent. These specific solvent mixtures have been chosen since they are commonly employed for chiral separations, although the amount of co-solvent varies depending on the analyte. The surfaces include end-caps, silanol groups, and a layer of immovable silicon with coverages that are consistent with experiment [222]. In practice, for the simulations, the interface distance is adjusted such that the solvent density in the center of the simulation cell is within 2% of the experimental density. This adjustment is required because the volume occupied by silanol groups, di-proline selectors, and end-caps cannot be exactly estimated. For the *n*-hexane/2-propanol simulations, the surfaces are 63.1 Å apart the simulation cell includes 231 *n*-hexane molecules and 169 2-propanol molecules, consistent with the experimental solvent density [156] at 298 K. The water/methanol is denser

with 645 water molecules and 155 methanol molecules in the simulation cell, the surfaces are 34.2 Å apart. In all cases, each surface includes 9 selectors and has a surface area of 1177.86 Å².

Ewald summations [146] have been used to treat the electrostatic forces between partially charged atoms. The empty space beyond the chiral surfaces along with a correction for the elongated shape of the simulation cell [147], allow the use of 3D Ewald sums. An Ewald convergence parameter of $\alpha = 1.756/L_x$ was chosen for all simulations, with a reciprocal space cutoff of $k^2 \leq 27$.

Simulations are performed in the canonical ensemble (NVT) with the use of two Nosé-Hoover thermostats [153,154]. One thermostat is applied to translational motion while the other maintains the rotational temperature of the rigid rings. All simulations are performed at 298 K. The conserved quantity within the simulations is:

$$H_{NH} = E^{full} + \frac{Q_t \zeta_t^2}{2} + g_t k_B T \ln s_t + \frac{Q_r \zeta_r^2}{2} + g_r k_B T \ln s_r, \quad 4-6$$

where the total energy, E^{full} , is given in Eq. 4-3. The two additional degrees of freedom, s_t for translation and s_r for rotation, have masses of Q_t and Q_r , respectively, velocities of ζ_t and ζ_r , and g_t and g_r degrees of freedom. The translational equations of motion are integrated following the algorithm of Martyna *et al.* [226] to preserve the Rattled [150] positions and velocities. The ring quaternions are advanced in time by coupling a leap-frog algorithm with a Newton-Raphson based scheme for the angular velocity and ζ_r .

For each solvent/surface combination, ten molecular dynamics simulations of roughly 0.8 ns duration have been performed for a total simulation time of around 8.0 ns. The time step in the simulations is 0.3 fs, and the simulations are performed with the MDMC program [227] which includes a parallel implementation of the electrostatic and LJ force calculations. The

equilibration period extends over the first 500,000 time steps of each simulation, leaving a 2,100,000 step collection period. With this time step, H_{NH} is relatively constant with long-term drift of less than 2.0 kJ/mol during the collection period of the simulation. The equilibrium properties reported in Section 4.3 are averages over twenty surfaces (2 surfaces/simulation and ten simulations per solvent).

Initial configurations are generated by placing each selector in the minimum energy configuration (TT for TMA-(Pro)₂-N(CH₃)-tether and CT for BOC-(Pro)₂-N(CH₃)-tether). As shown in Figure 4-2, the barrier for inter-conversion between *cis* and *trans* amide bonds is roughly 70 kJ/mol. Such a barrier is not expected to be overcome at 298K. We have allowed for equilibration of the three amide bonds in the selectors by introducing a time-dependent scaling function that temporarily lowers the barrier to amide *cis-trans* inter-conversion without altering the relative energies of the conformers. Specifically, the amide torsions are represented by

$$U^{torsion}(\varphi_t, t) = S(\varphi_t, t) \sum_{i=0}^6 c_{it} \left(\cos(\varphi_t + \varphi_t^{i,0}) \right)^i, \quad 4-7$$

where

$$S(\varphi_t, t) = \frac{(t - 3 \text{ ps}) + (60 \text{ ps} - t \cos^2 \varphi_t)}{(57 \text{ ps})},$$

for simulations between 3 *ps* and 60 *ps*. At other times, the scaling function is one. A further 90 *ps* of equilibration follows the “scaling in” of the amide torsional barrier. It is important to note that the amide conformational balance is obtained in the presence of solvent and other surface elements such as end-caps.

The results presented in Section 4.3 are obtained using the CCSD-corrected B3LYP-based force fields. Although B97-D force fields were also obtained, and simulations conducted, we do not report results here for the following reasons. The B97-D functional predicts TC

conformers of TMA-(Pro)₂-N(CH₃)₂ and BOC-(Pro)₂-N(CH₃)₂ that differ structurally from the B3LYP, MP2, and mPW2PLYP-D structures. In fact, the latter three methods always predict comparable minimum energy structures (dihedral angles are within five degrees for example). CCSD calculations indicate that the B97-D TC structures are roughly 2-4 kJ/mol higher in energy than the TC conformers predicted from the other methods. Thus, the B97-D functional is not consistent with other methods, and CCSD calculations confirm that B97-D predicts incorrect TC structures.

4.3 Results

In this section, results of molecular dynamics simulations are presented for the TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether chiral interfaces. Selector conformations at the interface are discussed in Section 4.3.1. The impact of the end-group, TMA or BOC, is evident from the comparison. Solvent distributions and hydrogen bonding at the TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether interfaces are discussed in Sections 4.3.2 and 4.3.3, respectively.

4.3.1 The structure of the TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether Chiral Interface

Snapshots of the TMA-(Pro)₂-N(CH₃)-tether and the BOC-(Pro)₂-N(CH₃)-tether interfaces, in the absence of solvent, in water/methanol, and in *n*-hexane/2-propanol are presented in Figure 4-4. Consider first the TMA-(Pro)₂-N(CH₃)-tether interface shown in Figure 4-4(e-h). In the absence of solvent, each selector adopts a compact structure that brings selector atoms closer to the surface. In the presence of water and methanol, some of the selectors “stretch out” into the fluid. The most extended selector conformations occur in hexane/2-propanol, as shown in Figure 4-4(f)

where a few selectors have their backbones roughly perpendicular to the surface. A more quantitative perspective on selector conformation is shown in Figure 4-5.

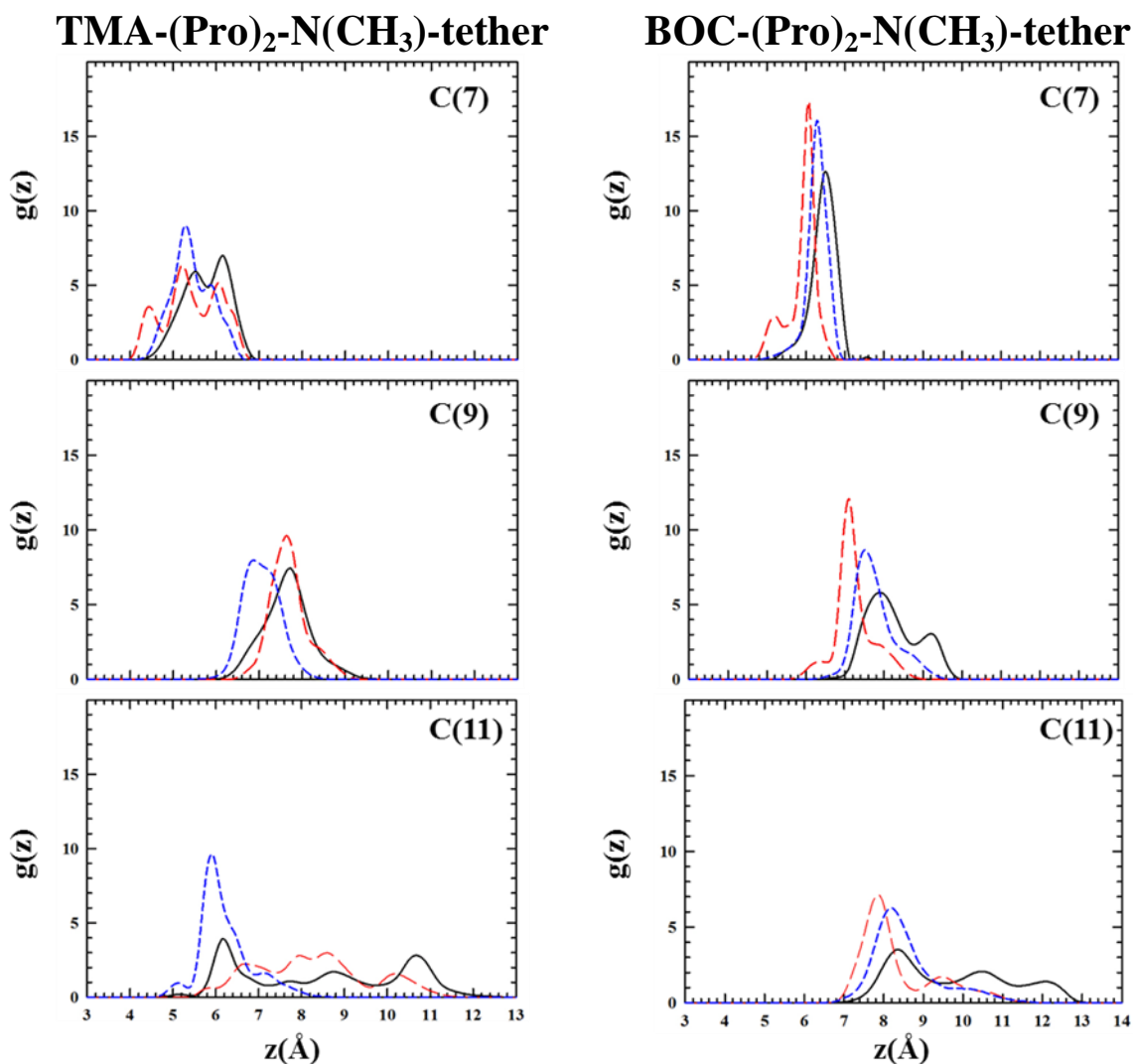


Figure 4-5. Surface distributions of carbonyl carbons (see Figure 4-1 for atom numbering) for the TMA-(Pro)₂-N(CH₃)-tether (left) and the BOC-(Pro)₂-N(CH₃)-tether (right) selector. Results are shown in the absence of solvent (blue, short dashes), in *n*-hexane/2-propanol (black, solid), and in water/methanol (red, long dashes).

The distribution of backbone carbonyl carbons [C(7), C(9), and C(11)] are shown as a function of distance from the underlying silicon layer. Even for C(7), the carbonyl carbon closest to the tether and to the surface, solvent impacts the distributions, with a higher probability of being further from the silicon layer in *n*-hexane/2-propanol. The presence of solvent broadens the distribution for C(11), the carbonyl carbon closest to the terminal group. Without solvent, this carbon is found only within 5-8 Å of the underlying silicon layer but, with solvent present, C(11) can be found up-to 11-12 Å from the silicon. The probability for this extended structure is highest in *n*-hexane/2-propanol.

Snapshots of the BOC-(Pro)₂-N(CH₃)-tether interface in the absence of solvent, in water/methanol, and in *n*-hexane/2-propanol are presented in Figure 4-4(a-d). The conformational distribution of these selectors is quite different from TMA-(Pro)₂-N(CH₃)-tether. In particular, the BOC terminated selectors have a higher probability of adopting a conformation that extends the selector along the surface. This is particularly evident in Figure 4-4(b) where the solvent is absent. The presence of water/methanol encourages the selectors to extend away from the surface, but *n*-hexane/2-propanol leads to significantly more extended selector structures. Figure 4-5 provides the distribution of backbone carbonyl carbons [C(7), C(9), and C(11)] as a function of distance from the underlying silicon layer. All three carbons indicate that the selectors extend furthest into the solvent for *n*-hexane/2-propanol.

A full conformational analysis of the di-proline selectors was performed and the results are summarized in Table 4-1. Not surprisingly, the observed conformers and their abundance are closely related to the expectations based on *ab initio* energies. Overall, we found that both selectors strongly favor a few conformers regardless of the solvent. However, the nature of the terminal group impacts the preference as does the solvent characteristics. We do not observe the

CC conformer in the simulations but this is expected since this structure is significantly higher in energy than others. The TT, TC, and CT conformers, on the other hand, are observed at the interface. For the TMA terminated selectors, the TT conformer is more probable than TC in accord with the expectation from *ab initio* energies. The CT conformer is higher in energy, relative to TC and TT, and is only encountered in water/methanol. The BOC-terminated selectors have different conformational preferences. The *ab initio* energies identify CT and TT as being very close in energy, with TC roughly 10 kJ/mol higher in energy. The simulations show that CT and TT account for roughly 86% of the conformations observed at the interface, with a strong preference for CT in the presence of solvent. The remaining conformers are predominantly TC.

Table 4-1. Conformational preferences of TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether at the interface. The percentage of selectors adopting the specific conformer is given.

Terminal group	Solvent	TT	TC	CT	TG(T)	TG(C)	Other
TMA	None	2	18	0	7	67	6
	water/methanol	43	30	0	2	25	0
	<i>n</i>-hexane/2-propanol	54	5	5	10	24	2
BOC	None	69	12	17	2	0	0
	water/methanol	14	1	85	0	0	0
	<i>n</i>-hexane/2-propanol	25	5	69	0	1	0

Two additional di-proline conformers, identified as TG(T) and TG(C), are also observed at the surface despite being absent for a single isolated selector (see Figure 4-3). TG(T) is similar to the TT conformer except that the inner amide torsion is roughly 210°, as opposed to 180° for TT. Similarly, for TG(C) the inner torsion is around 310°, as opposed to 340° for the TC conformer. These two conformers appear primarily for the TMA terminated selectors, only when the selectors are placed at the surface, and even in the absence of solvent. In fact over two-thirds

of the TMA terminated selectors adopt TG(C) when solvent is absent. The latter result signals that the TG(C) conformer results from interactions between the selectors and other surface components, such as the end-caps, silanol groups, and nearby selectors. The presence of solvent decreases the probability for TG(C), with roughly one-quarter of the TMA terminated selectors adopting this conformation in the presence of *n*-hexane/2-propanol or water/methanol.

Orientational probabilities for the carbonyl groups are provided in Figure 4-6. These distributions provide insight into steric hindrance that may result from the proximity of the carbonyls to the end-caps and other surface elements. Specifically, low angles correspond to a carbonyl group pointing towards the underlying surface. H-bonding to such a carbonyl may be difficult due to the presence of nearby surface elements. From Figure 4-6, the TMA terminated selectors have more variability in the orientation of the carbonyls. However, in *n*-hexane/2-propanol, the selectors clearly tend to orient O(10) towards the surface and O(8) and O(12) towards the bulk. Both selectors have similar carbonyl orientations except that the distributions are narrower for the BOC-terminated selector. In the more polar, water/methanol solvent, both selectors have a bimodal distribution for all three carbonyls with some probability of being directed at the surface and away from the surface. Again, the orientational preferences are more evident for the BOC-terminated selector.

Overall, the characteristics of the TMA terminated and BOC-terminated selectors at the interface differ in several ways. First, the distribution of the TMA terminated selectors is broader as demonstrated in Figure 4-5, where C(11) may be found between 5-12 Å above the Si layer. The analysis in Table 4-1 and the carbonyl orientations in Figure 4-6 also illustrate this conformational variability. In particular, the TG(T) and TG(C) conformers occur primarily at the TMA-(Pro)₂-N(CH₃)-tether interface. Second, the BOC-terminated selectors have a higher

probability of extending along the surface, particularly in the absence of solvent. Third, the TMA terminated selectors all prefer a *trans* amide bond joining the terminal TMA group to the remainder of the molecule but the amide between the rings may be either *cis* or *trans*. In contrast, the BOC-terminated selectors are consistently *trans* between the pyrrolidine rings but can adopt either *cis* or *trans* amides between the proline unit and the terminal BOC group.

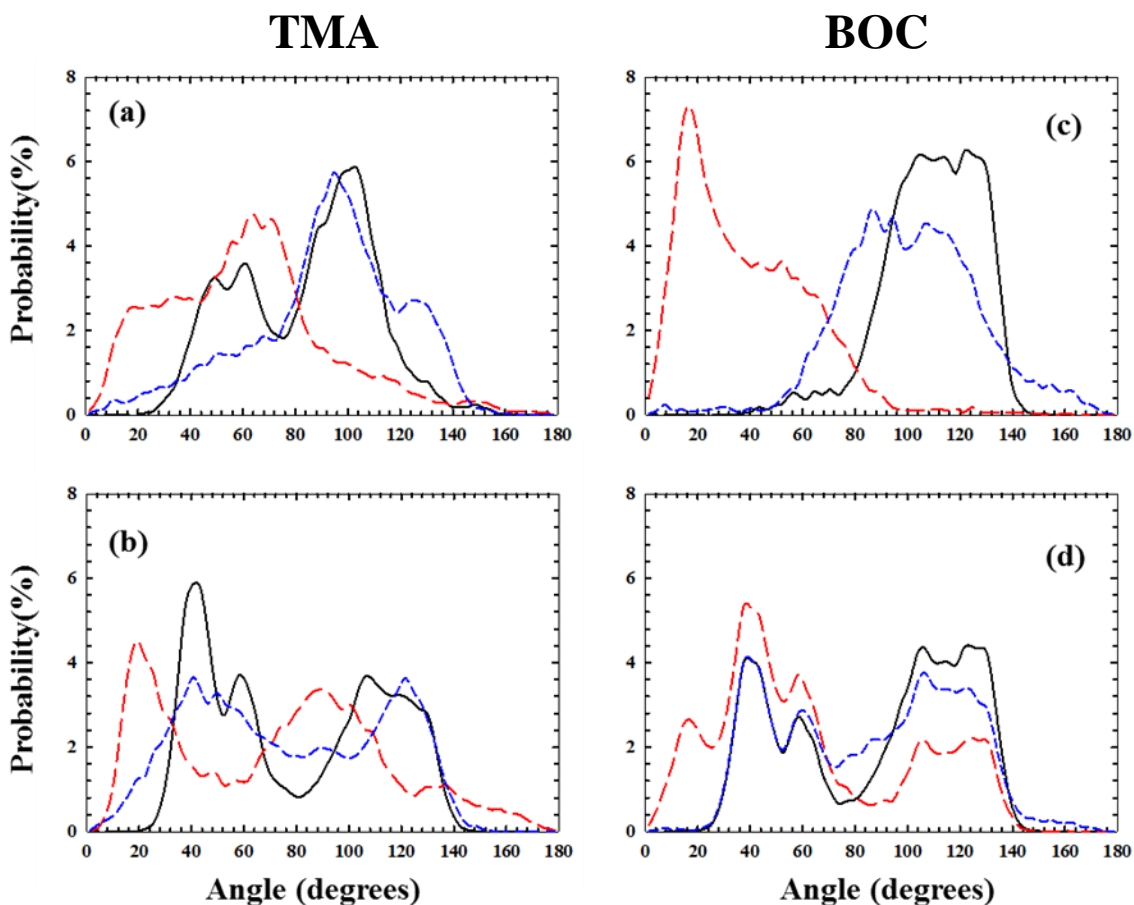


Figure 4-6. The distribution of carbonyl orientations at interface. Results are given for (a) *n*-hexane/2-propanol at the TMA-(Pro)₂-N(CH₃)-tether interface, (b) TMA-(Pro)₂-N(CH₃)-tether interface in water/methanol, (c) BOC-(Pro)₂-N(CH₃)-tether interface in *n*-hexane/2-propanol, and (d) BOC-(Pro)₂-N(CH₃)-tether interface in water/methanol. The percentage of carbonyls with the given angle relative to the Si layer is shown: An angle of zero degrees corresponds to the carbonyl pointing directly at the underlying Si layer. Distributions for O(8), O(10), and O(12) are represented by (black, solid), (red, long dashed), and (blue, short dashed) lines, respectively.

4.3.2 Solvation and Hydrogen Bonding at the TMA-(Pro)₂-N(CH₃)-tether Chiral Interface

Figure 4-7 provides surface distributions for the solvent at the TMA-(Pro)₂-N(CH₃)-tether interface. The distribution of *n*-hexane and 2-propanol are shown in Figure 4-7 (a). The alcohol distribution varies significantly at the interface, with a higher probability directly above the end-caps. At distances between 7 and 11 Å, where the bulky di-prolines and terminal groups are usually located, the alcohol concentration is lower but some *n*-hexane is found. 2-Propanol is again probable beyond 11 Å where hydrogen bonding to the upper amide and to other propanols can occur. Overall, *n*-hexane prefers the bulk to the interface.

The distribution of water and methanol above the TMA-(Pro)₂-N(CH₃)-tether interface is given in Figure 4-7(b). Both hydrogen-bonding solvents display similar distributions above the surface, with a higher density layer appearing at around 7 Å, followed by a low density region until around 10 Å. Beyond this distance, the solvents can hydrogen bond to an amide group or to other solvents. In the polar water/methanol solvent, the selectors have a greater propensity to adopt a TC conformation and this leads to a bent di-proline conformation (see Figure 3-2) that brings the proline units closer to the surface. Thus, the TC conformation excludes some solvent from the region directly above the end-caps but increases the possibility for hydrogen bonding further away from the silicon layer.

Snapshots were collected and analyzed to provide H-bonding statistics and our results are presented in Table 4-2 and Table 4-3. Table 4-2 reports the number of selectors, with a given configuration, that have between one and five simultaneous H-bonds. For instance, at any time, the probability of having a TMA terminated selector, in a TT conformation and with one H-bond, is 18.6% in *n*-hexane/2-propanol. Hydrogen bonding probability is reported as a function of the selector atom and backbone conformation in Table 4-3. From this table, a TMA terminated

selector in a TT conformation has a 4.0%, 8.3%, and 16.1% chance of having a H-bond at O(8), O(10), and O(12), respectively, in an *n*-hexane/2-propanol solvent. We do not report statistics for the nitrogen atoms in these tables, since H-bonding is almost negligible for these atoms. Table 4-3 also shows that H-bonding to the “extra” oxygen, O(13), in the BOC terminal group is very low. Thus, H-bonding occurs almost exclusively at the three carbonyl oxygens in the selectors.

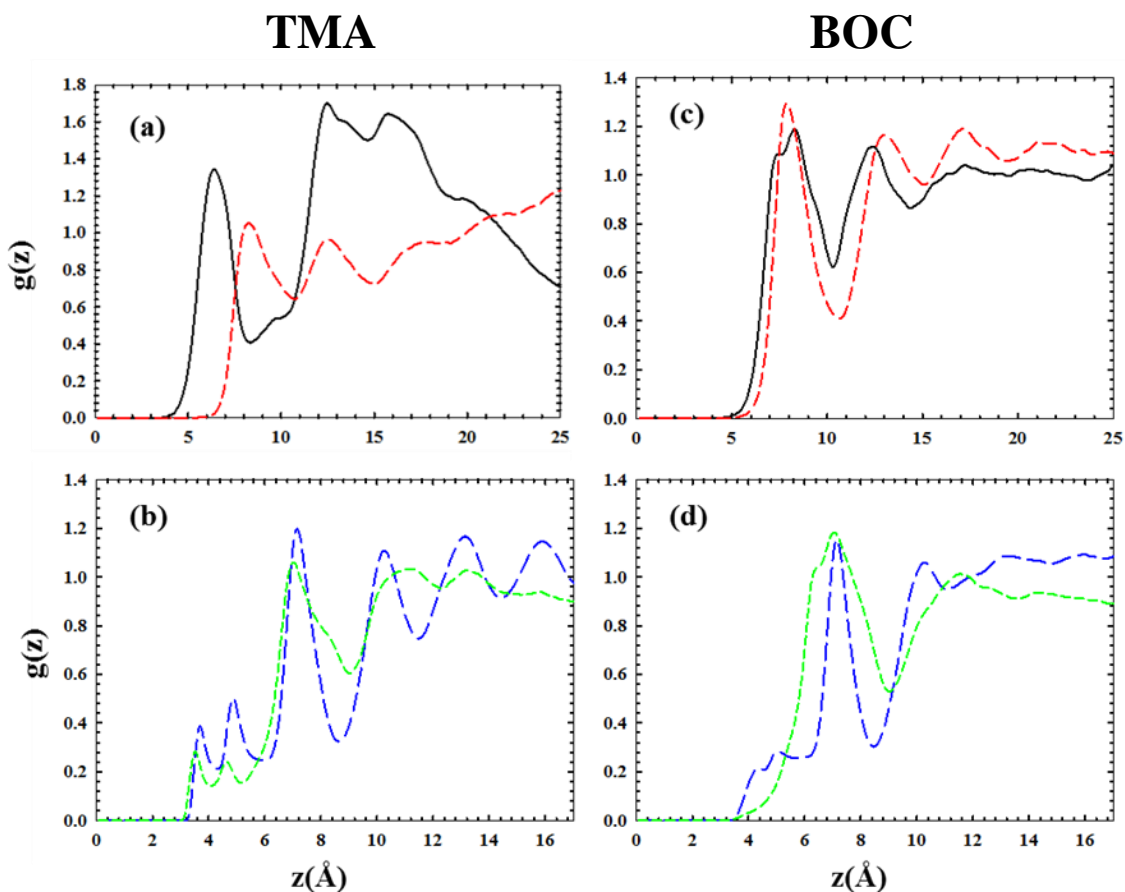


Figure 4-7. Solvent distributions above the TMA-(Pro)₂-N(CH₃)-tether and the BOC-(Pro)₂-N(CH₃)-tether interfaces. Results are shown for (a) hexane/2-propanol at the TMA-(Pro)₂-N(CH₃)-tether interface, (b) water/methanol at the TMA-(Pro)₂-N(CH₃)-tether interface, (c) *n*-hexane/2-propanol at the BOC-(Pro)₂-N(CH₃)-tether interface, and (d) water/methanol at the BOC-(Pro)₂-N(CH₃)-tether interface. The surface distribution for a central carbon of *n*-hexane (red, medium dashes), the O in 2-propanol (black, solid), the O in water (blue, long dashes), and the O in methanol (green, short dashes) are shown.

Table 4-2. Hydrogen bonding statistics for the TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether interfaces with *n*-hexane/2-propanol and water/methanol. The Percentage for one, or more, H-bonds as a function of conformation are given.

Terminal group	Solvent	Number of H-bonds	TT	TC	CT	TG(T)	TG(C)
TMA	water/methanol	1	14	11	0	1	10
		2	14	7	0	0	7
		3	7	3	0	0	2
		4	2	1	0	0	1
	<i>n</i> -hexane/2-propanol	1	19	1	2	3	5
		2	4	0	0	1	1
		3	1	0	0	0	0
BOC	water/methanol	1	5	1	26	0	0
		2	4	0	18	0	0
		3	2	0	8	0	0
		4	1	0	0	0	0
	<i>n</i> -hexane/2-propanol	1	8	2	16	1	1
		2	2	0	3	1	0
		3	0	0	0	0	0

Table 4-3. Hydrogen bonding statistics for the TMA-(Pro)₂-N(CH₃)-tether and BOC-(Pro)₂-N(CH₃)-tether interfaces with *n*-hexane/2-propanol and water/methanol. The H-bonding location and solvent identity, for water/methanol, are given. The percentage for a hydrogen bond at the specified carbon, in the given conformer, is presented. In water/methanol, the total probability is given with the probability for methanol given in brackets.

Terminal group	Solvent	H-bonding atom	TT	TC	CT	TG(T)	TG(C)
TMA	water/methanol	O(8)	23 (2)	6 (1)	0	1 (0)	3 (1)
		O(10)	19 (2)	20 (1)	0	2 (1)	15 (2)
		O(12)	32 (4)	13 (2)	0	1 (1)	15 (2)
	<i>n</i> -hexane/2-propanol	O(8)	4	1	1	1	1
		O(10)	8	1	1	1	2
		O(12)	16	1	1	3	4
BOC	water/methanol	O(8)	9 (1)	1 (0)	56 (7)	0	0
		O(10)	4(1)	1 (0)	19 (3)	0	0
		O(12)	8 (1)	1 (0)	40 (6)	0	0
		O(13)	1 (0)	1	1 (0)	0	0
	<i>n</i> -hexane/2-propanol	O(8)	4	2	12	1	0
		O(10)	2	1	2	1	0
		O(12)	7	1	8	1	0
		O(13)	1	0	1	0	0

As expected, water and methanol hydrogen bond extensively with the TMA terminated selector. Our simulations indicate that, at any time, roughly four-fifths of the selectors have at least one H-bond to solvent and over 45% have two-or-more H-bonds. Water accounts for over 87% of the H-bonds, as shown in Table 4-3 and this is due to three factors. First, the mixed solvent is 70% water. Second, water is smaller than methanol and better able to access the carbonyl oxygens in bent conformers. Finally, water can donate two H-bonds whereas methanol can only donate one. Table 4-2 shows that H-bonding is most probable for the TT conformer but the TC and TG(C) also contribute to H-bonding in this solvent. All three carbonyl oxygens have a significant probability to H-bond for the TT conformer. In water/methanol, the TC and TG(C) conformers also form H-bonds although O(8) is difficult to access for these conformers and, as shown in Table 4-3, has a lower probability to H-bond. 2D solvent distributions, $g(r,z)$, are shown in Figure 4-8 and Figure 4-9 for *n*-hexane/2-propanol and water/methanol, at the TMA-(Pro)₂-N(CH₃)-tether interface. In *n*-hexane/2-propanol, roughly one-third of all TMA-(Pro)₂-N(CH₃)-tether selectors have at least one hydrogen bond at any time. Only 6% of the selectors have more than one H-bond. This H-bonding occurs predominantly for TT conformers despite the fact that TC and TG(C) conformers account for roughly 30% of the observed selector conformers. The latter conformers are more compact and, for steric reasons, 2-propanol cannot access the carbonyl oxygens. The probabilities in Table 4-3 show that O(12), the oxygen nearest the terminal group, has the highest chances for H-bonding following by O(10), from the amide between the two rings.

Figures 4-8 and 4-9 show the distribution between the solvent H and the carbonyl O from the selector. The ridge at roughly $r^2 + z^2 \approx 2 \text{ \AA}$ corresponds to the H-bonding region. For O(8), the region is narrow with the H required to be equidistant ($z \approx 0 \text{ \AA}$) from the underlying Si layer.

For O(10), the solvent H tends to be somewhat closer to the surface indicating that H-bonding to O(10) occurs primarily “from below”. This is consistent with Figure 4-6(a) where C=O(10) is most often pointed towards the surface. The uppermost carbonyl oxygen, O(12), tends to H-bond with solvent located further from the Si layer ($z > 0 \text{ \AA}$). As shown in Figure 4-6, this oxygen is usually directed towards the bulk. The distribution for water/methanol, provided in Figure 4-9 shows that the H-bonding solvents have a bi-modal H-bonding distribution, with a peak for solvent located closer to the surface and a second peak, or a shoulder, for solvent located further away. The carbonyl orientational distributions in Figure 4-6 are also bimodal. Thus, the H-bonding solvent distribution is a direct result of the carbonyl orientation at the surface. It is interesting to note that the alcohols both have a depletion region beyond the H-bonding ridge. In contrast, secondary water-water interactions are evident in the 2D distributions with ridges at $r^2 + z^2 \approx 4 \text{ \AA}$ evident.

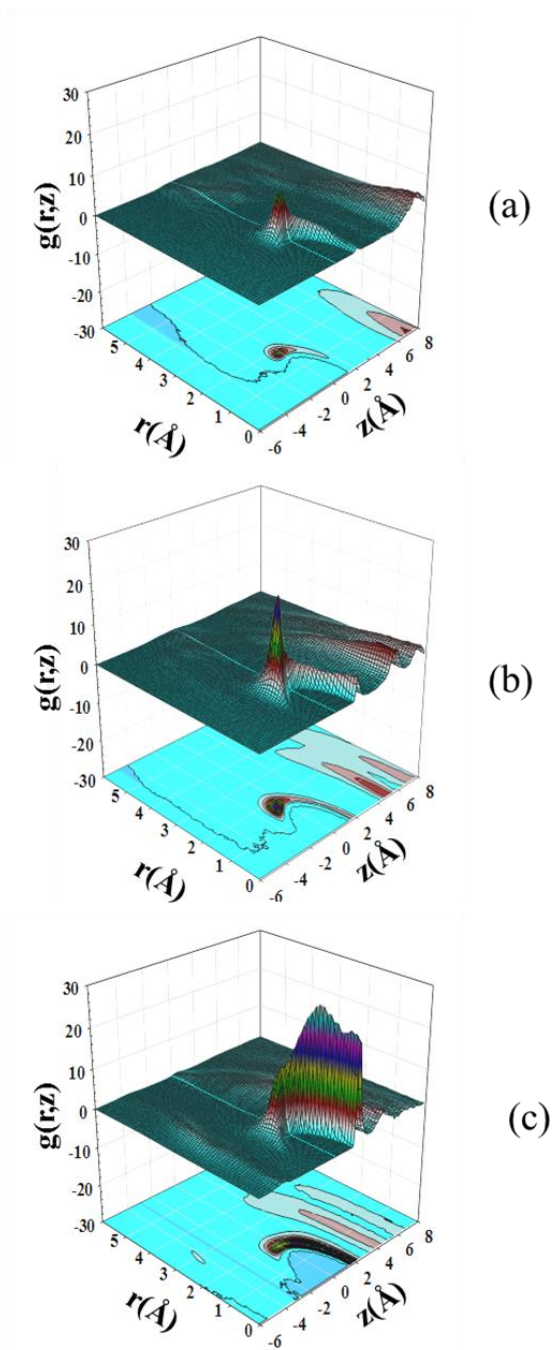


Figure 4-8. The 2D distribution between the hydrogen-bonding H of 2-propanol and the carbonyl oxygens of TMA-(Pro)₂-N(CH₃)-tether for the *n*-hexane/2-propanol solvated interface. Results are given for O(8), O(10), and O(12) in (a), (b), and (c) respectively. Contour plots are also provided to highlight the locations of the features.

4.3.3 Solvation and Hydrogen Bonding at the BOC-(Pro)₂-N(CH₃)-tether Chiral Interface

Figure 4-7(c-d) provide surface distributions for *n*-hexane/2-propanol and water/methanol at the BOC-(Pro)₂-N(CH₃)-tether interface. These selectors strongly prefer to adopt either CT or TT conformations, leading to a consistently *trans* amide torsion between the rings and, overall, the selectors are similarly oriented at the surface. For these two reasons, the solvent distributions are simpler and less variable for the BOC-terminated selectors. Overall, *n*-hexane and 2-propanol are more concentrated between 7 and 9 Å above the Si layer, while water and methanol are found preferentially between 6 and 8 Å. As shown in Figure 4-5, the selectors are more compact in water/methanol and this accounts for the shift to smaller distances. Both solvents show a depletion region at distances where the selector atoms are most often located. Directly beyond these distances, a higher solvent concentration is usually observed, but bulk distributions are quickly attained.

In the water/methanol solvent, water accounts for over 87% of the H-bonding to BOC terminated selector. Table 4-2 presents the probability of finding one or more H-bonds for each conformer. Unlike the TMA terminated selectors, H-bonding occurs most often for CT conformers. In water/methanol, there is a 26.2% chance of finding a CT conformer with one H-bond, and a 17.9% chance of finding this conformer with two H-bonds. From Table 4-3 all three carbonyl oxygens form H-bonds but the inter-ring oxygen, O(10), forms far fewer than O(8) and O(13).

2D solvent distributions, $g(r,z)$, are shown in Figure 4-10 and Figure 4-11 for *n*-hexane/2-propanol and water/methanol, respectively, at the BOC-(Pro)₂-N(CH₃)-tether interface. 2-Propanol H-bonds preferentially to O(8) and O(12) and this is evident from the peak intensity in Figure 4-10(b). The inter-ring carbonyl points towards the surface, as shown in Figure 4-6(c),

and the 2-propanol cannot be accommodated between the carbonyl and the surface. As well, H-bonding to O(8) occurs either from the side or from above. This is in contrast to the TMA terminated selectors where H-bonding occurs only from the side. The 2D solvent distribution in water/methanol shows a bimodal distribution for O(8) and O(12) but only weak H-bonding to O(10). In this solvent, over 85% of the selectors adopt a bent, CT, conformer with O(10) typically pointing towards the surface. With these two factors, even water cannot H-bond effectively to O(10).

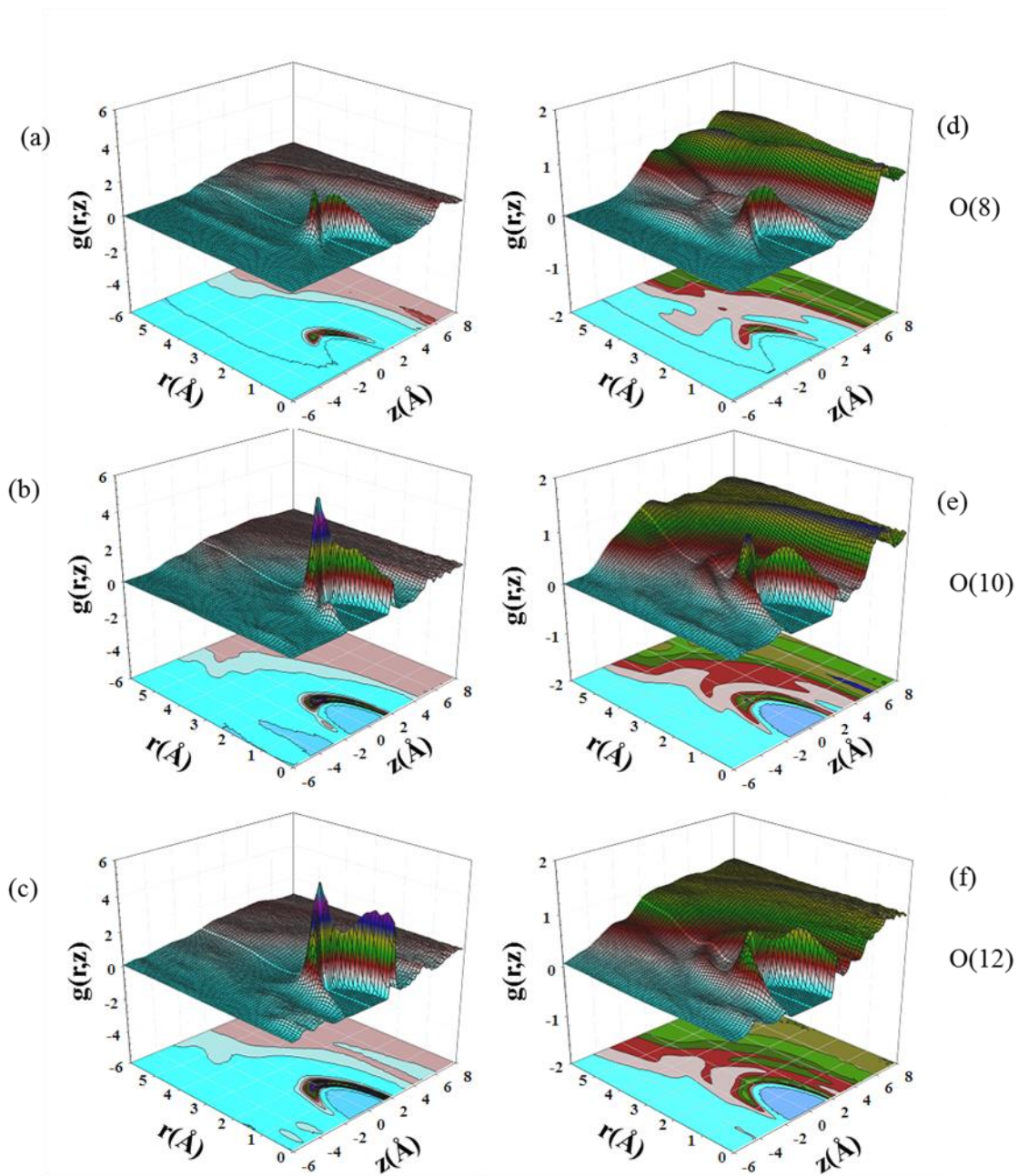


Figure 4-9. The 2D distribution between the hydrogen-bonding H of methanol (a-c) and of water (d-f) and the carbonyl oxygens of TMA-(Pro)₂-N(CH₃)-tether for the water/methanol solvated interface. Results are given for O(8) in (a) and (d), O(10) in (b) and (e), and O(12) in (c) and (f). Contour plots are also provided.

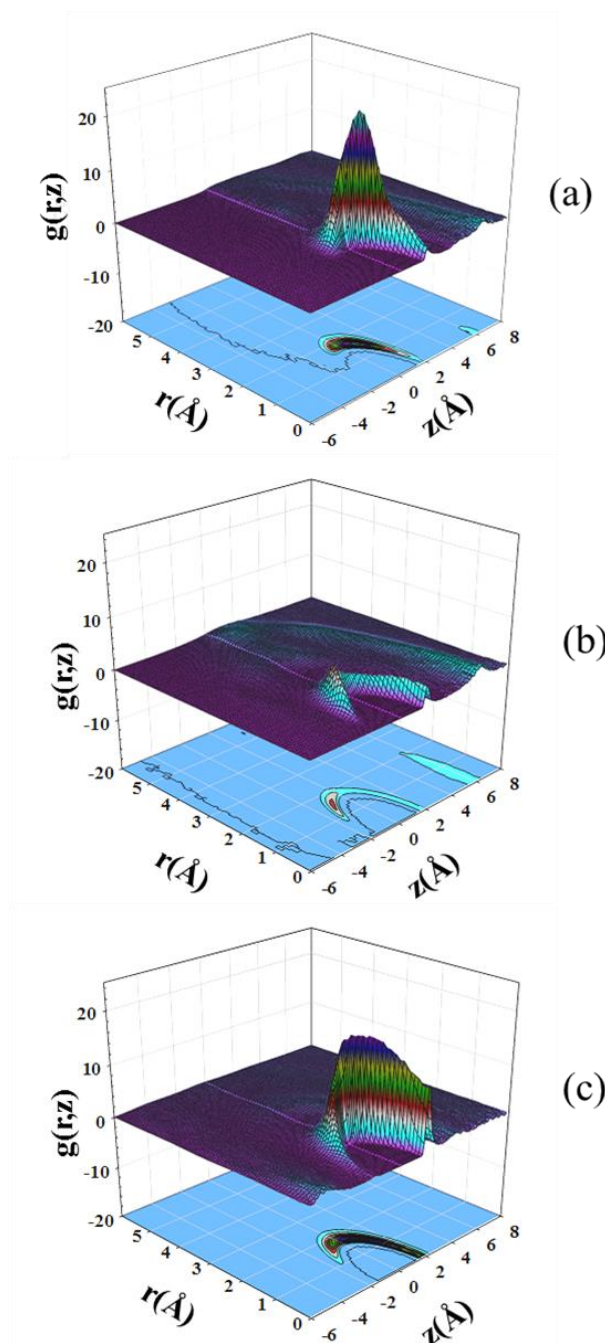


Figure 4-10. The 2D distribution between the hydrogen-bonding H of 2-propanol and the carbonyl oxygens of BOC-(Pro)₂-N(CH₃)-tether for the *n*-hexane/2-propanol solvated interface. Results are given for O(8), O(10), and O(12) in (a), (b), and (c) respectively. Contour plots are also provided to highlight the locations of the features.

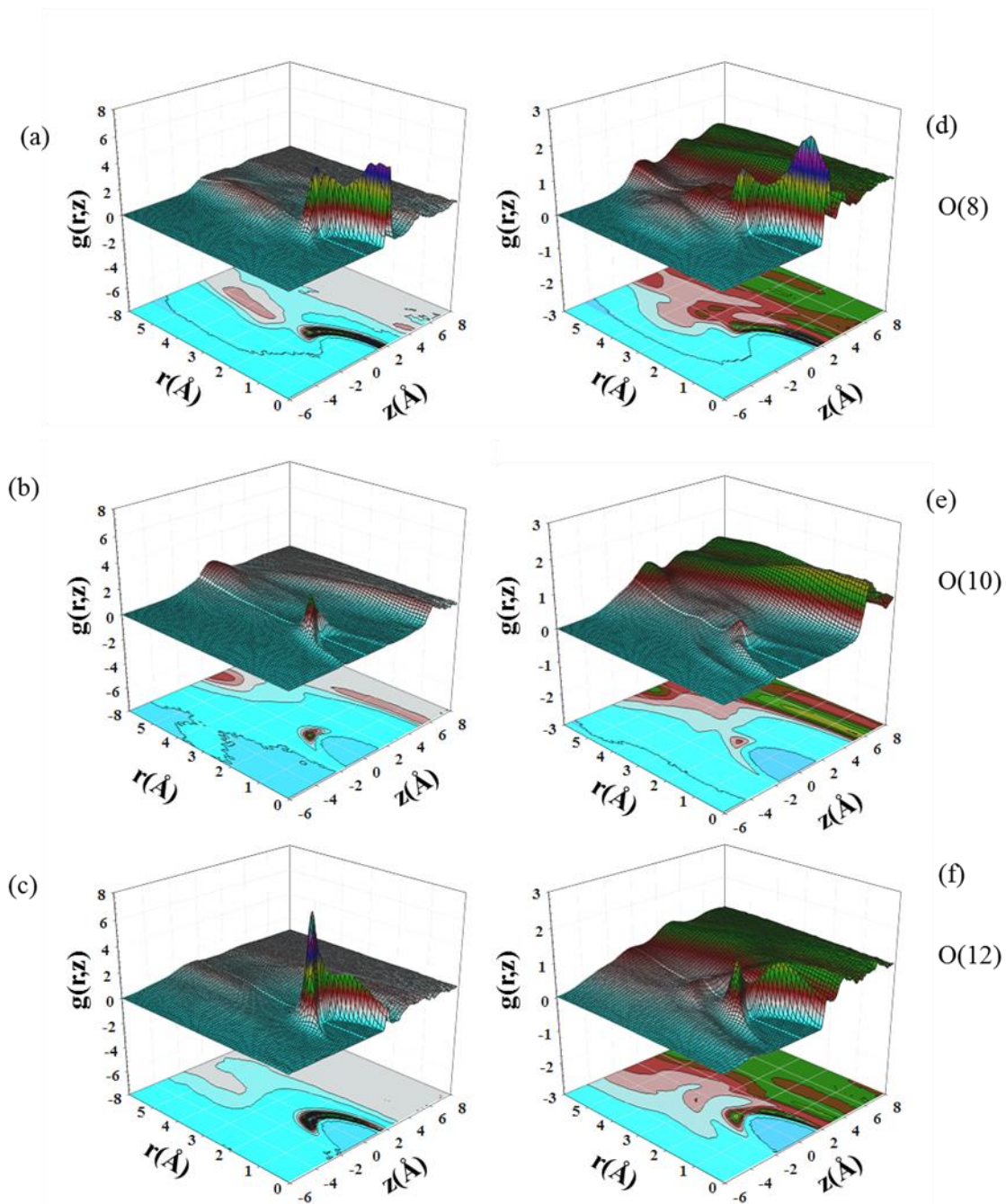


Figure 4-11. The 2D distribution between the hydrogen-bonding H of methanol (a-c) and of water (d-f) and the carbonyl oxygens of BOC-(Pro)₂-N(CH₃)-tether for the water/methanol solvated interface. Results are given for O(8) in (a) and (d), O(10) in (b) and (e), and O(12) in (c) and (f). Contour plots are also provided.

4.4 Conclusions

Molecular force fields have been developed for BOC-(Pro)₂-N(CH₃)₂ and TMA-(Pro)₂-N(CH₃)₂. The force fields have been adjusted to reproduce conformational energies from single point CCSD/6-311G(d,p) calculations.

Each chiral selector forms the basis of an interface that also includes trimethylsilyl end-caps and silanol groups. Simulations of these interfaces have been performed in the presence of a vacuum, a 70/30 v/v water/methanol, and a 70/30 v/v *n*-hexane/2-propanol. Simulations in the absence of solvent identify the conformational impact of the surface environment. For the TMA terminated selectors, surface interactions introduce two new conformations, TG(T) and TG(C), which are similar to TT and TC, but with a roughly 30° change in the inter-ring amide torsional angle. New surface-induced conformations do not appear for BOC terminated selectors. Overall, we found that the TMA terminated selectors display more conformational flexibility at the interface.

Solvent alters the conformational distributions of the selectors and introduces solvent-selector H-bonding. Virtually all of the H-bonding occurs at the carbonyl oxygens: Very few H-bonding events occur at the nitrogens or at the “extra” oxygen in the BOC terminal group. We find that roughly one-third of selectors have a hydrogen bond in *n*-hexane/2-propanol but roughly four-fifths H-bond in water/methanol. In the latter solvent, most of the H-bonding is between the selector and water.

For TMA terminated selectors, H-bonding is most important for the TT conformer and all three backbone carbonyl oxygens form H-bonds with solvent. BOC-terminated selectors, in the presence of solvent, prefer the CT conformer and H-bonding occurs predominantly at O(8) and

O(12). The inter-ring oxygen, O(10), is often directed towards the surface which sterically restricts H-bonding solvent.

Our results show that TMA-(Pro)₂-N(CH₃)-tether prefers a TT conformer but also displays a number of other conformers at the interface. This selector is most likely to hydrogen bond at O(12). BOC-(Pro)₂-N(CH₃)-tether, in contrast, preferentially adopts a bent CT conformer at the interface and is most likely to hydrogen bond at O(8). The bent shape of the latter, along with the overall surface-selector proximity which inhibits hydrogen bonding at O(10), would not be conducive to chiral selectivity. Thus, given the different conformational and H-bonding preferences of the two selectors, it is not surprising that the selectivity depends strongly on the nature of the terminal group. Experiments [94] show that the TMA terminated selector is superior and our simulations suggest that this is likely due to the greater accessibility of the selector to analyte. Specifically, the TT conformer of the TMA terminated selector is extended thereby providing more interaction opportunities, particularly for hydrogen-bonding, between TMA-(Pro)₂-N(CH₃)-tether and a chiral analyte.

Chapter 5

Solvation Studies of Tri-, Tetra-, Penta-, and Hexa-proline Chiral Stationary Phases

5.1 Introduction

In this chapter, longer proline chains are examined as the selector. We expect that increasing the conformational flexibility from di-proline to tetra-proline and also reduction of it from tetra- to hexa-proline will affect the solvation properties. As well, with increasing chain length, it is expected that helical conformers should become gradually dominated and it should affect the solvation and selectivity properties. Huang *et al.* [47] found that 31 out of 53 analytes were separated by TMA terminated di-, tetra-, hexa-, or deca-proline but 7 analytes were not resolved by any of the four selectors. The remaining 15 analytes were only resolved on some of the four stationary phases. Interestingly, di-proline and deca-proline resolved ten of the fifteen analytes but tetra-proline and hexa-proline resolved only five and four, respectively. Focusing on the ten analytes resolved by di-proline, five are also resolved by deca-proline but tetra-proline and hexa-proline resolve only one of the ten. All of these results point to a fundamental change in the characteristics of the interface with chain length.

Similar to Chapter 4, poly-proline conformers are identified by appeal to the correspondence between the number of proline residues and the number of amide linkages. The terminal TMA group and the topmost proline are joined by an amide, the proline moieties are joined by an amide, and an amide joins the tether to the poly-proline. In Figure 5-1(a), these torsions are defined by C(15)-C(13)-N(35)-C(31), C(31)-C(11)-N(29)-C(25), C(25)-C(9)-N(23)-C(19) and C(19)-C(7)-N(5)-C(4), respectively, for TMA-(Pro)₃-N(CH₃)-tether.

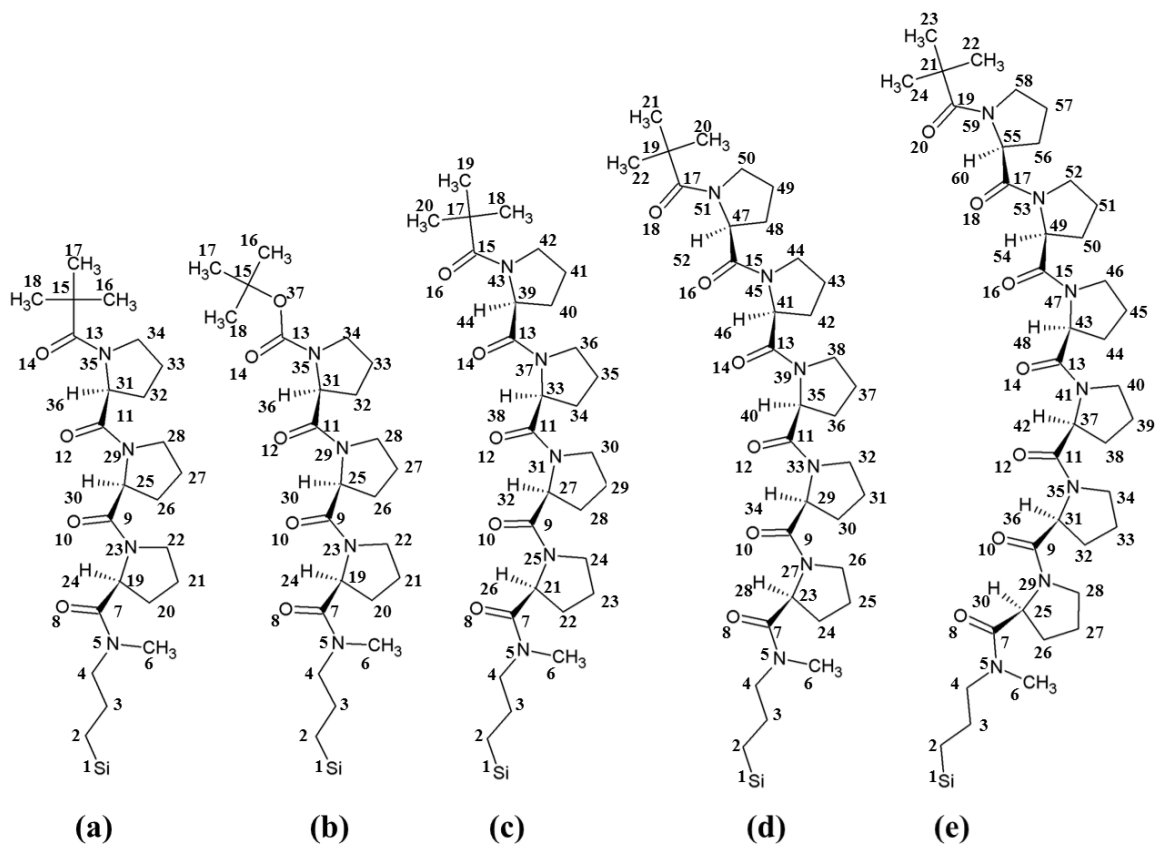


Figure 5-1. Poly-proline selectors: (a) TMA-(Pro)₃-N(CH₃)-tether; (b) BOC-(Pro)₃-N(CH₃)-tether; (c) TMA-(Pro)₄-N(CH₃)-tether; (d) TMA-(Pro)₅-N(CH₃)-tether; and (e) TMA-(Pro)₆-N(CH₃)-tether. The atom numbering shown above is used in this thesis. We represent methyl and methylene groups with a single number since a united-atom representation is used for these groups.

The number of possible conformers clearly grows with chain length, but the number of energetically accessible conformers is more relevant to interfacial properties. We focus on conformers that are within 10 kJ/mol of the global energy minimum. These are expected to dominate the population of conformers observed at the interface.

5.1.1 Force Field Development

Independent force fields have been developed for BOC-(Pro)₃-N(CH₃)₂ and TMA-(Pro)_n-N(CH₃)₂, n=3-6, poly-prolines. Each of the five peptides has a distinct conformational profile and a proper representation of the corresponding chiral interfaces requires individual force fields. Model development follows a multistep process as discussed in Section 4.1.2. All force field parameters are provided in Appendix A.

The B3LYP/6-311G(d,p) energies are not directly employed in the torsional potential fitting, in contrast to the fitting procedure for stretches, bends, and improper torsions. For twisting motions, several correction factors are required. First, B3LYP does not reproduce the correct relative energies of the conformers. To correct for this, for each scan over amide torsions, we introduce a scale factor so that the LPNO-CCSD (RI-MP2 for hexa-proline) relative conformer energies are reproduced but the barriers to inter-conversion are left largely unchanged. A similar scaling was previously adopted for di-proline in Chapter 4. Second, atoms that are separated by four-or-more bonds interact only via non-bonding terms. Torsional motion can lead to significant changes in the distance between atoms and, in particular, certain structures have high energies due to steric repulsion. The non-bonding terms account for this repulsion and, as a result, the torsional and nonbonding potentials are intertwined and must be optimized simultaneously.

The minimum energy structures predicted from the force fields are obtained similar to the procedure for di-proline. These structures are directly compared with the *ab initio* structures as shown for the penta-proline selector in Figure 5-2. The final intramolecular potentials are chosen based on the overall agreement with LPNO-CCSD (RI-MP2) structures and energies with an emphasis placed on conformers within 10 kJ/mol of the global minimum.

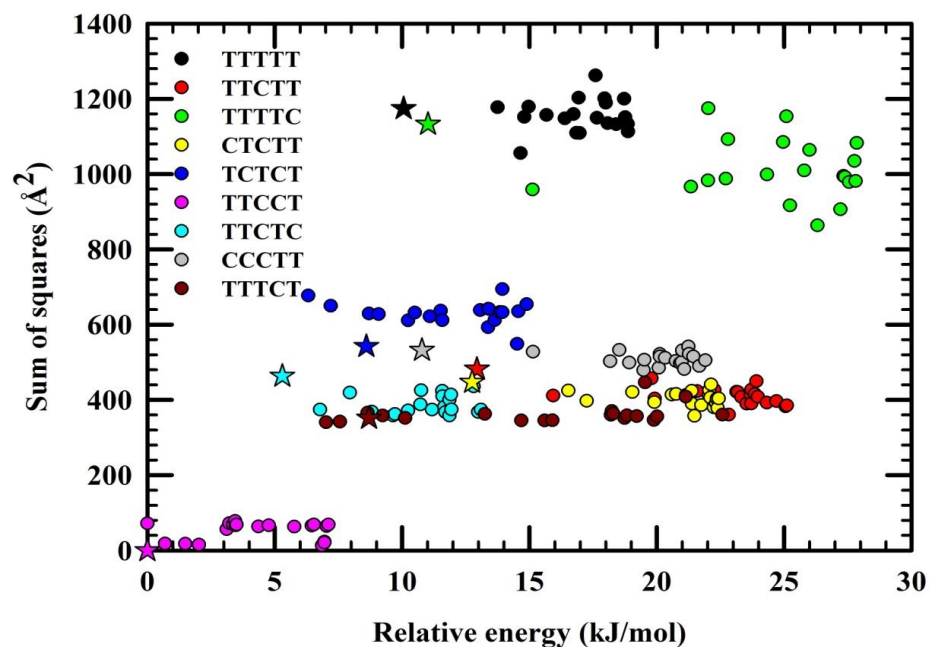


Figure 5-2. A comparison between the force field for TMA-(Pro)₅-N(CH₃)₂ and the results of LPNO-CCSD calculations. The relative conformer energies and structure (as assessed by a sum of squares over atomic positions, relative the lowest energy conformer) are shown. The filled stars identify the results from *ab initio* calculations. Each circle represents the lowest-energy structure from a single-molecule molecular dynamics simulation. Different conformers are indicated by color coding. The proximity between a sequence of circles and the corresponding star (yellow for CTCTT for example) indicates the level of agreement between the force field and the LPNO-CCSD results.

Scanning each torsion becomes impractical for the penta-proline and hexa-proline. Further, the backbone torsion energies are qualitatively similar for the shorter peptides. With this in mind, the torsional potentials for the penta-proline and the hexa-proline are adopted from the tetra-proline but re-scaled according to the LPNO-CCSD (RI-MP2 for hexa-proline) conformer energies. We find that the backbone torsions are all characterised by large energy barriers (at least 50 kJ/mol) and that they are qualitatively similar for di-proline to tetra-proline. The treatment of the pyrrolidine rings in the model fitting process was more difficult than expected.

For tri-proline and tetra-proline, the application of the scaling factors C and D could differ for groups within the pyrrolidine unit. That is, we found that an atomic representation of the ring was adequate and we were able to arrive at an intramolecular potential in reasonable agreement with LPNO-CCSD conformers. Despite considerable effort, this approach did not lead to a satisfactory model for penta-proline and hexa-proline: none of the parameter sets produced the right balance between structure and energy.

We revisited the issue of scaling factors for the pyrrolidine rings for the longer peptides. Given that the rings are rigid, we opted to consider the rings as a single unit in terms of the scaling factors. Here, all ring atoms adopt a single scaling factor. This greatly improved the quality of the intramolecular force field for the longer peptides. The relative energies from the intramolecular potentials and from LPNO-CCSD (RI-MP2) calculations are compared in Figure 5-3. The agreement between the force field and LPNO-CCSD (RI-MP2) is excellent. Before discussing the chiral interface, we note that the terminal groups, BOC and TMA, are considered part of the backbone and their force fields are developed for each poly-proline moiety. However, the *ab initio* calculations were performed on truncated selectors and the propyl tethers are not included in the force field fitting. Rather, the tether intramolecular potential is taken from literature [227-229].

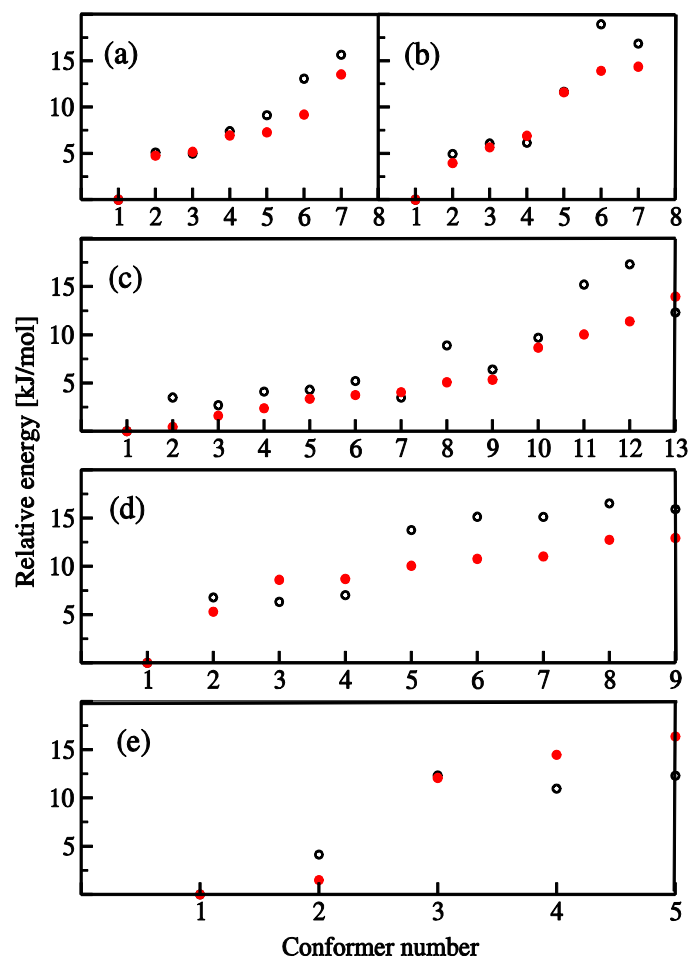


Figure 5-3. Assessment of the intramolecular potential for a) TMA-tri-proline, b) BOC-tri-proline, c) TMA-tetra-proline, d) TMA-penta-proline and e) TMA-hexa-proline. The open black circles show the relative energy of the force field conformers while the open red circles show the corresponding relative LPNO-CCSD (RI/MP2) energies. Conformers having relative energies of more than 20 kJ/mol are not shown.

5.2 The Chiral Interface

A snapshot of the TMA-(Pro)₅-N(CH₃)-tether interface with a water/methanol is shown in Figure 5-4. The simulation cell is similar to di-proline CSP. The total energy of the interfacial system is given by Eq. 4-3.

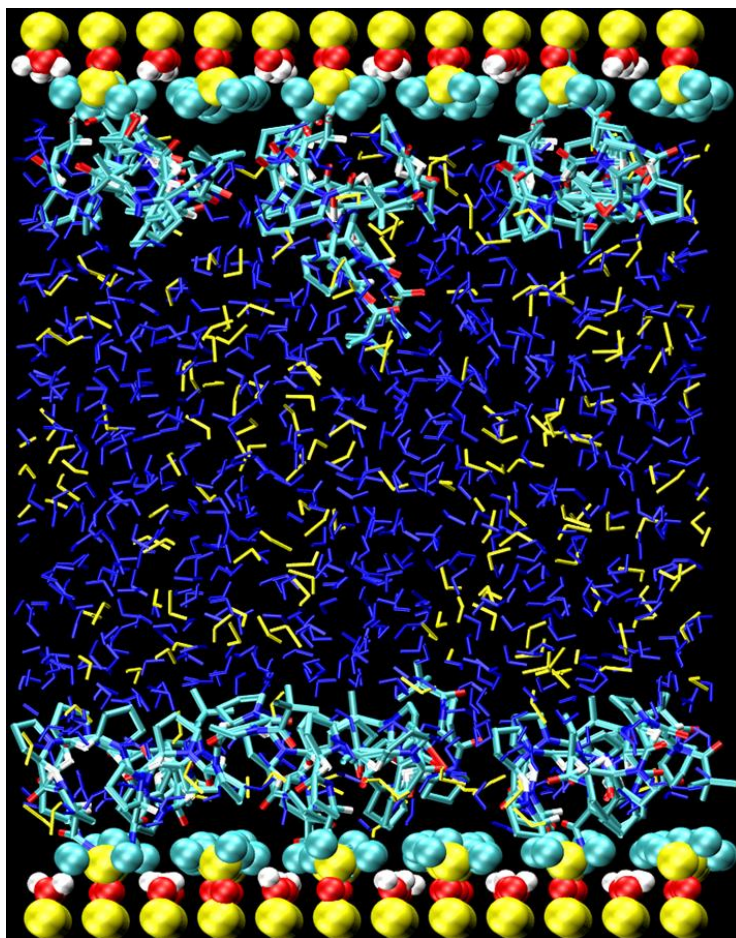


Figure 5-4. A snapshot of the TMA-(Pro)₅-N(CH₃)-tether interface with a water/methanol solvent. Silicon, carbon, oxygen, nitrogen and hydrogen atoms are shown in yellow, light blue, red, dark blue, and white, respectively. Water molecules are in blue and methanol molecules are yellow. The surfaces include silanol groups, trimethylsilyl end-caps, and the underlying silicon atoms.

5.3 Simulation Details

The distance between interfaces varies from 66.5 to 79.6 Å for *n*-hexane/2-propanol solvent and from 35.2 to 60.3 Å for water/methanol solvent. For all simulations involving *n*-hexane/2-propanol, 169 2-propanol and 231 *n*-hexane molecules are included in the simulation cell. For the

denser water/methanol, the number of molecules varies depending on the chiral selector: For tri-proline to hexa-proline, 645 to 806 water molecules and 155 to 266 methanol molecules are included in the cell.

Each distribution function is an average over ten independent 1.5 ns simulations, for a total simulation length of around 15 ns. The time step is 0.3 fs. The first 800 ps of each simulation constitute the equilibration period and the remaining 700 ps is the collection period.

All poly-proline selectors are initially placed on the surface in their lowest energy conformation. Several conformers are energetically close to the global energy minimum and may have a significant population at the surface. A high energy barrier hampers the transition between conformers and, on the time scale of the simulations, conformational equilibration will not occur without additional steps to expedite the inter-conversion. We allow *cis-trans* inter-conversion by temporarily lowering the torsion barrier for the amide and the adjacent C-C bond. Barrier lowering is implemented via a multiplicative scale function in the torsional potential

$$U^{torsion}(\varphi_t, \tau) = S(\varphi_t, \tau) \sum_{i=0}^6 c_{it} \left(\cos(\varphi_t + \varphi_t^{i;0}) \right)^i, \quad 5-2$$

here

$$S(\varphi_t, \tau) = \frac{(\tau - \tau_s) + (\tau_f - \tau) (\cos\varphi_t \times \cos 2\varphi_t)^2}{(\tau_f - \tau_s)}, \quad 5-3$$

and $\tau_s = 3.0 \times 10^{-5}$ ps and $\tau_f = 600$ ps. The scaling function lowers the energy barriers at $\tau = \tau_s$ and gradually increases then back to its normal value at $\tau = \tau_f$. For $\tau \geq \tau_f$ the scaling function is unity. Note that the relative energies of *cis* and *trans* isomers do not change during the scaling as $S(\varphi_t, \tau)$ is one when $\varphi_t = 0^\circ$ (*cis*) or 180° (*trans*). For this reason, the correct relative population is expected to be established even when the torsional barrier is lowered. A 600 ps

scaling-in period proved to be sufficient for establishing conformational equilibrium at the di-proline interface [230]. We have confirmed that this is also the case for the longer prolines by performing very long, 6 ns simulations for the hexa-proline interface where the torsions are scaled in over a 5 ns period. The 600 ps and 5 ns scaling-in periods led to conformer populations that agreed to within 3%. Note that the torsional energy barrier is scaled for all amide linkages and their corresponding coupled dihedral.

5.4 Simulation Results

In this section, the results of MD simulations of the TMA-(Pro)_n-N(CH₃)-tether, n=3-6, and BOC-(Pro)₃-N(CH₃)-tether interfaces with *n*-hexane/2-propanol and water/methanol solvents are reported. The effect of solvent, terminal group and chain length on the characteristics of the poly-proline interface is discussed in detail.

5.4.1 Interfacial Poly-proline

Experiments indicate [69,175,176,195,213,217, 232-236] that the conformational preferences of poly-proline are solvent dependent. These experiments are not specific to surface-bound peptides but a computational study of the role of surface environment for di-proline in Chapter 4 showed that interface crowding and solvent both alter the conformational balance. The combined impact of surface and solvent on the poly-proline interface is examined in this section. Specifically, we examine the role of solvent on the conformations of surface-bound TMA-(Pro)_n-N(CH₃)-tether with n=3-6 selectors.

In the gas phase, the tri-proline has six conformers within 10 kJ/mol, the tetra-proline has ten, the penta-proline has four, and the hexa-proline has two energetically competitive conformers. The average conformer populations at the interface are reported in Table 5-1 for conformers that account for at least 5% of the selector population. We find that solvent, either

normal phase or reverse phase, leads to a reduction in the number of conformers relative to the gas phase. Comparing both solvents, more conformers are observed in the presence of *n*-hexane/2-propanol than for the more polar water/methanol. The most stable conformer remains the same for the gas phase and the solvent phase with the exception of tetra-proline where the all *trans* (TTTT) conformer is slightly more probable than TTCT in water/methanol. Other higher energy conformers often appear in the gas phase and at the surface, but some are stabilized by the interface and others are destabilized. Consider the tri-proline where three conformers account for more than 90% of the selectors observed at the interface. The two lowest-energy conformers, TTC and TTT, are abundant as expected from *ab initio* results (Figure 3-3) but 8-9% of the conformers are CTC. *Ab initio* calculations alone predict CTT and TCT in higher abundance than CTC. The surface environment and interactions with solvent alter the conformational balance such that CTC is observed at the interface and CTT and TCT are not. For di-prolines, the interfacial environment was sufficient to introduce new *gauche* conformers but we did not find *gauche* amides for the tri-proline or the longer peptides.

The tetra-proline is unique among the poly-prolines because of the large number of energetically competitive conformers predicted in the gas phase. In water/methanol, we find that 90% of the selectors at the interface adopt one of three conformers. These have 3-4 *trans* amides and the most abundant conformer, the all *trans* TTTT, is not the global energy minimum predicted for the gas phase. In the normal phase solvent, *n*-hexane/2-propanol, a wider range of conformers is observed at the surface: six conformers appear. Two of these, TCCC and TCCT, have very little contribution in the gas phase but, together, they account for 15% of the selectors at the surface.

Table 5-1. Average conformer populations at the interface. Only conformers with relative population of more than 5% are listed.

Selector	Solvent	Conformer	Relative population (%)
TMA-tri-proline	<i>n</i> -hexane/ 2-propanol	TTC	42 ± 7
		TTT	38 ± 8
		CTC	9 ± 5
	water/ methanol	TTC	66 ± 6
		TTT	22 ± 4
		CTC	8 ± 3
BOC-tri-proline	<i>n</i> -hexane/ 2-propanol	CTC	54 ± 5
		TTC	25 ± 5
		CTT	10 ± 2
	water/ methanol	CTC	52 ± 8
		TTC	20 ± 4
		CTT	12 ± 4
		CCC	10 ± 1
		CCT	6 ± 2
TMA-tetra-proline	<i>n</i> -hexane/ 2-propanol	TTCT	33 ± 4
		TCTT	31 ± 5
		TTTT	16 ± 3
		TCCC	9 ± 2
		TCCT	6 ± 1
		TTTC	6 ± 1
	water/ methanol	TTTT	38 ± 6
		TTCT	35 ± 6
		TTTC	17 ± 5
TMA-penta-proline	<i>n</i> -hexane/ 2-propanol	TTCCT	48 ± 7
		TTCTC	34 ± 4
		TTCTT	11 ± 2
	water/ methanol	TTCCT	91 ± 5
		TTCCC	6 ± 2
TMA-hexa-proline	<i>n</i> -hexane/ 2-propanol	CTTCCT	86 ± 6
	water/ methanol	CTTCCT	90 ± 7

The tetra-proline results indicate that the selector environment is sufficient to alter the conformation balance, with some predicted conformers absent from the interface and unexpected conformers appearing.

The penta-proline and hexa-proline display only a few conformers at the interface. As with the shorter proline chains, the reverse phase solvent shifts the conformational balance toward a smaller number of competing structures: for the penta-proline and the hexa-proline in the presence of water and methanol, one conformer accounts for over 90% of all the chiral selectors.

Figure 5-5 and Figure 5-6 show the surface distribution of backbone carbons above the Si layer. All the selectors, with the exception of tetra-proline, tend to lie closer to the surface in water/methanol. This tendency increases with chain length: for the tri-proline, C(7) is roughly 0.5 Å closer to the surface in water/methanol but this difference increases to 1.2 Å for the hexa-proline. The backbone distributions are broader in the normal phase solvent, as expected since more poly-proline conformers are present at the interface with the less polar solvent (Table 5-1). For instance, the topmost carbonyl carbon, C(13), of the tri-proline is located between 6 and 12 Å from the Si layer when the solvent is water/methanol but, in the presence of *n*-hexane/2-propanol, C(13) can be found 7-16 Å away from the Si layer. Solvent impact is less pronounced for the penta-proline and hexa-proline, where the most important change is the overall shift of the poly-proline closer to the surface in water/methanol. Returning to the tetra-proline, the outermost carbonyls, C(13) and C(15), extend further into the bulk in water/methanol due to the presence of the TTTT conformer at the interface. The topmost carbonyl carbon, C(15), can be over 18 Å from the Si layer. The distribution of backbone positions is generally broad for the tetra-proline due to the larger number of conformers at the surface.

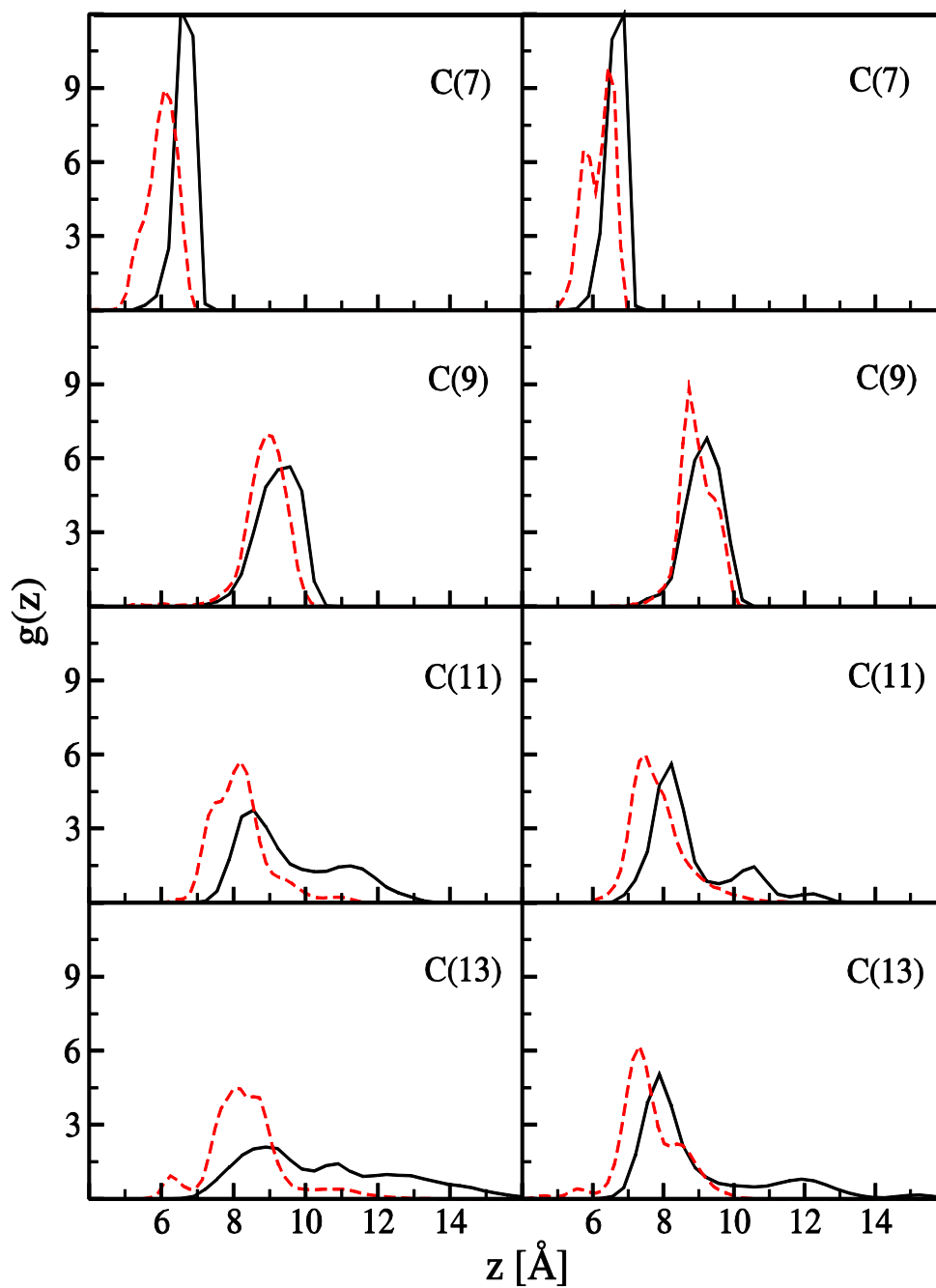


Figure 5-5. Surface distributions of carbonyl carbons for TMA-tri-proline-tether (left) and BOC-tri-proline-tether (right). Results are shown for *n*-hexane/2-propanol (solid black line) and water/methanol (dashed red line) solvents.

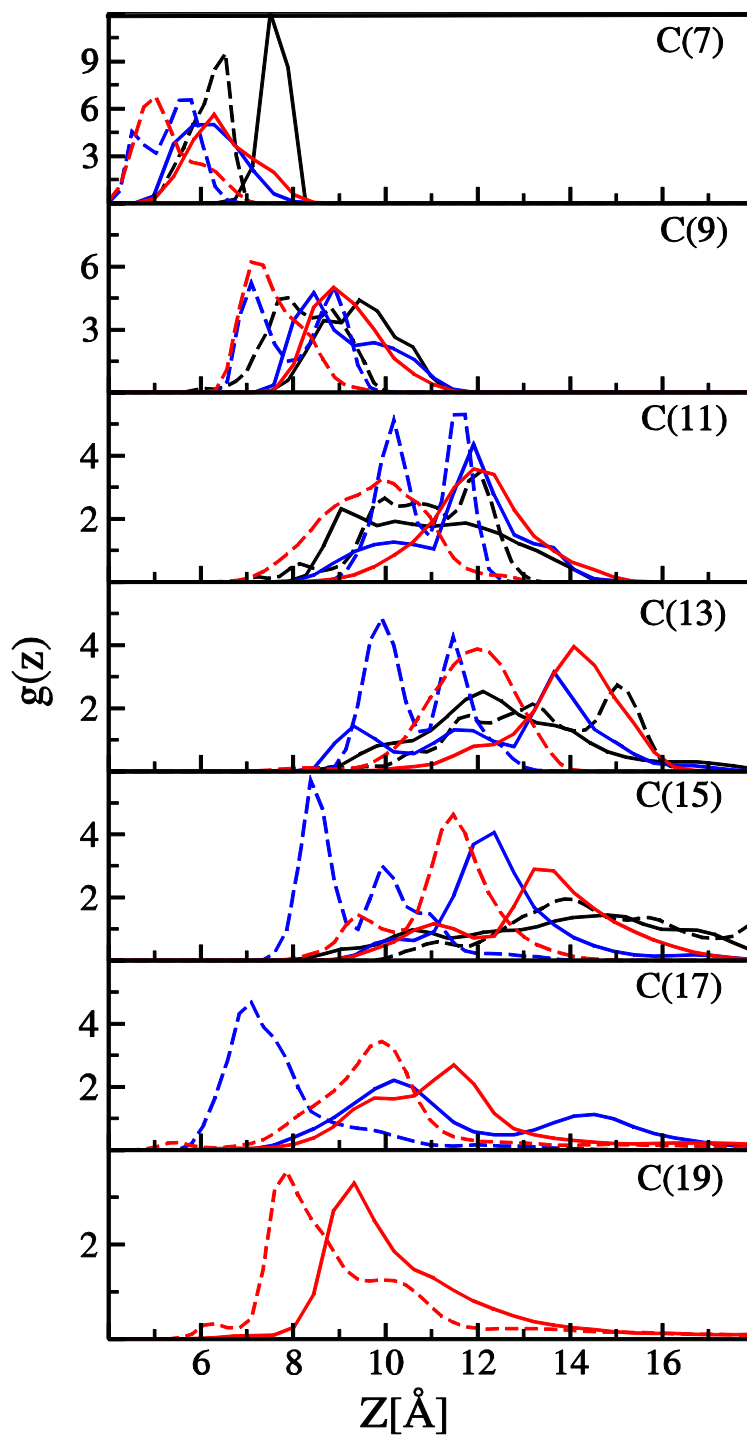


Figure 5-6. Surface distribution of carbonyl carbons. Results are shown for *n*-hexane/2-propanol (solid lines) and water/methanol (dashed lines). Black, blue, and red lines correspond to TMA-(Pro)₄-N(CH₃)-tether, TMA-(Pro)₅-N(CH₃)-tether, and TMA-(Pro)₆-N(CH₃)-tether, respectively.

The carbonyls further along the poly-proline backbone, from the middle of the chain to the terminal group, have a broader distribution of positions relative to those carbonyls located closer to the tether. Although longer poly-peptides are assumed to extend more towards the bulk, because of mixed *cis-trans* amide linkages they adopt bent structures. The first three or four proline residues extend away from the Si layer and towards the solvent. The remaining proline residues, however, tend to bend back towards the surface. The exception, once again, is tetraproline, due to the contribution from the extended, helical TTTT conformer.

The statistical analysis of the uncertainty for C(13) in the triproline selector and C(19) for the hexaproline selector in *n*-hexane/2-propanol are shown in Figure 5-7. Overall, the analysis of the convergence of the radial distributions has shown high consistency between independent simulations. For this reason and for the sake of simplicity, the error bars are removed from other radial distributions.

Figure 5-8 shows the carbonyl group orientation relative to the underlying surface for the two topmost carbonyls in each poly-proline selector. A carbonyl group that points directly into the surface or towards the inside of a bent conformer will be crowded and less likely to hydrogen bond. We find that the solvents differ markedly in the carbonyl orientations they promote. The systematic nature of the difference indicates an overall, solvent-induced rotation of the selector. This rotation is most pronounced for the longer poly-prolines and for the carbonyls closest to the surface. Specifically, if a carbonyl tends to point towards the bulk in water/methanol, it typically will point towards the surface in *n*-hexane/2-propanol. For the hexa-proline in *n*-hexane/2-propanol, the first three carbonyls are directly mainly towards the surface, the middle carbonyls point towards the bulk and, as the backbone bends, the topmost carbonyls point towards the

surface once again. The reverse applies in water/methanol where the carbonyls in the bend region point towards the surface and the others point toward the bulk.

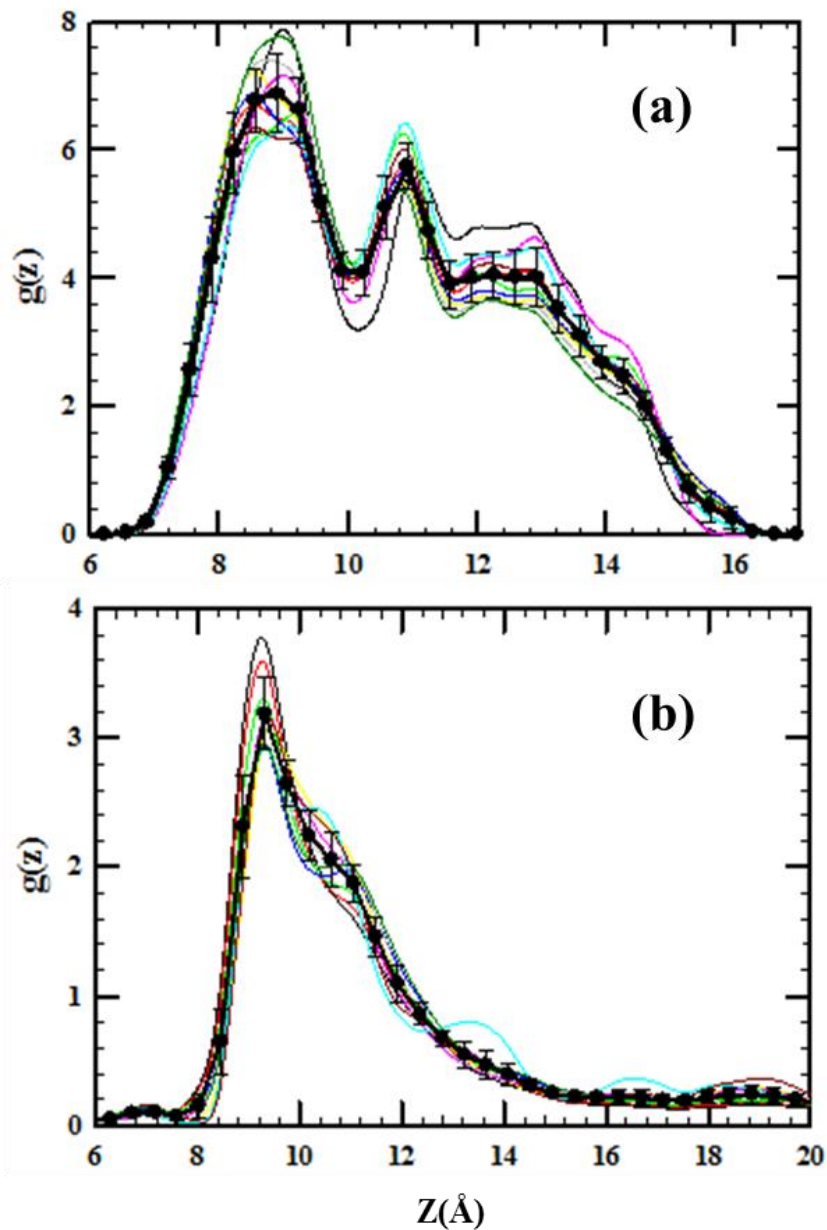


Figure 5-7. An illustration of the convergence of surface distributions. The surface distribution of a) C(13) in the triproline selector and b) C(19) for the hexaproline selector in *n*-hexane/2-propanol. Results for 10 independent simulations are shown. The thick black line is the average surface distribution and error bars are provided.

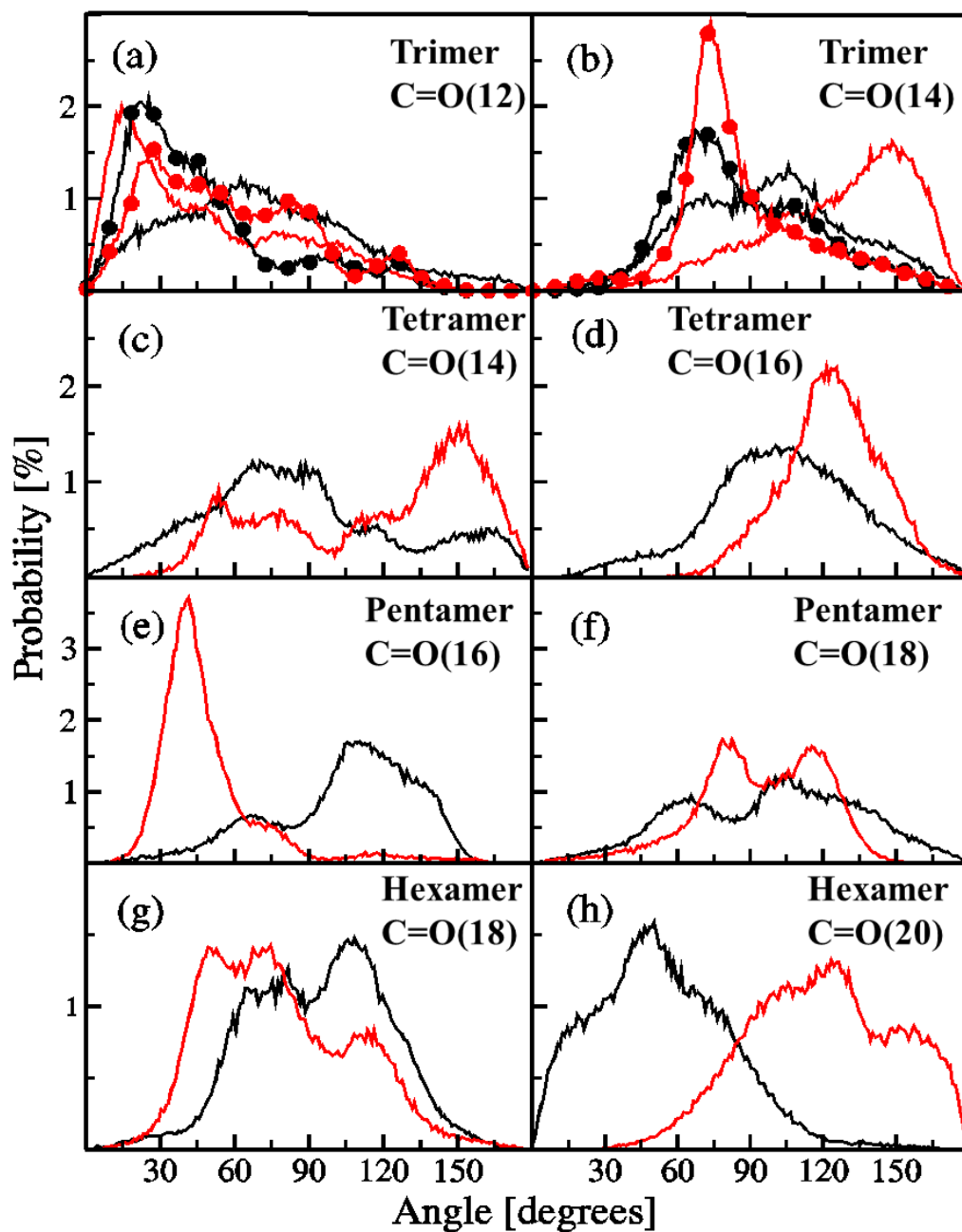


Figure 5-8. The distribution of the two topmost carbonyl groups at the interface. The probability of having a carbonyl group with a given angle relative to the underlying silicon layer is presented: an angle of 0° corresponds to a carbonyl group pointing directly towards the surface while an angle of 180° corresponds to a C=O group that points into the bulk. Results for TMA-(Pro) $_n$ -N(CH $_3$)-tether are shown for *n*-hexane/2-propanol (black line) and water/methanol (red line). Filled circles in panels (a) and (b) identify the corresponding distributions for BOC-(Pro) $_3$ -N(CH $_3$)-tether.

Snapshots of the interfaces are provided in Figure 5-9. Top-down views show the increase in surface crowding for the longer peptide chains. The surfaces in the presence of the normal phase solvent are more disordered and reflect the wider range of conformers present. Side views of the simulation cells illustrate the extended nature of the tetra-proline interface, relative to the others, and the compactness and order at the surface when the solvent is water/methanol.

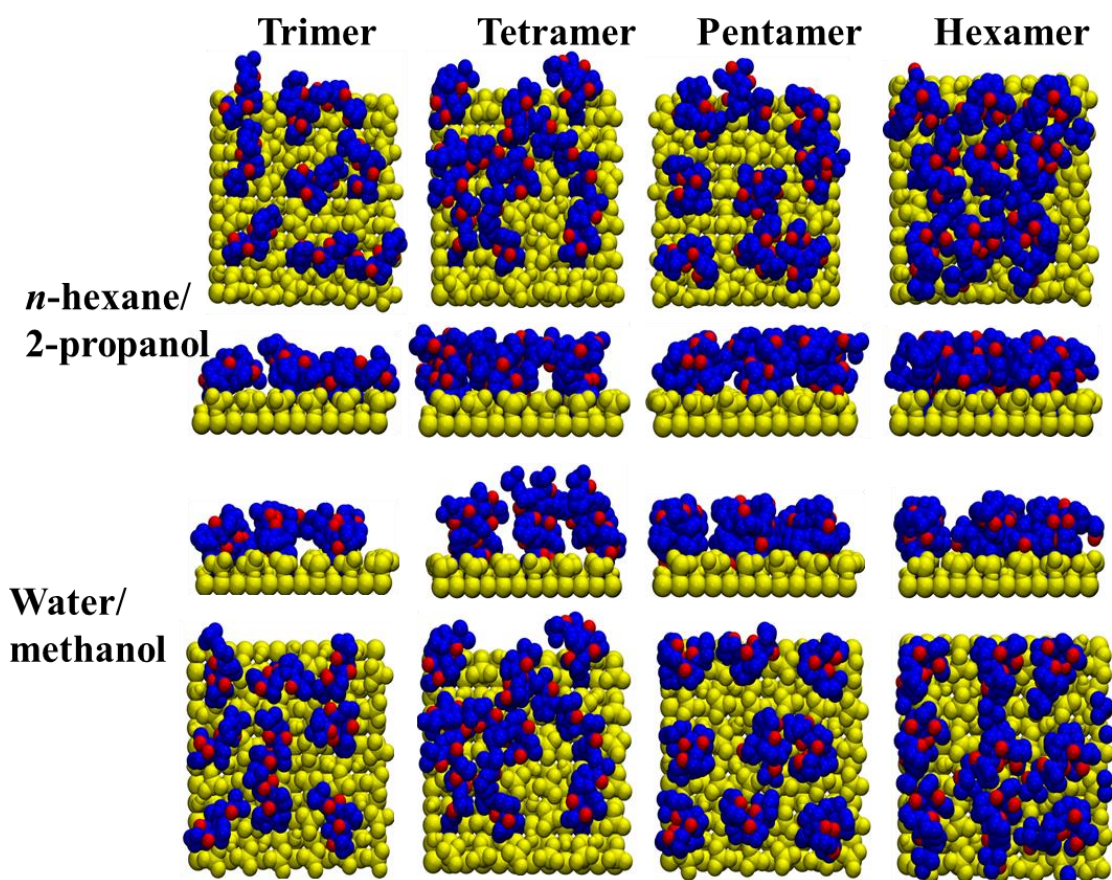


Figure 5-9. Top-down and side-view snapshots of the TMA-(Pro)_n-N(CH₃)-tether selective surface. All selector atoms are shown in blue, except for the carbonyl oxygens shown in red. The underlying Si layer, the end-caps and the silanol groups are represented in yellow. The top two rows show the selective surface in the *n*-hexane/2-propanol. The corresponding snapshots for water/methanol are provided in the bottom two rows.

Overall, we find that the presence of solvent reduces the number of conformers relative to the gas phase. The lowest-energy gas phase conformer is typically most frequently observed at the interface, but interfacial environment and solvent are sufficient to impact the nature and prevalence of the conformers. The conformational distribution is most well defined in water, particularly for the longer poly-prolines where a single conformer can account for over 90% of the selectors at the interface. The selectors are also more compact in the more polar, reverse phase solvent. Even when the solvents prefer the same backbone configuration, in water/methanol the selectors are closer to the surface and have different orientations so that a larger number of carbonyls are directed towards the bulk, to facilitate hydrogen bonding. The tetra-proline interface is unique among the peptides in that there are more conformers at the interface and the selectors extend further into the bulk, with the greatest extension in water/methanol.

5.4.2 Solvation and Hydrogen Bonding

The solvent distribution above the surface is presented in Figure 5-10. Consider first the *n*-hexane/2-propanol. The alcohol has a higher probability of being at the interface, as evidenced by the broad peaks between 6 and 15 Å from the Si layer. The backbone carbonyl oxygens are located in this region and hydrogen bonding brings the propanol into this region as well. In contrast, *n*-hexane does not interact strongly with the selectors and is less likely to be found at the interface and, in particular, the alkane is less probable within a 16 Å zone above the surface. The only exception is a small increase in probability between 6 and 9 Å, directly above the end-caps but below the bulk of the poly-proline.

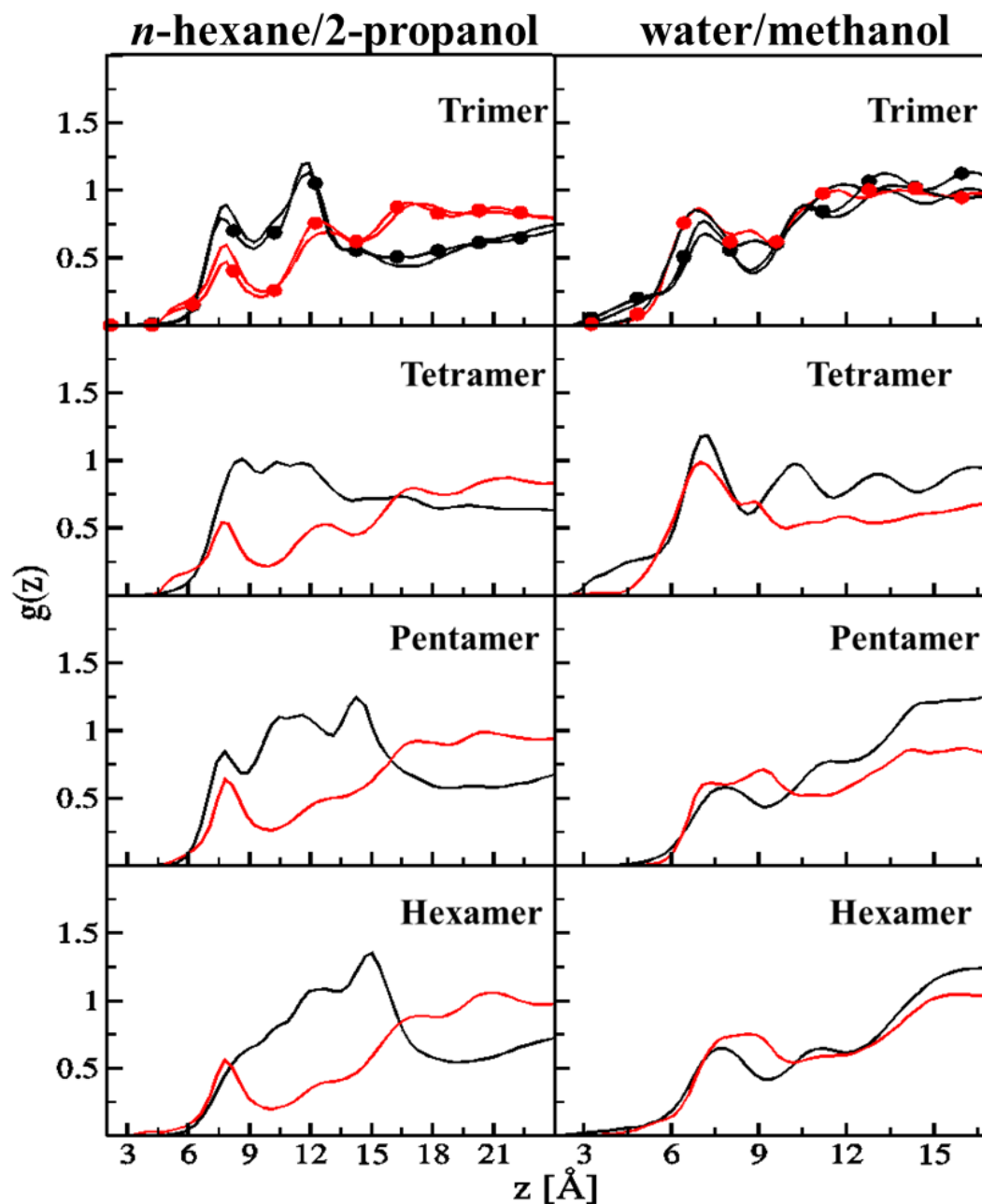


Figure 5-10. Solvent distribution above the interface for TMA-(Pro)_n-N(CH₃)-tether in *n*-hexane/2-propanol (left hand side) and water/methanol (right hand side). The results for the tri-proline, tetra-proline, penta-proline and hexa-proline are in the first, second, third, and fourth row, respectively, with the *n*-hexane and methanol distributions in red, and the 2-propanol and water distributions in black. The solvent distribution above the BOC-terminated tri-proline interface is given in the first row, with distributions identified by filled circles.

The presence of water and methanol at the surface leads to extensive hydrogen bonding with the poly-proline backbone. Both solvents are relatively small and can form hydrogen bonds with the poly-proline and with other solvent. For these reasons, they behave similarly at the interface. For instance, the region between 6 and 9 Å, above the end-caps and below the bulk of the selectors, shows an increased probability for either solvent. On the whole, these solvents are somewhat less likely to be found in the vicinity of the surface, relative to the bulk. This decrease is simply due to the steric bulk of the selectors.

H-bonding statistics for all selectors and solvents are provided in Table 5-2 and Table 5-3. Statistics are provided only for carbonyl oxygens since H-bonding to backbone nitrogens is infrequent [237]. A detailed structural examination of solvent-poly-proline H-bonding is possible from 2D distributions. These identify the solvent location around a specified selector atom, with the location subdivided into components parallel and perpendicular to the surface. Figure 5-11 provides an example for TMA-(Pro)₅-N(CH₃)-tether where several potential H-bonding pairs are examined. To be precise, the 2D distribution between backbone carbonyl oxygens and H-bonding hydrogens in the solvent are presented. A peak or ridge at an O···H separation of around 2 Å identifies a region where a solvent-peptide H-bond can occur. Secondary solvent-solvent hydrogen bonds and hydrogen-bonds to other selector atoms are often evident as ridges or peaks appearing at larger separations. Methanol and 2-propanol show secondary structures when they interact with O(10), for instance the methanol peak at $r=4$ Å results from a methanol-methanol or a methanol-water H-bond.

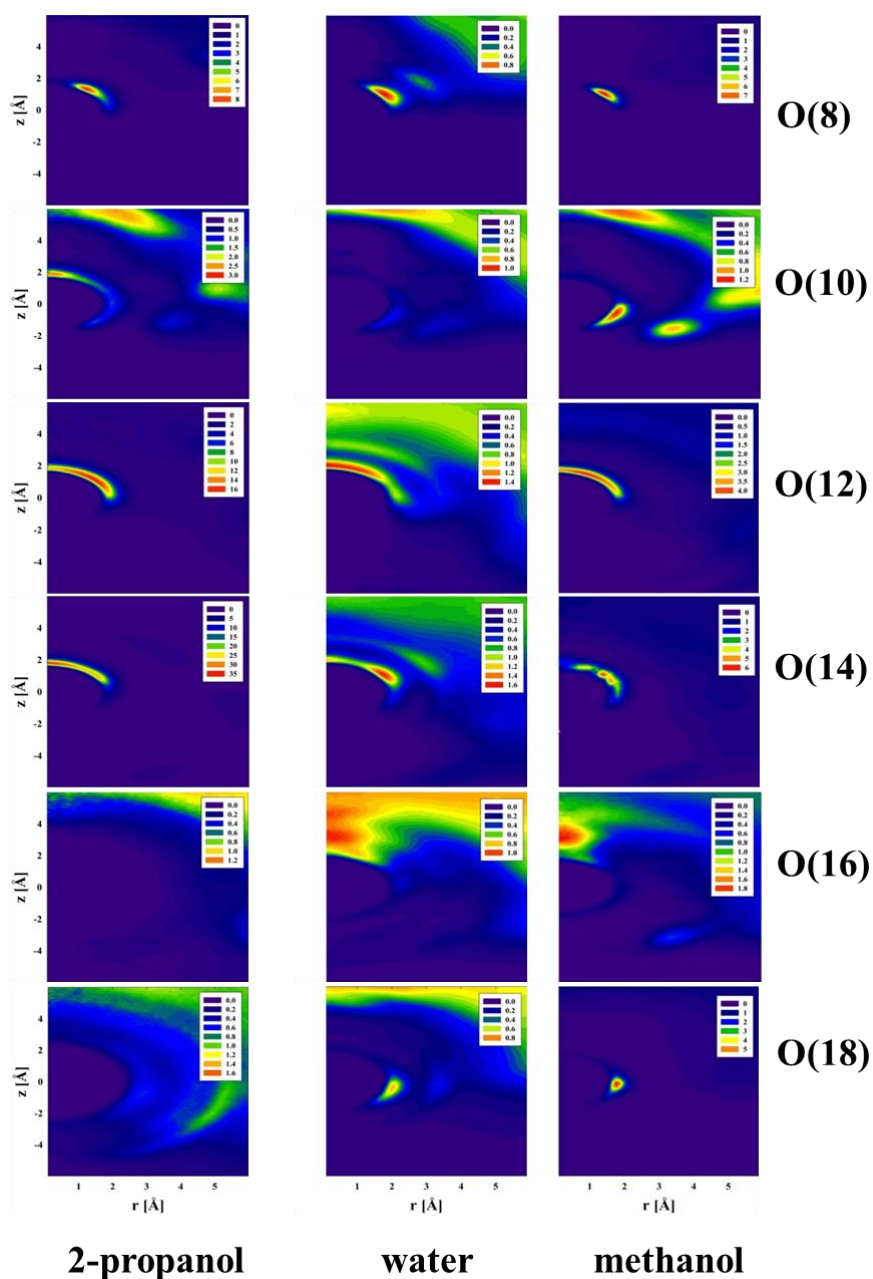


Figure 5-11. 2D distributions for the TMA-(Pro)₅-N(CH₃)-tether interface. Distributions between potential H-bonding pairs: Each row corresponds to a backbone carbonyl oxygen and each column corresponds to a solvent H. The first column shows the distribution for the alcohol H in *n*-hexane/2-propanol, the second and third correspond to H in water and the alcohol H, respectively, in water/methanol. The distance perpendicular to the surface is *z* and a positive value means that the solvent atom is further from the Si layer. *r* is the component of the O···H separation parallel to the surface.

Table 5-2. Hydrogen bonding statistics for carbonyl oxygens at the interface. The last column gives the percentage of selectors with an H-bond at the specified oxygen. In the water/methanol, the contribution from methanol is listed in brackets.

Selector	Solvent	Oxygen	H-bonding (%)
TMA-tri-proline	<i>n</i> -hexane/ 2-propanol	O(8)	4 ± 2
		O(10)	15 ± 3
		O(12)	7 ± 2
		O(14)	10 ± 2
	water/ methanol	O(8)	36 ± 5 (5 ± 2)
		O(10)	63 ± 5 (7 ± 1)
		O(12)	40 ± 7 (5 ± 1)
		O(14)	55 ± 6 (1 ± 1)
BOC-tri-proline	<i>n</i> -hexane/ 2-propanol	O(8)	5 ± 2
		O(10)	16 ± 2
		O(12)	2 ± 1
		O(14)	10 ± 2
		O(37)	0 ± 0
	water/ methanol	O(8)	21 ± 2 (4 ± 1)
		O(10)	37 ± 3 (3 ± 2)
		O(12)	17 ± 4 (2 ± 1)
		O(14)	28 ± 2 (3 ± 1)
O(37)	2 ± 1 (0 ± 0)		
TMA-tetra-proline	<i>n</i> -hexane/ 2-propanol	O(8)	7 ± 1
		O(10)	4 ± 1
		O(12)	5 ± 2
		O(14)	3 ± 1
		O(16)	5 ± 1
	water/ methanol	O(8)	40 ± 5 (3 ± 1)
		O(10)	37 ± 7 (3 ± 1)
		O(12)	29 ± 4 (1 ± 1)
		O(14)	40 ± 4 (3 ± 1)
O(16)	48 ± 6 (4 ± 1)		

Table 5-2. Continued

Selector	Solvent	Oxygen	H-bonding (%)
TMA-penta-proline	<i>n</i> -hexane/2-propanol	O(8)	4 ± 2
		O(10)	3 ± 2
		O(12)	10 ± 1
		O(14)	17 ± 3
		O(16)	2 ± 1
		O(18)	1 ± 1
	water/methanol	O(8)	22 ± 3 (5 ± 1)
		O(10)	13 ± 2 (2 ± 1)
		O(12)	42 ± 5 (5 ± 2)
		O(14)	53 ± 4 (7 ± 1)
		O(16)	18 ± 4 (2 ± 1)
		O(18)	27 ± 3 (4 ± 2)
TMA-hexa-proline	<i>n</i> -hexane/2-propanol	O(8)	1 ± 1
		O(10)	1 ± 1
		O(12)	10 ± 2
		O(14)	18 ± 2
		O(16)	6 ± 2
		O(18)	5 ± 1
		O(20)	2 ± 1
	water/methanol	O(8)	8 ± 3 (1 ± 1)
		O(10)	9 ± 2 (1 ± 1)
		O(12)	36 ± 4 (5 ± 2)
		O(14)	50 ± 3 (7 ± 1)
		O(16)	31 ± 5 (3 ± 1)
		O(18)	25 ± 2 (3 ± 1)
		O(20)	19 ± 2 (3 ± 1)

Consider the sterically hindered selector oxygen closest to the surface, O(8). An H-bonding peak, at positive z , is evident in Figure 5-11 indicating that solvent can only H-bond from above. 2-propanol shows strong, well defined H-bonding ridges for O(12) and O(14). These ridges appear at positive z indicating that solvent also approaches O(12) and O(14) from the bulk side. H-bonding is weak for the carbonyls nearest to the terminal group, O(16) and O(18), due to steric crowding as the selector turns towards back the surface: Ridges at around 2

Table 5-3. Hydrogen bonding statistics for poly-proline selectors. The last column gives the percentage of selectors with number of H-bonds. Only probabilities above 1% are listed.

Selector	Solvent	Number of H-bonds	Probability (%)
TMA-tri-proline	<i>n</i> -hexane/ 2-propanol	1	25 ± 5
		2	4 ± 2
	water/ methanol	1	27 ± 6
		2	30 ± 6
		3	19 ± 4
		4	8 ± 2
BOC-tri-proline	<i>n</i> -hexane/ 2-propanol	1	19 ± 5
		2	3 ± 2
	water/ methanol	1	18 ± 4
		2	18 ± 5
		3	10 ± 3
		4	4 ± 2
TMA-tetra-proline	<i>n</i> -hexane/ 2-propanol	1	19 ± 4
		2	7 ± 4
		3	3 ± 2
	water/ methanol	1	22 ± 5
		2	25 ± 5
		3	17 ± 3
TMA-penta-proline	<i>n</i> -hexane/ 2-propanol	1	20 ± 2
		2	7 ± 2
		3	2 ± 1
	water/ methanol	1	22 ± 5
		2	23 ± 3
		3	11 ± 4
TMA-hexa-proline	<i>n</i> -hexane/ 2-propanol	4	5 ± 2
		5	1 ± 1
		1	24 ± 4
		2	11 ± 2
	water/ methanol	3	4 ± 2
		4	2 ± 1
1		21 ± 4	
2		30 ± 6	
		3	16 ± 3
		4	12 ± 3
		5	4 ± 2
		6	1 ± 1

Å in the 2D distribution are barely visible. The broad, diffuse region at high z is due to solvent in the bulk.

In the presence of water/methanol, the peptide may H-bond to either component. From Table 5-2 most of the hydrogen bonding occurs with water: Water is smaller, more polar, and can form more H-bonds, per molecule, than methanol. The 2D distributions are qualitatively similar for the two solvents, with the strongest H-bonding regions observed for O(12) and O(14), the mid-chain carbonyls.

Comparing carbonyls on a poly-proline backbone, H-bonding frequency differs due to steric hindrance and orientation of the carbonyl relative to the surface. For the penta-proline in *n*-hexane/2-propanol, only 4% of carbonyls nearest the tether hydrogen bond (Table 5-2). Likewise, the two carbonyls nearest the TMA rarely hydrogen bond. On the other hand, over ten percent of the penta-prolines have a hydrogen bond to the middle two carbonyl oxygens, O(12) and O(14). These two carbonyl groups are optimally positioned at the top of the crowded selector region and pointing towards the bulk.

Table 5-3 shows that the probability for H-bonds. Hydrogen bonding is always much higher in the more polar water/methanol. Eighty-eight percent of the tri-prolines, for example, have at least one H-bond and 2% of the selectors have five. Since the backbone has only four carbonyls, in 2% of the selectors, one carbonyl has two H-bonds simultaneously. The same tri-proline selector, in the presence of *n*-hexane/2-propanol, has a 30% chance of having a least one H-bond but fewer than 1% have three or more. For the hexa-proline selector, around 45% have at least one H-bond in *n*-hexane/2-propanol but virtually all selectors have at least one H-bond water/methanol.

5.4.3 Terminal Group Effect

The role of the terminal group has been explored in detail for di-prolines [237] where it was found that TMA terminated selectors were significantly better at resolving analytes than BOC-terminated selectors. These selectors differ only by the addition of an O to the top of the backbone (Figure 5-1) and yet they are quite distinct at the interface. A molecular dynamics study of di-prolines (Chapter 4) found that the backbone conformational preferences differ for these two selectors and this, in turn, leads to differences in conformations and H-bonding at the interface. In this section, we examine the differences between TMA and BOC terminated tri-prolines to see whether the terminal group effect applies to the longer backbone.

BOC terminated tri-prolines show distinct conformational preferences even in the gas phase. Specifically, TMA-tri-proline adopts TTC as the most stable conformer, with TTT as a low-lying conformer. BOC-tri-proline prefers a *cis* conformer nearest the terminal group: it preferentially adopts a CTC conformation. The TTC conformer is somewhat higher in energy and TTT is unlikely as it lies 12 kJ/mol higher than CTC. There are five TMA terminated tri-prolines within 10 kJ/mol of the TTC conformer but, when the terminal group is BOC, only three conformers are within 10 kJ/mol of the CTC conformer. Despite these differences, it is interesting that the three lowest-energy conformers of both selectors adopt a middle *trans* amide linkage.

Table 5-1 summarizes the conformational preferences in the presence of *n*-hexane/2-propanol and water/methanol binary solvents. In the presence of solvent, the terminal group leads to different conformational preferences. The TMA terminated tri-proline has three conformers that are adopted by at least 5% of the selectors, the conformer prevalence follows expectations from *ab initio* energies, and the conformer population is most restricted in water/methanol. When

the tri-proline is terminated by BOC, the number of conformers increases in water/methanol. Of all the selectors examined, this is the only case where the poly-proline is more flexible in solution relative to the gas phase.

The backbones are, on average, more extended for the TMA terminated selectors, as shown in Figure 5-5. This is clear from the broad distribution for the topmost carbonyl carbon, C(13), which can be as far as 16 Å from the Si layer. The C=O(12) and C=O(14) orientational distributions are provided in Figure 5-8. On the whole, the carbonyls are more frequently directed towards the bulk when the terminal group is TMA and this, along with the extended backbone accounts, for the higher probability for hydrogen bonding in water/methanol (see Table 5-2).

Overall, the identity of the terminal group, BOC or TMA, impacts the interfacial properties for the tri-proline selector. First, BOC terminated tri-proline prefers different conformers than the TMA terminated selector and this leads to a slightly more compact structure at the interface with the outermost carbonyls more frequently directed towards the surface. We find that the likelihood of hydrogen-bonding is similar when the solvent is *n*-hexane/2-propanol but the BOC terminated selector H-bonds less in water/methanol.

5.5 Conclusions

The interfacial structure of poly-proline chiral selectors, in the presence of water/methanol and *n*-hexane/2-propanol, has been explored in detail. Five chiral selectors have been examined: TMA-(Pro)₃-N(CH₃)-tether; BOC-(Pro)₃-N(CH₃)-tether; TMA-(Pro)₄-N(CH₃)-tether; TMA-(Pro)₅-N(CH₃)-tether; and TMA-(Pro)₆-N(CH₃)-tether.

Following identification of major gas phase conformers, force field development proceeds via an extensive series of B3LYP/6-311G(d,p) calculations as each poly-proline chain is

systematically distorted. Least squares fits yield stretch, bend, improper torsion and torsional potentials along the backbone. The latter potentials are adjusted to reproduce, as best as possible, the LPNO-CCSD (RI-MP2) conformer energies. The final force fields are selected based on their ability to reproduce the structures and relative energies of the conformers predicted from *ab initio* calculations.

A series of molecular dynamics simulations of each chiral interface, in the presence of either a normal phase *n*-hexane/2-propanol or a reverse phase water/methanol, have been performed. The role of chain length has been examined for TMA terminated poly-prolines. We find that the number of conformers at the interface depends on solvent, but is most restricted in water/methanol. The added conformational flexibility in *n*-hexane/2-propanol leads to a more disordered interface (Figure 5-9). The number of conformers observed at the interface increases from the tri-proline to the tetra-proline but then decreases rapidly for longer poly-prolines. For the hexa-proline, a single backbone conformation accounts for over 83% of the selectors. The selectors are generally more compact in water/methanol and show a bent structure that brings the topmost portion of the selector nearer to the surface. The exception here is the tetra-proline where the peptide can extend up to 18 Å into water/methanol before bending back towards the surface.

Along with a more compact backbone, the nature of the solvent leads to a rotation of the selector such that more carbonyl groups are directed towards the bulk in water/methanol. Hydrogen bonding along the backbone is extensive in this solvent but the carbonyls in the middle of the backbone, near the bend, generally have the highest probability for H-bonding. The normal phase, *n*-hexane/2-propanol, solvent is less likely to form a H-bond with poly-proline, as expected. Nonetheless, at least 20% of selectors have at least one H-bond in this solvent.

Overall, we have found that the role of the surface and the solvent has a strong impact on the characteristics of the interface. The conformers tend to display a mix of *cis* and *trans* amides, in the gas phase and at the surface. Beyond the tetra-proline, conformational rigidity increases and this may explain the enhancement of selectivity for longer poly-prolines.

Chapter 6

Docking Studies of Poly-proline Chiral Stationary Phases

6.1 Introduction

Fifty three analytes have been experimentally tested on poly-proline CSPs and enantiomers with a hydrogen bond donor next to the chiral center tend to be separated [47,78,93-95]. In this Chapter, we examine the selectivity of several poly-proline selectors including TMA-(Pro)_n-N(CH₃)-tether with n=2,3,6 and BOC-(Pro)₂-N(CH₃)-tether using molecular dynamics simulations. Six chiral analytes in the presence of two binary solvents have been examined. These analytes have been chosen based on two factors: The success of resolution and structural similarities. These analytes are α -methyl-9-anthracenemethanol (AnC-CH₃), 1-anthracen-2-yl-ethanol (AnS-CH₃), α -(trifluoromethyl)benzyl alcohol (Ph-CF₃), 1-phenylethanol (Ph-CH₃), 1-phenyl-1-propanol (Ph-C₂H₅), and 1-(9-anthryl)-2,2,2-trifluoroethanol (AnC-CF₃) and are shown in Figure 6-1. They all exhibit hydrogen bonding donor site since analytes without hydrogen bonding sites are not resolved at all on poly-proline based selectors. Every analyte has an aromatic moiety, a hydroxyl group, and an allyl group attached to the chiral carbon. The selected enantiomers are also relatively small, with limited flexibility that reduces the number of possible docking arrangements. Solvents include 70/30 v/v *n*-hexane/2-propanol and 70/30 v/v water/methanol. Molecular dynamics simulations are undertaken for each combination of solvents, proline selectors, and analytes.

Despite the experimental work on poly-proline stationary phases, not much is known about the underlying mechanism of enantio-separation. According to the three point interaction model [36], for the enantio-separation to happen, three simultaneous interactions are needed to

differentiate between enantiomers. Such interactions often include hydrogen bonding, π - π stack, CH- π interaction, and steric interactions. For poly-proline stationary phases, however, the selectors lack any aromatic moiety and only accept hydrogen bonds. Also, the enantiomers under consideration only hydrogen bond through a single donating site. As a result, it is expected that the complex interactions of the selectors and enantiomers cannot be classified into a three point attachment model. In this Chapter we examine the selector-enantiomer interactions carefully.

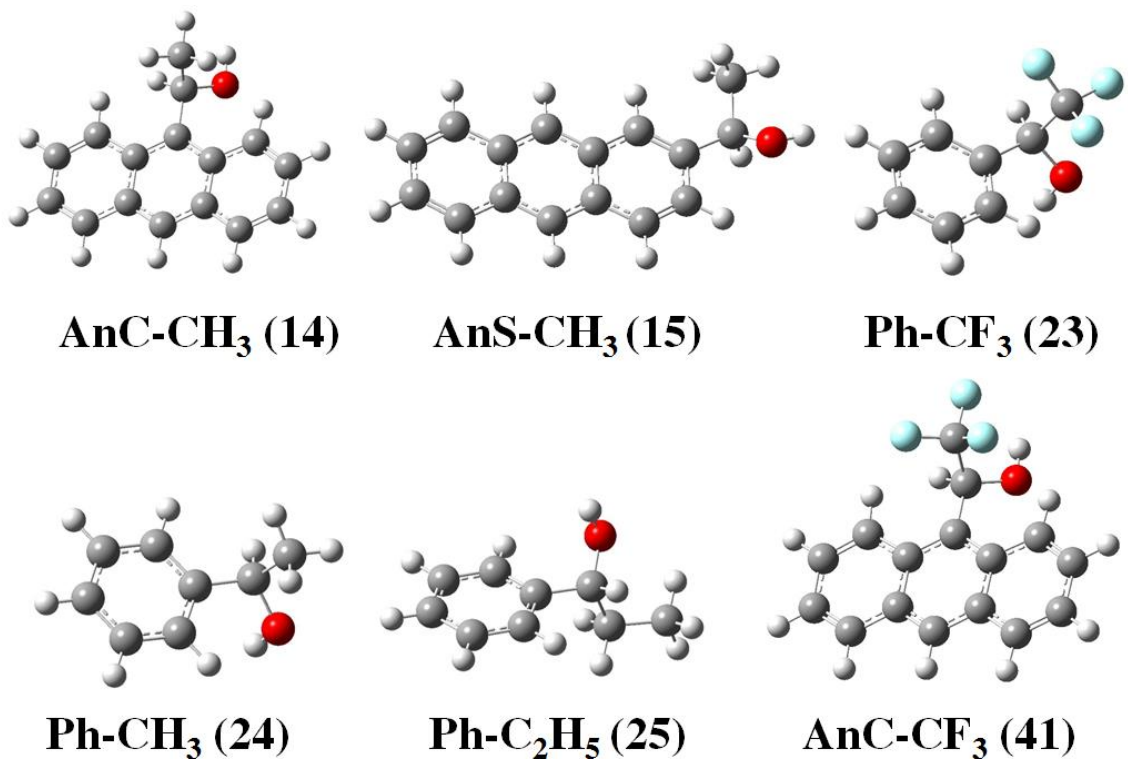


Figure 6-1. The B3LYP/6-311G(d,p) optimized structures of the analytes. “AnC” indicates an attachment to a central carbon in anthracyl, “AnS” indicates an attachment to a side carbon in anthracyl, and Ph represents a phenyl ring. Numbers in bracket indicate the analyte numbering used by the Li group [47].

The distinct force fields that were developed for the poly-proline selectors in Chapters 4 and 5 are used here. For the six analytes, force fields are developed that are described in Section 6.2.1. The details of the simulations are explained in Section 6.2.2. The distribution of the analytes at the interface, interaction statistics, nature of the interactions and binding modes are discussed in Section 6.3.

6.2 Theoretical Details

6.2.1 Potentials

All of analytes have aromatic rings and alcohol groups which are located next to the chiral center. A comparison between these analytes can isolate the effect of structural changes on selectivity. For instance, comparison between AnC-CH₃ and AnS-CH₃ shows the effect of aliphatic alcohol attachment position. Similarly, comparing AnC-CH₃ and AnC-CF₃ isolates the effect of bulky groups and halogenation. The size of the aromatic moiety and its impact on enantio-separation can be studied by comparing AnC-CH₃ and Ph-CH₃.

The analyte force fields are based on *ab initio* calculations performed at the B3LYP/6-311G(d,p) level following the method outlined in Chapters 4 and 5. Full details are not given here, but the force field parameters are given in the Appendix A.

For the analyte force field, an all atom approach is taken. A stretch potential (Eq. 2-26) is defined between the chiral carbon and the neighboring ring carbon. All other bonds are kept frozen during the simulation using RATTLE [150] algorithm. Bending potentials (Eq. 2-27) are defined for all three consecutive atoms except when all three atoms belong to the rigid aromatic moiety. Torsional dihedral potentials (Eq. 2-28) are employed to account for rotation of the molecule around single bonds. To keep the planarity of the ring, and to preserve the chirality of the chiral centers, improper torsion potentials (Eq. 2-29) are defined for the chiral carbon and the

ring carbon connected to it. OPLS parameters [217-220] are used for LJ potential and atomic charges are calculated using CHELPG algorithm [221].

Analyte force fields need to be developed for both R and S enantiomers. Here, all stretches and bending potentials are the same for both enantiomers. The differences in force field appear in the dihedral and improper torsions. Since the enantiomers are the mirror images of each other, the vectors defining the proper and improper torsional angles must be reversed. The effect of this is simply a sign change for $\varphi_t^{i;0}$ and ω_{eq} in Eqs. 2-28 and 2-29, respectively.

6.2.2 The Simulation Cell

A typical simulation cell is shown in Figure 6-2. The cell consists of two parallel interfaces with the solvent in between. Each interface includes 16 selectors, 48 trimethylsilyl end-caps, 64 silanols and 128 immovable silicon atoms. The simulation cell also consists of 8R and 8S enantiomers. The interfacial distance varies from 42.92 to 60.9 Å depending on the selector and the solvent. The solvent density is kept within $\pm 3\%$ of the experimental density by measuring the density in the mid-plane of the cell and adjusting the interfacial distance.

6.2.3 MD Simulations

The selectivities of TMA terminated di-, tri-, and hexa-proline and BOC terminated di-proline for six analytes are examined over 35-40 *ns* of simulation time in the presence of two binary solvents: *n*-hexane/2-propanol and water/methanol. For all simulations, the time step is 0.3 *fs*, chosen to ensure the Nosé-Hoover Hamiltonian [153,154] is conserved. Simulations are performed in NVT ensemble at 298 K. Ewald summation is used to treat the electrostatic interactions. For each combination of selector, analyte, and solvent, 15 independent simulations are performed to ensure convergence of the distributions and selectivity factors. Scaling of torsional energy barriers is undertaken using the outline discussed in Chapter 5.

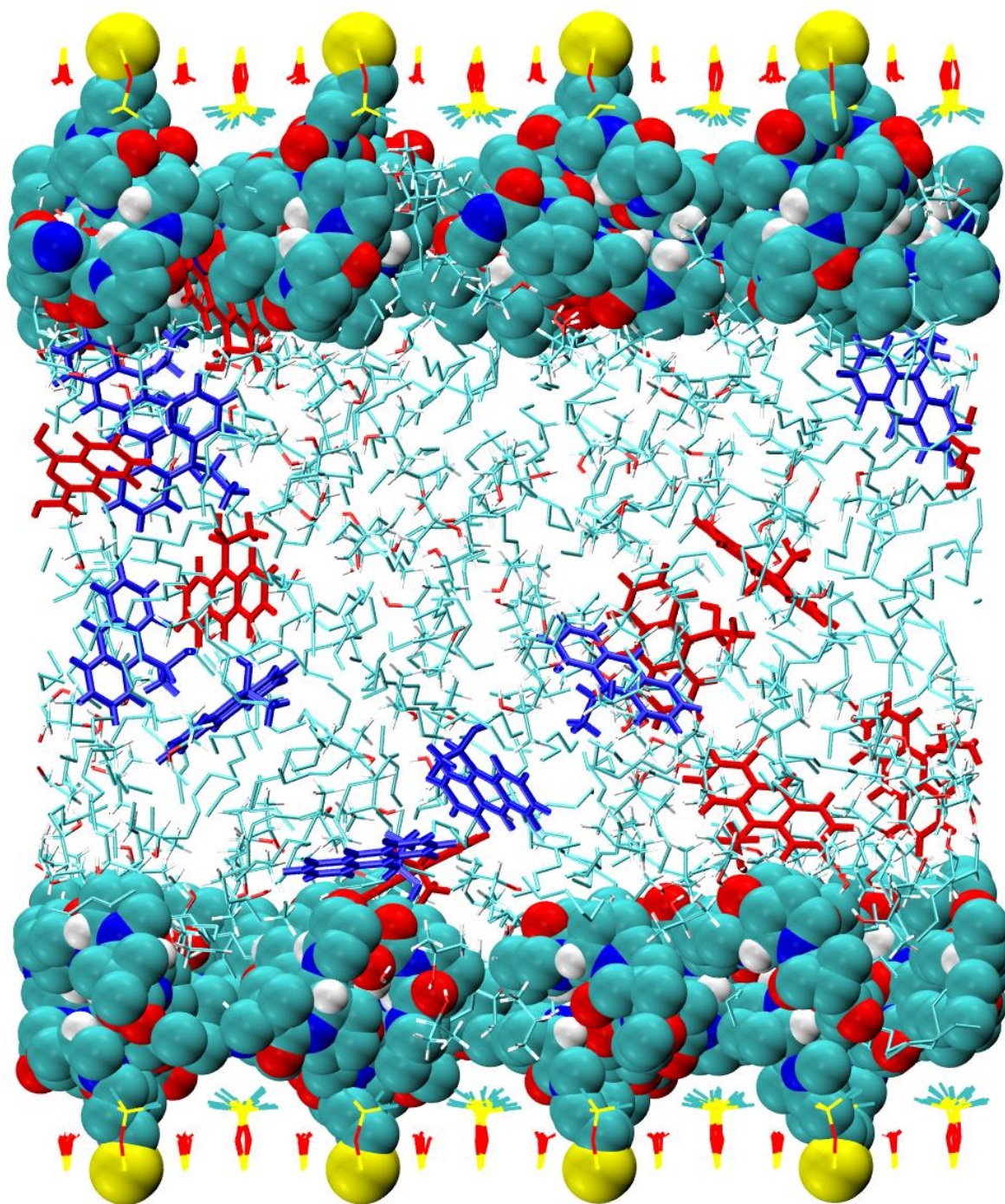


Figure 6-2. A snapshot of a typical simulation cell for TMA terminated hexa-proline in *n*-hexane/2-propanol with AnC-CH₃. The R and S enantiomers are shown in blue and red, respectively. Yellow, red, ocean-blue, blue and white correspond to silicon, oxygen, carbon, nitrogen and hydrogen atoms.

6.3 Results and Discussions

6.3.1 Distribution at the Interface

Analytes may hydrogen bond to the carbonyl oxygens of the selectors. With this in mind, we begin with the surface distribution of carbonyl oxygens. Figure 6-3 shows the $g(z)$ distribution of oxygen atoms for all selectors in the normal and reverse phase solvents. For TMA-di-proline selector O(8) and O(10) appear at the same distance from the Si layer while O(12) is further away. For BOC terminated di-proline, the selector accesses a wider range of conformations and O(12) appears from 8 to 14 Å from the Si layer in *n*-hexane/2-propanol while this O(12) peak is sharper for the TMA terminated di-proline. BOC terminated di-proline is closer to the surface in water/methanol compared to *n*-hexane/2-propanol where the backbone bends towards the surface so that O(10) is closer to the surface compared to the O(8). This is also true for TMA terminated di-proline. For the TMA terminated tri-proline, the oxygen distribution is roughly the same for both of the solvents with one peak at 7 Å for O(8), a double peak for O(10) and broad distributions for O(12) and O(14). The TMA terminated hexa-proline backbone adopts a bend towards the surface so that after O(14), all the backbone atoms move closer to the surface. This can be seen from the $g(z)$ distributions where O(10), O(14), O(18) and O(20) appear at 5-8, 10-16, 7-13, and 5-13 Å, respectively. This bend in the backbone results in a crowded interface for TMA terminated hexa-proline as the selector does not simply extend towards the bulk. The wide range of positions for O(16), O(18), and O(20) shows a relative flexibility in their movement to and from the interface. The distributions are in agreement with what has been observed in Chapter 5. The orientation of the carbonyl group brings the oxygen atoms closer to the Si layer in water/methanol solvents. So, it is expected that less hydrogen bonding opportunities will be provided for the analytes.

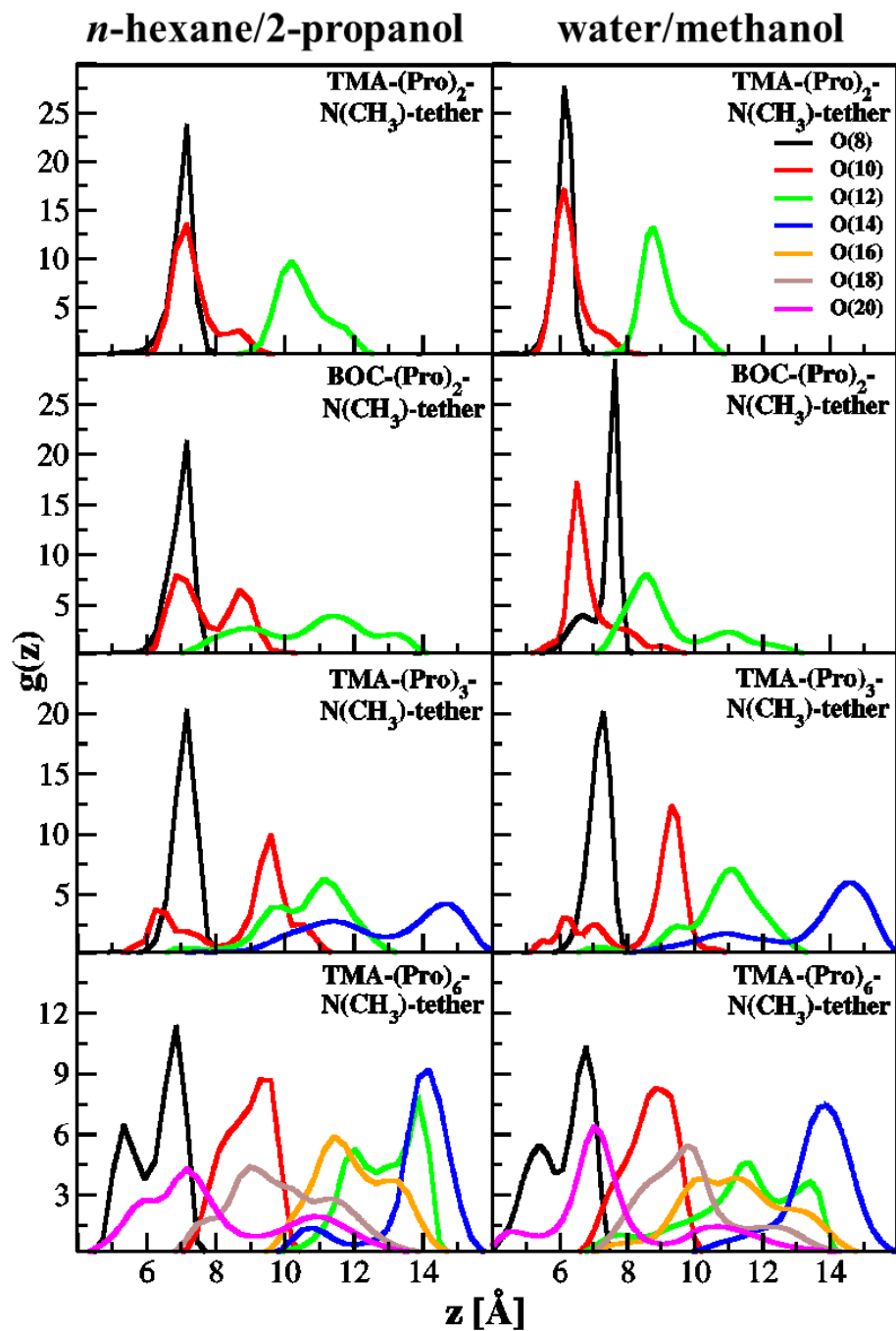


Figure 6-3. Surface distributions for backbone carbonyl oxygens. The poly-proline selectors are identical in each panel. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The color identification for backbone oxygens is shown on top right panel.

For AnC-CH₃ analytes, the $g(z)$ distributions are shown in Figure 6-4. These enantiomers can form a hydrogen bond between the alcohol group and the backbone carbonyl oxygens of selectors. Both enantiomers show similar distributions in water/methanol for all selectors: They are both concentrated and depleted from the same areas at the interface. In *n*-hexane/2-propanol, however, they behave differently. The S enantiomer interacts more strongly with the TMA terminated di-proline and is most likely to be found within 8-12 Å of the Si layer, where carbonyl oxygens are located (see Figure 6-3). So, it is retained more than the R enantiomer. For BOC terminated di-proline, the enantiomer distributions are similar with the exception that the S enantiomer is more concentrated at 8 Å. For the TMA terminated tri-proline, the distributions are distinct and the R enantiomer has a small chance of being found at around 8 Å where O(8) and O(10) also have peaks and a much stronger probability at around 12 Å where hydrogen bonding to O(10), O(12) or O(14) can occur. For the TMA terminated hexa-proline in *n*-hexane/2-propanol, strong peaks are observed for both of the enantiomers although the S enantiomer peaks are stronger. The S enantiomer shows a high chance to be found at about 7.5 Å

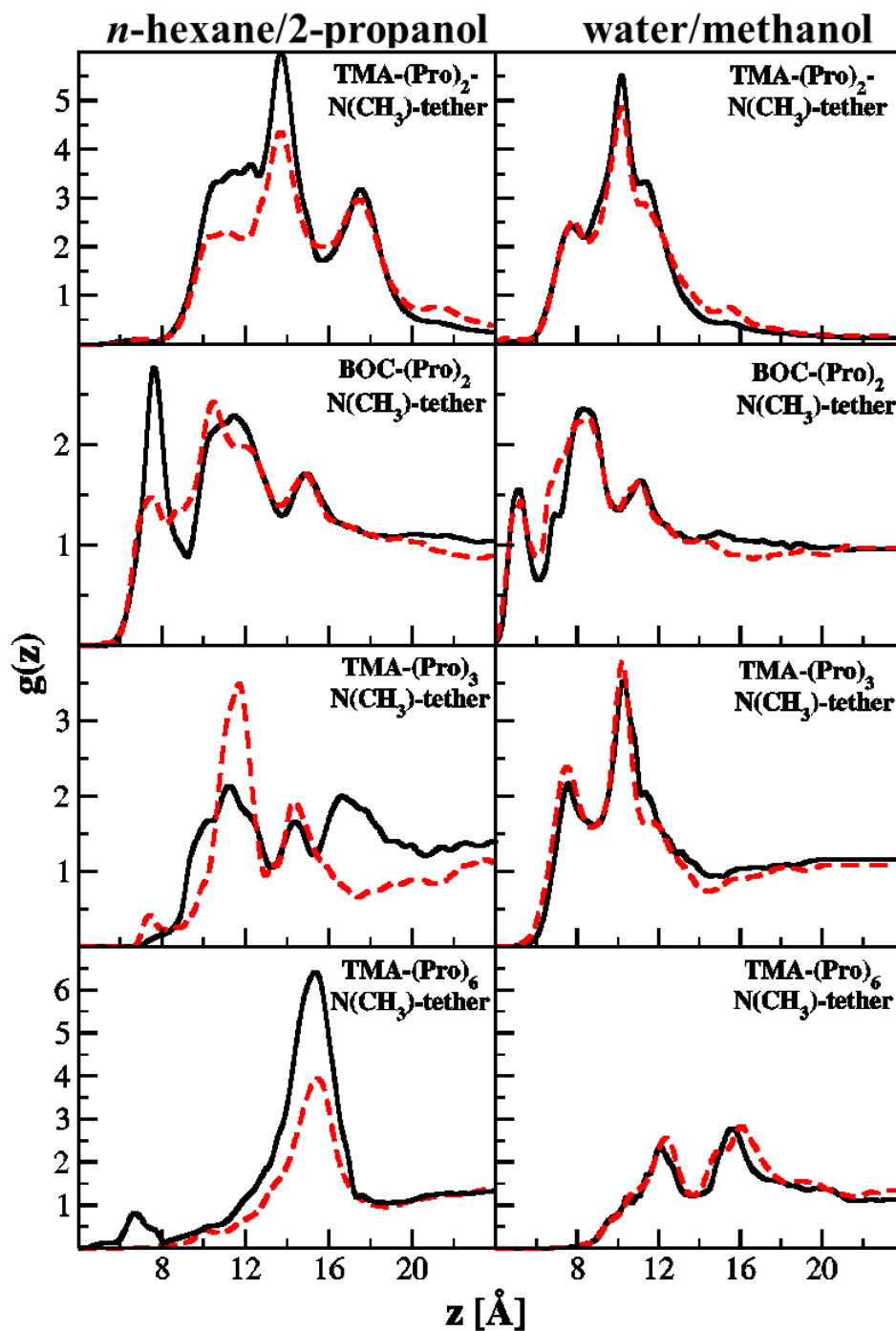


Figure 6-4. Surface distribution for the hydroxyl hydrogen of AnC-CH₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel.

where it can hydrogen bond to the O(8), O(18) or O(20) in the selector while this peak is totally absent for the R enantiomer. Both enantiomers have peaks at about 16 Å which corresponds to hydrogen bonding to any of O(12), O(14) and O(16) atoms.

AnS-CH₃ is very similar to AnC-CH₃ with a simple re-positioning of the aliphatic alcohol group to the side of the anthracene ring (see Figure 6-5). Experimentally, this reduces the selectivity factor of the TMA terminated di-proline from 1.61 to 1.1, but AnS-CH₃ can still be resolved. As shown in Figure 6-5, the hydroxyl hydrogen atoms of AnS-CH₃ have increased probabilities at 8 and 10 Å from the Si layer of the TMA terminated di-proline selector. The S enantiomer has stronger peaks at about 8 and 10 Å in *n*-hexane/2-propanol. This is the same for water/methanol with higher probability to be found closer to the surface. For the BOC terminated di-proline, the R enantiomer peak is slightly stronger at around 7.5 Å in *n*-hexane/2-propanol. The shapes of the peaks are totally different in water/methanol where at some distances the S enantiomer is more probable while for others the R enantiomer is more concentrated. For the TMA terminated tri-proline, the R enantiomer is less depleted from the surface and hydrogen bonds more with O(8) and O(10) oxygens in *n*-hexane/2-propanol. In water/methanol, both enantiomers behave very similarly. Enantiomers show distinct peaks at about 8.5 Å for TMA terminated hexa-proline in *n*-hexane/2-propanol. The R enantiomer hydrogen bonds to O(8), O(18), and O(20) more than the S enantiomer. This is not the case for water/methanol where the S enantiomer hydrogen bonds more to O(8), O(18), and O(20) compared to the R enantiomer.

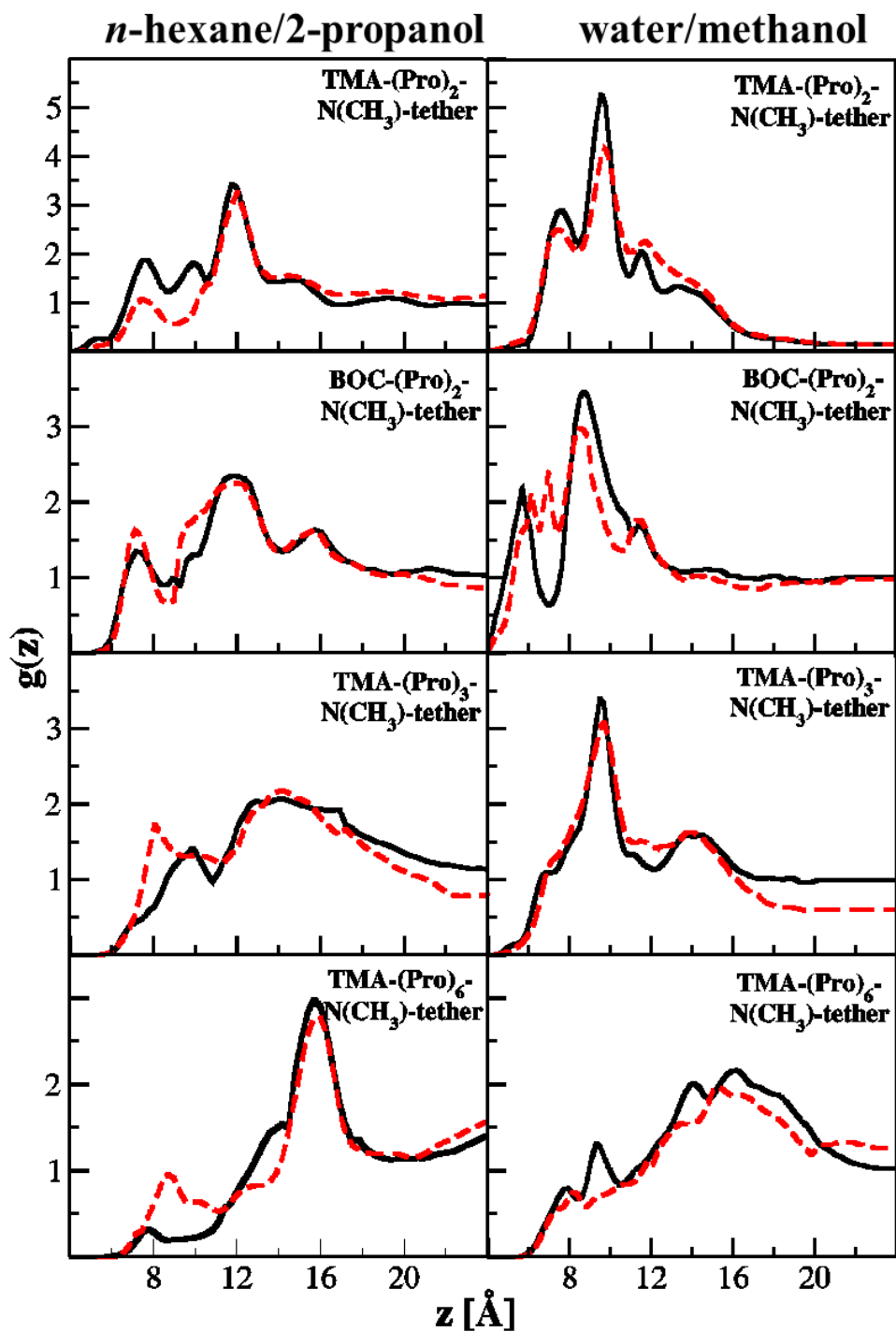


Figure 6-5. Surface distribution for the hydroxyl hydrogen of AnS-CH₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel.

Ph-CF₃ does not show any considerable difference between its enantiomer $g(z)$ distribution (see Figure 6-6). Both of the enantiomers are concentrated or depleted from the same regions. Regardless of the selector or the solvent environment, the $g(z)$ distributions are very similar for TMA and BOC terminated di-proline. The S enantiomer peak is a bit stronger at about 8 Å and it may hydrogen bond more to O(8) or O(10) compared to the R enantiomer. This difference is absent in water/methanol. For TMA terminated hexa-proline, in the presence of the *n*-hexane/2-propanol solvent, the R enantiomer is more concentrated at around 7.5 and 13 Å, but less concentrated at around 11.5 Å. In water/methanol, this difference can also be observed. For instance the R enantiomer is more likely to be found at about 7.5 Å while the S enantiomer is more concentrated at 10.5 Å from the Si layer. Overall, the observed differences are very small and do not show any preferred interactions between one of the enantiomers with the selector oxygens.

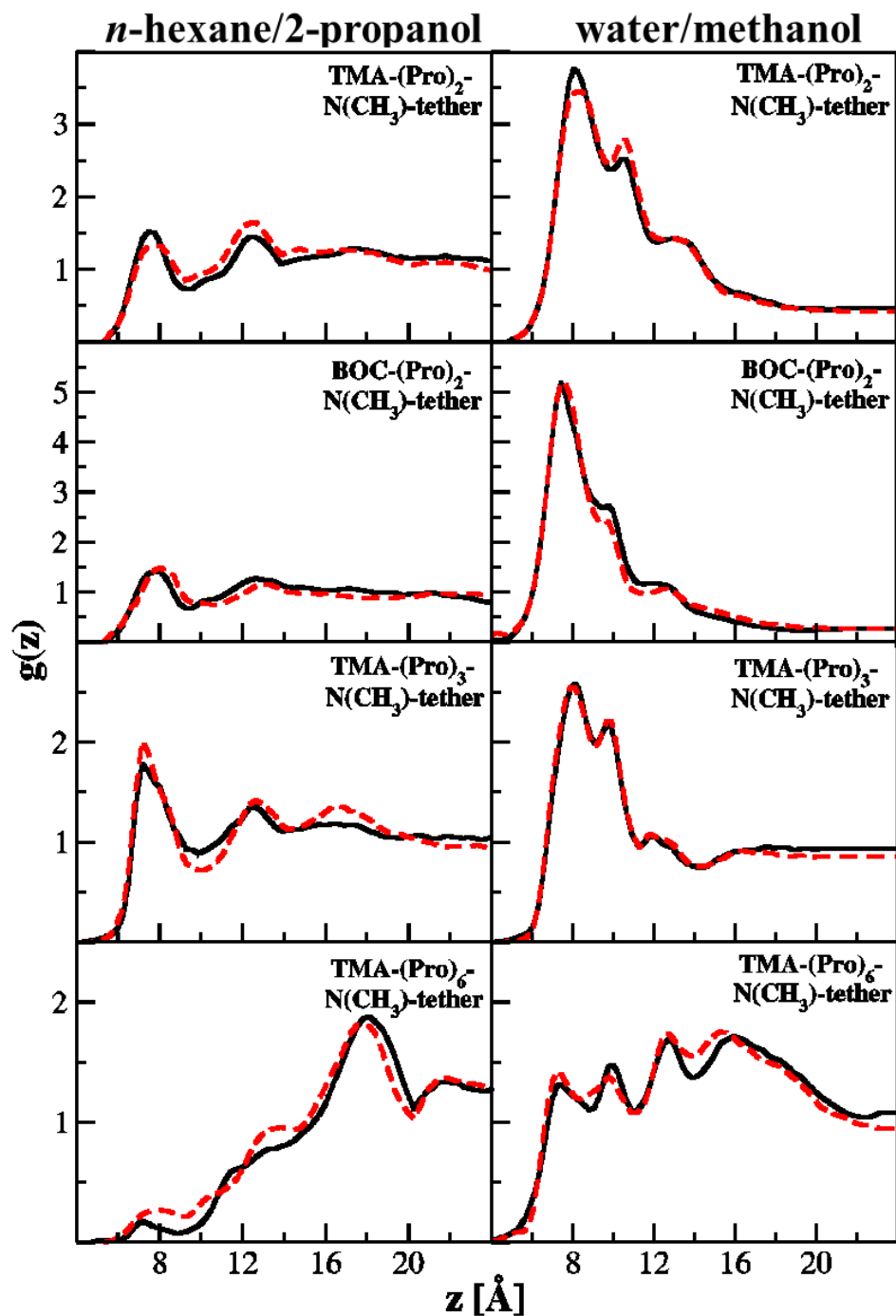


Figure 6-6. Surface distribution for the hydroxyl hydrogen of Ph-CF_3 . The left panels are for n -hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel.

Figure 6-7 shows the $g(z)$ distribution for analytes Ph-CH₃. These enantiomers differ by replacing the CF₃ group in the Ph-CF₃ by a methyl group. In *n*-hexane/2-propanol and with the TMA terminated di-proline, the R enantiomer has a higher probability to be found at about 12 Å from the surface, where O(12) is positioned. For water/methanol this has shifted to about 8-10 Å where the R enantiomer can possibly hydrogen bond to any of the carbonyl oxygens. For BOC terminated di-proline, the R enantiomer slightly prefers regions of 10-12 Å whereas the S enantiomer is more concentrated at 7-8 Å. In water/methanol, the R enantiomer is more concentrated at the interface. In the presence of the TMA terminated tri-proline, the S enantiomer has higher probability to be found at 7 and 11.5 Å while this is the reverse in the water/methanol. For TMA terminated hexa-proline, the distributions are very close at all distances from the Si layer. The same is true for water/methanol. Similar to the Ph-CF₃, the surface distributions of the enantiomers are very close.

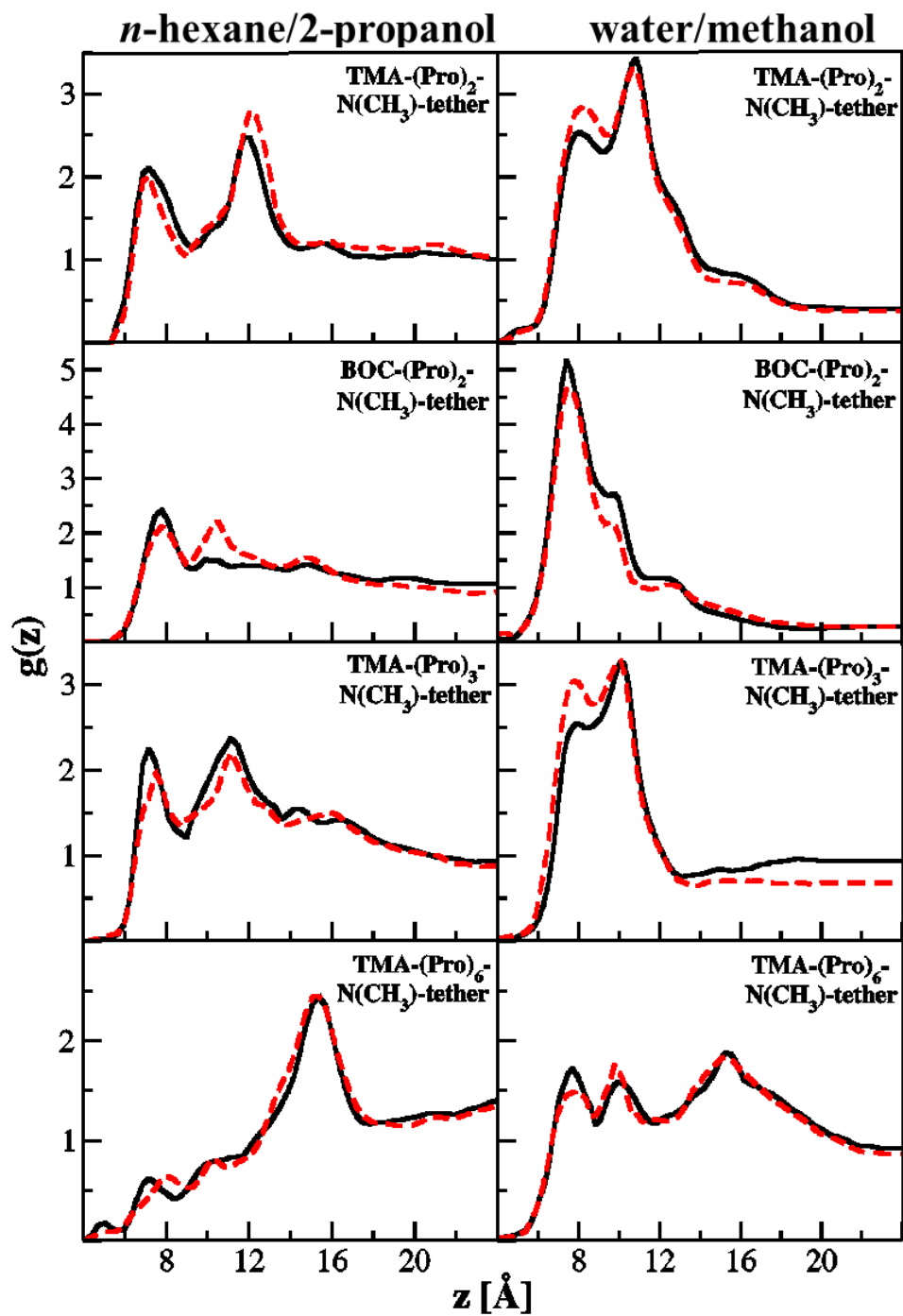


Figure 6-7. Surface distribution for the hydroxyl hydrogen of Ph-CH₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel.

For Ph-C₂H₅, the methyl group of Ph-CH₃ is replaced by an ethyl group. This replacement has very little impact as shown by the similar $g(z)$ distributions for TMA terminated di-proline in both solvents (see Figure 6-8). However, changing the terminal group to BOC brings the S enantiomer closer to the interface at 12 Å where it possibly hydrogen bonds to the O(10) and O(12) in *n*-hexane/2-propanol. In the water/methanol, the R enantiomer is more concentrated at all the regions close to the surface. For TMA terminated tri-proline the R enantiomer is slightly more probable at 12 Å in *n*-hexane/2-propanol. This changes in the water/methanol where the R enantiomer prefers regions at about 8 Å and the S enantiomer is more probable at 10 Å away from the Si layer. For TMA terminated hexa-proline, the S enantiomer has higher probability to be found at around 7.5 Å in *n*-hexane/2-propanol, but the R enantiomer is more likely at larger distances from the Si layer. The $g(z)$ distributions near the hexa-proline interface are very similar in the water/methanol solvent. Ph-C₂H₅ enantiomers do not interact differently with most CSPs at *n*-hexane/2-propanol while it offers resolution possibilities at water/methanol.

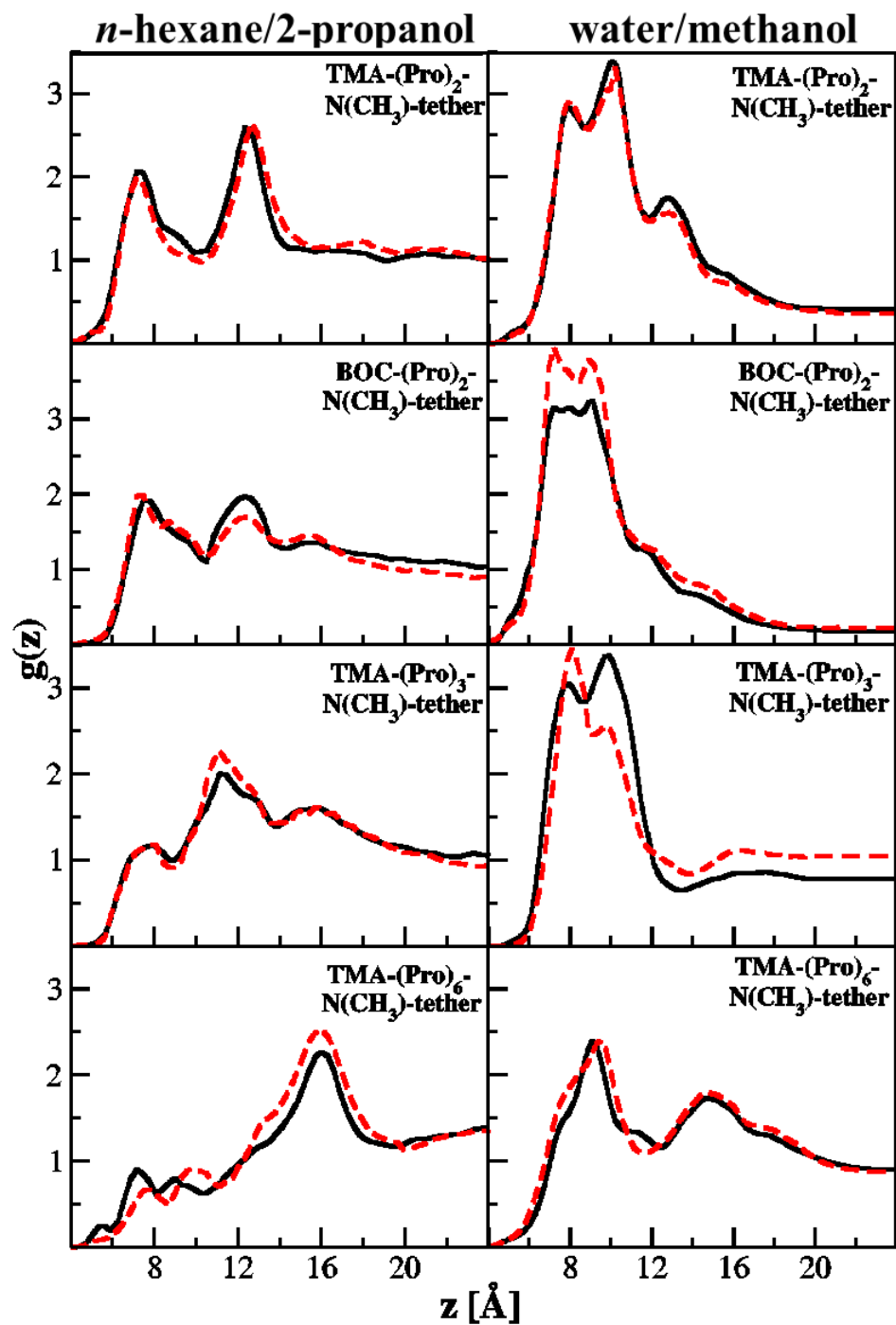


Figure 6-8. Surface distribution for the hydroxyl hydrogen of Ph-C₂H₅. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel.

The $g(z)$ distribution for AnC-CF₃ is illustrated in Figure 6-9. The analyte is structurally similar to AnC-CH₃ with replacing the methyl group with CF₃. They show distinct peaks at about 8 Å in *n*-hexane/2-propanol and at 10 Å in water/methanol with the TMA terminated di-proline. Here, the S enantiomer interacts stronger with O(8). The R enantiomer has a higher probability to be found at about 12 Å in *n*-hexane/2-propanol for TMA terminated tri-proline while both enantiomers distributions are essentially the same in water/methanol. For TMA terminated hexa-proline, the S enantiomer is strongly interacting with O(8), O(18), or O(20) at about 8 Å and with O(12), O(14) and O(16) at about 16 Å. It is expected that the S enantiomer is retained more in *n*-hexane/2-propanol. In water/methanol, the R enantiomer interacts stronger with the selector oxygens at around 8 Å.

Overall, the $g(z)$ distributions indicate that the regions of enantiomers concentration, highly match the selector carbonyl oxygen distributions, supporting the hydrogen bonding possibility between hydroxyl hydrogen of enantiomers and the carbonyl oxygens of the selector. Moreover, it shows that the interface is more crowded for TMA terminated hexa-proline as most of the enantiomers $g(z)$ distributions show a depleted regions with $z < 10$ Å.

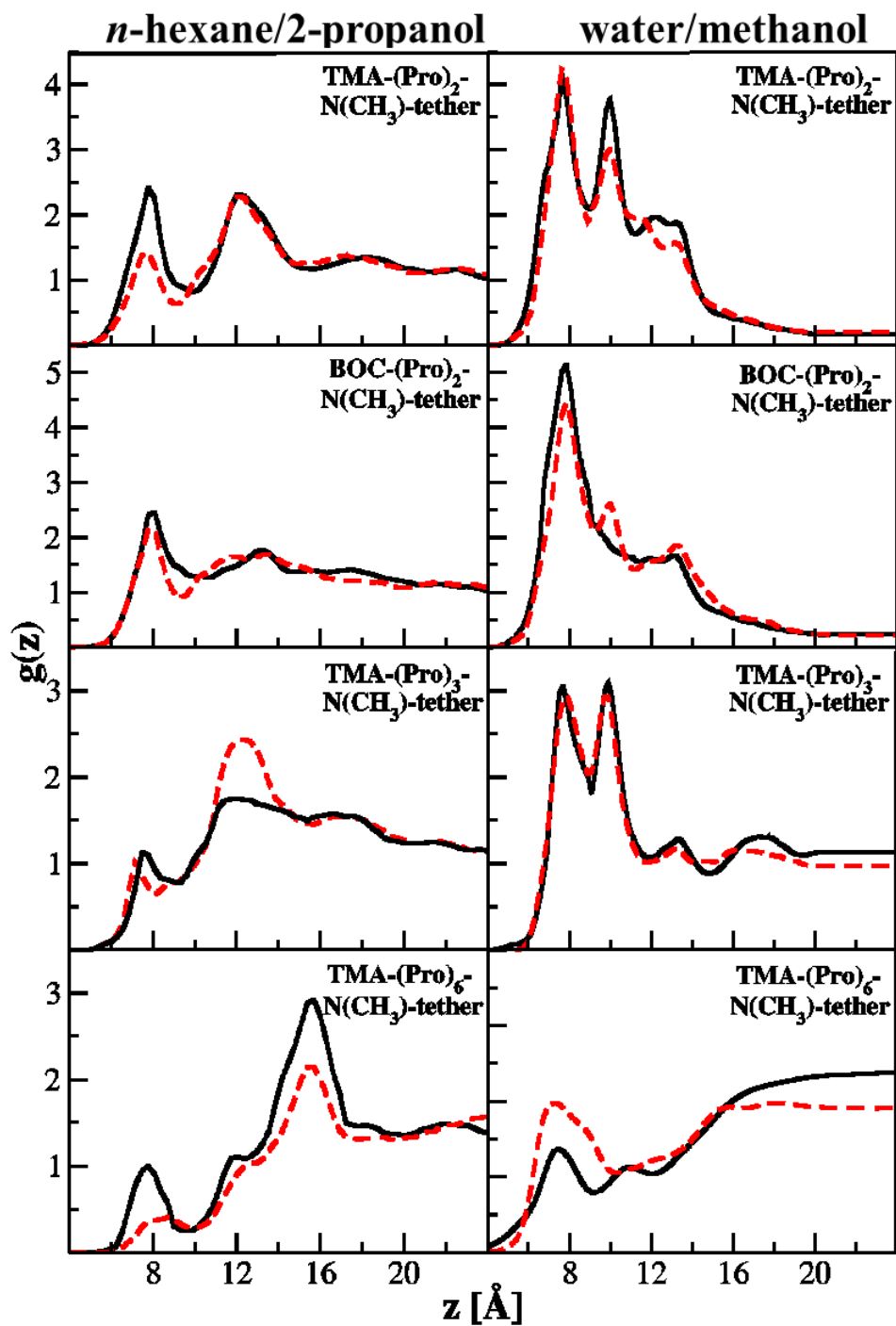


Figure 6-9. Surface distribution for the hydroxyl hydrogen of AnC-CF₃. The left panels are for *n*-hexane/2-propanol and the right panels are for water/methanol. The solid black and red-dashed lines show the S and R enantiomers. The stationary phase is indicated in each panel.

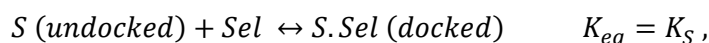
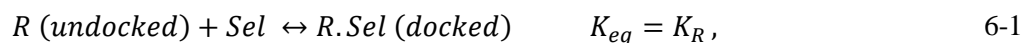
By comparing the surface distributions in *n*-hexane/2-propanol and water/methanol, it can be seen that the $g(z)$ distributions are more different for enantiomers closer to the Si layer in *n*-hexane/2-propanol. In water/methanol, the difference in enantiomers distribution is maximum when they H-bond to oxygens far from the Si layer. This is more pronounced for the analytes that are separated well. The analytes that are structurally similar show similar surface distributions. For instance, Ph-CF₃, Ph-CH₃, and Ph-C₂F₅, show close peaks at the interface.

While the surface distributions are informative, they cannot be solely used to infer about the docking mechanism and selectivity factors. The reason is that the stronger peak of one of the enantiomers at one region, can be cancelled by the stronger peak of the other enantiomer at some other regions.

6.3.2 The Selectivity

The three point attachment model is often invoked to explain enantioselectivity. For polyprolines, the selector has only hydrogen bond accepting sites and no aromatic rings. The analytes of interest can only donate one hydrogen bond, yet selectivity occurs. The enantio-selectivity occurs in the absence of three identifiable simultaneous interactions such as hydrogen bonding, π - π interactions, and steric interactions. For poly-proline selectors, hydrogen bonding of the carbonyl oxygens to the hydroxyl hydrogen is believed to be important for the resolution of enantiomers and we have chosen to use it for identification of docking events. Hydrogen bonding statistics are extracted from the snapshots collected during the simulation following the geometric criteria outlined in Sections 4.3.2 and 5.4.2.

Considering the interaction of R and S enantiomers with the selectors, the equilibria below govern enantio-separation



where K_R and K_S are the equilibrium constants for the docking process. The selectivity factor is then defined as $\alpha = \frac{K_R}{K_S}$, assuming that the R enantiomer is retained more. Then:

$$\alpha = \frac{[R]_{\text{docked}}/[R]_{\text{undocked}}}{[S]_{\text{docked}}/[S]_{\text{undocked}}}, \quad 6-2$$

Note that the selectivity factor is always greater than, or equal to, one so that the equilibrium constant for the most retained enantiomer is in the numerator. Therefore, by counting the number of “docked” arrangements from simulation snapshots, the selectivity factor can be calculated. The number of undocked enantiomers is calculated by the difference in the total number of enantiomers and the total number of docked enantiomers. A similar approach has been shown to be successful in predicting the selectivity factor for whlek-O1 stationary phase [109]. In our approach, we do not directly calculate $\Delta\Delta G$ for enantiomers to predict the selectivity factor, rather we target the equilibria in Eq. 6-1 directly. The major drawback of our method is that we need to define a criterion to differentiate between docked and undocked arrangements. Often, this criterion is based on geometric criteria for a specific type of well-known interaction such as hydrogen bonding or π - π stacking. The simultaneous presence of hydrogen bonding and π - π stack was used for whlek-O1 [109]. For poly-prolines, we use only one criterion, an H-bond, since others are not easily defined.

Table 6-1 summarizes the hydrogen bonding statistics for AnC-CH₃. The table shows the probability for specified selector oxygens to hydrogen bond with analyte. Note that the percentages of hydrogen bonding to the selector oxygens (columns 2-8 in Table 6-1) are scaled by 10 for simplicity. For instance, 0.63% of total O(8) of TMA-hexa-proline hydrogen bond to

the S enantiomer while this number is zero for the R enantiomer. The ninth column shows the percentage of analytes with a hydrogen bond. By our definition of a docking event, this corresponds to docked enantiomers. These statistics are then used to calculate the selectivity factor. According to the predicted selectivity factors, racemic AnC-CH₃ is resolved on all of the CSPs and in the presence of either solvent. As have been previously discussed in Chapters 4 and 5, the carbonyl oxygens hydrogen bond differently: the contribution of each carbonyl oxygen to the total hydrogen bonding depends on its relative orientation with respect to the Si layer. Carbonyl groups that are facing the surface hydrogen bond less due to the crowding. For carbonyl groups that are facing the bulk, and therefore are more accessible, hydrogen bonding is not sterically limited. The same pattern can be seen for analyte hydrogen bonding. For instance, O(8) and O(12) hydrogen bond much more than the O(10) for TMA-di-proline in *n*-hexane/2-propanol: Enantio-separation mostly occurs at O(8) where the S enantiomer hydrogen bonds almost twice as frequently as the R enantiomer. The predicted selectivity factor is 1.40 which is close to the experimental selectivity factor measured at 95/5 v/v solution of *n*-hexane/2-propanol. The predicted selectivity factor is closer to one in water/methanol due to the change in the conformational preferences of the TMA terminated di-proline and also water/methanol hydrogen bonding to the selector reduces the possibility for the enantiomers to hydrogen bond to the selector. Changing the terminal group to the BOC, reduces the selectivity factor to 1.28 in the *n*-hexane/2-propanol but it does not change the selectivity factor in water/methanol.

Table 6-1. Hydrogen bonding statistics and selectivity factors for analyte AnC-CH₃. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.69% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.28% of them hydrogen bond to the S enantiomer.. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].

	O(8)	O(10)	O(12)	O(14)	O(16)	O(18)	O(20)	%H	α_{pre} (α_{exp})
<i>n</i>-hexane/2-propanol									
TMA-di-proline									1.40 ± 0.13 (1.61)
S	0.69	0.04	0.58					0.52	
R	0.28	0.06	0.6					0.38	
BOC-di-proline									1.28 ± 0.10
S	0.43	0.2	0.68					0.54	
R	0.31	0.14	0.54					0.42	
TMA-tri-proline									1.86 ± 0.20
S	0.64	0.6	0.45	0.04				0.67	
R	0.68	1.41	0.73	0.06				1.23	
TMA-hexa-proline									1.92 ± 0.08 (2.60)
S	0.63	0.0	1.98	2.92	0.26	0.66	0.48	2.70	
R	0.0	0.0	0.95	2.26	0.07	0.25	0.08	1.41	
water/methanol									
TMA-di-proline									1.11 ± 0.19
S	0.14	0.04	0.32					0.19	
R	0.13	0.04	0.27					0.17	
BOC-di-proline									1.12 ± 0.13
S	0.19	0.16	0.2					0.22	
R	0.17	0.14	0.18					0.19	
TMA-tri-proline									1.08 ± 0.14
S	0.19	0.24	0.31	0.02				0.29	
R	0.21	0.25	0.34	0.02				0.32	
TMA-hexa-proline									1.32 ± 0.10
S	0.0	0.0	0.4	1.18	0.0	0.22	0.12	0.88	
R	0.0	0.0	0.31	0.72	0.71	0.16	0.53	0.66	

For BOC terminated di-proline in *n*-hexane/2-propanol, O(8) plays an important role in enantio-separation. O(8) hydrogen bonds more frequently with the S enantiomer. In water/methanol, though, both enantiomers interact almost equally with all carbonyl oxygens. For TMA terminated tri-proline, in *n*-hexane/2-propanol, O(10) and O(12) interact preferentially with

the R enantiomers. However, in water/methanol, all carbonyl oxygens interact almost equally with both enantiomers. For TMA terminated hexa-proline in *n*-hexane/2-propanol, O(8) selectively hydrogen bonds to the S enantiomer. O(10) is inactive and does not hydrogen bond to either of the analytes during the simulation. O(12)-O(16) hydrogen bond more to the S enantiomer. The predicted selectivity factor is 1.92 [47]. In water/methanol, O(8) and O(10) are inactive but other oxygens still H-bond to the AnC-CH₃ enantiomers. The predicted selectivity factor is reduced in water/methanol to 1.32.

The ninth column shows the percentage of each analyte that hydrogen bonds during the simulation. As expected, a lower hydrogen bonding probability is evident between the analyte and the selector in water/methanol for all analytes. Moreover, by increasing the number of selector oxygens, from di- to hexa-proline, more hydrogen bonding occurs between the analytes and the selectors.

Hydrogen bonding statistics for AnS-CH₃ are summarized in Table 6-2. For TMA terminated di-proline, the S enantiomer hydrogen bonds more to O(8) and O(12) and less to O(10). In total, though, the S enantiomer is retained more. Enantio-separation mainly occurs at O(8). The predicted selectivity factor is 1.21 which is close to the experimental value of 1.10 [47] measured in the 95/5 v/v solution of *n*-hexane/2-propanol. The selectivity factor increases considerably in water/methanol where the S enantiomer hydrogen bonds almost twice as frequently at both O(8) and O(12) compared to the R enantiomer. When using the BOC terminated di-proline selector, neither of the solvents results in a successful separation of the enantiomers. In *n*-hexane/2-propanol, for instance, the R enantiomer hydrogen bonds more to the O(8) and less to O(12) so that in total, both enantiomers have comparable hydrogen bond

Table 6-2. Hydrogen bonding statistics and selectivity factors for analyte AnS-CH₃. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.81% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.57% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].

	O(8)	O(10)	O(12)	O(14)	O(16)	O(18)	O(20)	%H	α_{pre} (α_{exp})
<i>n</i>-hexane/2-propanol									
TMA-di-proline									1.21 ± 0.11 (1.10)
S	0.81	0.07	0.87					0.7	
R	0.57	0.13	0.74					0.57	
BOC-di-proline									1.01 ± 0.10
S	0.24	0.09	0.5					0.34	
R	0.3	0.09	0.45					0.35	
TMA-tri-proline									1.16 ± 0.11
S	0.34	0.63	0.78	0.06				0.69	
R	0.64	0.77	0.69	0.04				0.82	
TMA-hexa-proline									1.09 ± 0.14 (1.00)
S	0.01	0.01	0.65	1.93	0.07	0.22	0.18	1.20	
R	0.07	0.01	0.55	1.72	0.05	0.17	0.22	1.08	
water/methanol									
TMA-di-proline									2.18 ± 0.16
S	0.42	0.6	0.42					0.35	
R	0.16	0.04	0.21					0.16	
BOC-di-proline									1.08 ± 0.12
S	0.14	0.03	0.26					0.17	
R	0.1	0.18	0.11					0.16	
TMA-tri-proline									1.05 ± 0.12
S	0.37	0.44	0.27	0.02				0.42	
R	0.39	0.46	0.27	0.04				0.45	
TMA-hexa-proline									1.58 ± 0.12
S	0.0	0.01	0.39	1.02	0.09	0.34	0.26	0.72	
R	0.0	0.01	0.03	0.34	0.09	0.14	0.11	0.44	

probabilities. In water/methanol, the S enantiomer hydrogen bonds more at O(8) and O(12) and less at O(10), resulting in the predicted selectivity factor of 1.0 to within statistics. TMA terminated tri-proline, does not perform as well as TMA terminated di-proline. In *n*-hexane/2-propanol, the predicted selectivity factor is 1.16 since the R enantiomer hydrogen bonds more to

O(8) and then O(10). The predicted selectivity factor with TMA terminated hexa-proline in *n*-hexane/2-propanol is 1.0 to within statistics which is close to the experimental value of 1.00 [47]. In water/methanol, the selectivity factor increases to 1.58, where O(12), O(14), O(18), and O(20) all hydrogen bond more to the S enantiomer. By comparing the total hydrogen bonding events of AnS-CH₃ with AnC-CH₃, it can be seen that overall hydrogen bonding is comparable when the alcohol group is attached to the middle carbon in the anthracene ring.

Ph-CF₃ enantiomers hydrogen bond less to the carbonyl oxygens (see Table 6-3). Intramolecular hydrogen bonding between fluorine and the hydroxyl hydrogen is possible and this effectively reduces the inter-molecular hydrogen bonding with the selector. As a result, hydrogen bonding counts are lesser than for the other analytes. This also causes the selectivity factor to be close to unity, regardless of the selector and the solvent.

By changing the CF₃ group to a methyl yielding the analyte Ph-CH₃ to methyl group, hydrogen bonding probability increases: Ph-CH₃ hydrogen bonds much more than Ph-CF₃ (see Table 6-4). For some combination of the selector and solvent, various backbone oxygens show difference in hydrogen bonding frequency to the R and S enantiomers of Ph-CH₃. Generally, to within our uncertainties, the selectivity factors are predicted to be unity. The one exception is the BOC terminated selector in water/methanol where the S enantiomer interacts more strongly with all three backbone oxygens leading to a predicted selectivity factor of 1.18.

Table 6-3. Hydrogen bonding statistics and selectivity factors for analyte Ph-CF₃. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.7% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.7% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].

	O(8)	O(10)	O(12)	O(14)	O(16)	O(18)	O(20)	%H	α_{pre} (α_{exp})
<i>n</i>-hexane/2-propanol									
TMA-di-proline									1.03 ± 0.16 (1.06)
S	0.07	0.0	0.04					0.04	
R	0.07	0.01	0.04					0.04	
BOC-di-proline									1.06 ± 0.14
S	0.08	0.04	0.04					0.06	
R	0.08	0.04	0.04					0.07	
TMA-tri-proline									1.15 ± 0.10
S	0.07	0.09	0.03	0.01				0.07	
R	0.05	0.06	0.04	0.01				0.06	
TMA-hexa-proline									1.00 ± 0.15 (1.10)
S	0.03	0.0	0.14	0.13	0.01	0.05	0.02	0.15	
R	0.02	0.0	0.12	0.13	0.01	0.07	0.03	0.15	
water/methanol									
TMA-di-proline									1.00 ± 0.18
S	0.07	0.0	0.01					0.03	
R	0.06	0.01	0.01					0.03	
BOC-di-proline									1.02 ± 0.18
S	0.06	0.02	0.01					0.03	
R	0.06	0.02	0.02					0.03	
TMA-tri-proline									1.06 ± 0.15
S	0.06	0.01	0.02	0.0				0.03	
R	0.06	0.01	0.02	0.0				0.04	
TMA-hexa-proline									1.02 ± 0.16
S	0.02	0.0	0.08	0.07	0.0	0.03	0.01	0.07	
R	0.01	0.0	0.07	0.07	0.01	0.04	0.02	0.07	

The hydrogen bonding statistics and selectivity factors are provided in Table 6-5 for Ph-C₂H₅. With one exception, the selectivity factors are predicted to be unity, within statistics. The exception is BOC-di-proline in water/methanol. Here, the R enantiomer interacts more strongly

with all three backbone oxygens. In fact, this combination also leads to a predicted distribution between the enantiomers of Ph-CH₃.

Table 6-4. Hydrogen bonding statistics and selectivity factors for analyte Ph-CH₃. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.66% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.66% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].

	O(8)	O(10)	O(12)	O(14)	O(16)	O(18)	O(20)	%H	$\alpha_{\text{pre}} (\alpha_{\text{exp}})$
<i>n</i>-hexane/2-propanol									
TMA-di-proline									1.13 ± 0.15 (1.10)
S	0.66	0.50	0.76					0.59	
R	0.66	0.13	0.9					0.68	
BOC-di-proline									1.09 ± 0.14
S	0.3	0.14	0.35					0.32	
R	0.27	0.18	0.41					0.35	
TMA-tri-proline									1.08 ± 0.19
S	0.58	0.70	0.67	0.10				0.79	
R	0.53	0.63	0.61	0.13				0.73	
TMA-hexa-proline									1.08 ± 0.16 (1.00)
S	0.21	0.02	0.89	1.06	0.07	0.22	0.24	1.05	
R	0.16	0.01	1.0	1.2	0.05	0.36	0.17	1.15	
water/methanol									
TMA-di-proline									1.12 ± 0.18
S	0.57	0.12	0.26					0.37	
R	0.5	0.09	0.25					0.31	
BOC-di-proline									1.18 ± 0.13
S	0.3	0.09	0.1					0.19	
R	0.28	0.07	0.07					0.16	
TMA-tri-proline									1.11 ± 0.12
S	0.31	0.34	0.2	0.01				0.38	
R	0.33	0.41	0.21	0.02				0.38	
TMA-hexa-proline									1.00 ± 0.17
S	0.26	0.03	0.26	0.51	0.01	0.44	0.3	0.62	
R	0.36	0.02	0.34	0.47	0.02	0.44	0.18	0.62	

Table 6-5. Hydrogen bonding statistics and selectivity factors for analyte Ph-C₂H₅. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.42% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.44% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].

	O(8)	O(10)	O(12)	O(14)	O(16)	O(18)	O(20)	%H	$\alpha_{\text{pre}} (\alpha_{\text{exp}})$
<i>n</i>-hexane/2-propanol									
TMA-di-proline									1.07 ± 0.12 (1.10)
S	0.42	0.17	0.83					0.57	
R	0.44	0.2	0.89					0.61	
BOC-di-proline									1.07 ± 0.10
S	0.36	0.21	0.22					0.33	
R	0.38	0.22	0.25					0.35	
TMA-tri-proline									1.02 ± 0.20
S	0.47	0.62	0.7	0.13				0.74	
R	0.54	0.57	0.75	0.1				0.75	
TMA-hexa-proline									1.12 ± 0.16 (1.0)
S	0.14	0.01	0.77	0.62	0.09	0.37	0.14	0.83	
R	0.05	0.0	1.08	0.73	0.10	0.29	0.18	0.95	
water/methanol									
TMA-di-proline									1.05 ± 0.13
S	0.6	0.06	0.17					0.32	
R	0.71	0.04	0.13					0.34	
BOC-di-proline									1.36 ± 0.13
S	0.34	0.08	0.04					0.18	
R	0.39	0.15	0.09					0.25	
TMA-tri-proline									1.04 ± 0.17
S	0.47	0.35	0.52	0.04				0.53	
R	0.53	0.31	0.41	0.07				0.51	
TMA-hexa-proline									1.07 ± 0.16
S	0.07	0.0	0.37	0.3	0.04	0.18	0.07	0.35	
R	0.07	0.0	0.38	0.31	0.09	0.21	0.05	0.34	

AnC-CF₃ enantiomers are resolved on most of the selectors in this study (see Table 6-6).

The predicted selectivity factor of 1.46 on TMA terminated di-proline and in *n*-hexane/2-propanol is in good agreement with the experimental value of 1.51 [47] is obtained. The enantio-separation occurs at O(8) where the S enantiomer hydrogen bonds almost twice as frequently as

the R enantiomer. In water/methanol, enantio-separation occurs at O(12) where the S enantiomer hydrogen bonds more. Both enantiomers hydrogen bond to O(8) more than the other carbonyl oxygens combined, but this interaction is not selective. TMA terminated tri-proline resolves the enantiomers well in the *n*-hexane/2-propanol where O(10) and O(12) hydrogen bonds better to the R enantiomer in *n*-hexane/2-propanol. In water/methanol all of the selector oxygens interact similarly with both enantiomers. The predicted selectivity factor for TMA terminated hexa-proline in *n*-hexane/2-propanol and water/methanol are 1.77 and 1.00, respectively. In *n*-hexane/2-propanol, O(8) and O(12) interacts more to the S enantiomer and O(10) is inactive. In water/methanol hydrogen bonding probabilities are quite close for all the backbone oxygens.

Figure 6-10 compares the predicted and the experimental selectivity factors. As can be observed, excellent agreement is achieved in most of the cases where the experimental selectivity factor is within the range of the error of the predicted selectivity factor. Highest deviations are observed for higher values of experimental selectivity factor. As will be discussed in Section 6.3.3, this originates from the effect of other mechanisms in the enantio-resolution where double hydrogen bonding occurs.

The uncertainties in the prediction of selectivity factors ($\Delta\alpha$) provide a guide and context for the calculated selectivity factors. The values of $\Delta\alpha$ are calculated following the block averaging procedure of Flyberg and Petersen [238,239]. Here each simulation trajectory is divided into 1, 3, 6, 9, 12, 15, 18, and 21 blocks and for each block, the standard deviation in the selectivity factor, $\Delta\alpha_b$, is calculated. The best estimate of the overall uncertainty, $\Delta\alpha$, is achieved when the plot of $(\Delta\alpha_b \times \sqrt{\# \text{ blocks}})$ against the number of blocks reaches a plateau as shown in Figure 6-11. These plots are employed to estimate the uncertainties listed in Tables 6-1 to 6-6. In some instances, the plateau is not well formed and the line still has a small but non-zero slope.

We have extracted our estimate of $\Delta\alpha$ from the small slope region. This occurs for combination of analytes, solvents, and selectors where the H-bonding probabilities are low.

Table 6-6. Hydrogen bonding statistics and selectivity factors for analyte AnC-CF₃. Columns 2-8 show hydrogen bonding statistics for each selector oxygen. For example, 0.46% of O(8) oxygens of TMA-di-proline in *n*-hexane/2-propanol hydrogen bond to the R enantiomer while 0.2% of them hydrogen bond to the S enantiomer. Column 9 shows the percentage of enantiomers that have a hydrogen bonding. The last column shows the predicted selectivity factors and, when available, experimental factors [47].

	O(8)	O(10)	O(12)	O(14)	O(16)	O(18)	O(20)	%H	α_{pre} (α_{exp})
<i>n</i>-hexane/2-propanol									
TMA-di-proline									1.46 ± 0.10 (1.51)
S	0.46	0.02	0.34					0.33	
R	0.2	0.03	0.33					0.22	
BOC-di-proline									1.13 ± 0.14
S	0.2	0.09	0.19					0.19	
R	0.17	0.1	0.14					0.17	
TMA-tri-proline									1.52 ± 0.11
S	0.27	0.32	0.3	0.01				0.35	
R	0.26	0.55	0.51	0.09				0.54	
TMA-hexa-proline									1.77 ± 0.09 (2.08)
S	0.45	0.0	0.9	0.68	0.11	0.01	0.01	0.84	
R	0.04	0.0	0.32	0.62	0.03	0.09	0.07	0.46	
water/methanol									
TMA-di-proline									1.15 ± 0.12
S	0.8	0.03	0.08					0.15	
R	0.29	0.01	0.04					0.13	
BOC-di-proline									1.16 ± 0.11
S	0.22	0.03	0.06					0.13	
R	0.17	0.06	0.04					0.11	
TMA-tri-proline									1.03 ± 0.18
S	0.23	0.26	0.15	0.03				0.26	
R	0.25	0.28	0.11	0.04				0.27	
TMA-hexa-proline									1.00 ± 0.13
S	0.03	0.04	0.31	0.38	0.30	0.14	0.01	0.41	
R	0.03	0.03	0.33	0.35	0.26	0.15	0.04	0.41	

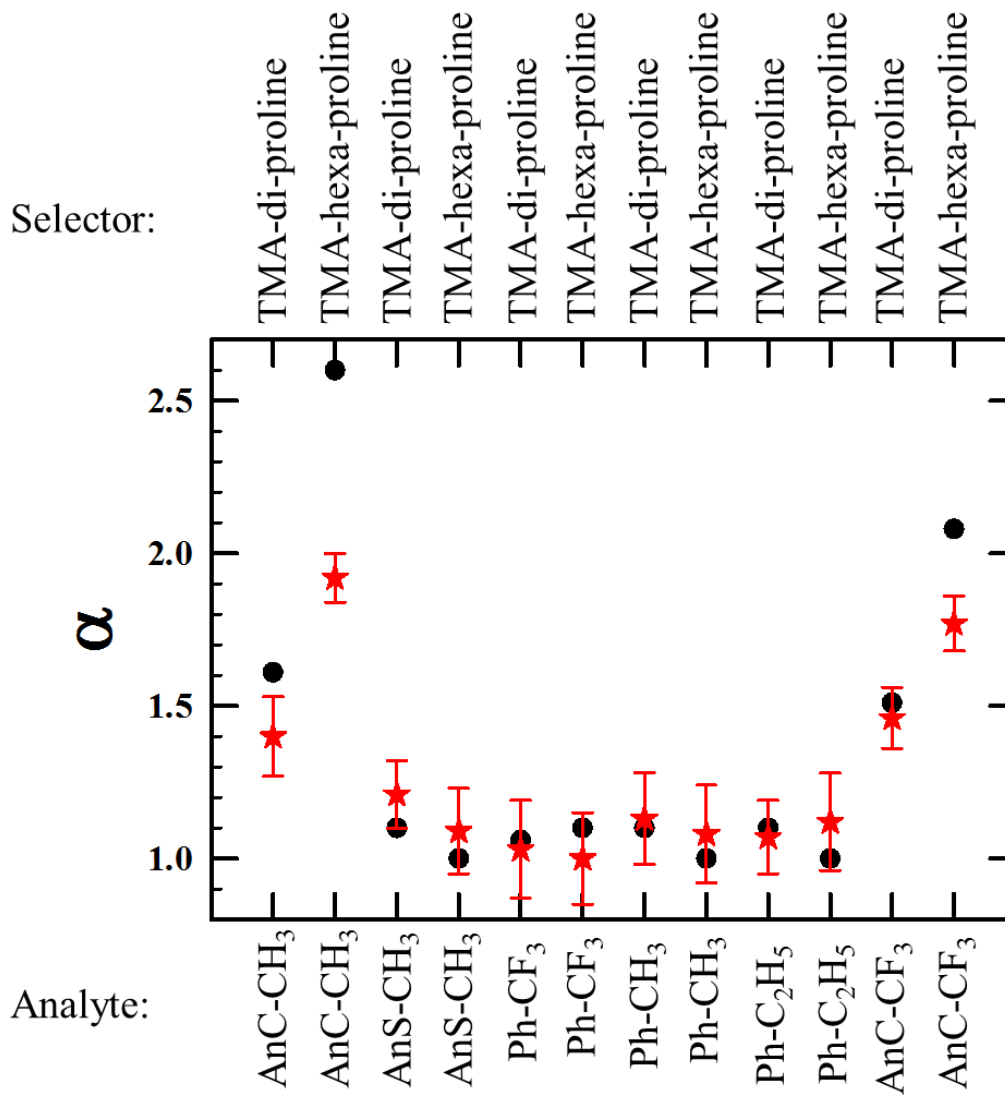


Figure 6-10. Predicted (red stars and error bars) and experimental (black circles) selectivity factors for di- and hexa-proline CSPs with six analytes. All selectivity factors are in *n*-hexane/2-propanol solvent. No experimental error is reported for the experimental selectivity factors. The selectors and analytes are shown on the top and bottom of the graph.

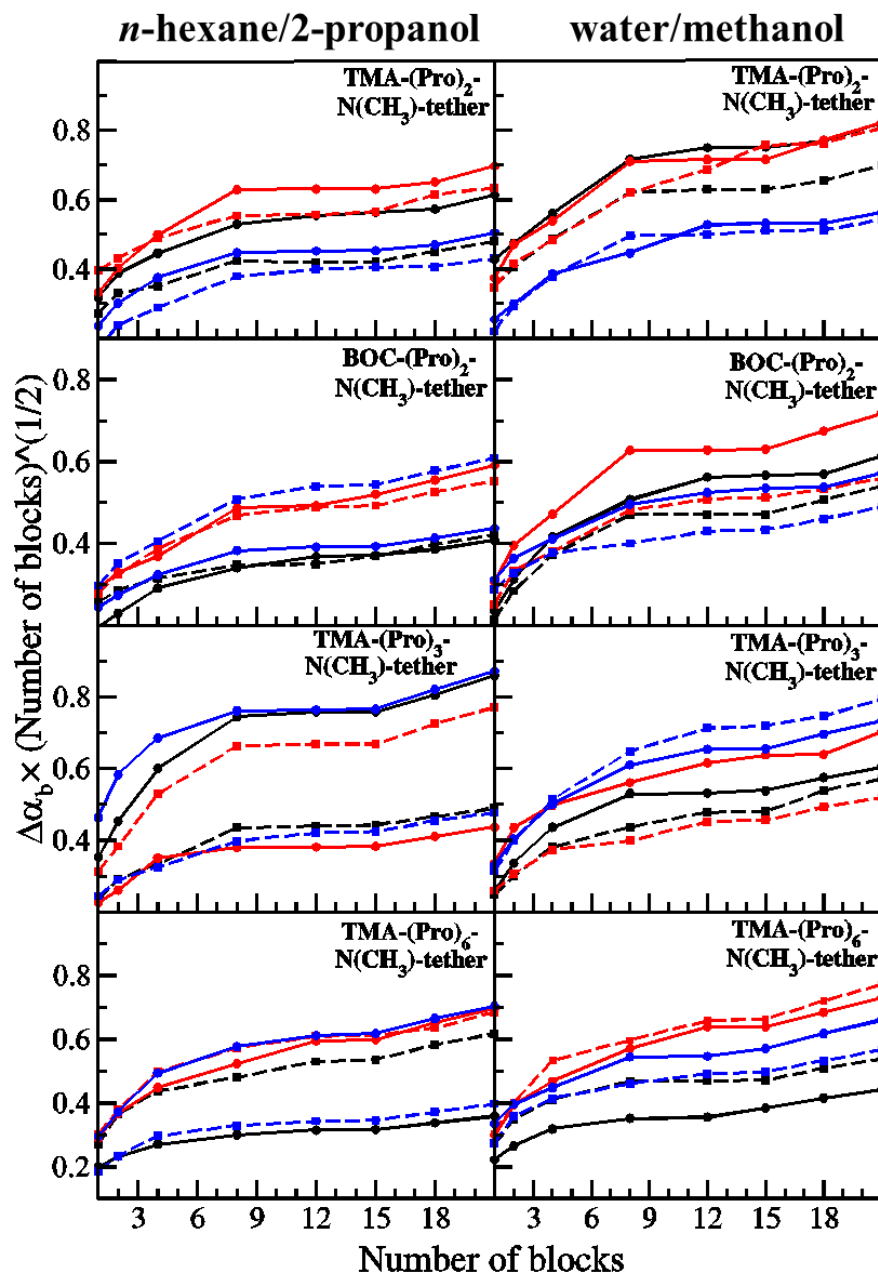


Figure 6-11. Uncertainty calculations for selectivity factors. The black solid line with filled circles is for AnC-CH₃, the black dashed line with filled squares is for AnS-CH₃, the red solid line with filled circles is for Ph-CF₃, the red dashed line with filled squares is for Ph-CH₃, the blue solid line with filled circles is for Ph-C₂H₅, the blue dashed line with filled squares is for AnC-CF₃. The selector is identified within each panel and the right and left hand panels correspond to water/methanol and *n*-hexane/2-propanol solvents.

6.3.3 The Docking Mechanism

The poly-proline selectors offer several hydrogen bond accepting sites that can hydrogen bond to the enantiomers. They also have distinct conformational preferences that differ in the presence of *n*-hexane/2-propanol and water/methanol. Hydrogen bonding is expected to be important for enantio-separation. Among the other factors playing role in the enantio-separation, steric hindrance is also assumed to play an important role. In this Section, we study the docking arrangements of the analytes and try to identify differences between enantiomers.

In this Section, we examine the mechanism for enantio-selectivity with an emphasis on the hexa-proline selector in *n*-hexane/2-propanol. Later in this section, we discuss solvent-selector combinations. Several mechanisms are responsible for enantio-resolution on poly-proline CSPs. Here, we focus on more important mechanisms that in the order of importance are: Hydrogen bonding close to the Si layer, double hydrogen bonding, and bulk hydrogen bonding.

By looking at hydrogen bonding statistics in Tables 6-1 to 6-6, it can be seen that different carbonyl oxygens hydrogen bond differently to the enantiomers. For instance, for the separation of AnC-CH₃ enantiomers on the hexa-proline selector in the presence of *n*-hexane/2-propanol, O(8) interacts efficiently with the S enantiomer while O(10) is inactive. The number of hydrogen bonds is high for O(12) and O(14) and is low for O(16) to O(20). This pattern is usually consistent for both enantiomers. To be more precise, O(12) and O(14) that are facing the bulk, easily hydrogen bond to both enantiomers, while hydrogen bonding with oxygens closer to the surface (such as O(8), O(10), O(18) and O(20)) is less probable because of interface crowding. However, in most cases, O(10), O(12), and O(14) do not selectively hydrogen bond to enantiomers. The oxygens closer to the surface, in contrast, tend to be more selective. This shows that surface crowding is an important issue in enantiomer docking. The interface is more crowded

for longer chains as discussed (see Figure 5-9). In order to hydrogen bond with carbonyls located near the Si layer, analytes must be positioned in between the bulky selectors.

Extensive snapshot analysis showed that the enantio-separation comes from two major interactions: hydrogen bonding and steric hindrance. Analyte flexibility can reduce the selectivity since more flexible analytes can adopt variety of conformations. To this end, the selectivity factors are understandable in terms of the torsional energy barrier between the ring and the chiral carbon. In this sense, AnC-CH₃ and AnC-CF₃ have the highest energy barriers (30 and 36 kJ/mol, respectively), AnS-CH₃, Ph-CH₃, and Ph-C₂H₅ have moderate energy barriers (15 kJ/mol) and Ph-CF₃ has the smallest energy barrier (4 kJ/mol). A low energy barrier for rotation may reduce the selectivity by providing conformational flexibility to the analyte.

To better understand the most important docking mechanisms, consider the separation of AnC-CH₃ on the TMA-hexa-proline CSP in *n*-hexane/2-propanol. The largest difference in H-bonding statistics is observed for O(20). Figure 6-12 shows docking arrangements for the R and S enantiomer when they hydrogen bond to O(20). We could have also shown docking to O(8) but the R enantiomer does not hydrogen bond at all to this carbonyl. From the top view in the right, it is clear that crowding at the interface only allows the analyte to approach the surface with the aromatic group perpendicular to the Si layer. Indeed, snapshot analysis shows that the S enantiomer approaches the surface with an angle of almost 90° (perpendicular) with respect to the Si layer.

Considering the $g(z)$ distribution of the selector oxygens (Figure 6-3) and the fact that other surface groups crowd the surface up until around 5 Å, hydrogen bonding only occurs if the aromatic ring is located above the alcoholic oxygen as can be seen from Figure 6-12. For the S enantiomer, by considering the angle of 120° for the torsion around the bond connecting the

chiral carbon to the ring, the hydrogen attached to the chiral carbon is positioned towards the surface, while the methyl group is roughly parallel to the surface, (as is the hydroxyl groups). For the R enantiomer, though, the methyl and the hydrogen are switched and this puts the methyl group very close to the trimethylsilyl end-caps. This results in unfavorable steric interactions for the R analyte. To overcome this repulsion, the torsional angle between the ring carbon and the chiral carbon adopts other values to the methyl group relative to the surface. As the surface is irregular, with different selector conformations and or orientation at various positions on the surface, surface repulsion varies with time along the surface. For these reasons, this torsion adopts a broad distribution from 0 to 360°. Another common possibility is that the S enantiomer interacts such that both the hydroxyl group and the hydrogen atom are directed towards the selector oxygen. In this case, the R enantiomer cannot hydrogen bond since replacing the positions of H and CH₃ groups will bring hydroxyl group and the selector oxygen far from each other. The same argument discussed here for O(20) also applies to other oxygens that are close to the surface, particularly for O(8).

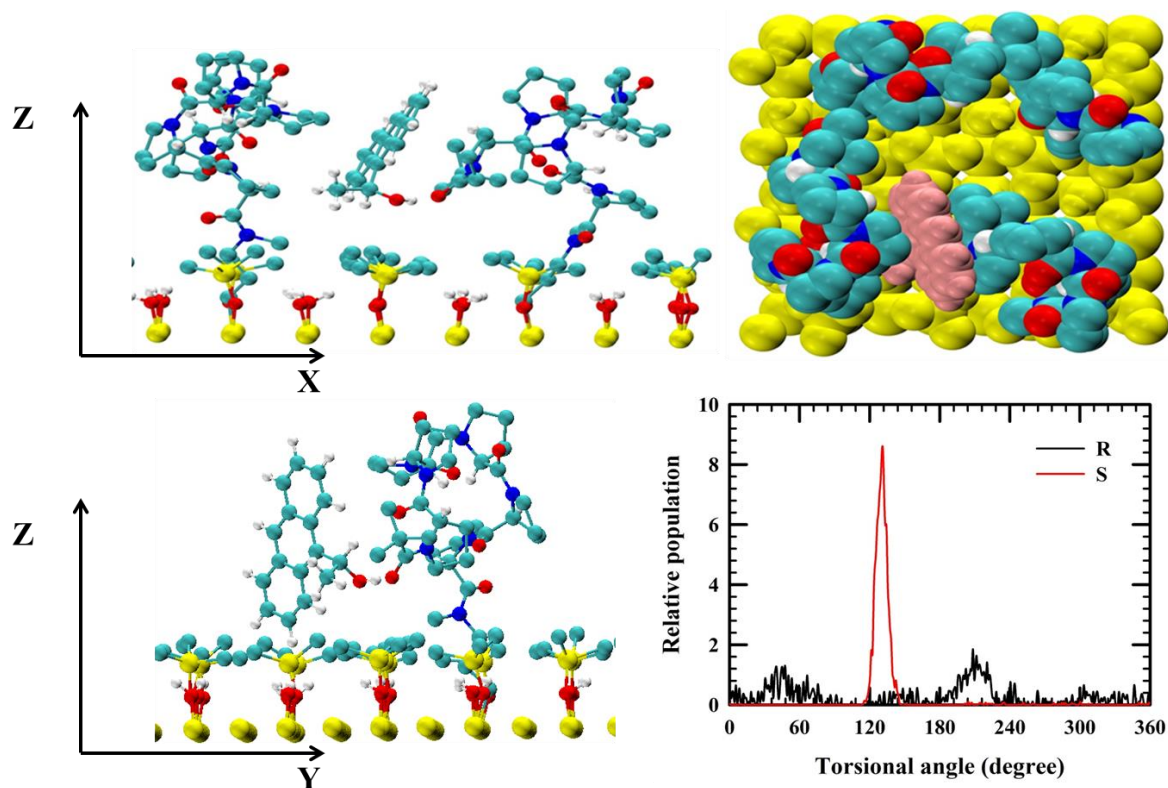


Figure 6-12. Side and top view for the docking of both enantiomers of AnC-CH₃ on TMA-hexaproline CSP. The top left picture shows the side view when the S analyte hydrogen bonds to O(20). The top right view clearly shows why the analyte (pink color) has steric constraints and can only “slide” into the surface. The lower left picture is showing the docking of the R enantiomer. The lower right graph shows the torsional distribution about the bond joining the aromatic group to the chiral carbon, for enantiomers that hydrogen bond to O(20). The equilibrium angles for the S and R enantiomers are 120 and 240°, respectively.

For AnS-CH₃, the situation is a bit different since the bulky aromatic group can be kept further from the Si layer. Snapshot analysis shows that for all docking arrangements, the hydroxyl group is close to the surface while the aromatic ring is further from the Si layer. A lower torsional barrier introduces analyte flexibility and acts to bring the overall hydrogen bonding statistics to agreement between enantiomers. In this case, the torsional distribution for the S enantiomer is the

same as most stable ab initio conformer, while for the S enantiomer it is roughly 50° shifted. The conformer has an energy that is only 9 kJ/mol higher. Again, this shows the unfavorable steric issues for the S enantiomer can be overcome for AnS-CH₃. The torsional distribution for the ring-chiral carbon dihedral at the events of hydrogen bonding is shown in Figure 6-13.

For Ph-CF₃, Ph-CH₃, and Ph-C₂H₅ the corresponding experimental selectivity factors are 1.0 with TMA terminated hexa-proline. The small energy barrier for the ring-chiral carbon dihedrals introduces conformational flexibility for these analytes. As well, the aromatic groups are less bulky, diminishing the steric constraints experienced by AnS-CH₃. For these reasons, the selector oxygens hydrogen bond comparably to both enantiomers. Figure 6-13 shows the distributions for the ring-chiral carbon dihedrals for hydrogen bonding events. For Ph-C₂H₅, the torsional energy barrier is moderate and this partially limits the flexibility at the interface. The torsional distribution is narrower in this case (see Figure 6-13).

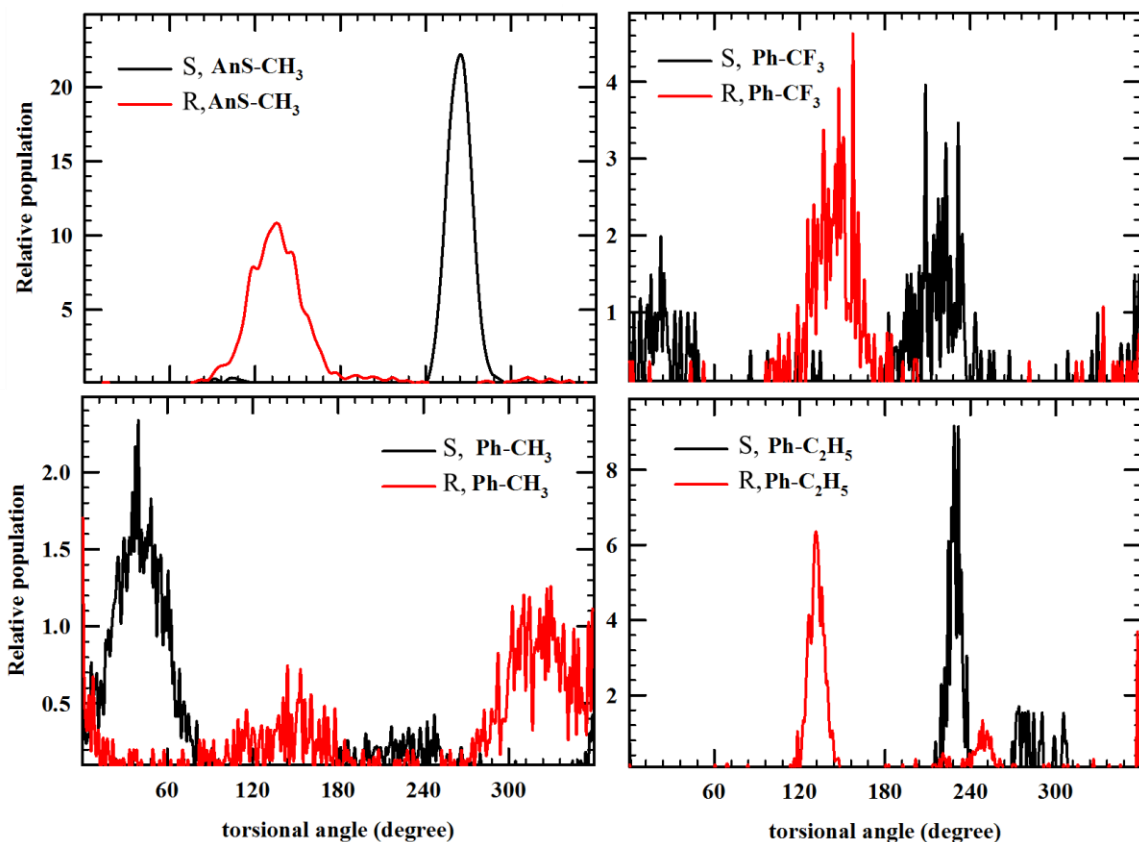


Figure 6-13. Torsional angle distributions for AnS-CH₃ (top left), Ph-CF₃ (top right), Ph-CH₃ (bottom left), and Ph-C₂H₅ (bottom right) at the events of hydrogen bonding with O(8). For AnS-CH₃, the equilibrium angles for S and R enantiomers are 140 and 220°, respectively. For Ph-CF₃, the equilibrium angles for S and R enantiomers are 35 and 315°, respectively. For Ph-CH₃, the equilibrium angles for S and R enantiomers are 35 and 315°, respectively. For Ph-C₂H₅, the equilibrium angles for S and R enantiomers are 155 and 205°, respectively.

The AnC-CH₃, AnS-CH₃ and AnS-CF₃ analytes are bulky and can only approach the surface if the rigid moieties are perpendicular to the Si layer. This can be observed from the higher population of AnC-CH₃ at the event of hydrogen bonding close to the interface where a strong peak is observed at 90° (see Figure 6-14). For smaller analytes, such as Ph-CF₃, although perpendicular arrangements are preferred, the distribution is broad.

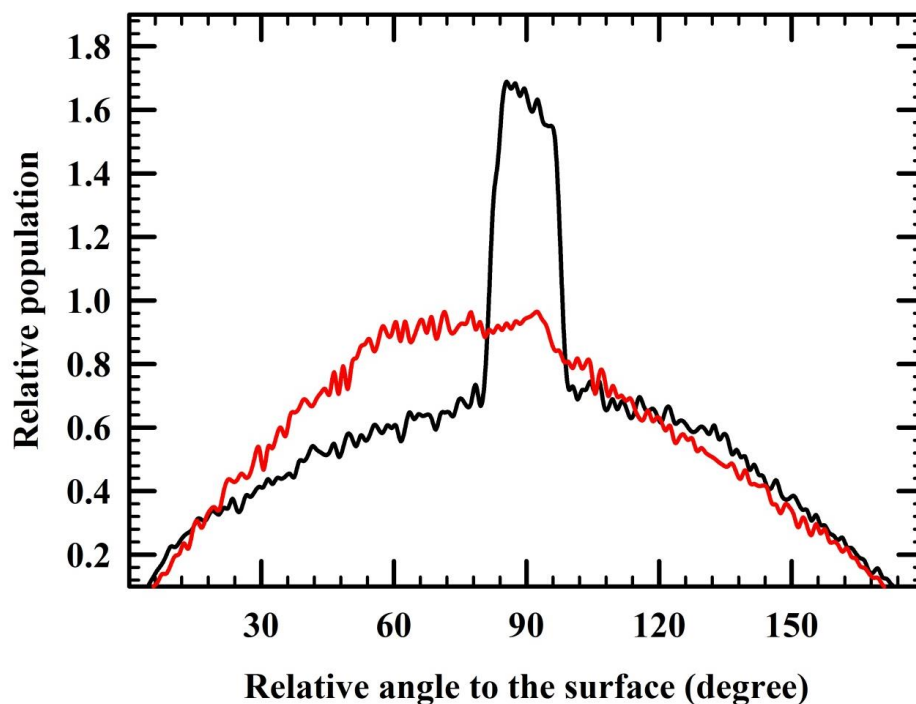


Figure 6-14. The relative angle between the plane of aromatic group and the Si layer. For AnC-CH₃ (black line) and Ph-CF₃ (red line) that hydrogen bond to O(20) and O(8), respectively, on hexa-proline and in *n*-hexane/2-propanol. An angle of 90° shows that the aromatic group is perpendicular to the Si layer and the angle of either 0 or 180° means that the aromatic group is perpendicular.

The steric hindrance coming from long poly-proline chains differ from the short chains. TMA terminated di-proline has only a few stable conformers, they are small and mostly extended towards the bulk. For the tri-proline selector, the interface is more irregular since there are several possible conformers at room temperature. Some of these di-proline and tri-proline conformers are bent at various backbone positions. For hexa-proline, the number of possible conformers is reduced according to the *ab initio* calculations in Chapter 3 and this is also the case in the presence of solvents since interface crowding and H-bonding do not easily allow re-orientation of the peptide bonds during the simulations. For hexa-proline, practically, there is only one

conformer available at the interface. The conclusion is that for di-proline the interface is moderately regular since there are only a few competing conformers at the interface and the more probable conformers do not bend back towards the Si layer. For tri-proline irregularity increases: There are more competing conformers and they are mostly bent towards the Si layer. Interface irregularity then decreases for hexa-proline since fewer conformers are available at the interface.

In addition to the mechanism discussed above, there is another mechanism that involves the simultaneous formation of two hydrogen bonds. Here, the alcohol H from the analyte is positioned such that an H-bond is formed with two carbonyl oxygens in the selectors. Figure 6-15 shows such a hydrogen bonding for tri-proline and AnC-CH₃ analyte. This minor mechanism can be observed along with the major mechanism. For instance, for AnC-CH₃ on hexa-proline the hydrogen bonding statistics are also different for O(12)-O(16) which is mostly due to this new mechanism. Interestingly, for these cases, the average selectivity factor is considerably lower compared to the major mechanism that occurs near the interface. It shows that this mechanism is weaker to differentiate between the enantiomers compared to the first mechanism and could be considered a minor mechanism.

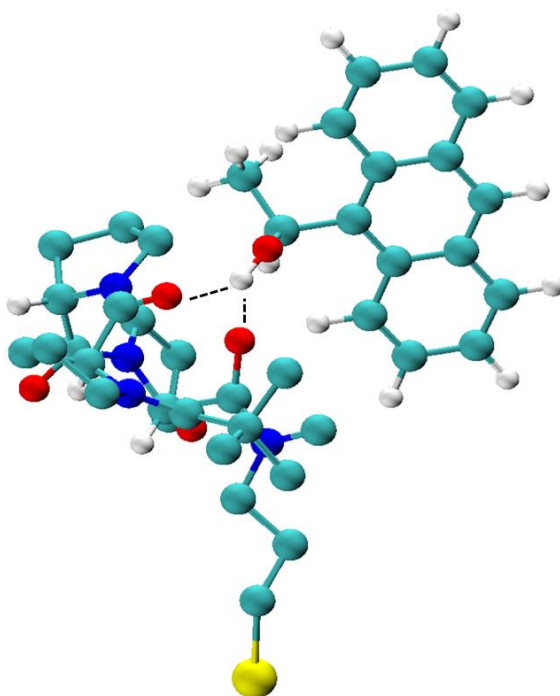


Figure 6-15. Double hydrogen bonding event for AnC-CH₃ with the tri-proline selector in *n*-hexane/2-propanol.

Enantio-separation in the presence of water/methanol is also different. This enantio-separation cannot be easily recognized by a few well-known interactions between the selector and enantiomers. In the presence of water/methanol, mainly the selector oxygens that face the bulk are responsible for enantio-separation, not the oxygens close to the interface. For instance, the predicted selectivity factor for AnS-CH₃ with hexa-proline at water/methanol is 1.32. This analyte cannot be separated in *n*-hexane/2-propanol. This is one example of ability of poly-proline chains to resolve enantiomers that cannot be separated, only by changing the solvent polarity.

For the separation of AnS-CH₃, O(14), O(18) and O(20) are responsible for selectivity with O(14) as the most important enantio-selective hydrogen bonding site (see Table 6-2) . This

selector oxygen is the furthest oxygen from the Si layer in water/methanol (see Figure 6-4) while in *n*-hexane/2-propanol, O(12) and O(14) both appear at around 14 Å from the Si layer. O(12) is closer to the Si layer. Hydrogen bonding events for both enantiomers with O(14) have been studied in detail. The O(14) is directed towards the bulk and H-bonds with both (see Figure 6-16). Snapshot analysis shows that both enantiomers approach the selector with the aromatic group almost parallel to the Si layer. Despite the observed difference in probability for enantiomers hydrogen bonding to O(14), we did not find any clear interactions that are preferred by one enantiomer and discouraged by the other. This example shows that the transient complexes are inherently different whether any well-known interaction can be found or not.

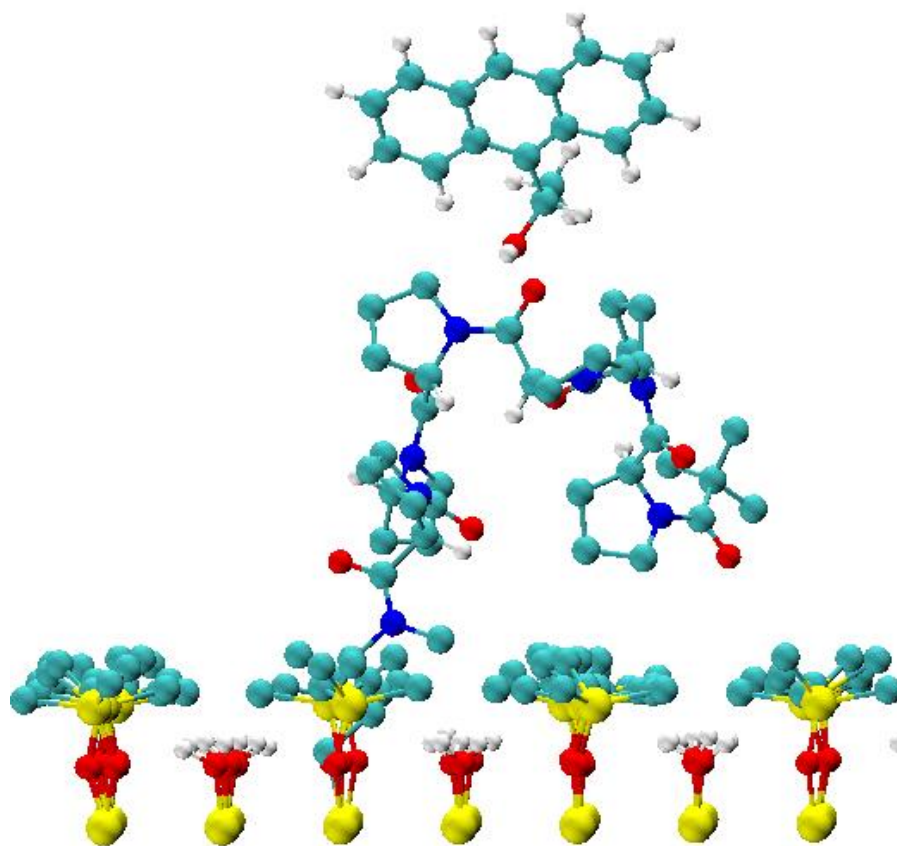


Figure 6-16. Hydrogen bonding of the R enantiomer of AnS-CH₃ with the hexa-proline selector at water/methanol.

6.4 Conclusions

The chiral stationary phases TMA-(Pro)_n-N(CH₃)-tether with n=2,3,6 and BOC-(Pro)₂-N(CH₃)-tether formed the basis of an enantio-selectivity study using molecular dynamics simulations. Six closely-related analytes have been chosen with the objective of identifying the interactions that lead to enantio-resolution. These analytes were previously studied experimentally [47] and showed differences in selectivity factors. The effect of hydrogen bonding, steric hindrance, and conformational flexibility of the analytes are studied. The role of solvent is also examined with two *n*-hexane/2-propanol and water/methanol solvents.

The detailed analysis of docking arrangements shows that enantio-separation often occurs at carbonyl oxygens that are located close to the Si layer. This effect is more obvious for hexaproline where highest differences in hydrogen bonding statistics belong to O(8), O(18) and O(20) that are the closest oxygens to the interface. Other oxygens that are located closer to the bulk may also show distinct interactions with enantiomers, and we have identified simultaneous hydrogen bonding that contributes to the selectivity. Crowding at the interface increases as the chain length increases and this is assumed to affect the arrangement of analytes docking into the surface: This surface crowding highly limits enantiomer movement when hydrogen bonding to the oxygens near the Si layer occurs so that the rigid aromatic groups are mostly perpendicular to the Si layer. Dihedral angle rotation around the bond connecting the aromatic group and the chiral carbon can rotate to avoid placing the methyl group in close contact to the surface-bound molecules for one of the conformers. Interestingly, for analytes with a low energy barrier for this dihedral angle, both enantiomers behave relatively similarly, while for analytes that are energetically limited in their flexibility, differences in hydrogen bonding are observed.

Simulation studies the water/methanol reveal the possibility of resolving analytes that cannot be resolved in *n*-hexane/2-propanol. This is due to different hydrogen bonding patterns between solvent-selector and solvent-analyte and also distinct conformational preferences of polyproline chains in polar solvents. On the whole hydrogen bonding between the analyte and the selector is much less likely in the protic solvent simply due to competition with the solvent. Polyproline based selectors clearly show that is not necessary to identify three or four simultaneous interactions for enantio-resolution. Here, two main types of interactions are identified; hydrogen bonding and steric hindrance to explain the selectivity.

Chapter 7

Conclusions and Outlook

The conformational preferences, solvation, and enantio-selective mechanism of the poly-proline selectors for HPLC were the main focus of this thesis. An extensive series of *ab initio* calculations has been employed to develop potential models for proline chains from mono- to hexa-proline and for six chiral analytes. All force fields were then tested for producing the correct energy-structure dependence. Then molecular dynamics simulations were used to examine the conformational preferences at the interface, solvation, and enantio-selectivity.

To better understand the relative stability of proline chains, several different levels of the theory were employed. Specifically, B3LYP, B97-D, mPW2PLYP-D, MP2, RI-MP2, LPNO-CCSD and CCSD methods are used with the 6-311G(d,p) basis set. We identified the relative conformer stability in the gas phase depends on the terminal group. For shorter chains, *trans* amides in the backbone are preferred, but with increasing the number of proline units, the contribution from *cis* amides increases. For instance, TT is the most stable conformer for TMA-(Pro)₂-N(CH₃)₂ while for the TMA-(Pro)₆-N(CH₃)₂ the most stable conformer is CTCCT. Also, the number of accessible conformations at room temperature increases from di-proline to tetra-proline and then decreases. This shows more flexibility for intermediate length chains and less flexibility for short and long chains.

Force field development was the most challenging part of this research. First, adjacent backbone dihedrals are strongly coupled together and they should be fitted simultaneously. In well-known transferable force fields [222], the adjacent dihedral torsions are assumed to be independent. However, *ab initio* calculations showed that when a dihedral torsion is changing, the

adjacent torsions also change considerably. The dihedral torsions further along the backbone do not change considerably and can be considered independent. By increasing the number of proline units, the potential energy surface becomes more complicated. For instance, for hexa-proline selector, the presence of 64 conformers that need to be modeled properly is a very challenging task. The reason is that by adding another proline unit to the chain, the relative stabilities are totally changed. Also, there are several conformers with close energies but very different structures. Incorporating structurally different structures into one force field is very hard. To overcome this, we have used scaling factors for atomic pairs that are four bonds apart. For hexa-proline, a semi-coarse grain approach is taken for scaling factors such that all ring atoms carry the same scaling factors with backbone atoms.

The solvation of di- to hexa-proline chains are studied in Chapters 4 and 5. Two solvents, *n*-hexane/2-propanol and water/methanol, are employed as the mobile phase. Conformational preferences are shown to be different in terms of solution phase and gas phase populations because of solvent and surface effects, but there is rough consistency in the relative energy order. In particular, the most probable conformer is usually the same in the gas phase and at the interface. However, the presence of the solvent and the surface tend to reduce the number of conformers relative to the gas phase. It has been observed that hydrogen bonding of the solvent to the carbonyl oxygens of the selector is highly dependent on the conformation of the selector at the interface. For some of the oxygens, the conformer limits amount of the solvent docking because of steric hindrance. As a result, some of the selector oxygens appear to be “inactive” in terms of the interaction with the solvent compared to others. Water and methanol hydrogen bonds more extensively with the selectors compared to *n*-hexane/2-propanol.

MD studies are also employed for understanding the docking mechanism and the results are presented in Chapter 6. For each simulation, the total length was about 35-40ns. Surface distributions and snapshots are employed to understand the docking mechanism. Specifically, hydrogen bonding statistics are extracted from saved snapshots of the simulations. It turns out that the hydrogen bonding, steric hindrance at the crowded interface, and the flexibility and size of the analytes all affect enantio-selectivity. Hydrogen bonding statistics are then used for prediction of selectivity factors: The selectivity factors are in excellent agreement with experimental values. In some cases, the solvent effect is predicted to affect the conformational balance at the interface such that the resolution becomes possible in water/methanol, while the separation is not present in *n*-hexane/2-propanol.

In spite of the success of the simulations, there are some issues that are worth noting. The first issue is the slow dynamics for docking of large analytes. The crowded interface usually slows down analyte docking into the surface, especially for larger analytes. The next challenge is surface equilibration. In our approach we started from the most stable conformer. During the equilibration period of the simulation the torsional energy barrier is lowered so that there is enough time for the selector to re-orient and adopt other conformers. For longer selectors, however, this is a challenge. For instance, for hexa-proline, we have observed that almost no re-orientation occurs at the interface and interfacial crowding may be playing a role. It is a concern that the correct equilibration is not achieved for longer chains. The next challenge is the force field development. Our proline selectors are designed to produce the correct energy-structure relationship the gas phase. Since the relative stability of the conformers is not known in solution (due to very high computational cost of real solvent calculation and also insufficient accuracy of implicit models), the model is tested and verified in the gas phase.

Our research indicates the possibility of molecular dynamics simulations to model and understand enantio-separation of chiral compounds. This method is broadly applicable to any type of selector and chiral compound. The next step would be modifying the selectors to increase the enantio-selectivity. Now that the mechanism and the factors favoring or discouraging the separation are known, it is in principle possible to modify the selectors to improve the separation.

Bibliography

- [1] V.A. Davankov, *Pure Appl. Chem.* **69** (1997) 1469.
- [2] [Anon], *Chirality* **4** (1992) 338.
- [3] A. M. Thayer, *Chem. Eng. News August* (**6**) (2007) 11.
- [4] K. M. Rentsch, *J. Biochem. Bioph. Methods* **54** (2002) 1.
- [5] D. J. Cordato, L. E. Mather, G. K. Herkes, *J. Clin. Neurosci.* **10** (2003) 649.
- [6] H. Caner, E. Groner, L. Levy, I. Agranat, *Drug Discovery Today* **9** (2004) 105.
- [7] A. J. Hutt, S.C. Tan, *Drugs* **52** (1996) 1.
- [8] M. R. Islam, J. G. Mahdi, I. D. Bowen, *Drug Safety* **17** (1997) 149.
- [9] A. G. Rauws, K. Groen, *Chirality* **6** (1994) 72.
- [10] W. S. Knowles, M. J. Sabacky, *Chem. Commun.* (1968) 1445.
- [11] R. Noyori, T. Ohkuma, M. Kitamura, H. Takaya, N. Sayo, H. Kumobayashi, S. Akutagawa, *J. Am. Chem. Soc.* **109** (1987) 5856.
- [12] T. Katsuki, K. B. Sharpless, *J. Am. Chem. Soc.* **102** (1980) 5974.
- [13] G. Subramanian, *Chiral Separation Techniques: A Practical Approach*, Wiley-VCH, Weinheim, 2001.
- [14] I. Ali, K. Kumerer, H. Y. Aboul-Enein, *Chromatographia* **63** (2006) 295.
- [15] 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, *J. Med. Chem.* **43** (2000) 4126.
- [16] M. P. Grella, R. Danso-Danquah, M. K. Safo, G. S. Joshi, J. Kister, M. Marden, S. J. Hoffman, D. J. Abraham, *J. Med. Chem.* **43** (2000) 4726.
- [17] L. Pasteur, *C. R. Hebd. Acad. Sci.* **26** (1848) 535.

- [18] J. Debowski, D. Sybilska, J. Jurczak, J. Chromatogr. **237** (1982) 303.
- [19] S. M. Han, Biomed. Chromatogr. **11** (1997) 259.
- [20] W. Lindner, J. N. Lepage, G. Davies, D. E. Seitz, B. L. Karger, J. Chromatogr. **185** (1979) 323.
- [21] S. Allenmark, B. Bomgren, H. Boren, J. Chromatogr. **237** (1982) 473.
- [22] C. Zhao, N. M. Cann, Anal. Chem. **80** (2008) 2426.
- [23] H. Y. Aboul-Enein, I. Ali, Chiral Separations by Liquid Chromatography and Related Technologies, Marcel Dekker, Inc., 2003.
- [24] Y. Okamoto, Y. Kaida, J. Chromatogr. A **666** (1994) 403.
- [25] Y. Okamoto, M. Kawashima, K. Hatada, J. Am. Chem. Soc. **106** (1984) 5357.
- [26] Y. Okamoto, M. Kawashima, K. Hatada, J. Chromatogr. **363** (1986) 173.
- [27] Y. Okamoto, R. Aburatani, T. Fukumoto, K. Hatada, Chem. Lett. (1987) 1857.
- [28] T. E. Beesley, R. P. W. Scott, Chiral Chromatography, Wiley, 1999.
- [29] F. Mikes, G. Boshart, E. Gilav, J. Chromatogr. **122** (1976) 205.
- [30] N. Oi, M. Nagase, T. Doi, J. Chromatogr. **257** (1983) 111.
- [31] W. H. Pirkle, C. J. Welch, J. Liq. Chromatogr. **15** (1992) 1947.
- [32] W. H. Pirkle, C. J. Welch, B. Lamm, J. Org. Chem. **57** (1992) 3854.
- [33] Y. Machida, H. Nishi, K. Nakamura, Chirality **11** (1999) 173.
- [34] M. H. Hyun, J. S. Jin, W. J. Lee, J. Chromatogr. A **822** (1998) 155.
- [35] V. A. Davankov, Chirality **9** (1997) 99.
- [36] L. H. Easson, E. Stedman, Biochem. J. **27** (1933) 1257.
- [37] A. G. Ogston, Nature **162** (1948) 963.
- [38] W. H. Pirkle, T. C. Pochapsky, Chem. Rev. **89** (1989) 347.

- [39] W. Lindner, *Mikrochim. Acta* **2** (1991) 113.
- [40] C. J. Welch, *J. Chromatogr. A* **666** (1994) 3.
- [41] S. Topiol, M. Sabio, *J. Am. Chem. Soc.* **111** (1989) 4109.
- [42] A. D. Mesecar, D. E. Koshland, *Nature* **403** (2000) 614.
- [43] V. Sundaresan, R. Abrol, *Chirality* **17** (2005) S30.
- [44] R. Dappen, H. R. Karfunkel, F. J. J. Leusen, *J. Comput. Chem.* **11** (1990) 181.
- [45] A. Del Rio, J. M. Hayes, M. Stein, P. Piras, C. Roussel, *Chirality* **16** (2004) S1.
- [46] N. M. Maier, S. Schefzick, G. M. Lombardo, M. Feliz, K. Rissanen, W. Lindner, K. B. Lipkowitz, *J. Am. Chem. Soc.* **124** (2002) 8611.
- [47] J. Huang, H. Chen, T. Li, *J. Chromatogr. A* **1113** (2006) 109.
- [48] K. B. Lipkowitz, *J. Chromatogr. A* **906** (2001) 417.
- [49] K. B. Lipkowitz, *J. Chromatogr. A* **694** (1995) 15.
- [50] K. B. Lipkowitz, G. Pearl, B. Coner, M.A. Peterson, *J. Am. Chem. Soc.* **119** (1997) 600.
- [51] K. B. Lipkowitz, B. Baker, *Anal. Chem.* **62** (1990) 770.
- [52] K. B. Lipkowitz, *J. Chromatogr. A* **666** (1994) 493.
- [53] K. B. Lipkowitz, *Acc. Chem. Res.* **33** (2000) 555.
- [54] K. Mihlbachler, M. A. De Jesus, K. Kaczmariski, M. J. Sepaniak, A. Seidel-Morgenstern, G. Guiochon, *J. Chromatogr. A* **1113** (2006) 148.
- [55] C. Temporini, E. Calleri, G. Fracchiolla, G. Carbonara, F. Loiodice, A. Lavecchia, P. Tortorella, G. Brusotti, G. Massolini, *J. Pharm. Biomed. Anal.* **45** (2007) 211.
- [56] U. Norinder, J. Hermansson, *Chirality* **3** (1991) 422.
- [57] R. Kaliszan, T. A. G. Noctor, I. W. Wainer, *Chromatographia* **33** (1992) 546.

- [58] V. Andrisano, R. Gotti, M. Recanatini, A. Cavalli, L. Varoli, C. Bertucci, *J. Chromatogr.* **768** (2002) 137.
- [59] T. D. Booth, I. W. Wainer, *J. Chromatogr. A* **737** (1996) 157.
- [60] Z. Chilmonczyk, H. Ksycinska, M. Lisowaka-Kuzmicz, A. Jonczyk, M. Mazgajska, M. Jaronczyk, A. Strzelczyk, H. Y. Aboul-Enein, *Anal. Chim. Acta* **536** (2005) 7.
- [61] W. M. F. Fabian, W. Stampfer, M. Mazur, G. Uray, *Chirality* **15** (2003) 271.
- [62] A. Del Rio, P. Piras, C. Roussel, *Chirality* **18** (2006) 498.
- [63] M. Karelson, V. S. Lobanov, A. R. Katritzky, *Chem. Rev.* **96** 1027 (1996).
- [64] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, B. P. Feuston, *J. Chem. Inf. Model.* **43**, (2003) 1947.
- [65] S. P. Gupta, *Chem. Rev.* **87** (1987) 1183.
- [66] G. Paola, *QSAR Comb. Sci.* **26** (2007) 694.
- [67] B. W. Clare, *Theor. Chem. Acc.* **87** (1994) 415.
- [68] M. V. Cubellis, F. Caillez, T. L. Blundell, S. C. Lovell, *Proteins: Struct. Funct. Bioinf.* **58** (2005) 880.
- [69] B. Bochicchio, A. M. Tamburro, *Chirality* **14** (2002) 782.
- [70] J. S. Richardson, D. C. Richardson, *Prediction of Protein Structure and the Principles of Protein Conformations*, Springer, 1989.
- [71] M. A. Kelly, B. W. Chellgren, A. L. Rucker, J. M. Troutman, M. G. Fried, A. F. Miller, T. P. Creamer, *J. Biochem.* **40** (2001) 14376.
- [72] A. Meister, *Biochemistry of Amino Acids*, Academic Press Inc., 1956.
- [73] F. J. Brandts, R. H. Halvorson, M. Brennan, *J. Biochem.* **14** (1975) 4953.

- [74] F. X. Schmid, L. M. Mayr, M. Mucke, E. R. Schonbrunner, *Adv. Protein Chem.* **44** (1993) 25.
- [75] W. J. Wedemeyer, E. Welker, H. A. Scheraga, *J. Biochem.* **41** (2002) 14637.
- [76] C. Dugave, L. Demange, *Chem. Rev.* **103** (2003) 2475.
- [77] S. R. C. Lummis, D. L. Beene, L. W. Lee, *Nature* **438** (2005) 248.
- [78] J. M. Huang, P. Zhang, H. Chen, T. Y. Li, *Anal. Chem.* **77** (2005) 3301.
- [79] E. Yashima, C. Yamamoto, Y. J. Okamoto, *J. Am. Chem. Soc.* **118** (1996) 4036.
- [80] A. Berthod, X. Chen, J. P. Kullman, D. W. Armstrong, F. Gasparrini, I. D'Acquarica, C. Villani, A. Carotti, *Anal. Chem.* **72** (2000) 1767.
- [81] K. H. Ekborg-Ott, X. Wang, D. W. Armstrong, *Microchem. J.* **62** (1999) 26.
- [82] C. J. Welch, T. Szczerba, S. R. Perinn, *J. Chromatogr. A* **785** (1997) 93.
- [83] C. D. Haurou, G. Declercq, P. Ramiandrasoa, J. L. Millet, *J. Chromatogr. A* **547** (1991) 31.
- [84] W. H. Pirkle, P. G. Murray, *J. Chromatogr. A* **719** (1996) 299.
- [85] E. Veigl, W. Lindner, *J. Chromatogr. A* **660** (1994) 255.
- [86] R. Sancho, A. M. Perze, C. Minguillon, *J. Sep. Sci.* **29** (2006) 914.
- [87] B. Delgado, E. Perze, M. C. Santano, *J. Chromatogr. A* **1092** (2005) 36.
- [88] T. Takeuchi, H. Asai, D. Ishii, *J. Chromatogr.* **107** (1987) 151.
- [89] F. H. Ling, V. Lu, F. Svec, J. M. J. Frechet, *J. Org. Chem.* **67** (2002) 1993.
- [90] R. Sancho, C. Minguillon, *J. Chromatogr. B* **875** (2008) 93.
- [91] W. J. Lao, J. Gan, *J. Chromatogr. A* **1216** (2009) 5020.
- [92] W. J. Lao, J. Gan, *J. Sep. Sci.* **32** (2009) 2359.
- [93] Y. Bao, J. Huang, T. Li, D. W. Armstrong, *J. Chromatogr. Suppl.* **67** (2008) S13.
- [94] J. Huang, H. Chen, P. Zhang, T. Li, *J. Chromatogr. A* **1109** (2006) 307.

- [95] M. Li, J. Huang, T. Li, *J. Chromatogr. A* **1191** (2008) 199.
- [96] H. Y. Aboul-Enein, I. W. Wainer, *The impact of Stereochemistry on Drug Development and Use*, Wiley-Interscience, 1997.
- [97] I. Ali, H. Y. Aboul-Enein, *Chiral Pollutants: Distribution, Toxicity and Analysis*, Wiley, 2004.
- [98] M. Lammerhofer, *J. Chromatogr. A* **1217** (2010) 814.
- [99] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **21** (1953) 1087.
- [100] B. J. Alder, T. E. Wainwright, *J. Chem. Phys.* **31** (1959) 459.
- [101] L. J. Rafferty, J. I. Siepmann, M. R. Schure, *Top. Curr. Chem.* **307** (2012) 181.
- [102] J. L. Rafferty, I. J. Siepmann, M. R. Schure, *J. Chromatogr. A* **1218** (2011) 2203.
- [103] P. Szabelski, *Appl. Surf. Sci.* **253** (2007) 5387.
- [104] P. Szabelski, D. S. Sholl, *J. Chem. Phys.* **126** (2007) 144709.
- [105] K. B. Lipkowitz, R. Coner, M. A. Peterson, *J. Am. Chem. Soc.* **119** (1997) 11269.
- [106] Z. Hu, J. Jiang, *J. Phys. Chem. B* **113** (2009) 15851.
- [107] R. B. Kasat, N. H. L. Wang, E. I. J. Franses, *J. Chromatogr. A* **1190** (2008) 110.
- [108] R. B. Kasat, S. Y. Wee, J. X. Loh, N. H. L. Wang, E. I. J. Franses, *J. Chromatogr. B* **875** (2008) 81.
- [109] C. Zhao, S. Dimert, N. M. Cann, *J. Chromatogr. A* **1216** (2009) 5968.
- [110] C. Zhao, N. M. Cann, *J. Chromatogr. A* **1149** (2007) 197.
- [111] C. Zhao, N. M. Cann, *J. Chromatogr. A* **1131** (2006) 110.
- [112] S. Nita, N. M. Cann, J. H. Horton, *J. Phys. Chem. B* **108** (2004) 3512.
- [113] S. Nita, J. H. Horton, N. M. Cann, *J. Phys. Chem. B* **110** (2006) 9511.

- [114] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comp. Chem.* **26** (2005) 1781.
- [115] I. N. Levine, *Quantum Chemistry*, Longman Higher Education, 1983.
- [116] P. Hohenberg, W. Kohn, *Phys. Rev. B* **136** (1964) B864.
- [117] W. Kohn, L. J. Sham, *Phys. Rev.* **140** (1965) 1133.
- [118] A. D. Becke, *Phys. Rev. A* **38** (1988) 3098.
- [119] C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **37** (1988) 785.
- [120] R. G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, 1992.
- [121] J. P. Perdew, Y. Wang, *Phys. Rev. B* **45** (1992) 13244.
- [122] A. D. Becke, *J. Chem. Phys.* **98** (1993) 5648.
- [123] S. Grimme, *J. Comp. Chem.* **27** (2006) 1787.
- [124] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **8** (2006) 4398.
- [125] S. Grimme, F. Neese, *J. Chem. Phys.* **126** (2007) 124115.
- [126] S. Grimme, *J. Comput. Chem.* **25** (2004) 1463.
- [127] S. Ye, F. Neese, *Inorg. Chem.* **49** (2010) 772.
- [128] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **132** (2010) 154104.
- [129] F. Neese, T. Schwabe, S. Kossman, B. Schirmer, S. Grimme, *J. Chem. Theory Comput.* **5** (2009) 3060.
- [130] R. J. Bartlett, *Ann. Rev. Phys. Chem.* **32** (1981) 359.
- [131] S. Kossmann, F. Neese, *J. Phys. Chem.* **114** (2010) 11768.
- [132] R. A. Kendall, H. A. Fruchtl, *Theor. Chem. Acc.* **97** (1997) 158.
- [133] S. Kossmann, F. Neese, *J. Chem. Theory Comput.* **6** (2010) 2325.

- [134] J. Cizek, J. Chem. Phys. **45** (1966) 4256.
- [135] F. Neese, D. Liakos, A. Hansen, J. Chem. Phys. **131** (2009) 64103.
- [136] F. Neese, F. Wennmohs, A. Hansen, J. Chem. Phys. **130** (2009) 114108.
- [137] D. Frenkel, B. Smit, Understanding Molecular Simulation: From Algorithms to Applications, Academic Press, 2002.
- [138] J. P. Ryckaert, A. Bellemans, Faraday Discussions (1978) 95.
- [139] C. J. Cramer, Essentials of Computational Chemistry: Theories and Models, Wiley, 2004.
- [140] H. A. Lorentz, Ann. Phys. **248** (1881) 127.
- [141] D. Berthelot, Comptes rendus hebdomadaires des séances de l'Académie des sciences **126** (1898).
- [142] S. W. Rick, S. J. Stuart, B. J. Berne, J. Chem. Phys. **101** (1994) 6141.
- [143] G. Lamoureux, A. D. MacKerell, B. Roux, J. Chem. Phys. **119** (2003) 5185.
- [144] M. P. Allen, D. J. Tildesley, Computer Simulation of Liquids, Clarendon Press, 1989.
- [145] D. M. York, T. A. Darden, L.G. Pedersen, J. Chem. Phys. **99** (1993) 8345.
- [146] P. P. Ewald, Ann. Phys. **64** (1921) 253.
- [147] I. C. Yeh, M. L. Berkowitz, J. Chem. Phys. **111** (1999) 3155.
- [148] L. Verlet, Phys. Rev. **159** (1967) 98.
- [149] R. W. Hockney, Methods in Computational Physics (1970) 136.
- [150] H. C. Andersen, J. Comput. Phys. **52** (1983) 24.
- [151] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, J. Comput. Phys. **23** (1977) 327.
- [152] A. Patriciu, G. S. Chirikjian, R. V. Pappu, J. Chem. Phys. **121** (2004) 12708.
- [153] S. Nosé, J. Chem. Phys. **81** (1984) 511.
- [154] W. G. Hoover, Phys. Rev. A **31** (1985) 1695.

- [155] D. J. Evans, *Mol. Phys.* **34** (1977) 317.
- [156] J. Nawrocki, *J. Chromatogr. A* **779** (1997) 29.
- [157] B. D. Smith, R. Srivastava, *Thermodynamic Data for Pure Compounds*, Elsevier, 1986.
- [158] I. C. Yeh, M. L. Berkowitz, *Chem. Phys. Lett.* **301** (1999) 81.
- [159] D. Bostick, M. L. Berkowitz, *Biophys. J.* **85** (2003) 97.
- [160] E. J. W. Wensink, A. C. Hoffmann, P. J. van Maaren, D. van der Spoel, *J. Chem. Phys.* **119** (2003) 7308.
- [161] E. R. Smith, *Proc. R. Soc. London, Ser. A* **375** (1981) 475.
- [162] M. Moradi, V. Babin, C. Roland, C. Sagui, *J. Chem. Phys.* **133** (2010) 125104.
- [163] D. E. Stewart, A. Sarkar, E. Wampler, *J. Mol. Bio.* **214** (1990) 253.
- [164] E. J. Milner-White, L. H. Bell, P. H. Maccallum, *J. Mol. Bio.* **228** (1992) 725.
- [165] K.Y. Kang, J. S. John, H. S. Park, *J. Phys. Chem. B* **110** (2006) 17645.
- [166] R. Zhang, J. S. Madalengoitia, *Tetrahedron Lett.* **37** (1996) 6235.
- [167] T. Isemura, H. Okabayashi, S. Sakakibara, *Biopolymers* **6** (1968) 307.
- [168] H. Okabayashi, T. Isemura, B. Sakakibara, *Biopolymers* **6** (1968) 323.
- [169] C. M. Deber, F. A. Bovey, J. P. Carver, E. R. Blout, *J. Am. Chem. Soc.* **92** (1970) 6191.
- [170] W. B. Rippon, J. L. Koenig, A. G. Walton, *J. Am. Chem. Soc.* **92** (1970) 7455.
- [171] M. Rothe, H. Rott, *Angew. Chem. Int. Ed. Engl.* **15** (1976) 770.
- [172] N. Helbecque, M. H. Loucheux-Lefebvre, *Int. J. Pept. Protein Res.* **19** (1982) 94.
- [173] R. K. Dukor, T. A. Keiderling, V. Gut, *Int. J. Pept. Protein Res.* **38** (1991) 198.
- [174] R. K. Dukor, T. A. Keiderling, *Biopolymers* **31** (1991) 1747.
- [175] S. Kakinoki, Y. Hirano, M. Oka, *Polym. Bull.* **53** (2005) 109.
- [176] M. Mutter, T. Wöhr, S. Gioria, M. Keller, *Pept. Sci.* **51** (1999) 121.

- [177] W. Traub, U. Shmueli, *Nature* **198** (1963) 1165.
- [178] P. M. Cowan, S. McGavin, *Nature* **176** (1955) 501.
- [179] Y. Y. H. Chao, R. Bersohn, *Biopolymers* **17** (1978) 2761.
- [180] I. Z. Steinberg, A. Berger, E. Katchalski, *Biochim. Biophys. Acta* **28** (1958) 647.
- [181] J. Kurtz, A. Berger, E. Katchalski, *Nature* **178** (1956) 1066.
- [182] E. G. Robertson, J. P. Simons, *Phys. Chem. Chem. Phys.* **3** (2001) 1.
- [183] V. Madison, J. Schellman, *Biopolymers* **11** (1972) 1041.
- [184] S. Doose, H. Neuweiler, H. Barsch, M. Sauer, *Proc. Natl. Acad. Sci. U.S.A.* **104** (2007) 17400.
- [185] T. Higashijima, M. Tasumi, T. Miyazawa, *Biopolymers* **16** (1977) 1259.
- [186] H. Zhong, H. A. Carlson, *J. Chem. Theory Comput.* **2** (2006) 342.
- [187] M. Kuemin, S. Schweizer, C. Ochsenfeld, H. Wennemers, *J. Am. Chem. Soc.* **131** (2009) 15474.
- [188] P. Bour, J. Kubelka, T. A. Keiderling, *Biopolymers* **65** (2002) 45.
- [189] V. Profant, V. Baumruk, X. Li, M. Safarik, P. Bour, *J. Phys. Chem. B* **115** (2011) 15079.
- [190] S. G. Stepanian, I. D. Reva, E. D. Radchenko, L. Adamowicz, *J. Phys. Chem.* **105** (2001) 10664.
- [191] B. Ishimoto, K. Tonan, S. Ikawa, *Spectrochim. Acta* **56** (1999) 201.
- [192] M. Montagut, B. Lemanceau, A. M. Bellocq, *Biopolymers* **13** (1974) 1259.
- [193] M. Brenzik, S. G. Grdadolink, G. Giester, I. Leban, D. Kikelj, *J. Org. Chem.* **66** (2001) 7044.
- [194] M. Sugawara, K. Tonan, S. Ikawa, *Spectrochim. Acta Part A* **57** (2001) 1305.

- [195] I. Z. Steinberg, W. F. Harrington, A. Berger, M. Sela, E. J. Katchalski, *J. Am. Chem. Soc.* **82** (1960) 5263.
- [196] F. Gornick, L. Mandelkern, A. F. Diorio, D. E. Roberts, *J. Am. Chem. Soc.* **86** (1964) 2549.
- [197] H. Strassmair, J. Engel, G. Zundel, *Biopolymers* **8** (1969) 237.
- [198] S. Tanaka, H. A. Scheraga, *Macromolecules* **8** (1975) 504.
- [199] S. Tanaka, H. A. Scheraga, *Macromolecules* **8** (1975) 516.
- [200] J. L. Kofron, P. Kuzmic, V. Kishore, E. Colon-Bonilla, D. H. Rich, *J. Biochem.* **30** (1991) 6127.
- [201] C. Grathwohl, K. Wuthrich, *Biopolymers* **15** (1976) 2043.
- [202] E. Czinki, A. G. Csaszar, *Chem. Eur. J.* **9** (2003) 1008.
- [203] S. Michael, L. A. Marshall, A. Burns, C. David Sherrill, *J. Chem. Phys.* **135** (2011) 194102.
- [204] E. G. Hohenstein, H. M. Jaeger, E. J. Carrell, G. S. Tschumper, C. David Sherrill, *J. Chem. Theory Comput.* **7** (2011) 2842.
- [205] L. A. Burns, Á. Vázquez-Mayagoitia, B. G. Sumpter, C. David Sherrill, *J. Chem. Phys.* **134** (2011) 084107.
- [206] Y. K. Kang, B. J. Byun, *J. Comp. Chem.* **31** (2009) 2915.
- [207] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J.

Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Gaussian, Inc., Wallingford, CT, 2009.

[208] F. Neese, ORCA, An Ab initio, Density Functional and Semiempirical Program Package, Version 2.8-00, University of Bonn, Bonn, Germany, 2009.

[209] S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen, F. Neese, *J. Phys. Chem. A* **110** (2006) 2235.

[210] C. Cox, T. Lectka, *J. Am. Chem. Soc.* **120** (1998) 10660.

[211] L. Vitagliano, R. Berisio, A. Masterangelo, L. Mazzarella, A. Zagari, *Protein Sci.* **10** (2001) 2627.

[212] S. Donnini, G. Groenhof, R. K. Wierenga, A. H. Juffer, *Proteins* **64** (2006) 700.

[213] J. Kapitan, V. Baumruk, P. Bour, *J. Am. Chem. Soc.* **128** (2006) 2438.

[214] D. A. Torchia, *Macromolecules* **4** (1971) 440.

[215] Y. K. Kang, *J. Phys. Chem. B.* **111** (2007) 10550.

[216] Y. K. Kang, H. Y. Choi, *Biophys. Chem.* **111** (2004) 135.

[217] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **118** (1996) 11225.

[218] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **110** (1988) 1657.

[219] M. L. P. Price, D. Ostrovsky, W. L. Jorgensen, *J. Comput. Chem.* **22** (2001) 1340.

[220] E. M. Duffy, P. J. Kowalczyk, W. L. Jorgensen, *J. Am. Chem. Soc.* **115** (1993) 9271.

[221] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **11** (1990) 361.

- [222] P. E. M. Lopes, V. Murashov, M. Tazi, E. Demchuk, A.D. MacKerell Jr., *J. Phys. Chem. B* **110** (2006) 2782.
- [223] M. Levitt, M. Hirshberg, R. Sharon, K. E. Laidig, V. Daggett, *J. Phys. Chem. B* **101** (1997) 5051.
- [224] M. Levitt, R. Sharon, *Proc. Natl. Acad. Sci.* **85** (1988) 7557.
- [225] A. Luzar, D. Chandler, *Phys. Rev. Lett.* **76** (1996) 928.
- [226] G. J. Martyna, D. J. Tobias, M. L. Klein, *J. Chem. Phys.* **101** (1994) 4177.
- [227] E. Cressman, B. Das, J. Dunford, R. Ghenea, Y. Huh, S. Nita, I. Paci, S. Wang, C. Zhao, M. Ashtari, N.M. Cann, unpublished.
- [228] W. L. Jorgensen, J. D. Madura, C. J. Swenson, *J. Am. Chem. Soc.* **106** (1984) 6638.
- [229] J. M. Briggs, T. Matsui, W. L. Jorgensen, *J. Comput. Chem.* **11** (1990) 958.
- [230] W. L. Jorgenson, *J. Phys. Chem.* **90** (1896) 1276.
- [231] K. Hall, M. Ashtari, N. M. Cann, *J. Chem. Phys.* **136** (2012) 114705.
- [232] B. W. Chellgren, T. P. Creamer, *J. Am. Chem. Soc.* **126** (2004) 14734.
- [233] Z. Liu, K. Chen, A. Ng, Z. Shi, R. W. Woody, N. R. Kallenbach, *J. Am. Chem. Soc.* **126** (2004) 15141.
- [234] J. C. Horng, N. T. Raines, *Protein Sci.* **15** (2006) 74.
- [235] S. Knof, J. Engel, *Isr. J. Chem.* **12** (1974) 165.
- [236] J. A. Villa, H. A. Baldoni, D. R. Ripoll, A. Ghosh, H. A. Scheraga, *Biophys. J.* **86** (2004) 731.
- [237] M. Ashtari, N. M. Cann, *J. Chromatogr. A* **1218** (2011) 6331.
- [238] H. Flyvbjerg, H. G. Petersen, *J. Chem. Phys.* **91** (1989) 461.
- [239] K. Wilson, *Recent Developments in Gauge Theories*, Springer, 1980.

Appendix A

Interaction Potentials for Chiral Selectors and Analytes

This section lists the parameters for developed force fields. Individual force field parameters are listed for TMA-(Pro)_n-N(CH₃)₂ with n=2-6 and BOC-(Pro)_m-N(CH₃)₂ with m=2,3 and six chiral analytes: α -methyl-9-anthracenemethanol, 1-anthracen-2-yl-ethanol, α -(trifluoromethyl)benzyl alcohol, 1-phenylethanol, 1-phenyl-1-propanol, and 1-(9-anthryl)-2,2,2-trifluoroethanol. For each molecule, five tables are given.

The first table includes the atomic labels, atomic charges, and Lennard-Jones parameters. For atom numbering of poly-proline selectors, please see Figures 4-1 and 5-1. All atomic charges are obtained using CHELPG algorithm and B3LYP/6-311G(d,p) method and basis set (unless it is stated) and are based on the optimized structure of the most stable conformer. Lennard-Jones parameters are taken from reference [216-219] and [221]. All of the charges have the dimension of $|e|$. Lennard-Jones parameters ϵ and σ (see Eq. 2-30) have the units of kJ/mol and nm , respectively.

The second table lists the stretching potentials (Eq. 2-26). The atoms involved in each stretching potential are noted in the first column. Stretching potentials are obtained using B3LYP/6-311G(d,p) method and basis set (unless it is stated). The r_{eq} (equilibrium bond length) has the dimension of nm and the k_r (stretching force constant) has the dimension of $kJ/(mol \cdot nm^2)$.

The third table lists the bending potentials (Eq. 2-27). The atoms involved in each bending potential are noted in the first column. Bending potentials are obtained using B3LYP/6-311G(d,p) method and basis set (unless it is stated). The θ_{eq} (equilibrium angle value) has the dimension of $degree$, and the k_θ (bending force constant) has the dimension of $kJ/(mol \cdot rad^2)$.

The fourth table lists the torsional potentials (Eq. 2-28). The atoms involved in each torsional potential are noted in the first column. Torsional potentials are obtained using B3LYP/6-311G(d,p) method and basis set (unless it is stated). The c_i (linear fitting parameter) has the dimension of kJ/mol , and the $\varphi^{i;0}$ (nonlinear phase shift) has the dimension of rad .

The fifth table lists the improper torsional potentials (Eq. 2-29). The atoms involved in each improper torsional potential are noted in the first column. Improper torsional potentials are obtained using B3LYP/6-311G(d,p) method and basis set (unless it is stated). The ω_{eq} (equilibrium angle value) has the dimension of *degree*, and the k_ω (bending force constant) has the dimension of $kJ/(mol \cdot rad^2)$.

Potentials for TMA-(Pro)₂-N(CH₃)₂

Numbering, partial charge, and LJ parameters				
Atom label and number	Charge	Charge B97-D	ϵ	σ
Si(1)	0.0	0.0	2.510	0.392
CH2(2)	0.0	0.0	0.276	0.350
CH2(3)	0.0	0.0	0.276	0.350
CH2(4)	0.102	0.070	0.276	0.350
N(5)	-0.080	0.009	0.711	0.325
CH3(6)	0.067	0.036	0.276	0.350
C(7)	0.364	0.378	0.439	0.375
O(8)	-0.486	-0.475	0.878	0.296
C(9)	0.560	0.483	0.439	0.375
O(10)	-0.551	-0.499	0.878	0.296
C(11)	0.402	0.218	0.439	0.375
O(12)	-0.529	-0.458	0.879	0.296
C(13)	0.486	0.426	0.276	0.350
CH3(14)	-0.128	-0.114	0.276	0.350
CH3(15)	-0.128	-0.114	0.276	0.350
CH3(16)	-0.128	-0.114	0.276	0.350
C(17)	0.158	0.061	0.276	0.350
CH2(18)	-0.052	0.015	0.276	0.350
CH2(19)	0.116	0.033	0.276	0.350
CH2(20)	0.077	0.146	0.276	0.350
N(21)	-0.278	-0.291	0.711	0.325
H(22)	0.029	0.045	0.125	0.250
C(23)	0.084	0.050	0.276	0.350
CH2(24)	-0.007	0.003	0.276	0.350
CH2(25)	0.077	0.017	0.276	0.350
CH2(26)	0.075	0.101	0.276	0.350
N(27)	-0.264	-0.067	0.711	0.325
H(28)	0.035	0.031	0.125	0.250

Stretching potentials				
Atoms/Groups involved in Stretch	r_{eq}	k_r	r_{eq} B97-D	k_r B97-D
C(7)-C(17)	0.154	9.0×10^4	0.155	9.0×10^4
C(9)-N(21)	0.135	1.6×10^5	0.136	1.6×10^5
C(9)-C(23)	0.154	9.4×10^4	0.156	9.4×10^4
C(11)-N(27)	0.137	1.8×10^5	0.137	1.8×10^5

Bending potentials				
Atoms/Groups involved in bend	θ_{eq}	k_{θ}	θ_{eq} B97-D	k_{θ} B97-D
Si(1)-C(2)-C(3)	112.70	259.81	112.70	259.81
C(2)-C(3)-C(4)	115.51	259.81	115.51	259.81
C(3)-C(4)-N(5)	112.70	259.81	112.70	259.81
C(4)-N(5)-C(6)	112.68	574.39	114.82	515.06
C(6)-N(5)-C(7)	118.49	659.58	119.59	693.72
C(4)-N(5)-C(7)	125.41	861.58	124.73	821.39
N(5)-C(7)-C(17)	118.57	1097.34	116.19	1152.67
N(5)-C(7)-O(8)	122.37	675.72	122.75	720.97
O(8)-C(7)-C(17)	119.03	786.61	120.96	800.95
N(21)-C(9)-C(23)	117.75	4496.14	117.67	4589.01
O(10)-C(9)-N(21)	122.46	1069.58	121.09	1085.11
O(10)-C(9)-C(23)	119.70	868.16	121.22	809.53
C(13)-C(11)-N(27)	121.61	1263.00	119.80	1293.76
O(12)-C(11)-N(27)	118.62	923.99	120.41	903.26
O(12)-C(11)-C(13)	119.75	731.77	119.78	733.93
C(11)-C(13)-C(14)	109.81	623.19	110.01	633.23
C(11)-C(13)-C(15)	109.81	623.19	110.01	633.23
C(11)-C(13)-C(16)	109.81	623.19	110.01	633.23
C(14)-C(13)-C(15)	109.07	481.69	108.83	459.09
C(14)-C(13)-C(16)	109.07	481.69	108.83	459.09
C(15)-C(13)-C(16)	109.07	481.69	108.83	459.09
C(7)-C(17)-C(18)	110.84	769.58	109.74	744.94
C(7)-C(17)-N(21)	110.32	733.24	111.03	754.72
C(7)-C(17)-H(22)	112.21	371.20	110.48	358.43
C(9)-N(21)-C(20)	126.32	1569.58	119.39	1524.39
C(9)-N(21)-C(17)	119.44	2454.76	128.42	2435.89
C(9)-C(23)-H(28)	110.84	375.37	109.94	333.59
C(9)-C(23)-C(24)	110.74	1069.58	110.24	1070.46
C(9)-C(23)-N(27)	110.26	1293.49	112.24	1293.14
C(11)-N(27)-C(23)	116.72	1029.74	120.63	1073.39
C(11)-N(27)-C(26)	131.18	1269.58	128.61	1238.23

Torsional potentials				
Atoms/ Groups involved in torsion	C₀,C₁,C₂,C₃,C₄,C₅,C₆	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0},$ $\varphi^{5;0}, \varphi^{6;0}$	C₀,C₁,C₂,C₃,C₄,C₅,C₆ B97-D	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0},$ $\varphi^{5;0}, \varphi^{6;0}$ B97-D
C(4)-C(3)- C(2)-Si(1)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0
N(5)-C(4)- C(3)-C(2)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0
C(7)-N(5)- C(4)-C(3)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0
C(4)-N(5)- C(7)-C(17)	26.0, -7.3, 22.3, -47.5, -4.8, 59.0, 25.0	-5.6,-4.7, 4.5, 4.6, -1.8, -1.5	73.8,26.2,-79.9, -97.0,-28.1,-64.2, -13.7	-0.49,-6.31, 18.31,-20.94, -9.88,-0.97
N(5)-C(7)- C(17)-N(21)	0.0, -38.5, 106.0, -41.7, -101.1, -19.2, -51.6	17.3,-9.3, 1.0, -3.2, 2.3, 7.0	0.0,-27.1,89.3, 56.8,-58.5,-28.8, 32.6	3.81,-16.40, -2.41,8.94, -2.43,-2.17
C(7)-C(17)- N(21)-C(9)	56.7, -211.5, -86.4, -122.7, 325.8, 110.7,-233.2	-30.7, -15.7, -26.9, 59.9, 6.3, -31.2	-30.3,-177.1,65.2, 257.1,74.1,-134.2, 88.9	-58.4,29.9, -33.6,31.5, 14.5,-11.7
C(17)-N(21)- C(9)-C(23)	51.9, -44.0, -68.1, 39.0, 84.5, 33.3, 37.2	0.8, 10.1, 0.5, -52.1, 1.8, 3.5	-37.9,-88.4,76.5, 103.3,50.0,-48.1, 30.8	2.3,-1.9,2.1,5.8 ,0.0,-4.3
N(21)-C(9)- C(23)-N(27)	0.0, -57.1, -52.7, 28.4, 18.4, 25.4, 26.8	23.5, -15.0, -1.4, -5.0, 3.9, 3.5	0.0,-4.38,6.43, -1.26,-44.6,11.2, 15.3	17.5,-1.2,7.4, -10.7,-1.3,3.4
C(9)-C(23)- N(27)-C(11)	124.1, 209.0, -86.8, -116.5, -162.1, 103.9, 151.9	-52.5, 21.2, -33.1, 35.3, 12.7, -24.3	39.5,29.9,-97.4, -144.2,106.6, -98.1,131.8	-52.0,44.2, -30.2,18.9, 10.8,-17.9
C(23)-N(27)- C(11)-C(13)	-16.0, 80.8,-38.9, 93.3, 69.9,-49.5, 39.8	5.4, -1.7,-4.2, 10.6, 0.0,-4.3	115.0,-44.9, -85.4,44.5, -168.9,30.8,83.6	-10.8,-21.2, 14.8,30.9,-5.2, -12.8
N(27)-C(11)- C(13)-C(14)	0.0,2.5,25.1, -6.0,23.3,8.5,15.3	-5.6,2.1,5.2,-5.4, -1.4,0.0	0.0,-50.2,48.8, 79.6,-112.3, -16.3,67.9	2.0,-6.4, -10.4,15.5,2.1, -9.5

Improper torsional potentials				
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}	ω_{eq} B97-D	k_{ω} B97-D
C(4)-C(7)-C(6)-N(5)	-0.68	42.92	6.15	48.73
C(17)-N(5)-O(8)-C(7)	1.39	244.51	2.12	234.11
N(20)-C(17)-C(9)-H(21)	-8.80	71.12	-2.02	69.15
H(21)-C(23)-O(10)-C(9)	-1.93	350.90	-0.87	341.93
C(26)-C(23)-C(11)-N(27)	2.59	70.56	-12.70	85.54
N(27)-O(12)-C(13)-C(11)	0.77	322.49	-0.42	354.78
C(7)-C(18)-H(21)-C(17)	-35.14	852.62	-35.67	810.31
N(27)-C(24)-C(9)-C(23)	37.78	823.43	37.64	819.59

Potentials for BOC-(Pro)₂-N(CH₃)₂

Numbering, partial charge, and LJ parameters			
Atom label and number	Charge	ϵ	σ
Si(1)	0.0	0.0	2.510
CH2(2)	0.0	0.0	0.276
CH2(3)	0.0	0.0	0.276
CH2(4)	0.096	0.059	0.276
N(5)	-0.101	0.002	0.711
CH3(6)	0.079	0.052	0.276
C(7)	0.415	0.366	0.439
O(8)	-0.499	-0.463	0.878
C(9)	0.476	0.512	0.439
O(10)	-0.520	-0.513	0.878
C(11)	0.755	0.579	0.439
O(12)	-0.583	-0.514	0.879
C(13)	0.801	0.846	0.276
CH3(14)	-0.137	-0.141	0.276
CH3(15)	-0.137	-0.141	0.276
CH3(16)	-0.137	-0.141	0.276
C(17)	0.085	0.043	0.276
CH2(18)	-0.024	0.006	0.276
CH2(19)	0.103	0.044	0.276
CH2(20)	0.082	0.137	0.276
N(21)	-0.228	-0.286	0.711
H(22)	0.040	0.043	0.125
C(23)	0.097	0.036	0.276
CH2(24)	0.004	0.012	0.276
CH2(25)	0.066	0.018	0.276
CH2(26)	0.102	0.102	0.276
N(27)	-0.288	-0.088	0.711
H(28)	0.013	0.017	0.125
O(29)	-0.561	-0.587	0.711

Stretching potentials				
Atoms/Groups involved in Stretch	r_{eq}	k_r	r_{eq} B97-D	k_r B97-D
C(7)-C(17)	0.154	9.0×10^4	0.155	9.0×10^4
C(9)-N(21)	0.135	1.6×10^5	0.136	1.6×10^5
C(9)-C(23)	0.154	9.4×10^4	0.156	9.4×10^4
C(11)-N(27)	0.136	1.8×10^5	0.137	1.8×10^5

Bending potentials				
Atoms/Groups involved in bend	θ_{eq}	k_θ	θ_{eq} B97-D	k_θ B97-D
C(4)-N(5)-C(6)	116.10	521.59	115.05	589.56
C(6)-N(5)-C(7)	125.38	517.24	124.70	456.32
C(4)-N(5)-C(7)	118.49	730.89	119.18	749.98
N(5)-C(7)-C(17)	118.55	919.23	116.15	890.25
N(5)-C(7)-O(8)	122.44	712.55	122.91	802.30
O(8)-C(7)-C(17)	118.96	765.45	120.92	753.26
N(21)-C(9)-C(23)	117.92	3915.42	118.05	3720.19
O(10)-C(9)-N(21)	122.37	1120.05	121.09	1007.66
O(10)-C(9)-C(23)	119.67	850.88	120.81	901.00
C(27)-C(11)-N(29)	110.78	563.00	110.83	657.08
O(12)-C(11)-N(27)	123.54	879.94	124.22	897.79
O(12)-C(11)-O(29)	125.67	533.77	124.91	517.34
O(29)-C(13)-C(14)	107.51	226.63	107.86	276.10
O(29)-C(13)-C(15)	107.51	226.63	107.86	276.10
O(29)-C(13)-C(16)	107.51	226.63	107.86	276.10
C(14)-C(13)-C(15)	111.29	348.35	110.97	320.28
C(14)-C(13)-C(16)	111.29	348.35	110.97	320.28
C(15)-C(13)-C(16)	111.29	348.35	110.97	320.28
C(7)-C(17)-C(18)	110.87	744.51	110.16	704.49
C(7)-C(17)-N(21)	112.17	701.24	110.74	810.59
C(7)-C(17)-H(22)	110.34	409.85	110.76	430.75
C(9)-N(21)-C(20)	126.71	1358.69	119.41	1510.63
C(9)-N(21)-C(17)	119.41	2123.12	128.20	2320.21
C(9)-C(23)-H(28)	111.10	412.08	108.97	489.07
C(9)-C(23)-C(24)	111.06	987.56	110.02	879.44
C(9)-C(23)-N(27)	110.65	1185.32	113.78	1098.51
C(11)-N(27)-C(23)	119.92	1009.43	120.35	1220.43
C(11)-N(27)-C(26)	126.25	1312.85	124.83	1350.52
C(11)-O(29)-C(13)	121.18	512.23	120.69	517.99

Torsional potentials				
Atoms/ Groups involved in torsion	C₀,C₁,C₂,C₃,C₄,C₅,C₆	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0},$ $\varphi^{5;0}, \varphi^{6;0}$	C₀,C₁,C₂,C₃,C₄,C₅,C₆ B97-D	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0},$ $\varphi^{5;0}, \varphi^{6;0}$ B97-D
C(4)-C(3)- C(2)-Si(1)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0
N(5)-C(4)- C(3)-C(2)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0
C(7)-N(5)- C(4)-C(3)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0
C(6)-N(5)- C(7)-C(17)	-17.3,-108.5, -157.6,115.8, 151.4,-66.8, -66.5	3.8,4.6,-8.4, -7.9,6.1,3.0	75.4,-58.8,-55.7, 45.4,-80.5,60.4, 54.3	1.1,-0.2,7.0, -8.9,-4.3,6.6
N(5)-C(7)- C(17)-N(21)	0.0,-5.8,83.9, 0.6,88.5,-26.3, -56.2	2.6,-14.6,10.1,5.4, -7.2,0.0	-96.2,-88.2, 138.7,-70.7, 111.2,-54.3,46.4	-4.0,10.6, -13.1,-19.4, 11.0,12.7
C(7)-C(17)- N(21)-C(9)	340.5,-170.8, -314.9,-110.8, -215.5,-96.2, -150.7	-50.1,33.8,-28.9, 28.7,15.9,-23.7	29.8,-45.0,73.6, 40.3,-60.5,79.8, 79.8	-49.2,36.6, -44.0,8.2, 10.1,-11.8
C(17)-N(21)- C(9)-C(23)	168.1,-30.5, -251.2,-23.6, 136.2,2.1,-102.1	-6.1,-12.0,3.2, 25.7,-0.2,-13.4	41.9,-135.6, -58.3,122.8, 37.7,64.8,12.9	-12.3,-9.4, 32.0,-3.4, -12.9,-1.1
N(21)-C(9)- C(23)-N(27)	0.0,-31.0,-23.0, 46.0,-17.3,29.4, -13.1	7.8,-3.6,39.7,3.6, -5.2,-1.9	-67.9,14.1, 104.0,20.5,94.7, 17.9,57.9	9.8,-11.9,8.4, 5.0,2.4,-0.5
C(9)-C(23)- N(27)-C(11)	104.7,174.1, -210.5,110.1, 207.6,-79.6,-87.1	-45.5,32.2, -29.1,19.6,10.8, -13.3	41.9,-175.3,12.4, -253.7,184.8, 140.3,-162.8	-54.9,29.4, -51.4,31.8, 24.0,-21.7
C(23)-N(27)- C(11)-C(13)	157.1,-36.3, -152.4,-30.7, -92.4,-17.0,-59.3	0.0,21.7,-2.9, -21.0,2.3,8.1	27.1,-41.2,-22.3, -32.4,98.6,18.2, -66.1	6.5,27.83-9.4, -26.5,1.3,8.1
N(27)-C(11)- O(29)-C(13)	2.5,70.3,19.6, -42.9,-23.0,14.2, 14.4	-5.9,-3.7, -5.4,2.1,2.7,-1.3	97.3,47.5, -50.8,10.5, -33.0,10.5,-26.4	19.3,-5.3,2.0, -7.3,-1.3,6.0
C(11)-O(29)- C(13)-C(14)	0.0,31.1,40.3,37.0, 33.4,6.3,24.2	34.6,47.2, -0.0,1.3,0.5,2.1	11.7,-35.0,-35.5, -53.5,33.2,7.9, 20.8	1.9,5.2,-1.1, -3.9,-1.5,1.4

Improper torsional potentials				
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}	ω_{eq} B97-D	k_{ω} B97-D
C(4)-C(7)-C(6)-N(5)	-1.03	68.15	7.48	72.35
C(17)-N(5)-O(8)-C(7)	1.44	317.09	1.18	352.80
N(20)-C(17)-C(9)-H(21)	-7.62	102.03	1.84	73.00
H(21)-C(23)-O(10)-C(9)	-1.11	310.54	-1.56	380.44
C(26)-C(23)-C(11)-N(27)	0.75	66.37	-16.65	77.59
N(27)-O(12)-C(13)-C(11)	-0.27	380.42	1.14	338.61
C(7)-C(18)-H(21)-C(17)	-35.12	912.58	-35.58	905.84
N(27)-C(24)-C(9)-C(23)	38.01	786.19	36.75	840.73

Potentials for TMA-(Pro)₃-N(CH₃)₂

Numbering, partial charge, and LJ parameters			
Atom label and number	Charge	ϵ	σ
Si(1)	0.000	2.510	0.392
CH2(2)	0.000	0.276	0.350
CH2(3)	0.000	0.276	0.350
CH2(4)	0.100	0.276	0.350
N(5)	-0.241	0.711	0.325
CH3(6)	0.110	0.276	0.350
C(7)	0.616	0.439	0.375
O(8)	-0.512	0.878	0.296
C(9)	0.473	0.439	0.375
O(10)	-0.506	0.878	0.296
C(11)	0.469	0.439	0.375
O(12)	-0.478	0.878	0.296
C(13)	0.380	0.439	0.375
O(14)	-0.520	0.878	0.296
C(15)	0.503	0.276	0.350
CH3(16)	-0.132	0.276	0.350
CH3(17)	-0.132	0.276	0.350
CH3(18)	-0.132	0.276	0.350
C(19)	-0.023	0.276	0.350
CH2(20)	0.029	0.276	0.350
CH2(21)	0.024	0.276	0.350
CH2(22)	0.152	0.276	0.350
N(23)	-0.294	0.711	0.325
H(24)	0.057	0.125	0.250
C(25)	0.121	0.276	0.350
CH2(26)	-0.040	0.276	0.350
CH2(27)	0.102	0.276	0.350
CH2(28)	0.094	0.276	0.350
N(29)	-0.258	0.711	0.325
H(30)	0.018	0.125	0.250
C(31)	0.120	0.276	0.350
CH2(32)	-0.016	0.276	0.350
CH2(33)	0.096	0.276	0.350
CH2(34)	0.037	0.276	0.350
N(35)	-0.247	0.711	0.325
H(36)	0.029	0.125	0.250

Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(7)-C(19)	0.154	9.0×10^4
C(9)-N(23)	0.135	1.6×10^5
C(9)-C(25)	0.154	9.0×10^4
C(11)-N(29)	0.135	1.6×10^5
C(11)-C(31)	0.154	9.0×10^4
C(13)-N(35)	0.137	1.8×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
Si(1)-C(2)-C(3)	115.5	259.81	C(11)-N(29)-C(28)	125.5	551.0
C(2)-C(3)-C(4)	115.2	259.81	C(11)-C(31)-C(32)	110.3	605.2
C(3)-C(4)-N(5)	112.6	259.81	C(11)-C(31)-N(35)	110.7	407.7
C(4)-N(5)-C(6)	115.2	382.6	C(11)-C(31)-H(36)	110.6	278.4
C(4)-N(5)-C(7)	118.0	479.7	O(12)-C(11)-N(29)	122.7	579.6
N(5)-C(7)-O(8)	122.9	594.1	O(12)-C(11)-C(31)	120.1	574.4
C(5)-C(7)-C(19)	116.7	467.6	C(13)-C(15)-C(16)	109.8	239.5
C(6)-C(5)-C(7)	123.3	361.3	C(13)-C(15)-C(17)	109.8	239.5
C(7)-C(19)-C(20)	109.4	593.9	C(13)-C(15)-C(18)	109.8	239.5
C(7)-C(19)-N(23)	111.3	330.4	C(13)-N(35)-C(31)	116.7	471.5
C(7)-C(19)-H(24)	111.3	272.9	C(13)-N(35)-C(34)	131.3	553.3
O(8)-C(7)-C(19)	119.9	461.4	O(14)-C(13)-C(15)	119.7	560.3
C(9)-N(23)-C(19)	129.5	453.4	O(14)-C(13)-N(35)	118.6	670.7
C(9)-N(23)-C(22)	118.6	552.0	C(15)-C(13)-N(35)	121.5	566.2
C(9)-C(25)-C(26)	111.0	585.0	C(16)-C(15)-C(17)	109.0	350.3
C(9)-C(25)-N(29)	110.2	354.6	C(16)-C(15)-C(18)	109.0	350.3
C(9)-C(25)-H(30)	112.0	262.8	C(17)-C(15)-C(18)	109.0	350.3
O(10)-C(9)-N(23)	120.6	602.2	N(23)-C(9)-C(25)	119.4	498.3
O(10)-C(9)-C(25)	119.8	370.7	N(29)-C(11)-C(31)	116.9	497.2
C(11)-N(29)-C(25)	119.6	458.3			

Torsional potentials					
Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ , C ₅ ,C ₆	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0},$ $\varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ ,C ₅ ,C ₆	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0},$ $\varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
Si(1)-C(2)- C(3)-C(4)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	N(23)-C(9)- C(25)-N(29)	0.0,8.8,-98.2, 19.8,-46.0,-19.8, 32.0	21.9,-12.0, -4.70,2.1, 3.3,0.5
C(2)-C(3)- C(4)-N(5)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(9)-C(25)- N(29)-C(11)	-212.9,-176.6, 415.5,110.3, 339.8,105.6, -168.4,	-56.3,24.2, -32.0,41.5, 12.6,-25.9
C(3)-C(4)- N(5)-C(6)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(25)-N(29)- C(11)-C(31)	17.6,-17.3,37.4, -43.7,77.7,-22.7, 28.7	5.4,6.9,-3.2, -7.8,0.0,2.4
C(4)-N(5)- C(7)-C(19)	166.8, -61.7, -117.8,71.5, -69.1,-37.7, -42.8	8.5,12.5, -10.6,-13.7, 6.1,4.0	N(29)-C(11)- C(31)-N(35)	0.0,-36.0,-18.6, 30.5,-22.3,11.9, 7.0	23.2,-12.7, -1.9,-.5,1.4, 5.4
N(5)-C(7)- C(19)-N(23)	0.0,-23.1, -58.3,37.3, -33.3,-15.9, 41.1	30.44,-13.85, -25.29,16.06, -0.13,-4.44	C(11)-C(31)- N(35)-C(13)	24.3,135.3,28.1, 231.3,200.6, -112.5,-148.0	-55.2,22.2, -33.6,38.8, 14.5,-23.8
C(7)-C(19)- N(23)-C(9)	-212.8, -176.6, 415.5,110.3, 339.8,105.6,- 168.4	-56.3,24.2, -32.0,41.5, 12.6,-25.9	C(31)-N(35)- C(13)-C(15)	-13.1,-27.2, -45.1,-78.7, 81.2,-51.2, -81.2	-0.0,12.6, -9.5,-16.9, 6.1,5.1
C(19)- N(23)-C(9)- C(25)	107.8,-8.3, -33.1,55.2, 96.3,-65.6, -60.9	4.2, -6.7,11.3, 5.1,-7.4,-1.0	N(35)-C(13)- C(15)-C(16)	0.0,5.6,61.2, -4.9,52.4, -3.5,36.5	-19.8,2.1, 5.3,-2.8, -0.9,1.0

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}
N(5)-O(8)-C(19)-C(7)	4.52	426.77
C(7)-C(20)-N(23)-C(19)	-36.03	808.19
C(19)-C(22)-C(9)-N(23)	-3.72	84.21
N(23)-C(25)-O(10)-C(9)	0.23	471.85
C(9)-N(29)-C(26)-C(25)	35.29	913.98
C(25)-C(11)-C(28)-N(29)	-15.81	89.38
N(29)-O(12)-C(31)-C(11)	2.12	425.47
C(11)-N(35)-C(32)-C(31)	-35.58	907.30
C(31)-C(13)-C(34)-N(35)	-7.96	74.12
N(35)-O(14)-O(37)-C(13)	-0.80	802.85

Potentials for BOC-(Pro)₃-N(CH₃)₂

Numbering, partial charge, and LJ parameters			
Atom label and number	Charge	ϵ	σ
Si(1)	0.000	2.510	0.392
CH2(2)	0.000	0.276	0.350
CH2(3)	0.000	0.276	0.350
CH2(4)	0.140	0.276	0.350
N(5)	-0.286	0.711	0.325
CH3(6)	0.102	0.276	0.350
C(7)	0.634	0.439	0.375
O(8)	-0.517	0.878	0.296
C(9)	0.489	0.439	0.375
O(10)	-0.510	0.878	0.296
C(11)	0.312	0.439	0.375
O(12)	-0.369	0.878	0.296
C(13)	0.770	0.439	0.375
O(14)	-0.563	0.878	0.296
C(15)	0.774	0.276	0.350
CH3(16)	-0.139	0.276	0.350
CH3(17)	-0.139	0.276	0.350
CH3(18)	-0.139	0.276	0.350
C(19)	-0.027	0.276	0.350
CH2(20)	0.004	0.276	0.350
CH2(21)	0.041	0.276	0.350
CH2(22)	0.142	0.276	0.350
N(23)	-0.286	0.711	0.325
H(24)	0.077	0.125	0.250
C(25)	0.079	0.276	0.350
CH2(26)	-0.025	0.276	0.350
CH2(27)	0.088	0.276	0.350
CH2(28)	0.064	0.276	0.350
N(29)	-0.148	0.711	0.325
H(30)	0.004	0.125	0.250
C(31)	0.233	0.276	0.350
CH2(32)	-0.027	0.276	0.350
CH2(33)	0.073	0.276	0.350
CH2(34)	0.104	0.276	0.350
N(35)	-0.390	0.711	0.325
H(36)	-0.005	0.125	0.250
O(37)	-0.564	0.711	0.300

Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(7)-C(19)	0.154	9.0×10^4
C(9)-N(23)	0.135	1.6×10^5
C(9)-C(25)	0.154	9.0×10^4
C(11)-N(29)	0.135	1.6×10^5
C(11)-C(31)	0.154	9.0×10^4
C(13)-N(35)	0.136	1.8×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
Si(1)-C(2)-C(3)	115.5	259.81	C(11)-C(31)-C(32)	111.1	614.0
C(2)-C(3)-C(4)	115.2	259.81	C(11)-C(31)-N(35)	111.0	420.9
C(3)-C(4)-N(5)	112.6	259.81	C(11)-C(31)-H(36)	109.5	286.1
C(4)-N(5)-C(6)	113.9	400.2	O(12)-C(11)-N(29)	122.4	603.9
C(4)-N(5)-C(7)	117.6	478.6	O(12)-C(11)-C(31)	121.9	576.5
N(5)-C(7)-O(8)	123.4	593.5	C(13)-N(35)-C(31)	122.4	384.3
C(5)-C(7)-C(19)	116.4	476.4	C(13)-N(35)-C(34)	121.0	364.7
C(6)-C(5)-C(7)	122.4	342.9	C(13)-C(37)-C(15)	120.7	324.1
C(7)-C(19)-C(20)	109.4	617.3	O(14)-C(13)-C(35)	124.7	354.0
C(7)-C(19)-N(23)	110.8	331.1	O(14)-C(13)-N(37)	126.0	387.6
C(7)-C(19)-H(24)	111.0	257.8	C(15)-C(37)-N(13)	120.7	351.4
O(8)-C(7)-C(19)	119.9	446.7	C(16)-C(15)-C(17)	110.8	349.4
C(9)-N(23)-C(19)	129.8	445.7	C(16)-C(15)-C(18)	110.8	374.6
C(9)-N(23)-C(22)	118.5	534.5	C(17)-C(15)-C(18)	110.8	340.2
C(9)-C(25)-C(26)	111.3	591.1	N(23)-C(9)-C(25)	119.4	523.0
C(9)-C(25)-N(29)	109.0	362.9	N(29)-C(11)-C(31)	115.6	513.3
C(9)-C(25)-H(30)	112.3	269.4	N(35)-C(13)-O(37)	109.3	591.5
O(10)-C(9)-N(23)	120.8	624.4	O(37)-C(15)-C(16)	110.3	361.2
O(10)-C(9)-C(25)	119.7	382.9	O(37)-C(15)-C(17)	110.3	361.2
C(11)-N(29)-C(25)	120.6	477.0	O(37)-C(15)-C(18)	110.3	361.2
C(11)-N(29)-C(28)	126.2	571.2			

Torsional potentials					
Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ ,C ₅ , C ₆	$\varphi_{0}^{1;0}, \varphi_{0}^{2;0}, \varphi_{0}^{3;0},$ $\varphi_{0}^{4;0}, \varphi_{0}^{5;0}, \varphi_{0}^{6;0}$	Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ ,C ₅ , C ₆	$\varphi_{0}^{1;0}, \varphi_{0}^{2;0}, \varphi_{0}^{3;0},$ $\varphi_{0}^{4;0}, \varphi_{0}^{5;0}, \varphi_{0}^{6;0}$
Si(1)-C(2)- C(3)-C(4)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(9)-C(25)- N(29)-C(11)	-212.9,-176.6, 415.5,110.3, 339.8,105.6, -168.4,	-56.3,24.2, -32.0,41.5, 12.6,-25.9
C(2)-C(3)- C(4)-N(5)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(25)-N(29)- C(11)-C(31)	316.5,72.4, -243.6,-64.7, -219.0,42.7, -137.6	7.4,-14.7, 0.8,11.8, -1.2,-6.2
C(3)-C(4)- N(5)-C(6)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	N(29)-C(11)- C(31)-N(35)	0.0,-23.3, -66.9,-3.0, 39.6,8.7,24.7,	23.6,223.4, -5.9,3.8, 0.0,-3.6
C(4)-N(5)- C(7)-C(19)	81.5,-6.3,174.8 ,-33.6,-129.6, -45.8,-92.1	23.7,8.5, -8.2,-16.1, 7.3,8.2	C(11)-C(31)- N(35)-C(13)	24.3,135.3,28. 1,231.3,200.6, -112.5,-148.0	-55.2,22.2, -33.6,38.8, 14.5,-23.8
N(5)-C(7)- C(19)-N(23)	0.0,27.2,-54.4 ,-20.8,-37.3, -9.0,-20.5	33.1,-8.9, -3.4,2.0, -0.7,-1.7	C(31)-N(35)- C(13)-O(37)	-48.7,-37.7, 65.8,-34.2, 66.5,-15.1,63.9	0.1,-10.1, 3.0,10.2, -2.2,-4.6
C(7)-C(19)- N(23)-C(9)	-212.9,-176.6, 415.5,110.3, 339.8,105.6, -168.4	-56.3,24.2, -32.0,41.5, 12.6,-25.9	N(35)-C(13)- O(37)-C(15)	64.5,-46.4, -28.9,-6.6, -19.6,-16.0,8.1	3.4,-1.2, -5.1,0.6, -11.6,-1.3
C(19)-N(23)- C(9)-C(25)	62.3,-2.2, -29.5,-13.7, 137.8,-0.4, -96.5	8.9,2.6,-1.1, -3.7,0.1,8.2	C(13)-O(37)- C(15)-C(16)	0.0,2.43,5.3, -42.9,-51.8, -4.3,-32.9	-3.0,2.1,3.2, -4.3,-5.5,2.7
N(23)-C(9)- C(25)-N(29)	0.0,14.8,-21.5 ,-20.0,-27.1, -18.3,13.5	12.4,5.2, -1.5,0.4, -0.6,-1.1			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}
N(5)-O(8)-C(19)-C(7)	3.42	401.95
C(7)-C(20)-N(23)-C(19)	-35.92	793.48
C(19)-C(22)-C(9)-N(23)	-1.48	101.29
N(23)-C(25)-O(10)-C(9)	0.00	476.10
C(9)-N(29)-C(26)-C(25)	35.69	893.91
C(25)-C(11)-C(28)-N(29)	-7.50	112.34
N(29)-O(12)-C(31)-C(11)	1.26	413.14
C(11)-N(35)-C(32)-C(31)	34.56	905.94
C(31)-C(13)-C(34)-N(35)	12.37	91.56
N(35)-O(14)-O(37)-C(13)	1.08	832.56

Potentials for TMA-(Pro)₄-N(CH₃)₂

Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
Si(1)	0.00	2.510	0.392	CH2(23)	0.086	0.276	0.350
CH2(2)	0.00	0.276	0.350	CH2(24)	0.139	0.276	0.350
CH2(3)	0.00	0.276	0.350	N(25)	-0.287	0.711	0.325
CH2(4)	0.092	0.276	0.350	H(26)	0.093	0.125	0.250
N(5)	-0.099	0.711	0.325	C(27)	0.271	0.276	0.350
CH3(6)	0.065	0.276	0.350	CH2(28)	-0.021	0.276	0.350
C(7)	0.422	0.439	0.375	CH2(29)	0.074	0.276	0.350
O(8)	-0.497	0.878	0.296	CH2(30)	0.129	0.276	0.350
C(9)	0.506	0.439	0.375	N(31)	-0.370	0.711	0.325
O(10)	-0.490	0.878	0.296	H(32)	-0.032	0.125	0.250
C(11)	0.541	0.439	0.375	C(33)	0.069	0.276	0.350
O(12)	-0.512	0.878	0.296	CH2(34)	-0.120	0.276	0.350
C(13)	0.487	0.439	0.375	CH2(35)	0.088	0.276	0.350
O(14)	-0.523	0.878	0.296	CH2(36)	0.091	0.276	0.350
C(15)	0.320	0.439	0.375	N(37)	-0.245	0.711	0.325
O(16)	-0.518	0.878	0.296	H(38)	0.045	0.125	0.250
C(17)	0.587	0.276	0.350	C(39)	0.143	0.276	0.350
CH3(18)	-0.146	0.276	0.350	CH2(40)	-0.047	0.276	0.350
CH3(19)	-0.146	0.276	0.350	CH2(41)	0.086	0.276	0.350
CH3(20)	-0.136	0.276	0.350	CH2(42)	0.035	0.276	0.350
C(21)	0.034	0.276	0.350	N(43)	-0.218	0.711	0.325
CH2(22)	0.032	0.276	0.350	H(44)	0.019	0.125	0.250

Stretching potentials					
Atoms/Groups involved in Stretch	r_{eq}	k_r	Atoms/Groups involved in Stretch	r_{eq}	k_r
C(7)-C(21)	0.153	9.0×10^4	C(11)-C(33)	0.153	9.0×10^4
C(9)-N(25)	0.135	1.6×10^5	C(13)-N(37)	0.135	1.6×10^5
C(9)-C(27)	0.153	9.0×10^4	C(13)-C(39)	0.153	9.0×10^4
C(11)-N(31)	0.136	1.6×10^5	C(15)-N(43)	0.136	1.8×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_{θ}	Atoms/Groups involved in bend	θ_{eq}	k_{θ}
Si(1)-C(2)-C(3)	115.5	259.81	O(12)-C(11)-N(31)	121.5	572.84
C(2)-C(3)-C(4)	115.2	259.81	O(12)-C(11)-C(33)	119.1	357.91
C(3)-C(4)-N(5)	112.6	259.81	C(13)-N(37)-C(33)	120.7	476.01
C(4)-N(5)-C(6)	115.8	403.30	C(13)-N(37)-C(36)	125.3	567.48
C(4)-N(5)-C(7)	118.7	480.66	C(13)-C(39)-C(40)	109.0	577.25
N(5)-C(7)-O(8)	122.6	610.02	C(13)-C(39)-N(43)	111.2	368.67
N(5)-C(7)-C(21)	118.0	432.21	C(13)-C(39)-C(44)	110.7	277.43
C(6)-N(5)-C(7)	124.9	347.93	O(14)-C(13)-N(37)	123.5	584.54
C(7)-C(21)-C(22)	111.1	602.99	O(14)-C(13)-C(39)	120.3	365.97
C(7)-C(21)-N(25)	110.0	358.10	C(15)-N(43)-C(39)	116.4	486.27
C(7)-C(21)-C(26)	112.8	262.93	C(15)-N(43)-C(42)	131.6	556.94
O(8)-C(7)-C(21)	119.1	446.48	C(15)-C(17)-C(18)	109.7	208.15
C(9)-N(25)-C(21)	118.6	473.18	C(15)-C(17)-C(19)	109.7	208.15
C(9)-N(25)-C(24)	124.4	540.77	C(15)-C(17)-C(20)	109.7	208.15
C(9)-C(27)-C(28)	108.9	604.29	O(16)-C(15)-C(17)	120.3	538.19
C(9)-C(27)-N(31)	110.5	343.07	O(16)-C(15)-N(43)	118.7	563.67
C(9)-C(27)-C(32)	111.6	277.71	C(17)-C(15)-N(43)	120.7	566.41
O(10)-C(9)-N(25)	123.1	565.42	C(18)-C(17)-C(19)	109.1	341.79
O(10)-C(9)-C(27)	119.7	375.98	C(18)-C(17)-C(20)	109.1	341.79
C(11)-N(31)-C(27)	126.0	453.22	C(19)-C(17)-C(20)	109.1	341.79
C(11)-N(31)-C(30)	118.6	562.10	N(25)-C(9)-C(27)	116.1	513.83
C(11)-C(33)-C(34)	110.5	629.16	N(31)-C(11)-C(33)	118.6	480.23
C(11)-C(33)-N(37)	108.4	351.00	N(37)-C(13)-C(39)	115.7	469.42
C(11)-C(33)-C(38)	112.2	283.92			

Torsional potentials					
Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ ,C ₅ , C ₆	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}$; $\varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ ,C ₅ , C ₆	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}$; $\varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
Si(1)-C(2)- C(3)-C(4)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(27)-N(31)- C(11)-C(33)	-51.6,-34.3, 55.6,25.9,60.5 , -20.7,66.7	-8.1,0.6, -8.3,-3.9,2.3 ,1.5
C(2)-C(3)- C(4)-N(5)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	N(31)-C(11)- C(33)-N(37)	0.0,19.9,42.3,1 6.4,32.5,8.2,21 .9	26.3,-8.0, 1.7,-3.2, -1.4,2.2
C(3)-C(4)- N(5)-C(6)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(11)-C(33)- N(37)-C(13)	10.6,-24.1, -137.5,-68.6, 137.1,110.9,15 0.9	-58.3,32.6, -29.7,32.8, 10.1,-24.4
C(4)-N(5)- C(7)-C(21)	110.2,-20.0, -88.5,-37.0, -108.7,-46.9, 46.9	10.1,12.9, -12.4,-16.4, 7.3,5.6	C(33)-N(37)- C(13)-C(39)	29.1,-25.3, 132.2,-10.3, -101.6,14.4, -63.1	12.0,0.8, -2.9,6.7, -0.5,-5.1
N(5)-C(7)- C(21)-N(25)	0.0,-3.6,22.5 ,20.4,-39.7,7.3, 21.3	41.7,-7.1, -4.6,9.0, 0.5,-3.3	N(37)-C(13)- C(39)-C(43)	0.0,33.5,29.4,- 15.0,7.6,-12.3, 5.3	38.5,-7.4, 0.6,0.8, -0.6,0.0
C(7)-C(21)- N(25)-C(9)	38.1,204.6, -249.1,-107.4, 294.6,-110.2, 147.9	-58.8,32.6, -26.9,32.3, 9.5,-23.3	C(13)-C(39)- N(43)-C(15)	-115.9,-56.5, 235.8,217.8,29 3.8,116.2, -142.7	-56.5,22.7, -33.8,43.1, 15.2,-22.8
C(21)-N(25)- C(9)-C(27)	-24.1,11.9, 40.3,-3.4, 78.9,15.6, 25.4	12.7,2.5, -4.5,1.4, 1.8,-2.5	C(39)-N(43)- C(15)-C(17)	120.6,-6.2, -96.9,48.5, -51.8,54.8, -56.4	5.8,13.0, -14.0,-16.1, 8.0,5.1
N(25)-C(9)- C(27)-N(31)	0.0,-35.5,35.1 , -18.1,15.9, -15.8,10.8	29.0,-7.2, 0.6,-1.9, -0.6,0.4	N(43)-C(15)- C(17)-C(18)	0.0,7.1,56.9,5. 8,-49.7, -4.6,-32.2	-0.9,-0.0, 2.3,0.2,-1.0, -0.5
C(9)-C(27)- N(31)-C(11)	84.6,62.0,335. 6,-110.4, -344.6,105.1, -171.5	-58.6,21.8, -31.0,40.7, 12.7,-25.9			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}
N(5)-O(8)-C(21)-C(7)	0.74	355.98
C(7)-C(22)-N(25)-C(21)	35.06	787.24
C(21)-C(24)-C(9)-N(25)	15.29	67.77
N(25)-C(27)-O(10)-C(9)	-2.86	438.10
C(9)-N(31)-C(28)-C(27)	35.92	794.57
C(27)-C(11)-C(30)-N(31)	10.59	104.03
N(31)-C(33)-O(12)-C(11)	-1.37	474.03
C(11)-N(37)-C(34)-C(33)	36.26	826.09
C(33)-C(13)-C(36)-N(37)	-10.42	106.98
N(37)-C(39)-O(14)-C(13)	-3.95	466.44
C(13)-N(43)-C(40)-C(39)	35.40	801.00
C(39)-C(15)-C(42)-N(43)	0.63	74.71
N(43)-C(17)-O(16)-C(15)	-1.03	436.73

Potentials for TMA-(Pro)₅-N(CH₃)₂

Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
Si(1)	0	2.510	0.392	N(27)	-0.121	0.711	0.325
CH2(2)	0	0.276	0.350	H(28)	-0.05	0.125	0.250
CH2(3)	0	0.276	0.350	C(29)	0.167	0.276	0.350
CH2(4)	0.047	0.276	0.350	CH2(30)	-0.029	0.276	0.350
N(5)	0.018	0.711	0.325	CH2(31)	0.083	0.276	0.350
CH3(6)	0.061	0.276	0.350	CH2(32)	0.082	0.276	0.350
C(7)	0.311	0.439	0.375	N(33)	-0.321	0.711	0.325
O(8)	-0.563	0.878	0.296	H(34)	-0.009	0.125	0.250
C(9)	0.385	0.439	0.375	C(35)	0.056	0.276	0.350
O(10)	-0.426	0.878	0.296	CH2(36)	0.002	0.276	0.350
C(11)	0.578	0.439	0.375	CH2(37)	0.042	0.276	0.350
O(12)	-0.498	0.878	0.296	CH2(38)	0.142	0.276	0.350
C(13)	0.592	0.439	0.375	N(39)	-0.361	0.711	0.325
O(14)	-0.541	0.878	0.296	H(40)	0.018	0.125	0.250
C(15)	0.468	0.439	0.375	C(41)	0.176	0.276	0.350
O(16)	-0.449	0.878	0.296	CH2(42)	-0.009	0.276	0.350
C(17)	0.455	0.439	0.375	CH2(43)	0.037	0.276	0.350
O(18)	-0.525	0.878	0.296	CH2(44)	0.122	0.276	0.350
C(19)	0.447	0.276	0.350	N(45)	-0.301	0.711	0.325
CH3(20)	-0.105	0.276	0.350	H(46)	-0.089	0.125	0.250
CH3(21)	-0.109	0.276	0.350	C(47)	0.215	0.276	0.350
CH3(22)	-0.11	0.276	0.350	CH2(48)	0.008	0.276	0.350
C(23)	0.202	0.276	0.350	CH2(49)	0.043	0.276	0.350
CH2(24)	-0.002	0.276	0.350	CH2(50)	0.135	0.276	0.350
CH2(25)	-0.008	0.276	0.350	N(51)	-0.381	0.711	0.325
CH2(26)	0.111	0.276	0.350	H(52)	0.003	0.125	0.250

Stretching potentials					
Atoms/Groups involved in Stretch	r_{eq}	k_r	Atoms/Groups involved in Stretch	r_{eq}	k_r
C(7)-C(23)	0.154	9.0×10^4	C(13)-N(39)	0.136	1.6×10^5
C(9)-N(27)	0.136	1.6×10^5	C(13)-C(41)	0.154	9.0×10^4
C(9)-C(29)	0.154	9.0×10^4	C(15)-N(45)	0.136	1.6×10^5
C(11)-N(33)	0.136	1.6×10^5	C(15)-C(47)	0.154	9.0×10^4
C(11)-C(35)	0.155	9.0×10^4	C(17)-N(51)	0.137	1.8×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_{θ}	Atoms/Groups involved in bend	θ_{eq}	k_{θ}
Si(1)-C(2)-C(3)	115.5	259.81	C(13)-C(41)-C(42)	110.03	574.38
C(2)-C(3)-C(4)	115.2	259.81	C(13)-C(41)-N(45)	108.61	383.56
C(3)-C(4)-N(5)	112.6	259.81	C(13)-C(41)-H(46)	112.81	269.17
C(4)-N(5)-C(6)	115.50	391.52	C(14)-C(13)-N(39)	122.07	611.20
C(4)-N(5)-C(7)	118.38	452.48	C(14)-C(13)-C(41)	119.72	367.12
N(5)-C(7)-O(8)	123.02	617.04	C(15)-N(45)-C(41)	115.63	439.73
N(5)-C(7)-C(23)	116.93	452.20	C(15)-N(45)-C(44)	127.49	549.60
C(6)-N(5)-C(7)	124.46	383.51	C(15)-C(47)-C(48)	108.02	601.02
C(7)-C(23)-C(24)	110.71	577.68	C(15)-C(47)-N(51)	109.50	361.92
C(7)-C(23)-N(27)	138.44	339.75	C(15)-C(47)-H(52)	110.92	270.32
C(7)-C(23)-H(28)	112.15	295.10	C(16)-C(15)-N(45)	121.36	562.27
C(8)-C(7)-C(23)	120.02	431.01	C(16)-C(15)-C(47)	121.51	357.07
C(9)-N(27)-C(23)	115.74	445.31	C(17)-N(51)-C(47)	115.22	464.22
C(9)-N(27)-C(26)	126.16	570.98	C(17)-N(51)-C(50)	131.73	560.30
C(9)-C(29)-C(30)	109.86	615.49	C(17)-C(19)-C(20)	109.74	219.08
C(9)-C(29)-N(33)	109.42	337.63	C(17)-C(19)-C(21)	109.74	219.08
C(9)-C(29)-H(34)	112.63	276.50	C(17)-C(19)-C(22)	109.74	219.08
C(10)-C(9)-N(27)	121.56	613.38	C(18)-C(17)-N(51)	120.99	569.31
C(10)-C(9)-C(29)	120.03	332.43	C(18)-C(17)-C(19)	118.03	603.44
C(11)-N(33)-C(29)	127.65	484.58	C(19)-C(17)-N(51)	120.97	553.92
C(11)-N(33)-C(32)	119.46	543.23	C(20)-C(19)-C(21)	109.14	355.10
C(11)-C(35)-C(36)	108.61	614.93	C(20)-C(19)-C(22)	109.14	355.10
C(11)-C(35)-N(39)	110.80	349.91	C(21)-C(19)-C(22)	109.14	355.10
C(11)-C(35)-H(40)	112.49	286.24	N(27)-C(9)-C(29)	118.34	353.82
C(12)-C(11)-N(33)	122.33	590.54	N(33)-C(11)-C(35)	116.58	356.85
C(12)-C(11)-C(35)	120.54	340.35	N(39)-C(13)-C(41)	123.79	383.71
C(13)-N(39)-C(35)	127.00	483.76	N(45)-C(15)-C(47)	116.83	357.02
C(13)-N(39)-C(38)	119.67	542.89			

Torsional potentials					
Atoms/ Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4,$ c_5, c_6	$\varphi_{0;0}^{1;0}, \varphi_{4;0}^{2;0}, \varphi_{6;0}^{3;0},$ $\varphi_{4;0}^{5;0}, \varphi_{6;0}^{5;0}, \varphi_{6;0}^{6;0}$	Atoms/ Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5,$ c_6	$\varphi_{4;0}^{1;0}, \varphi_{5;0}^{2;0}, \varphi_{6;0}^{3;0},$ $\varphi_{4;0}^{4;0}, \varphi_{5;0}^{5;0}, \varphi_{6;0}^{6;0}$
Si(1)-C(2)- C(3)-C(4)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	N(33)-C(11)- C(35)-N(39)	0.0, 13.3, -32.7 , -26.0, -14.9, -15.8, -15.0	-6.5, -9.1, -1.5, 5.1, -0.6, -1.9
C(2)-C(3)- C(4)-N(5)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(11)-C(35)- N(39)-C(13)	57.0, -128.7, -287.1, -181.9, -220.3, 97.0, -156.8	-62.4, 26.6, -28.6, 39.6, 8.8, -23.8
C(3)-C(4)- N(5)-C(6)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(35)-N(39)- C(13)-C(41)	47.5, 15.4, -35.3 , 23.7, 10.9, 6.9, 16.9	10.6, -266.8, -4.8, -0.0, 5, 4.6
C(4)-N(5)- C(7)-C(23)	151.1, -3.4, -119.5, 14.9, -58.9, -12.8, 25.4	11.2, 3.4, -4.8, -0.9, 1.1, 0.5	N(39)-C(13)- C(41)-C(45)	0.0, -23.86, 31.1, 12.4, -10.4, -5.9, 9.6	17.3, -10.8, -0.3, 1.4, -0.7, 0.0
N(5)-C(7)- C(23)- N(27)	0.0, 38.1, -35.9, 34.6, -16.6, 13.5, -3.7	20.0, -11.8, -1.9, 2.5, 1.4, 0.0	C(13)-C(41)- N(45)-C(15)	-136.7, -118.1, 364.9, -219.7, 187.8, 104.6, -151.1	-58.4, -65.6, -17.9, 36.4, 11.4, -21.8
C(7)- C(23)- N(27)-C(9)	424.7, 56.3, -339.1, 39.7, -348.4, 89.5, -175.2	-15.7, 23.4, -28.8, 43.9, 10.1, -26.0	C(41)-N(45)- C(15)-C(47)	0.7, -15.8, -78.1, 19.5, -74.4, -7.4, 56.6	8.2, 1.6, -4.6, 0.0, -1.0, -1.5
C(23)- N(27)- C(9)-C(29)	76.7, -15.8, -78.1, 19.5, -74.4, -7.7, 56.6	8.2, 1.6, - 4.6, 5.4, - 1.0, -1.5	N(45)-C(15)- C(47)-N(51)	0.0, 43.3, 56.0, 29.6, 24.8, 13.3, -24.4	20.8, -10.0, -0.4, 4.1, -1.6, -0.5
N(27)- C(9)- C(29)- N(33)	0.0, 43.3, 56.0, 29.6, 24. 8, 13.3, -24.4	20.7, -10.1, -0.4, 4.1, -1.6, -0.5	C(15)-C(47)- N(51)-C(17)	-136.7, -118.1, 364.9, -219.7, 187.8, 104.6, -151.1	-58.4, -65.6, -17.9, 36.3, 11.4, -21.8
C(9)- C(29)- N(33)- C(11)	57.1, -128.8, -287.1, -181.9 , -220.3, 97.0, -156.83	-62.4, 26.6, -28.6, 39.6, 8.8, -23.8	C(47)-N(51)- C(17)-C(19)	20.4, -78.9, 23.6, 8.69, 6.14, 47.7, -49.4	1.4, 1.5, -4.5, -4.4, 0.6, 2.0
C(29)- N(33)- C(11)- C(35)	135.9, 16.07, -57.3, 7.5, -38.2, 21.4, -40.9	3.1, 0.5, -4.9, 2.9, -6.4, -4.0	N(51)-C(17)- C(19)-C(20)	0.0, -3.3, 66.5, -7.4, -56.4, 11.0, -37.4	-2.8, -3.3, 5.1, 8.9, -1.7, -5.9

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}
N(5)-O(8)-C(23)-C(7)	1.09	417.11
C(7)-C(24)-N(27)-C(23)	-36.33	751.74
C(23)-C(26)-C(9)-N(27)	-18.85	26.05
N(27)-C(29)-O(10)-C(9)	1.43	418.38
C(9)-N(33)-C(34)-C(29)	36.33	856.35
C(29)-C(11)-C(32)-N(33)	7.96	60.00
N(33)-C(35)-O(12)-C(11)	-4.87	487.22
C(11)-C(36)-N(39)-C(35)	-36.38	832.25
C(35)-C(38)-C(13)-N(39)	-9.51	51.33
N(39)-C(41)-O(14)-C(13)	-3.67	464.00
C(13)-N(45)-C(42)-C(41)	36.67	817.44
C(41)-C(44)-C(15)-N(45)	-14.44	83.92
N(45)-C(47)-O(16)-C(15)	-3.67	395.60
C(15)-C(48)-N(51)-C(47)	-37.13	846.32
C(47)-C(50)-C(17)-N(51)	-9.17	105.40
N(51)-C(19)-O(18)-C(17)	0.57	369.79

Potentials for TMA-(Pro)₆-N(CH₃)₂

Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
Si(1)	0.000	2.510	0.392	C(31)	0.144	0.276	0.350
CH2(2)	0.000	0.276	0.350	CH2(32)	-0.047	0.276	0.350
CH2(3)	0.000	0.276	0.350	CH2(33)	0.099	0.276	0.350
CH2(4)	0.069	0.276	0.350	CH2(34)	0.076	0.276	0.350
N(5)	-0.018	0.711	0.325	N(35)	-0.364	0.711	0.325
CH3(6)	0.069	0.276	0.350	H(36)	-0.008	0.125	0.250
C(7)	0.398	0.439	0.375	C(37)	0.039	0.276	0.350
O(8)	-0.516	0.878	0.296	CH2(38)	0.004	0.276	0.350
C(9)	0.540	0.439	0.375	CH2(39)	0.052	0.276	0.350
O(10)	-0.400	0.878	0.296	CH2(40)	0.119	0.276	0.350
C(11)	0.614	0.439	0.375	N(41)	-0.268	0.711	0.325
O(12)	-0.517	0.878	0.296	H(42)	0.010	0.125	0.250
C(13)	0.488	0.439	0.375	C(43)	0.378	0.276	0.350
O(14)	-0.543	0.878	0.296	CH2(44)	-0.050	0.276	0.350
C(15)	0.541	0.439	0.375	CH2(45)	0.033	0.276	0.350
O(16)	-0.520	0.878	0.296	CH2(46)	0.095	0.276	0.350
C(17)	0.575	0.439	0.375	N(47)	-0.287	0.711	0.325
O(18)	-0.542	0.878	0.296	H(48)	-0.203	0.125	0.250
C(19)	0.528	0.439	0.375	C(49)	0.085	0.276	0.350
O(20)	-0.548	0.878	0.296	CH2(50)	0.024	0.276	0.350
C(21)	0.449	0.276	0.350	CH2(51)	0.039	0.276	0.350
CH3(22)	-0.138	0.276	0.350	CH2(52)	0.136	0.276	0.350
CH3(23)	-0.138	0.276	0.350	N(53)	-0.375	0.711	0.325
CH3(24)	-0.138	0.276	0.350	H(54)	0.027	0.125	0.250
C(25)	0.196	0.276	0.350	C(55)	0.168	0.276	0.350
CH2(26)	-0.059	0.276	0.350	CH2(56)	0.191	0.276	0.350
CH2(27)	0.061	0.276	0.350	CH2(57)	-0.174	0.276	0.350
CH2(28)	0.061	0.276	0.350	CH2(58)	0.311	0.276	0.350
N(29)	-0.230	0.711	0.325	N(59)	-0.512	0.711	0.325
H(30)	-0.010	0.125	0.250	H(60)	-0.013	0.125	0.250

Stretching potentials					
Atoms/Groups involved in Stretch	r_{eq}	k_r	Atoms/Groups involved in Stretch	r_{eq}	k_r
C(7)-C(25)	0.155	9.0×10^4	C(13)-C(43)	0.154	9.0×10^4
C(9)-N(29)	0.136	1.6×10^5	C(15)-N(47)	0.135	1.6×10^5
C(9)-C(31)	0.154	9.0×10^4	C(15)-C(49)	0.154	9.0×10^4
C(11)-N(35)	0.136	1.6×10^5	C(17)-N(53)	0.136	1.6×10^5
C(11)-C(37)	0.155	9.0×10^4	C(17)-C(55)	0.154	9.0×10^4
C(13)-N(41)	0.136	1.6×10^5	C(19)-N(59)	0.137	1.8×10^5

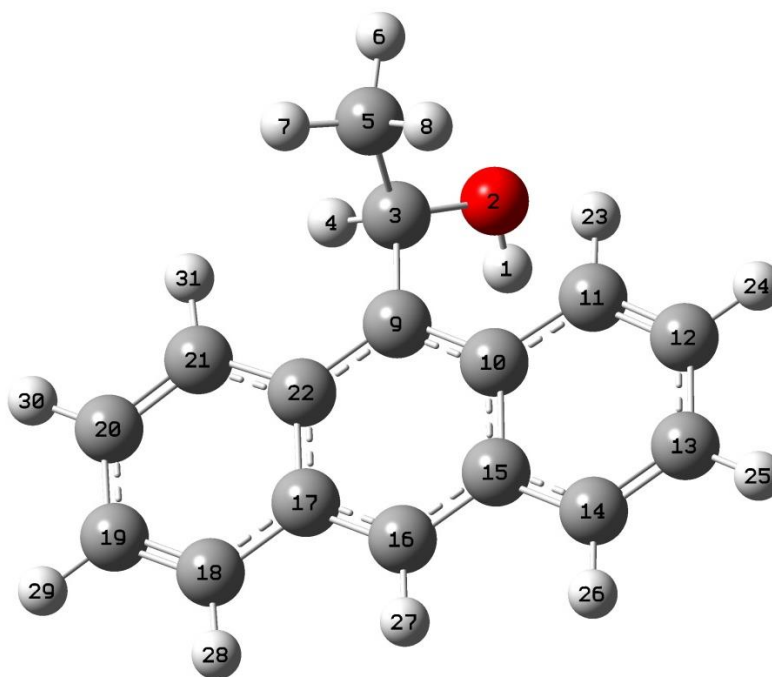
Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_{θ}	Atoms/Groups involved in bend	θ_{eq}	k_{θ}
Si(1)-C(2)-C(3)	115.5	259.81	C(14)-C(12)-C(43)	121.81	310.59
C(2)-C(3)-C(4)	115.2	259.81	C(15)-N(47)-C(43)	117.40	603.49
C(3)-C(4)-N(5)	112.6	259.81	C(15)-N(47)-C(46)	128.17	631.02
C(4)-N(5)-C(6)	115.45	393.39	C(15)-C(49)-C(50)	110.47	399.76
C(4)-N(5)-C(7)	118.37	481.65	C(15)-C(49)-N(53)	107.37	314.65
N(5)-C(7)-C(8)	123.01	631.79	C(15)-C(49)-H(54)	111.50	608.49
N(5)-C(7)-C(25)	116.94	491.90	C(16)-C(15)-N(47)	121.75	389.73
C(6)-N(5)-C(7)	124.45	378.52	C(16)-C(15)-C(49)	120.84	609.24
C(7)-C(25)-C(26)	109.66	648.29	C(17)-N(53)-C()	116.43	530.99
C(7)-C(25)-N(29)	108.00	330.27	C(17)-N(53)-C()	126.68	577.32
C(7)-C(25)-H(30)	111.90	235.51	C(17)-C(55)-C()	107.26	362.19
C(8)-C(7)-C(25)	119.35	432.20	C(17)-C(55)-N()	110.98	227.94
C(9)-N(29)-C(25)	116.65	414.03	C(17)-C(55)-H()	110.98	609.82
C(9)-N(29)-C(28)	125.08	588.04	C(18)-C(17)-N(53)	121.70	383.36
C(9)-C(31)-C(32)	108.35	551.73	C(18)-C(17)-C(55)	121.58	546.63
C(9)-C(31)-N(35)	110.29	372.08	C(19)-C(21)-C(22)	109.95	205.96
C(9)-C(31)-H(36)	112.59	239.32	C(19)-C(21)-C(23)	109.95	205.96
C(10)-C(9)-N(29)	122.33	567.24	C(19)-C(21)-C(24)	109.95	205.96
C(10)-C(9)-C(31)	120.66	403.82	C(19)-N(59)-C(55)	128.46	626.59
C(11)-N(35)-C(31)	126.28	478.35	C(19)-N(59)-C(58)	119.35	596.74
C(11)-N(35)-C(34)	119.06	533.16	C(20)-C(19)-C(21)	120.09	566.86
C(11)-C(37)-C(38)	108.80	591.20	C(20)-C(19)-N(59)	120.03	364.38
C(11)-C(37)-N(41)	111.10	348.51	C(21)-C(19)-N(59)	119.52	560.83
C(11)-C(37)-H(42)	112.64	248.59	C(22)-C(21)-C(23)	109.49	343.78
C(12)-C(12)-N(35)	122.44	603.51	C(22)-C(21)-C(24)	109.49	343.78
C(12)-C(12)-C(37)	120.21	346.64	C(23)-C(21)-C(24)	109.49	343.78
C(13)-N(41)-C(37)	126.11	451.00	N(29)-C(9)-C(31)	116.94	498.92
C(13)-N(41)-C(40)	120.95	567.62	N(35)-C(11)-C(37)	116.77	510.11
C(13)-C(43)-C(44)	110.64	595.12	N(41)-C(13)-C(43)	114.65	471.60
C(13)-C(43)-N(47)	110.70	376.44	N(47)-C(15)-C(49)	117.34	489.22
C(13)-C(43)-H(48)	110.81	282.87	N(53)-C(17)-C(55)	116.60	515.32
C(14)-C(12)-N(41)	123.24	584.06			

Torsional potentials					
Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ ,C ₅ , C ₆	$\varphi^{1:0}, \varphi^{2:0}, \varphi^{3:0},$ $\varphi^{4:0}, \varphi^{5:0}, \varphi^{6:0}$	Atoms/ Groups involved in torsion	C ₀ ,C ₁ ,C ₂ ,C ₃ ,C ₄ , C ₅ ,C ₆	$\varphi^{1:0}, \varphi^{2:0}, \varphi^{3:0},$ $\varphi^{4:0}, \varphi^{5:0}, \varphi^{6:0}$
Si(1)-C(2)- C(3)-C(4)	7.9, -7.8, 2.7, 18.5, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(37)-N(41)- C(13)-C(43)	69.4,11.3, -104.0,-22.7, 102.5,-9.2, -52.8	10.0,1.6, -8.0,1.5, 1.1,-3.0
C(2)-C(3)- C(4)-N(5)	6.9, -17.7, 0.8, 25.6, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	N(41)-C(13)- C(43)-C(47)	0.0,-17.0, 44.8,-10.5, -19.8,3.4, 12.2	23.2,-10.2, -1.1,2.4, 0.0,-2.2
C(3)-C(4)- N(5)-C(6)	8.2, -16.9, 1.13, 26.3, 0.0, 0.0, 0.0	0.0, 0.0, 0.0, 0.0, 0.0, 0.0	C(13)-C(43)- N(47)-C(15)	424.6,56.3, -339.1,39.7, -348.4,89.5, -175.2	-15.7,23.3, -28.8,43.9, 10.1,-26.0
C(4)-N(5)- C(7)-C(25)	109.9,7.4, -9.8,6.8,-36.3, 15.6,-25.4	10.1,3.0, -5.7,-5.2, 1.7,2.0	C(43)-N(47)- C(15)-C(49)	-76.3,-53.1, 135.7,5.2,80. 7,25.4,44.4	8.6,1.5, -10.6,2.9, 279.3,-2.5
N(5)-C(7)- C(25)-N(29)	0.0,20.6,71.04, 31.6,-50.8, -19.2,-23.48	20.7,-16.4, -0.3,-0.4, -0.6,5.0	N(47)-C(15)- C(49)-N(53)	0.0,38.8, 61.3,19.8, -32.6,-3.8, 21.9	13.6,-7.8, -2.1,4.5, -0.0,-3.3
C(7)-C(25)- N(29)-C(9)	424.7,56.3, -339.1,39.7, -348.4,89.5, -175.2	-15.7,23.3, -28.8,43.9, 10.1,-26.0	C(15)-C(49)- N(53)-C(17)	424.6,56.3, -339.1,39.7, -348.4,89.5, -175.2	15.7,23.3, -28.8,43.9, 10.1,-26.0
C(25)-N(29)- C(9)-C(31)	158.8,15.64, -63.2,22.2, -37.9,-13.5, -38.1	12.8,-5.6, -4.1,12.2, 2.6,-4.0	C(49)-N(53)- C(17)-C(55)	28.4,19.1, 17.85,3.15,4 8.17,-37.2, -23.6	7.1,-1.2, -7.7,1.3, 4.8,-2.0
N(29)-C(9)- C(31)-N(35)	0.0,14.9, -117.8,-6.9, 62.7,17.9, -32.5	26.9,-9.0, 42.7,-2.7, -2.0,1.9	N(53)-C(17)- C(55)-N(59)	0.0,22.4,15.8 , -9.3,15.5, -7.9,-8.8	-23.9,-7.6, -7.6,-245.6, 1.1,-0.6
C(9)-C(31)- N(35)-C(11)	424.6,56.3, -339.1,39.7, -348.4,89.5, -175.29	-15.7,23.3, -28.8,43.9, 10.1,-26.0	C(17)-C(55)- N(59)-C(19)	424.6,56.3, -339.1,39.7, -348.4, 89.5, -175.2	-15.7,23.3, -28.8,43.9, 10.1,-26.0
C(31)-N(35)- C(11)-C(37)	101.9,38.7, -108.0,50.2, 53.7,2.89,14.9	6.7,3.1,-8.8, -3.3,2.6,0.4	C(55)-N(59)- C(19)-C(21)	85.1,12.3, -85.6,31.4, -42.8,-37.3, -20.8	-2.4,-0.1, -11.5,-5.3, 7.2,2.0

N(35)-C(11)- C(37)-N(41)	0.0,-17.0, -38.6,-10.4, 19.2,3.8,-11.6	23.1,2.1, -1.1,-0.2, 0.0,-1.8	N(59)-C(19)- C(21)-C(22)	0.0,-8.9, -17.6,14.6, 14.8,3.8,9.2	-4.0,-0.1, 2.4,1.1, -0.2,-1.1
C(11)-C(37)- N(41)-C(13)	424.6,56.3, -339.1,39.7, -348.4,89.5, -175.2	-15.7,23.4, -28.8,43.9, 10.1,-26.0			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_{ω}
N(5)-O(8)-C(25)-C(7)	1.83	390.05
C(7)-C(26)-N(29)-C(25)	-37.36	817.03
C(25)-C(28)-C(9)-N(29)	-18.33	59.98
N(29)-C(31)-O(10)-C(9)	0.00	466.64
C(9)-N(35)-C(32)-C(31)	36.67	818.41
C(31)-C(11)-C(34)-N(35)	11.34	56.30
N(35)-C(37)-O(12)-C(11)	-45.55	468.82
C(11)-C(38)-N(41)-C(37)	-36.04	855.62
C(37)-C(40)-C(13)-N(41)	-2.06	103.05
N(41)-C(43)-O(14)-C(13)	-3.04	432.85
C(13)-N(47)-C(44)-C(43)	35.12	821.47
C(43)-C(46)-C(15)-N(47)	-6.47	63.30
N(47)-C(49)-O(16)-C(15)	-0.34	485.42
C(15)-C(50)-N(53)-C(49)	-37.07	819.98
C(49)-C(52)-C(17)-N(53)	-16.21	83.77
N(53)-C(55)-O(18)-C(17)	-2.12	411.60
C(17)-N(59)-C(56)-C(55)	36.55	829.84
C(55)-C(58)-C(19)-N(59)	-11.00	86.10
N(59)-O(20)-C(21)-C(19)	3.38	393.61

Potentials for α -methyl-9-anthracenemethanol (AnC-CH₃)



Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
H(1)	0.389	0.000	0.000	C(17)	0.216	0.293	0.355
O(2)	-0.672	0.711	0.312	C(18)	-0.203	0.293	0.355
C(3)	0.467	0.276	0.350	C(19)	-0.061	0.293	0.355
H(4)	-0.034	0.126	0.250	C(20)	-0.098	0.293	0.355
C(5)	-0.163	0.276	0.350	C(21)	-0.170	0.293	0.355
H(6)	0.018	0.126	0.250	C(22)	0.065	0.293	0.355
H(7)	0.040	0.126	0.250	H(23)	0.127	0.126	0.242
H(8)	0.013	0.126	0.250	H(24)	0.093	0.126	0.242
C(9)	-0.185	0.293	0.355	H(25)	0.086	0.126	0.242
C(10)	0.100	0.293	0.355	H(26)	0.105	0.126	0.242
C(11)	-0.110	0.293	0.355	H(27)	0.174	0.126	0.242
C(12)	-0.131	0.293	0.355	H(28)	0.106	0.126	0.242
C(13)	-0.050	0.293	0.355	H(29)	0.089	0.126	0.242
C(14)	-0.209	0.293	0.355	H(30)	0.090	0.126	0.242
C(15)	0.200	0.293	0.355	H(31)	0.109	0.126	0.242
C(16)	-0.402	0.293	0.355				

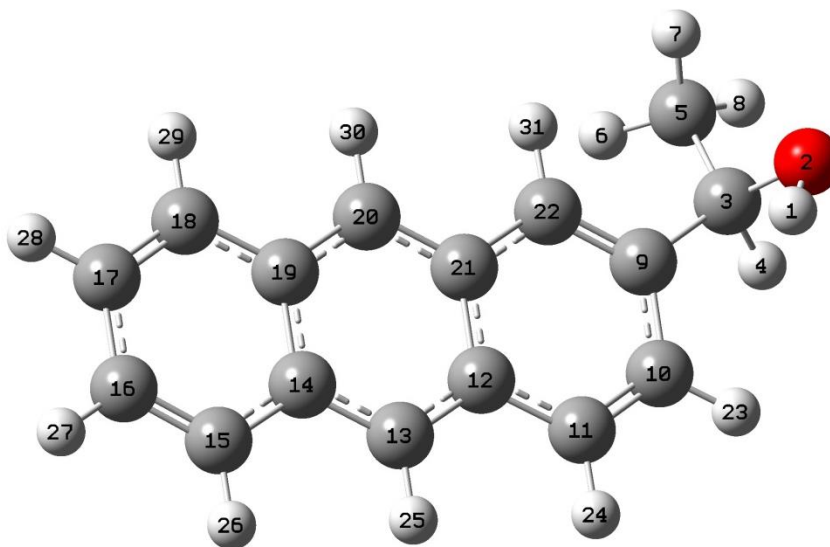
Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(3)-C(9)	0.153	1.9×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
H(1)-O(2)-C(3)	107.12	234.5	C(3)-C(5)-H(8)	110.53	285.4
O(2)-C(3)-H(4)	106.87	384.9	C(4)-C(3)-C(9)	107.99	360.6
O(2)-C(3)-C(5)	111.60	417.9	H(4)-C(3)-C(5)	107.81	293.8
O(2)-C(3)-C(9)	109.83	603.1	C(5)-C(3)-C(9)	112.48	450.1
C(3)-C(9)-C(22)	118.89	573.2	H(6)-C(5)-H(7)	108.43	215.9
C(3)-C(9)-C(10)	121.20	568.1	H(6)-C(5)-H(8)	108.43	215.9
C(3)-C(5)-H(6)	110.53	285.4	H(7)-C(5)-H(8)	108.43	215.9
C(3)-C(5)-H(7)	110.53	285.4			

Torsional potentials					
Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
C(22)-C(9) -C(3)-C(2)	42.3, -9.3, -100.3, 9.3, 127.8, -6.5, -68.4	-42.5, 11.5, 1.0, 8.3, -1.0, -1.0	C(9)-C(3) -C(5)-H(8)	3.6, -20.0, 3.2, 25.8, 2.9, 1.0, 1.8	0.0, 1.2, 0.0, -0.6, -2.0, 0.2
C(9)-C(3) -O(2)-H(1)	13.1, -9.7, -6.9, 6.4, 2.0, 0.9, -0.9	5.6, 2.7, 0.0, -0.1, 0.2, -0.1			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_ω
C(22)-C(3)-C(10)-C(9)	1.81	390.05
C(2)-C(5)-C(9)-C(3)	-32.12	883.44

Potentials for 1-anthracen-2-yl-ethanol (AnS-CH₃)



Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
H(1)	0.389	0.000	0.000	C(17)	-0.074	0.293	0.355
O(2)	-0.659	0.711	0.312	C(18)	-0.198	0.293	0.355
C(3)	0.480	0.276	0.350	C(19)	0.193	0.293	0.355
H(4)	-0.093	0.126	0.250	C(20)	-0.394	0.293	0.355
C(5)	-0.178	0.276	0.350	C(21)	0.242	0.293	0.355
H(6)	0.034	0.126	0.250	C(22)	-0.344	0.293	0.355
H(7)	0.030	0.126	0.250	H(23)	0.101	0.126	0.242
H(8)	0.038	0.126	0.250	H(24)	0.109	0.126	0.242
C(9)	0.084	0.293	0.355	H(25)	0.159	0.126	0.242
C(10)	-0.112	0.293	0.355	H(26)	0.104	0.126	0.242
C(11)	-0.183	0.293	0.355	H(27)	0.089	0.126	0.242
C(12)	0.130	0.293	0.355	H(28)	0.087	0.126	0.242
C(13)	-0.341	0.293	0.355	H(29)	0.106	0.126	0.242
C(14)	0.156	0.293	0.355	H(30)	0.166	0.126	0.242
C(15)	-0.187	0.293	0.355	H(31)	0.142	0.126	0.242
C(16)	-0.079	0.293	0.355				

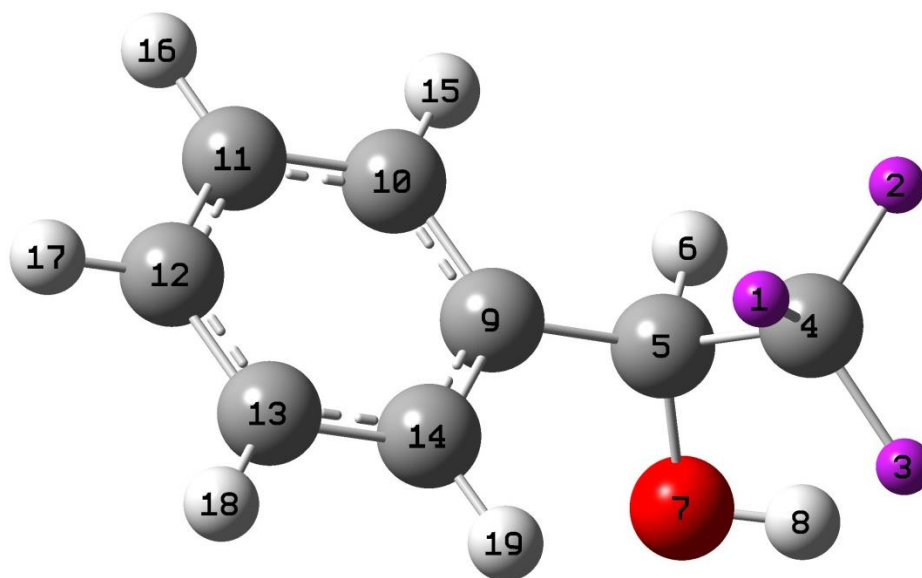
Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(3)-C(9)	0.153	1.9×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
H(1)-O(2)-C(3)	107.63	227.29	C(3)-C(5)-H(8)	110.68	283.66
O(2)-C(3)-H(4)	108.71	374.43	C(4)-C(3)-C(9)	107.05	355.03
O(2)-C(3)-C(5)	110.40	422.84	H(4)-C(3)-C(5)	108.38	288.59
O(2)-C(3)-C(9)	106.82	612.55	C(5)-C(3)-C(9)	115.28	460.59
C(3)-C(9)-C(22)	123.06	558.54	H(6)-C(5)-H(7)	108.22	210.17
C(3)-C(9)-C(10)	117.98	561.32	H(6)-C(5)-H(8)	108.22	210.17
C(3)-C(5)-H(6)	110.68	283.66	H(7)-C(5)-H(8)	108.22	210.17
C(3)-C(5)-H(7)	110.68	283.66			

Torsional potentials					
Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
C(22)-C(9) -C(3)-C(2)	-19.7,-0.1,2.2, 9.5,26.2,-5.0, 12.7	-4.2,3.38,2.0 , -4.7,-.7,14.8	C(9)-C(3) -C(5)-H(6)	6.4,-19.2,3.4, 26.1,-8.2,1.2, 5.2	-2.0,-0.0,-2.1, 0.0,-0.1,-0.05
C(9)-C(3) -O(2)-H(1)	10.9,-1.2,-18.4 ,16.9,23.0,-6.4, -11.5	-2.3,8.2,-2.1, -1.2,-0.3,1.8			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_ω
C(22)-C(3)-C(10)-C(9)	-1.67	312.19
C(2)-C(5)-C(9)-C(3)	-34.58	815.71

Potentials for α -(trifluoromethyl)benzyl alcohol (Ph-CF₃)



Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
F(1)	-0.230	0.222	0.295	C(11)	-0.072	0.293	0.355
F(2)	-0.230	0.222	0.295	C(12)	-0.103	0.293	0.355
F(3)	-0.230	0.222	0.295	C(13)	-0.054	0.293	0.355
C(4)	0.751	0.276	0.350	C(14)	-0.169	0.293	0.355
C(5)	-0.138	0.276	0.350	H(15)	0.110	0.126	0.242
H(6)	0.122	0.126	0.250	H(16)	0.101	0.126	0.242
O(7)	-0.510	0.711	0.312	H(17)	0.095	0.126	0.242
H(8)	0.375	0.000	0.000	H(18)	0.089	0.126	0.242
C(9)	0.167	0.293	0.355	H(19)	0.123	0.126	0.242
C(10)	-0.184	0.293	0.355				

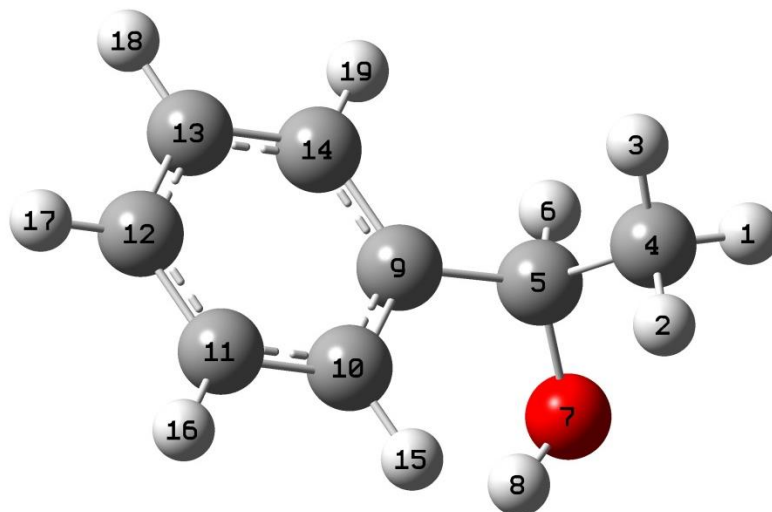
Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(5)-C(9)	0.152	1.9×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
F(1)-C(4)-F(2)	107.59	401.20	C(4)-C(5)-C(9)	112.19	394.75
F(1)-C(4)-F(3)	107.59	401.20	C(5)-O(7)-C(8)	107.55	594.72
F(2)-C(4)-F(3)	107.59	401.20	C(5)-C(9)-C(10)	120.48	287.24
F(1)-C(4)-C(5)	111.28	379.85	C(5)-C(9)-C(11)	120.15	298.06
F(2)-C(4)-C(5)	111.28	379.85	H(6)-C(5)-O(7)	111.21	295.41
F(3)-C(4)-C(5)	111.28	379.85	H(6)-C(5)-C(9)	109.33	346.27
C(4)-C(5)-H(6)	104.85	233.36	O(7)-C(5)-C(9)	113.40	433.22
C(4)-C(5)-O(7)	105.49	588.59			

Torsional potentials					
Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
C(10)-C(9) -C(5)-C(4)	-1.0, 4.2, 4.9, 4.1, 2.9, 2.5, -3.1	2.5, -1.1, -0.8, 0.2, 0.1, -1.2	C(9)-C(5) -O(7)-H(8)	6.2, 6.6, 4.9, 4.2, -11.8, -5.1, 4.6	0.1, -0.2, 4.1, -0.0, -0.9, -0.0
C(9)-C(5) -C(4)-F(1)	0.59, -28.0, 15.00, -37.63, 9.03, 0.83, -6.80	2.04, -0.600, -1.08, 0.95, 0.22, -0.58			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_ω
C(10)-C(5)-C(14)-C(9)	7.23	359.20
C(4)-C(7)-C(9)-C(5)	-36.0	791.54

Potentials for 1-phenylethanol (Ph-CH₃)



Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
H(1)	0.020	0.126	0.250	C(11)	-0.058	0.293	0.355
H(2)	0.020	0.126	0.250	C(12)	-0.110	0.293	0.355
H(3)	0.020	0.126	0.250	C(13)	-0.077	0.293	0.355
C(4)	-0.076	0.276	0.350	C(14)	-0.125	0.293	0.355
C(5)	0.313	0.276	0.350	H(15)	0.100	0.126	0.242
H(6)	-0.017	0.126	0.250	H(16)	0.085	0.126	0.242
O(7)	-0.577	0.711	0.312	H(17)		0.126	0.242
H(8)	0.333	0.000	0.000	H(18)		0.126	0.242
C(9)	0.037	0.293	0.355	H(19)		0.126	0.242
C(10)	-0.174	0.293	0.355				

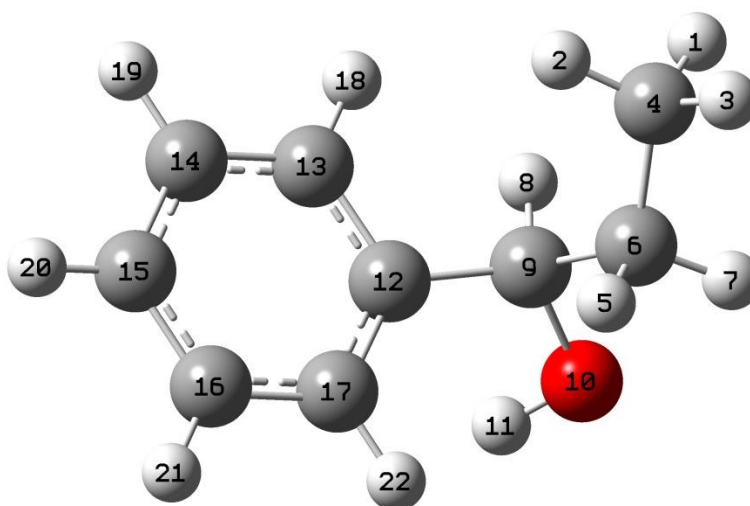
Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(5)-C(9)	0.152	1.9×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
H(1)-C(4)-H(2)	108.72	400.05	C(4)-C(5)-C(9)	112.35	503.75
H(1)-C(4)-H(3)	108.72	400.05	C(5)-O(7)-C(8)	107.75	596.86
H(2)-C(4)-H(3)	108.72	400.05	C(5)-C(9)-C(10)	120.59	289.21
H(1)-C(4)-C(5)	110.21	381.73	C(5)-C(9)-C(11)	120.59	302.06
H(2)-C(4)-C(5)	110.21	366.15	H(6)-C(5)-O(7)	109.62	308.23
H(3)-C(4)-C(5)	110.21	371.69	H(6)-C(5)-C(9)	108.1	454.35
C(4)-C(5)-H(6)	108.49	244.11	O(7)-C(5)-C(9)	112.14	423.16
C(4)-C(5)-O(7)	106.06	386.27		112.35	

Torsional potentials					
Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
C(12)-C(9) -C(5)-C(4)	-12.0,-0.3, 26.7,2.8, 2.7,-2.5,-12.5	-0.1,1.0, -0.0,-0.4, -0.0,-1.9	C(9)-C(5) -O(7)-H(8)	9.0,-9.2, -19.2,8.1, 13.3,3.39,7.4	-0.1,1.8, -0.0,-1.1,0.1, 1.1
C(9)-C(5) -C(4)-H(3)	6.5,-20.7,-3.9, 2.7,-9.6,1.0,6.1	0.0,0.0,0.0, 0.0,0.0,0.0			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_ω
C(10)-C(5)-C(14)-C(9)	0.74	348.27
C(4)-C(7)-C(9)-C(5)	-35.6	827.46

Potentials for 1-phenyl-1-propanol (Ph-C₂H₅)



Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
H(1)	0.017	0.126	0.250	C(12)	0.048	0.293	0.355
H(2)	0.017	0.126	0.250	C(13)	-0.150	0.293	0.355
H(3)	0.017	0.126	0.250	C(14)	-0.089	0.293	0.355
C(4)	-0.104	0.276	0.350	C(15)	-0.092	0.293	0.355
H(5)	-0.064	0.126	0.250	C(16)	-0.089	0.293	0.355
C(6)	0.258	0.276	0.350	C(17)	-0.116	0.293	0.355
H(7)	-0.064	0.126	0.250	H(18)	0.095	0.126	0.242
H(8)	0.012	0.126	0.250	H(19)	0.093	0.126	0.242
C(9)	0.154	0.276	0.350	H(20)	0.087	0.126	0.242
O(10)	-0.563	0.711	0.312	H(21)	0.088	0.126	0.242
H(11)	0.346	0.000	0.000	H(22)	0.098	0.126	0.242

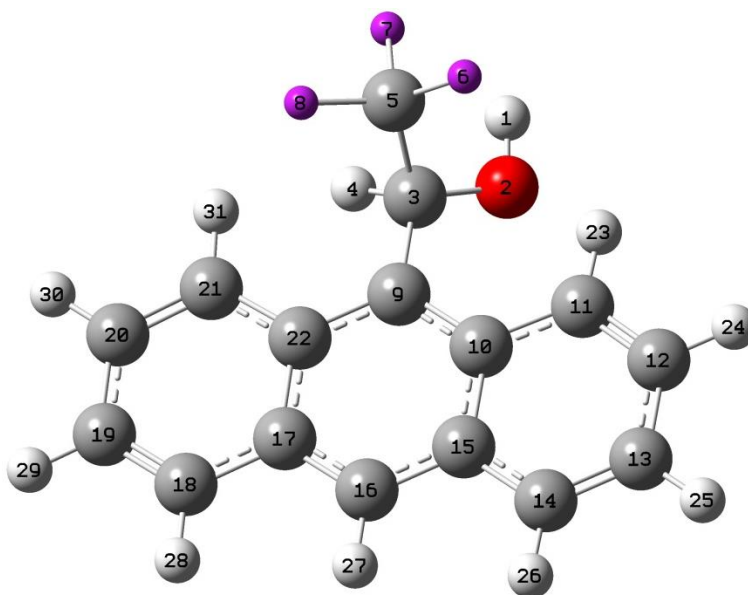
Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(9)-C(12)	0.152	1.9×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
H(1)-C(4)-H(2)	108.23	255.37	C(6)-C(9)-C(8)	108.19	344.21
H(1)-C(4)-H(3)	108.23	255.37	C(6)-O(9)-C(10)	106.54	370.47
H(2)-C(4)-H(3)	108.23	255.37	C(6)-C(9)-C(13)	112.14	340.89
H(1)-C(4)-C(6)	110.91	369.12	C(7)-C(6)-C(9)	107.98	386.15
H(2)-C(4)-C(6)	110.91	369.12	H(8)-C(9)-O(10)	109.59	296.94
H(3)-C(4)-C(6)	110.91	369.12	H(8)-C(9)-C(13)	108.26	295.62
C(4)-C(6)-H(5)	110.34	330.97	O(9)-C(10)-C(11)	107.87	322.06
C(4)-C(6)-O(7)	110.34	330.97	O(9)-C(13)-C(14)	120.59	283.72
C(4)-C(6)-O(9)	113.21	335.14	O(9)-C(13)-C(15)	120.59	285.25
C(5)-C(6)-O(7)	106.72	241.25	O(10)-C(9)-C(13)	112.04	349.74
C(5)-C(6)-O(9)	107.98	391.48			

Torsional potentials					
Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$	Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi^{1;0}, \varphi^{2;0}, \varphi^{3;0}, \varphi^{4;0}, \varphi^{5;0}, \varphi^{6;0}$
C(17)-C(12)-C(9)-C(6)	1.0,-4.3,10.0,-2.8,8.4,-3.8,2.8	-4.2,-.3,0.5,0.6,-0.8,-0.1	C(9)-C(6)-O(4)-H(3)	5.4,-18.0,1.9,-23.5,-5.5,-0.8,3.7	2.0,-0.0,-1.0,-0.0,-0.2,-0.0
C(12)-C(9)-C(6)-H(4)	-3.93,-24.05,16.25,36.09,13.03,-1.05,5.60	-0.07,0.94,-0.02,-0.73,-0.16,-0.02	C(12)-C(9)-O(10)-H(11)	9.5,9.1,-20.6,8.1,14.8,2.6,7.9	2.8,-1.3,0.0,-1.1,0.2,1.2

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_ω
C(13)-C(9)-C(17)-C(12)	-0.86	299.13
C(6)-C(10)-C(12)-C(9)	-34.0	784.77

Potentials for 1-(9-anthryl)-2,2,2-trifluoroethanol (AnC-CF₃)



Numbering, partial charge, and LJ parameters							
Atom label and number	Charge	ϵ	σ	Atom label and number	Charge	ϵ	σ
H(1)	0.376	0.000	0.000	C(17)	0.275	0.293	0.355
O(2)	-0.576	0.711	0.312	C(18)	-0.225	0.293	0.355
C(3)	0.226	0.276	0.350	C(19)	-0.056	0.293	0.355
H(4)	0.035	0.126	0.250	C(20)	-0.111	0.293	0.355
C(5)	0.642	0.276	0.350	C(21)	-0.124	0.293	0.355
F(6)	-0.237	0.222	0.295	C(22)	0.016	0.293	0.355
F(7)	-0.237	0.222	0.295	H(23)	0.120	0.126	0.242
F(8)	-0.237	0.222	0.295	H(24)	0.090	0.126	0.242
C(9)	-0.150	0.293	0.355	H(25)	0.091	0.126	0.242
C(10)	0.090	0.293	0.355	H(26)	0.112	0.126	0.242
C(11)	-0.133	0.293	0.355	H(27)	0.180	0.126	0.242
C(12)	-0.087	0.293	0.355	H(28)	0.114	0.126	0.242
C(13)	-0.066	0.293	0.355	H(29)	0.093	0.126	0.242
C(14)	-0.214	0.293	0.355	H(30)	0.101	0.126	0.242
C(15)	0.229	0.293	0.355	H(31)	0.093	0.126	0.242
C(16)	-0.430	0.293	0.355				

Stretching potentials		
Atoms/Groups involved in Stretch	r_{eq}	k_r
C(3)-C(9)	0.152	1.9×10^5

Bending potentials					
Atoms/Groups involved in bend	θ_{eq}	k_θ	Atoms/Groups involved in bend	θ_{eq}	k_θ
H(1)-O(2)-C(3)	107.66	590.98	C(3)-C(5)-F(8)	381.30	363.79
O(2)-C(3)-H(4)	108.68	309.18	C(4)-C(3)-C(9)	108.43	343.99
O(2)-C(3)-C(5)	109.79	600.35	H(4)-C(3)-C(5)	103.99	232.29
O(2)-C(3)-C(9)	111.62	418.10	C(5)-C(3)-C(9)	113.93	388.82
C(3)-C(9)-C(22)	118.03	292.77	F(6)-C(5)-F(7)	106.87	367.11
C(3)-C(9)-C(10)	121.57	296.76	F(6)-C(5)-F(8)	106.87	367.11
C(3)-C(5)-F(6)	109.01	381.30	F(7)-C(5)-F(8)	106.87	367.11
C(3)-C(5)-F(7)	109.01	381.30			

Torsional potentials					
Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi_{0,0}^{1;0}, \varphi_{0,0}^{2;0}, \varphi_{0,0}^{3;0}, \varphi_{0,0}^{4;0}, \varphi_{0,0}^{5;0}, \varphi_{0,0}^{6;0}$	Atoms/Groups involved in torsion	$c_0, c_1, c_2, c_3, c_4, c_5, c_6$	$\varphi_{0,0}^{1;0}, \varphi_{0,0}^{2;0}, \varphi_{0,0}^{3;0}, \varphi_{0,0}^{4;0}, \varphi_{0,0}^{5;0}, \varphi_{0,0}^{6;0}$
C(22)-C(9) -C(3)-C(2)	33.6, -4.4, 58.0, -10.2, -65.8, -5.9, -32.4	-11.5, -38.9, -14.8, -11.1, 0.7, 5.0	C(9)-C(3) -C(5)-F(8)	11.9, -6.6, -3.9, -5.2, -9.4, -3.31, -3.8	4.2, 1.5, 1.0, -0.1, -8.0, 0.6
C(9)-C(3) -O(2)-H(1)	2.9, -23.1, 5.1, 31.4, 11.3, -0.6, -6.3	-0.0, 5.5, 0.0, 0.8, -0.1, 0.8			

Improper torsional potentials		
Atoms/Groups involved in improper torsion	ω_{eq}	k_ω
C(22)-C(3)-C(10)-C(9)	0.51	373.47
C(2)-C(5)-C(9)-C(3)	-31.6	854.63