

MOLECULAR DYNAMICS SIMULATIONS AND VIRTUAL SCREENING TO IDENTIFY
POTENT INHIBITORS OF HUMAN ASPARAGINE SYNTHETASE

By

LAMEES HEGAZY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2013

© 2013 Lamees Hegazy

To the martyrs and wounded of the Egyptian revolution

ACKNOWLEDGMENTS

First and foremost, I thank the almighty God, for giving me so many blessings and perseverance to complete this journey. Special thanks to the martyrs and wounded of the Egyptian revolution. Your sacrifices for the sake of justice and freedom will always be remembered.

I would like to express my deepest gratitude to my advisor, Dr. Nigel Richards, for his continuous help and support even after leaving the University of Florida. I could not accomplish this work without his encouragement, guidance, and effort throughout the course of this work. I am grateful to Dr. Adrian Roitberg for his support and acceptance to chair the committee after Dr. Nigel Richards' move to Indiana University–Purdue University. I would like to extend my gratitude to all my committee members: Dr. David Ostrov, Dr. Jon Stewart, Dr. Steve Bruner, Dr. Rebecca Butcher, and former member Dr. Kenneth Merz for their assistance and all of their contribution. I would also like to thank former and current members of the Richards group. Special thanks to Dr. Kathryn Williams for her help with thesis editing.

I cannot express my appreciation to my father and my mother for their love, prayers, unconditional support, and believing in me. I also thank my brothers Sherief and Muhammad and my sisters Alaa and Sarah for their continuous love and support. I have been blessed with two little angels Albaraa and Darine who brought joy and happiness to my life. Finally, I would like to thank my husband Bahaa Elgendy. His love, support, patience and encouragement were definitely the foundation upon which the past five years of my life have been built. Thank you will never be enough.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	7
LIST OF FIGURES.....	9
ABSTRACT	15
CHAPTER	
1 INTRODUCTION	17
Asparagine Synthetase Structure and Function.....	17
Asparagine synthetase kinetic mechanism	21
Asparagine Synthetase Inhibitors	24
Research Objectives.....	26
2 OPTIMIZED CHARMM GENERAL FORCE FIELD PARAMETERS FOR ACYLPHOSPHATE AND N-PHOSPHONOSULFONIMIDOYL FUNCTIONAL GROUPS	29
Background.....	29
Molecular Mechanics Force Fields	30
CGenFF Potential Energy Function.....	31
CGenFF Parametr Optimization Protocol	35
Computational Details.....	36
Results and Discussion.....	38
Identification of Missing Parameters.....	38
Charge Optimization.....	40
Optimization of Equilibrium Values and Force Constants for Bond Lengths and Angles	44
Optimization of Dihedral Terms	50
Molecular Dynamics Simulations for Model Compounds in Aqueous Solution	54
Summary	57
3 MOLECULAR DYNAMICS SIMULATIONS OF UNLIGANDED ASPARAGINE SYNTHETASE AND ASNS/BASPAMP COMPLEX.....	60
Background.....	60
Molecular Dynamics Simulation	62
Computational Details.....	66
Results and Discussion.....	72
Root Mean Square Deviation (RMSD)	72

Root Mean Square Fluctuations (RMSF).....	72
Mg ²⁺ Coordination	74
Hydrogen Bonds.....	79
Behavior of Important Residues in the Active Site of Unliganded ASNS and the ASNS/ β AspAMP/PPi/Mg ²⁺ Complex.....	86
Glu-348	86
Glu-352	89
Asp-384.....	92
Lys-376	92
Lys-429	97
Ser-346	97
Ser-234 and Ser-239	97
Lys-449	103
Loop 443-452 Mobility	108
Summary	112
4 DISCOVERY OF A NEW INHIBITOR OF HUMAN ASPARAGINE SYNTHETASE IDENTIFIED BY STRUCTURE-BASED VIRTUAL SCREENING.	115
Background.....	115
Structure-Based Virtual Screening.....	117
Computational Details.....	120
Virtual Screening Using GLIDE	120
Virtual Screening Using DOCK.....	122
Results and Discussion.....	123
Summary	136
5 CONCLUSIONS AND FUTURE WORK	138
Concluding Remarks.....	138
Future Work: Free Energy Calculations to Estimate the Difference in Binding Free Energy between (<i>R</i>)- and (<i>S</i>)- Isomers of an <i>N</i> -phosphonosulfonimidoyl Based Inhibitor.....	142
APPENDIX	
A QM AND MM OPTIMIZED VIBRATIONAL SPECTRA FOR <i>N</i> - PHOSPHONOSULFONIMIDOYL MODEL COMPOUND.	151
B CHARMM TOPOLOGY AND PARAMETER FILE FOR ACYLPHOSPHATE AND <i>N</i> -PHOSPHONOSULFONIMIDOYL MODEL COMPOUNDS.....	154
CHARMM Topology and Parameter File for Acylphosphate Model	154
CHARMM Topology and Parameter File for <i>N</i> -phosphonosulfonimidoyl Model....	156
LIST OF REFERENCES	158
BIOGRAPHICAL SKETCH.....	167

LIST OF TABLES

<u>Table</u>	<u>page</u>
2-1	Comparison of HF/6-31G(d) and CHARMM interaction energies and interaction distances before and after optimization for the acylphosphate molecule. 42
2-2	Comparison of HF/6-31G(d) and CHARMM interaction energies and interaction distances before and after optimization for the <i>N</i> -phosphonosulfonimidoyl molecule. 42
2-3	Optimized atomic partial charges for atoms in the acylphosphate and <i>N</i> -phosphonosulfonimidoyl molecules. 43
2-4	Comparison of the CHARMM optimized geometry for missing parameters with the QM geometry for the acylphosphate molecule. 46
2-5	Comparison of the CHARMM optimized geometry for missing parameters with the QM geometry for the <i>N</i> -phosphonosulfonimidoyl molecule. 47
2-6	Vibrational spectra computed for the acylphosphate molecule at the scaled MP2 level and with optimized CGenFF parameters 48
2-7	New bonded interaction parameters assigned for the acylphosphate moiety. 58
2-8	New bonded interaction parameters assigned for the <i>N</i> -phosphonosulfonimidoyl moiety 59
3-1	Hydrogen bonding interactions of β -aspartyl-AMP (BAA) with active site residues. 82
3-2	Hydrogen bonding interactions of the pyrophosphate group (PPi) with amino acid residues in the active site. 83
3-3	Hydrogen bonding interactions between the residues in the active site in MD simulations of ASNS/ β AspAMP. 84
3-4	Hydrogen bonding interactions between the residues in the active site in MD simulations of unliganded ASNS. 85
4-1	Docking score of the ten selected hits from GLIDE virtual screening, their Lipinski properties and their residual activity values. 124
4-2	pKas of amide hydrogens of some sulfonamide derivatives in water 132
4-3	Docking score of some selected hits from DOCK virtual screening 135

A-1 Vibrational spectra computed for the *N*-phosphonosulfonimidoyl molecule at the scaled MP2 level and with optimized CGenFF parameters 151

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Overview of the reaction catalyzed by glutamine-dependent asparagine synthetase.	20
1-2 Crystal structure of AS-B.....	20
1-3 Proposed kinetic model for glutamine- dependent asparagine synthetase.	23
1-4 Chemical structures of hASNS identified inhibitors.	25
1-5 Conjugate bases of the functionalized acylsulfamate 2 and sulfamide 3	25
1-6 Pairwise sequence alignment between the <i>E. coli</i> and human asparagine synthetases.....	28
2-1 CGenFF parameters optimization process.....	34
2-2 Parts of the molecules that require extensive optimization are enclosed in boxes.	39
2-3 The equilibrium geometry and atomic labels used in the text and tables.	39
2-4 Water–model compound complexes.....	41
2-5 Potential energy scans on the acylphosphate molecule dihedrals.....	51
2-6 Potential energy scans on the <i>N</i> -phosphonosulfonimidoyl molecule dihedrals.....	52
2-7 MD trajectory data showing that the phosphate moiety in the acylphosphate model undergoes rotation during the simulation.....	55
2-8 MD trajectory data showing that the phosphate moiety in the acylphosphate model undergoes rotation during the simulation.	56
3-1 Structure based sequence alignment of β -LS and ASNS.....	67
3-2 Stereoview of the AS-B and β -LS active sites.....	68
3-3 Comparison between Chemical reactions of AS-B and β -LS.	70
3-4 Root mean square deviation of backbone atoms for unliganded and ligand bound asparagine synthetase enzyme.....	73
3-5 RMSF of backbone atoms for first independent trajectory simulations for unliganded and ligand bound asparagine synthetase enzyme.	75

3-6	RMSF of backbone atoms for the second independent trajectory simulations for unliganded and ligand bound asparagine synthetase enzyme	76
3-7	RMSF of backbone atoms for third independent trajectory simulations for unliganded and ligand bound asparagine synthetase enzyme	77
3-8	Enzyme regions with changes in RMSF of backbone atoms greater than 0.5 Å are colored.....	78
3-9	Focused view of the Mg ²⁺ ions' coordinating residues	80
3-10	Hydrogen bonding interactions of βAspAMP with active site residues.....	80
3-11	Interactions of PPI with active site residues..	83
3-12	Glu-348 χ ₁ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	87
3-13	Glu-348 χ ₂ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	88
3-14	Glu-352 χ ₁ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	90
3-15	Glu-352 χ ₂ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	91
3-16	Asp-384 χ ₁ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	93
3-17	Asp-384 χ ₂ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	94
3-18	Lys 376 χ ₁ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	95
3-19	Lys-376 χ ₂ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	96
3-20	Lys-429 χ ₁ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	98
3-21	Lys-429 χ ₂ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	99
3-22	Ser346 χ ₁ torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	100

3-23 Ser234 χ_1 torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	101
3-24 Ser239 χ_1 torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	102
3-25 Lys-449 χ_1 torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	104
3-26 Lys-449 χ_2 torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	105
3-27 Lys-449 χ_3 torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	106
3-28 Lys-449 χ_4 torsion change during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme	107
3-29 The closed and open forms of loop 443-452 in ASNS.	110
3-30 Structural overlay of the conserved lysine residue, β AspAMP intermediate in AS-B, CMA-AMP intermediate analog in β -LS and the product DGPC in β -LS ...	110
3-31 Distance between K449-N and D351-C α during MD simulations of the unliganded and ligand bound asparagine synthetase enzyme.	111
4-1 Overlay of the re-docked β AspAMP intermediate using DOCK and the native binding pose of β AspAMP	124
4-2 2D predicted binding poses of the ten selected hits from GLIDE virtual screening.	125
4-3 Glutamine-dependent production of PPI in the presence of the inhibitor NSC 605322.....	128
4-4 Ammonia-dependent production of PPI in the presence of the inhibitor NSC 605322.....	129
4-5 Putative interactions of compound NSC605322 with ASNS active site residues.	131
4-6 Overlay of the docking pose of ASNS-inhibitor complex and the ASNS- β AspAMP complex	131
4-7 Conjugate base of ionized sulfonamide functional group of compound NSC605322.	132
4-8 2D predicted binding poses of some hits identified by virtual screening using DOCK.	134

5-1	Chemical structures of hASNS <i>N</i> -phosphonosulfonimidoyl-based inhibitors.	141
5-2	Overview of the reaction catalyzed by glutamine-dependent asparagine synthetase.	141
5-3	Putative interactions of the four diastereoisomers with key residues in the ASNS active site.	143
5-4	Electrostatic potential in the synthetase active site of the AS-B model.	145
5-5	Thermodynamic cycle used for the calculation of relative free energy of binding of (<i>R</i>)- and (<i>S</i>)- diastereoisomers of the phosphonosulfonimidoyl inhibitor	147
5-6	Pathway for the transmutation between the (<i>R</i>)- and (<i>S</i>)- phosphonosulfonimidoyl inhibitor	150

LIST OF ABBREVIATIONS

ALL	Acute lymphoblastic leukemia
AMP-CPP	α,β -methyleneadenosine-5'-triphosphate
AS-B	<i>Escherichia coli</i> glutamine dependent ASNS
ASNase	L-asparaginase
ASNS	Glutamine dependent asparagine synthetase
ATP	Adenosine triphosphate
CEA	<i>N</i> ² -(carboxyethyl)-L-arginine
CGenFF	CHARMM general force field
DGPC	Deoxyguanidinoproclavaminic acid
hASNS	Human asparagine synthetase
HTS	High-throughput screening
LBVS	Ligand based virtual screening
LJ	Lennard-Jones
MD	Molecular Dynamics
MM	Molecular mechanics
PES	Potential energy surface
PPi	Inorganic pyrophosphate
QM	Quantum mechanics
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuations
SBVS	Structure based virtual screening
UB	Urey-Bradley distances
vdW	van der Waals interactions
VS	Virtual screening

β AspAMP

Beta-aspartyl adenosine monophosphate

β -LS

Beta-lactam synthetase

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MOLECULAR DYNAMICS SIMULATIONS AND VIRTUAL SCREENING TO IDENTIFY
POTENT INHIBITORS OF HUMAN ASPARAGINE SYNTHETASE

By

Lamees Hegazy

December 2013

Chair: Adrian Roitberg
Major: Chemistry

Asparagine synthetase (ASNS), which mediates the biosynthesis of L-asparagine, has become of increasing interest as a drug target. Overexpression of ASNS was observed in ASNase-resistant leukemia cell lines and inhibitors of human asparagine synthetase were able to suppress the proliferation of ASNase resistant leukemia cells. Human asparagine synthetase is also involved in other forms of cancer, and was shown to be a predictive biomarker of L-asparaginase activity in ovarian cancer cell lines. In addition, recent work suggested that ASNS is up-regulated in the castration-resistant stage of prostate cancer (CRPC) where knockdown of ASNS in asparagine-depleted media led to growth inhibition in in both androgen-responsive lymph nodes and castration-resistant prostate cancer cell lines. Thus, depletion of asparagine using ASNS inhibitors may be a new approach for treatment of CRPC.

In this study, new parameters that describe the acylphosphate and *N*-phosphonosulfonimidoyl functional groups were developed following the standard protocols of CHARMM general force field. In addition, molecular dynamics simulations were performed on ASNS complexed with a key reaction intermediate. The ligand-free

ASNS was also simulated to detect related conformational changes as a result of intermediate binding. A model representative of the average structure was further used in virtual screening calculations to identify potent ASNS inhibitors. A total of 500,000 compounds from the NCI chemical database were docked to the ASNS active site. A new lead inhibitor of hASNS was identified by these docking studies. This lead inhibitor has a new structural motif that should permit the development of compounds with improved cell permeability than those previously reported.

CHAPTER 1 INTRODUCTION

Asparagine Synthetase Structure and Function

Asparagine synthetase is an ATP-dependent enzyme that catalyzes the synthesis of L-asparagine from L-aspartate with glutamine or ammonia as the nitrogen source (Figure 1-1).¹ Asparagine synthetase catalyzes three reactions:

1. $\text{Gln} + \text{H}_2\text{O} \longrightarrow \text{Glu} + \text{NH}_3$
2. $\text{ATP} + \text{Asp} + \text{NH}_3 \longrightarrow \text{AMP} + \text{PPi} + \text{Asn}$
3. $\text{ATP} + \text{Asp} + \text{Gln} \longrightarrow \text{AMP} + \text{PPi} + \text{Asn} + \text{Glu}$

Reaction 1 is the glutaminase reaction in which glutamine is hydrolyzed to glutamate and ammonia. Reactions 2 and 3 are the ammonia and glutamine dependent reactions of asparagine synthetase respectively. The *Escherichia coli* enzyme, asparagine synthetase B (AS-B), and the human enzyme (hASNS) employ glutamine as the primary source of nitrogen. Glutamine-dependent asparagine synthetase catalyzes the transfer of glutamine amide nitrogen to another substrate (Figure 1-1) and is classified as Ntn or class II amidotransferase, because it has an N-terminal conserved cysteine residue.² Other members of this class include glutamine fructose 6-phosphate amidotransferase (GFAT),³ glutamine-5-phosphoribosyl pyrophosphate amidotransferase (GPAT),⁴ and glutamine synthetase.⁵

Asparagine synthetase is important as a drug target. Overexpression of ASNS is observed in ASNase-resistant leukemia cell lines,⁶ and several studies suggest that inhibiting ASNS activity represents a possible approach for treating ASNase-resistant acute lymphoblastic leukemia (ALL).^{7,8,9} The enzyme L-asparaginase (ASNase), in combination with other drugs, has been commonly used in the chemotherapy of ALL.^{10,11} The molecular basis of ASNase chemotherapy is not completely understood,

but it is believed that the malignant lymphocytes produce asparagine at lower levels than normal lymphocytes. Therefore, ASNase treatment causes asparagine reduction in both normal and malignant cells. Normal cells are able to produce asparagine in adequate amount to respond to this asparagine depletion. However, cancer cells produce asparagine in lower rate, leading to suppression of the growth of leukemia cells,^{7,12} and continuous treatment with ASNase leads to remission.¹³ Chemotherapy of ALL using *Escherichia coli* L-asparaginase in combination with other drugs has been very successful with remission rates of 95%.^{10,14} However, 12% of patients suffer from development of drug resistance during the treatment, leading to the failure of treatment protocol.^{15,16,17} Several studies have reported a correlation between ASNase resistance and elevated expression levels of glutamine-dependent asparagine synthetase (ASNS), which catalyzes the biosynthesis of L-asparagine (Figure 1-1).¹⁸⁻²¹ Kilberg *et al* have shown that asparagine synthetase expression alone is sufficient to induce L-asparaginase resistance in MOLT-4 human leukaemia cells. In their study, the drug resistance phenotype was induced when parental MOLT-4 cells were transduced by ASNS-expressing virus.⁶ In addition, a potent inhibitor of human asparagine synthetase (hASNS), compound **1**, suppressed the proliferation of ASNase resistant leukemia cells (Figure 1-4).²² This finding supports the hypothesis that inhibitors of ASNS can be used in drug discovery for treatment of ALL.

Human asparagine synthetase is also involved in other forms of cancer, and was shown to be a predictive biomarker of L-asparaginase activity in ovarian cancer cell lines.^{23,24} In addition, recent work suggested that ASNS is up-regulated in the castration-resistant stage of prostate cancer (CRPC) in which knockdown of ASNS by

small-interfering RNAs in asparagine-depleted media leads to growth inhibition in both androgen-responsive lymph nodes and castration-resistant prostate cancer cell lines. Thus depletion of asparagine using ASNS inhibitors may be a new approach for treatment of CRPC.²⁵

Studying the structural and kinetic properties of the hASNS has been difficult due to low availability and purification problems. Alternatively, studies on *Escherichia coli* glutamine dependent ASNS (AS-B) have provided detailed information about the structure and reaction mechanism of the enzyme. The N-terminal cysteine residue is conserved among the Ntn amidotransferase family and is the nucleophile for hydrolysis of glutamine. Therefore, the C1A mutant of the enzyme AS-B does not have glutaminase activity. It does, however, have high glutaminase affinity and was crystallized with glutamine bound in the N-terminal glutaminase domain and AMP bound in the C-terminal synthase domain, with resolution of 2.0 Å (Figure 1-2).¹

As expected from previous studies using monoclonal antibodies of bovine ASNS and from sequence alignment, the crystal structure of C1A mutant of AS-B (PDB 1CT9) showed that AS-B has two distinct domains.¹ Each domain has a separate active site, the N-terminal glutaminase domain and the C-terminal synthetase domain. The N-terminal active site catalyzes the glutamine hydrolysis to glutamate and ammonia. The N-terminal active site catalyzes the glutamine hydrolysis to glutamate and ammonia. The N-terminal domain's tertiary structure is similar to that of the N-terminal amidohydrolase superfamily.^{26,27}

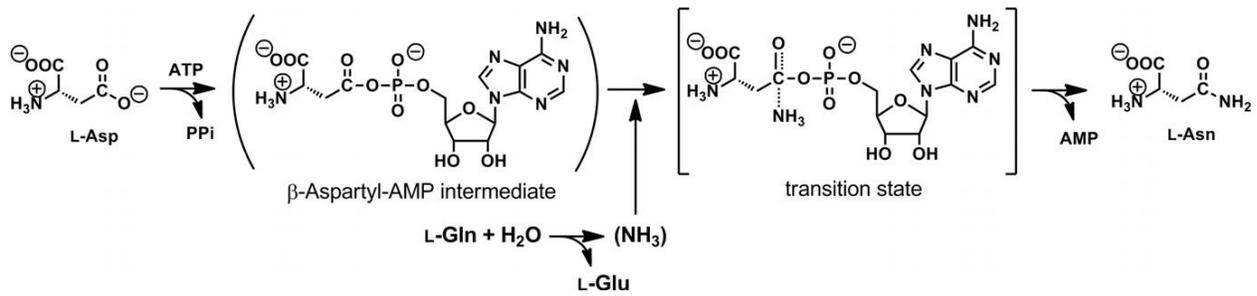


Figure 1-1. Overview of the reaction catalyzed by glutamine-dependent asparagine synthetase.

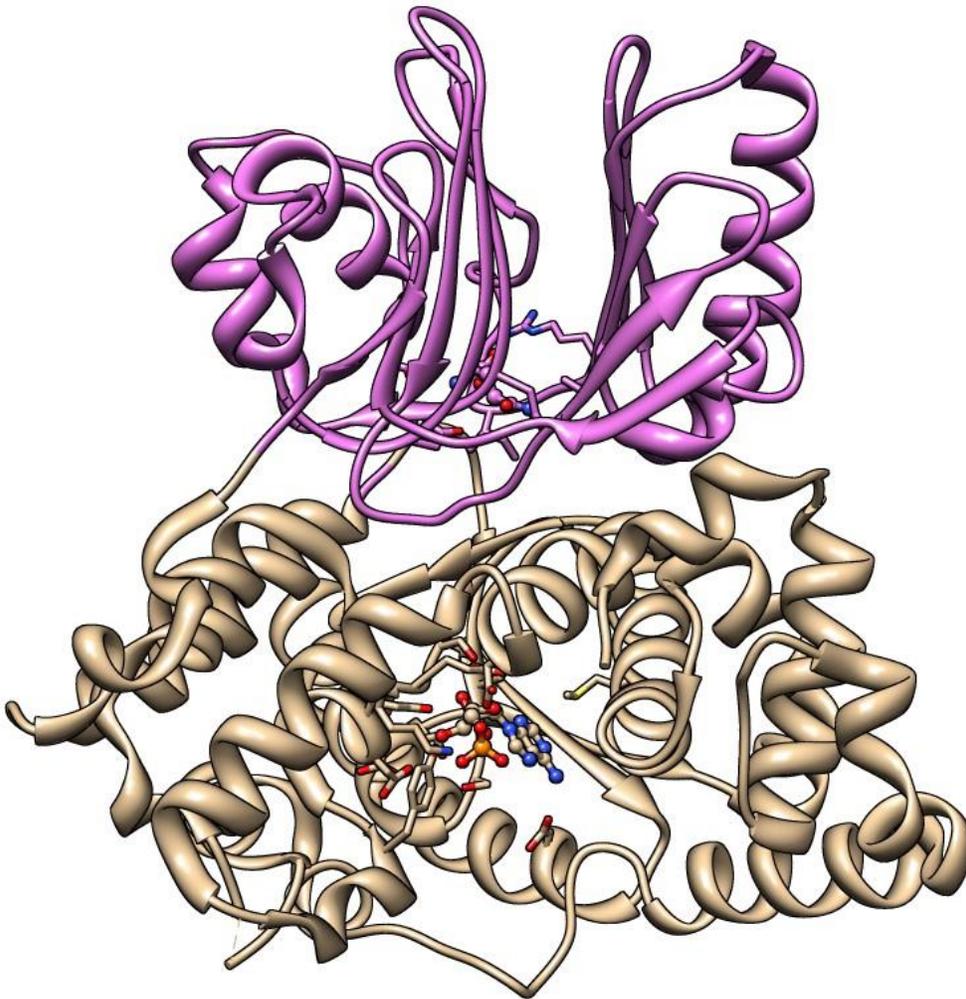


Figure 1-2. Crystal structure of AS-B (PDBID; 1CT9). Glutaminase domain is shown in purple and synthase domain in brown. Glutamine and AMP are shown in ball and stick representation. Image rendered in CHIMERA.²⁸

The C-terminal domain uses ATP and Mg^{2+} as cofactors and catalyzes the conversion of aspartate to asparagine via an acyl intermediate β -aspartyl-AMP (β AspAMP). The C-Terminal active site tertiary structure is observed in ATP pyrophosphatases such as GMP synthetase,^{29,30} arginosuccinate synthetase,²⁹ β -lactam synthetase,³¹ and carbenam synthetase.³² The two active sites are separated by a solvent inaccessible tunnel, through which ammonia moves from the glutaminase domain to the synthetase domain.¹

Asparagine synthetase kinetic mechanism

An understanding of the steady state kinetics of the glutamine-dependent ASNS, and identification of key residues involved in the binding of the intermediate β AspAMP, will aid in the design of selective inhibitors of ASNS.

Asparagine synthetase is characterized by high glutaminase activity in the absence of aspartate. In addition, glutamine production is not strictly coordinated with asparagine production.^{33,34} Studies using human asparagine synthetase (hASNS) incubated with glutamine in the absence of aspartate showed that the glutaminase activity increases.³⁴ This behavior was also observed in the *Escherichia coli* ASNS (AS-B), where the Glu/Asn ratio increased with increasing concentration of L-glutamine, reaching a limiting value of approximately 1.8.³⁵ The weak coordination between the glutaminase and synthase domains is unusual and contrasts with other glutamine-dependent amidotransferases, such as glutamine 5'-phosphoribosylpyrophosphate amidotransferase (GPATase), in which higher glutaminase activity is observed in the presence of substrates or reactive intermediates in the synthase domain.^{36,37} A notable important characteristic of ASNS behavior is the lack of ATP/PPi exchange.³⁸ Some of class II glutamine amidotransferase subfamily of enzymes such as GPATase³⁹ and

xanthylate synthetase exhibit this behavior.^{40,41} The absence of ATP/PPi exchange leads to the assumption that PPi is released as the final product from the enzyme. This hypothesis is supported by the crystal structures of the structurally similar enzyme β -lactam synthetase bound with the adenylated intermediate α,β -methyleneadenosine-5'-triphosphate (β LS/CMA-AMP/PPi) and of β -lactam synthetase complexed with the product deoxyguanidinoproclavaminic acid (DGPC) (β LS/DGPC/PPi), both of them show that PPi is located deep in the active site.⁴² This contrasts with the ASNS from *Vibrio cholerae*, where PPi release was reported to occur before glutamine binding.⁴³ The amino acid residues lining the binding pocket of PPi are also conserved in AS-B and other ATP pyrophosphatase enzymes. The characteristic sequence SGGXDS is referred to as the PP-motif.⁴⁴

Several kinetic mechanisms have been proposed for AS-B. However, kinetic simulations, showed that only one kinetic model was consistent with all observed experimental data for AS-B (Figure 1-3), and was also able to reproduce the glutamine-dependent Glu/Asn stoichiometry at saturating levels of aspartate and ATP.⁴⁵ Rate constants for each of the steps in this kinetic model were assigned from literature data on the steady state kinetics of AS-B.^{34,46} The production of β AspAMP exhibits burst kinetics in the absence of a nitrogen source,⁴⁷ where the rate of formation exceeds that of its hydrolytic breakdown and product release. Therefore, the turnover number was assigned to k_7 which includes the steps involved in the release of asparagine, AMP, PPi, and glutamate from the enzyme active sites (Figure 1-3). The rate constant k_3 was adjusted by numerical simulation to imitate the experimental Glu/Asn stoichiometry.

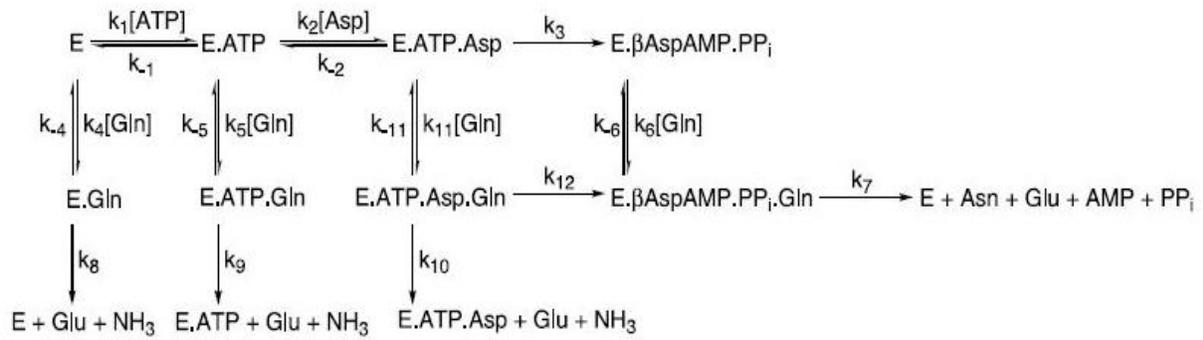


Figure 1-3. Proposed kinetic model for glutamine- dependent asparagine synthetase. Reprinted with permission from Archive of Biochemistry and Biophysics vol 413. Tesson, A.R, Soper, T.S., Ciusteau, M., Richards, N.G.J. Pages 23-31. Copyright 2003.

In this kinetic model, glutamine binds to the ternary complex E.ATP.ASP to yield the quaternary complex E.ATP.Asp.Gln, from which glutamine can be hydrolyzed to ammonia and glutamate before β AspAMP formation. Therefore, this model suggests that the two active sites are weakly coupled before the formation of β AspAMP, which commits the enzyme to asparagine formation, in agreement with the previously proposed model.^{35,48} The PPI is released as the final product, thereby rationalizing the lack of ATP/PPi exchange. This kinetic model also indicates that ASNS must bind β AspAMP with high affinity and stable analogs of this intermediate and the transition state may be potent ASNS inhibitors (Figure 1-1).

Asparagine Synthetase Inhibitors

Kinetic studies suggested that the β AspAMP intermediate must be stabilized within the active site, and that analogues for the ASNS intermediate or the transition state should act as tight binding inhibitors.⁴⁵ An *N*-phosphonosulfonimidoyl inhibitor **1** that is a stable analogue of the ASNS transition state inhibited hASNS with nanomolar potency (Figure 1-4).²² In addition, this inhibitor slowed the growth of ASNase-resistant leukemia cells when used in high concentrations (100–1000 μ M) and when L-asparagine was reduced in the external medium by ASNase.⁴⁹ These findings validated the hypothesis that inhibitors of human asparagine synthetase (hASNS) can be used in the discovery of therapeutic drugs to treat ALL. According to one hypothesis, *N*-phosphonosulfonimidoyl **1** is needed in high concentrations to suppress the growth of ASNase-resistant leukemia cells because the charged groups prevent its entry into the cell.²² Further studies with acyl-adenylate analogue **2**, in which the phosphate group is replaced by a sulfamate group, inhibited hASNS with micromolar potency.⁵⁰ Further kinetic studies to investigate their inhibitory effect of compounds **3** and **4** showed that

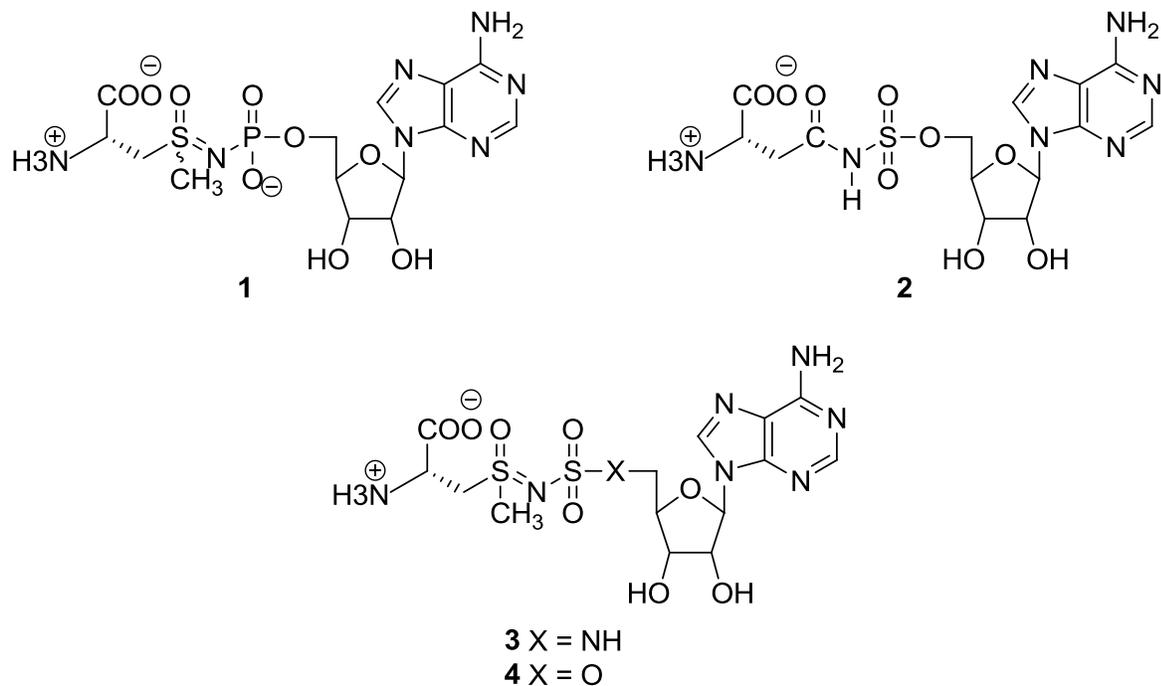


Figure 1-4. Chemical structures of hASNS identified inhibitors.

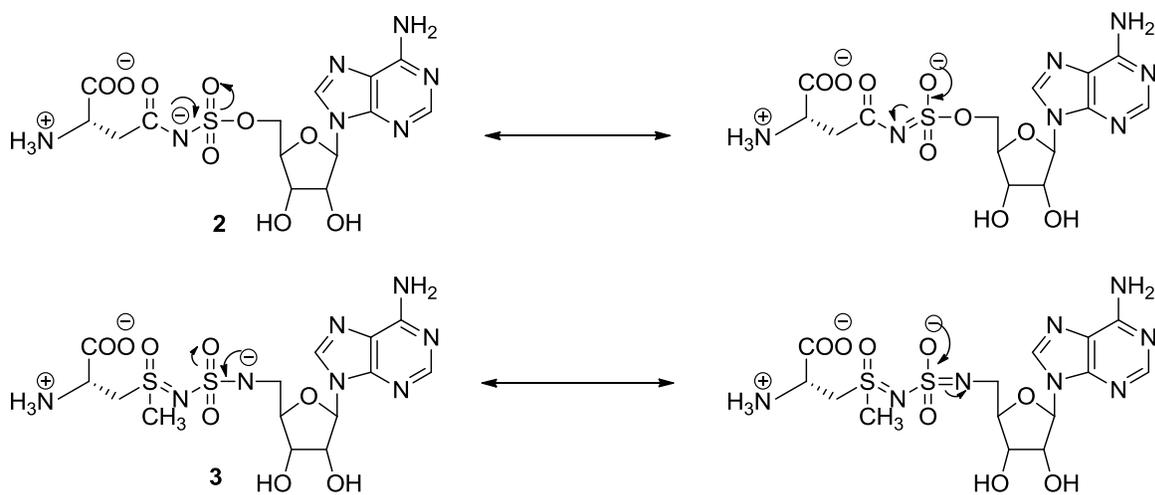


Figure 1-5. Conjugate bases of the functionalized acylsulfamate **2** and sulfamide **3**.

compound **3** has much reduced inhibitor activity compared to compound **1** while compound **4** did not show any inhibitory effect at all.⁵¹ These results suggested that the negatively charged phosphate group is essential to inhibit hASNS and it was hypothesized that the catalytically important lysine (Lys 449 in AS-B and Lys 466 in hASNS) makes a critical electrostatic interaction with the negatively charged oxygen of the phosphate group.⁵¹ This suggestion is supported by the presence of resonance forms of compounds **2** and **3** (Figure 1-5), which stabilize the negative charge on the sulfamate oxygen when the acidic hydrogen of NH dissociates, thereby allowing compound **2** and **3** to inhibit hASNS.⁵¹ Therefore, a negatively charged group that mimic the phosphate group should be considered in the future design of new asparagine synthetase inhibitors.

Research Objectives

The main goal of this research is identification of new lead inhibitors that can be further optimized as drugs to treat acute lymphoblastic leukemia (ALL). Prior identification of the nanomolar *N*-phosphonosulfonimidoyl based inhibitor (compound **1**) is an important step in this process. However, *N*-phosphonosulfonimidoyl **1** is limited in its clinical utility, due to the high concentration necessary to exert its biological activity and the difficulty of its synthesis. Therefore, it is necessary to discover new leads that are specific and easier to synthesize and which have higher potency. Availability of the structure of ASNS bound with the intermediate or transition state will aid in understanding structure-activity relationships between the enzyme and bound intermediate and will facilitate efforts towards the design of new leads. Currently, there is no available x-ray crystal structure of human asparagine synthetase, and the bacterial X-ray crystal structure of AS-B with AMP bound in the active site is the best available

structure. The active site residues involved in binding the substrate and intermediate are conserved in both enzymes (Figure 1-6).

The goal of this work is to refine a model of the *E. Coli* ASNS- β AspAMP complex using CHARMM software. However, there are no accurate CHARMM general force field (CGenFF) parameters to represent the acyl phosphate functional group of the β AspAMP intermediate. In order to refine this model, new CGenFF parameters for the acylphosphate functional group of the β AspAMP intermediate must be developed. This model should be further refined using computational methods, such as molecular dynamics simulations. The refined model will then be used as the docking receptor in virtual screening of chemical libraries to identify new lead inhibitors of hASNS.

The specific goals of this research are the following: 1) Development of new CGenFF parameters for the acyl phosphate functional group of the β AspAMP intermediate. 2) Simulations of the ASNS- β AspAMP complex by molecular dynamics simulations to generate a refined model of ASNS with the bound intermediate and to understand the structure-function relationship between ASNS and the bound intermediate. 3) Structure-based virtual screening using the refined model of ASNS- β AspAMP complex as the docking receptor to identify new lead inhibitors of human asparagine synthetase.

ASNB_ECOLI/1-553 1 CSIFGVFDIKTDAVELRKKALELSRLMRHRGPD---WSGI 37
 ASNS_HUMAN/1-560 1 CGIWALFG-SDDCLSV--QCLSAMKI-AHRGPD AFRFENV 36

ASNB_ECOLI/1-553 38 YASDNAILAHERLSIVDVNAGAQP LYNQKQTHV-LAVNGE 76
 ASNS_HUMAN/1-560 37 NGYTNCCFGFHLAVVDPLFGMQRIRVKKYPYLWLCYNGE 76

ASNB_ECOLI/1-553 77 IYNHQALRAEYGDRYQFQTGSDCEVILALYQEKGPE-FLD 115
 ASNS_HUMAN/1-560 77 IYNHKKMQQHF--EFEYQTKVDGEIILHLYDKGGIEQTIC 114

ASNB_ECOLI/1-553 116 DLQGMFAFALYDSEKDAYLIGROHLGIIPLYMGYDEHGQL 155
 ASNS_HUMAN/1-560 115 MLDGVFAFVLLDTANKKVFLGRDTYGVRPLFKAMTEDGFL 154

ASNB_ECOLI/1-553 156 YVASEMKALVPV--CRIT----IKEFPAGSY--L-WSQDGE 186
 ASNS_HUMAN/1-560 155 AVCSEAKGLVTLKHSATPFLKVEPFLPGHYEVLDLKPNGK 194

ASNB_ECOLI/1-553 187 IRS-----YYHRDWFYDAVKDNV-----TDKNEEL 211
 ASNS_HUMAN/1-560 195 VASVEMVKYHHCRDVPLHALYDNVEKLPFGFEIETVKNNL 234

ASNB_ECOLI/1-553 212 RQALEDSVRSKSHLMSDVPYGVLLSGGLDSSII SAITKKYAA 251
 ASNS_HUMAN/1-560 235 RILFNNAVKKRLMTDRRIGCLL SGGLDSSLV-----AA 267
 ** *

ASNB_ECOLI/1-553 252 RRVEDQERS EAWWPQLHSFAVGLPGSPDLKAAQEVANHLG 291
 ASNS_HUMAN/1-560 268 TLLKQLKEAQVQYP-LQTF AIGMEDSPDLLAARKVADHIG 306
 *

ASNB_ECOLI/1-553 292 TVHHEIHFTVQEGLD AIRDVIYHIETYDVTTIRASTPMYL 331
 ASNS_HUMAN/1-560 307 SEHYEVLFNSEEGIQALDEVIFSL ETYDITTVRASVGMYL 346

ASNB_ECOLI/1-553 332 MSRKIKAMG IKMVL-SGEGSDEVFGGYLYFHKAPNAKELH 370
 ASNS_HUMAN/1-560 347 ISKYIRKNTDSVVI FSGEGSDEL TQGYIYFHKAPSPEKAE 386
 *** *

ASNB_ECOLI/1-553 371 EETVRKLLALHMYDCARANKAMSAWGVEARVPFLDKKFLD 410
 ASNS_HUMAN/1-560 387 EESERLLRELYLFDVLRADRTTAAHGLELRVPFLDHRFSS 426
 *

ASNB_ECOLI/1-553 411 VAMRINPQDKMCGNGKMEKHILRECFE-A-YLPASVAWRQ 448
 ASNS_HUMAN/1-560 427 YYLSLPPPEMRIPKNG-IEKHLRET FEDSNLIPKEILWRP 465
 *

ASNB_ECOLI/1-553 449 KEQFSDG---VGYSWIDTLKEVAAQQVSDQQLETARFRFP 485
 ASNS_HUMAN/1-560 466 KEAFSDGITSVKNSWFKILQEYVEHQVDDAMMANAAQKFP 505
 *

ASNB_ECOLI/1-553 486 YNTPTSKEAYLYREIFEELFPLPSAAECVPGGPSVACSSA 525
 ASNS_HUMAN/1-560 506 FNTPKTKEGYYYRQVFERHYP-----GRADW 531

ASNB_ECOLI/1-553 526 KAIEWDEAFKKMDDPSGRAVG VHQSAVK 553
 ASNS_HUMAN/1-560 532 LSHYWMPKWINATDPSARTLTHYKSAVK 559

Figure 1-6. Pairwise sequence alignment between the *E. coli* and human asparagine synthetases. Residues involved in binding the substrate or the intermediate are marked by a star.

CHAPTER 2

OPTIMIZED CHARMM GENERAL FORCE FIELD PARAMETERS FOR ACYLPHOSPHATE AND *N*-PHOSPHONOSULFONIMIDOYL FUNCTIONAL GROUPS*

Background

Computational chemistry uses mathematical models implemented via computer programs to investigate physical and chemical properties such as energy, dipole moment, charge distribution, vibrational frequencies and reactivity. The accuracy of the calculated properties depends on the quality of the computational methods used. Accurate computational methods include quantum mechanical (QM) methods, where applying theory directly to compute chemical properties of interest. However, QM methods require extensive computer time; therefore their use is limited to chemical systems with no more than 100 atoms. On the other hand molecular mechanical (MM) methods use empirical equations, which implement experimental parameters with simple mathematical equations that can be applied on large systems such as proteins.

Empirical energy functions, unlike quantum mechanics, do not include the nuclei and electrons in the energy function used to describe the system. Rather the atoms are the smallest particles in the empirical energy equation. In addition, compared to QM, the empirical energy equations use simpler intra- and inter-molecular terms with highly optimized parameters to describe the molecular interactions. Therefore, empirical energy calculations can be performed on very large biological systems in reasonable computer time and with high accuracy. Examples of widely used empirical energy functions are CHARMM,⁵³ AMBER,^{54,55} and OPLS-AA.⁵⁶⁻⁵⁸

* Reprinted in part with permission from Hegazy, L.; Richards, N. G. J. *J. Mol. Model.* **2013**, doi: 10.1007/s00894-013-1990-x

Glutamine-dependent asparagine synthetase (ASNS) was identified recently as a biomarker for ovarian cancer, and it is also involved in the development of prostate cancer.^{24,25} In addition, elevated levels of ASNS have been observed in ASNase resistant acute lymphoblastic leukemia.¹⁸⁻²¹ Glutamine-dependent asparagine synthetase catalyzes the synthesis of asparagine from aspartate in the presence of glutamine as the predominant nitrogen source. An *N*-phosphonosulfonimidoyl based inhibitor, which mimics the ASNS transition state, inhibited hASNS with nanomolar potency (Figure 1-4, Chapter 1),⁵⁹ and was able to suppress the proliferation of asparaginase-resistant MOLT-4 cells.

Computational studies of ASNS in complex with the intermediate β -aspartyl-AMP (β AspAMP) and the *N*-phosphonosulfonimidoyl inhibitor have been hindered by the lack of parameters that accurately describe both functional groups. In this research, parameters for β AspAMP and the *N*-phosphonosulfonimidoyl moieties were developed following the CHARMM general force field protocol. The newly developed parameters will facilitate the detailed computational studies of inhibitors with *N*-phosphonosulfonimidoyl derivatives and acyladenylation-catalyzing enzymes such as tRNA aminoacyl synthetases,⁶⁰ glutamine-dependent NAD⁺ synthetase,^{61,62} and aminoacyl-tRNA transamidating enzymes.⁶³⁻⁶⁵

Molecular Mechanics Force Fields

Molecular mechanics uses potential energy functions to describe molecular interactions of large molecules. Parameters are included in the mathematical equations and the accuracy of the molecular interactions derived from these equations depends mainly on the quality of the derived parameters. Together the mathematical equation and the parameters comprise a force field. There are many force fields available for

different types of chemical systems, depending on the way the parameters were developed and optimized. For example, there are extended atom force fields and additive all-atom force fields.⁶⁶ In extended atom force fields, the hydrogen atoms are not represented explicitly but are included with the atoms to which they are connected, and the van der Waals parameters and charges are modified to account for the deleted hydrogen atoms. The omitted hydrogen atoms are mainly bonded to elements with low electron negativities, while polar hydrogen atoms involved in hydrogen bonding are treated explicitly. This approach reduces the number of atoms and hence computer time significantly. However because of improvements of computer resources, the additive all-atom force fields, which represent all hydrogen atoms explicitly, are more commonly used to treat chemical and biological systems. CHARMM,⁵² AMBER^{54,55} and OPLS-AA^{56,57} are widely used additive all-atom force fields.

CGenFF Potential Energy Function

In this work, parameterization process followed the CHARMM general force field (CGenFF) protocol.⁶⁷ CGenFF is an organic force field developed mainly to simulate drug-like molecules in a biological environment, represented by the CHARMM additive force fields. The potential energy function of CGenFF is the same as that of the CHARMM force field for biological molecules, thereby providing properly balanced electrostatic and van der Waals (vdW) interactions between the ligand and the biomolecule. The potential energy function is a mathematical equation that describes the various intramolecular and intermolecular physical interactions of a particular system. The potential energy function, $V(r)$ used in CGenFF is⁶⁷

$$\begin{aligned}
V(r) = & \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{UB} K_{UB} (x - x_0)^2 + \sum_{dihedrals} K_\phi (1 + \cos(n\phi - \delta)) \\
& + \sum_{impropers} K_\varphi (\varphi - \varphi_0)^2 + \sum_{nonbond} \frac{q_i q_j}{4\pi D r_{ij}} + \epsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right]
\end{aligned}$$

The CGenFF potential energy function contains intramolecular or internal terms and intermolecular or external terms. The internal terms describe bonds, valence angles and dihedral energies, while external terms describe nonbonded interactions, such as electrostatic and van der Waals (vdW) interactions. The potential energy function contains terms that are obtained from the known structure (X-ray crystal structure, NMR structure or computationally modeled structures) such as bond lengths, b , valence angle, θ , dihedral angle, ϕ , improper dihedral angle, φ , Urey-Bradley or (1,3) distance, x , and the distance between atoms, r_{ij} . The remaining terms are called parameters and, they are specific for each atom based on its type and the types of the other atoms covalently bonded to it. Therefore, it is the quality of those parameters that affect the accuracy of the results calculated by the potential energy function. The parameters are accurate if they are able to reproduce target data obtained experimentally or calculated by quantum mechanical methods. These parameters are b_0 , θ_0 , φ_0 and x_0 , which are the equilibrium values for bond lengths, angles, improper dihedral angles and Urey-Bradley distances, respectively, K 's are their respective force constants; K_ϕ , n , δ are the force constant, multiplicity and phase of the dihedral angle; q_i and q_j are the partial atomic charges of atoms i and j , respectively; D is the dielectric constant; ϵ_{ij} is the well depth (The geometric mean between ϵ_i and ϵ_j). $R_{min,i,j}$ is the radius in the Lennard-Jones (LJ)

6–12 term (The arithmetic mean between $R_{\min,i}$ and $R_{\min,j}$); and r_{ij} is the distance between i and j .

Bond stretching and angle bending are represented by harmonic potentials, since the biological simulations were performed at room temperature, where the bonds and angles are close to equilibrium. The Urey-Bradley term (1-3 distance) is a quadratic function of the distance, x , between the first and third atoms connected by two bonds. The improper dihedral angle term is also a quadratic function, which is used at “branch points”, for example, in the case where three connected atoms l, i, k and fourth atom j connected to the central atom i , the improper dihedral angle is $iljk$. Both Urey-Bradley and improper dihedral terms are used in limited cases when better vibrational spectra are needed such as out of plane motion.⁶⁷

The dihedral or torsion function is a cosine function, which models the barriers and minima encountered for a rotation around the central bond of four atoms connected by three covalent bonds (1, 4 interactions). The force constant, K_ϕ , indicates the height of the barrier; the periodicity, n , is the number of barriers or minima per rotation; and the phase, δ , is the location of the minima. The dihedral angle can be represented by a sum of dihedral terms with different force constants, multiplicities and phases (Fourier series), thereby allowing accurate reproduction of the dihedral profile of the target data.

The nonbonded interactions are represented by Lennard Jones (LJ) and electrostatic potentials. The LJ potential describes the interatomic potential between two atoms; where ϵ_{ij} is the depth of the potential well and is the geometric mean of ϵ_i and ϵ_j . $R_{\min,i,j}$ is the distance at which the LJ potential is minimum, evaluated as the arithmetic mean between $R_{\min,i}$ and $R_{\min,j}$; and r_{ij} is the distance between i and j .

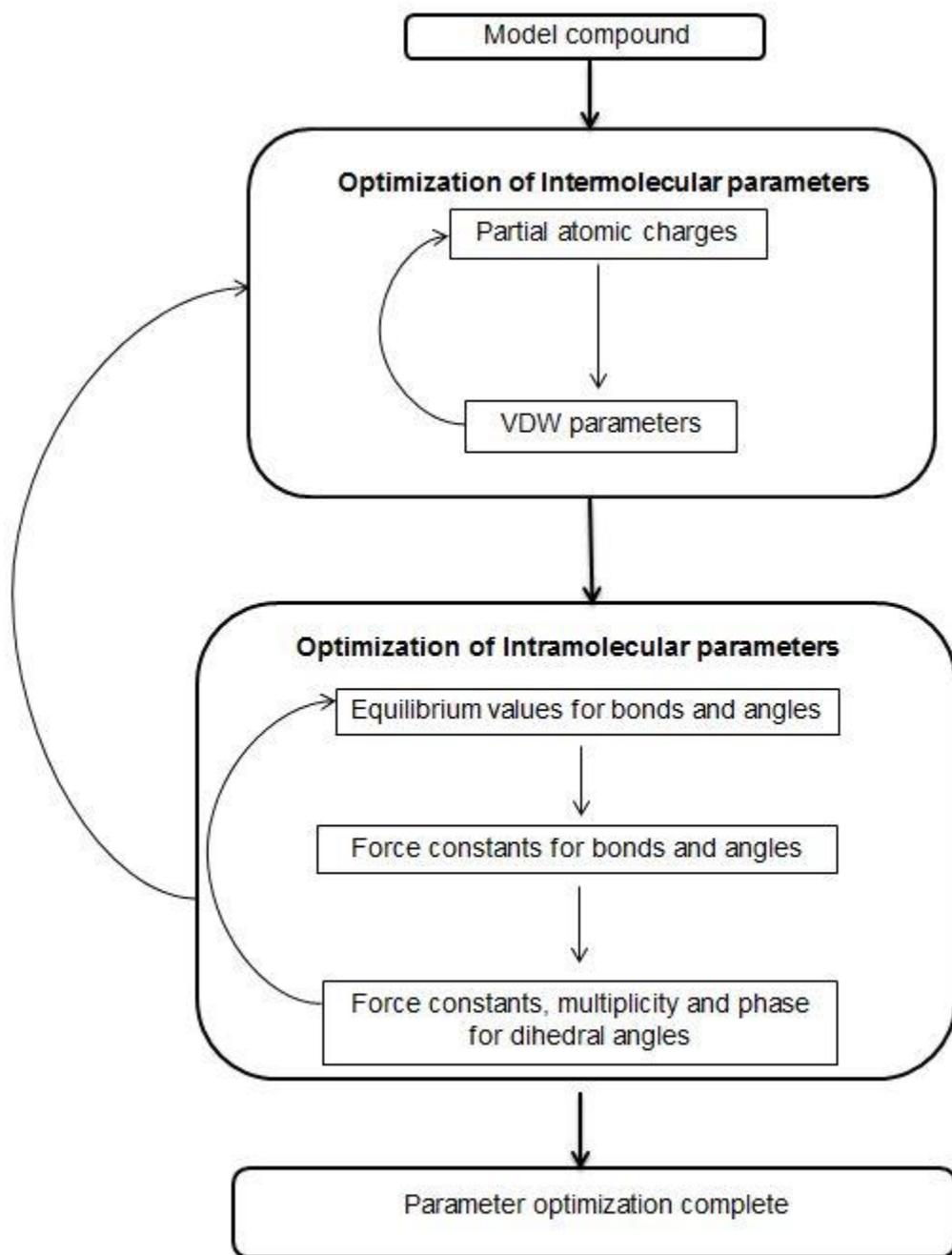


Figure 2-1. CGenFF parameters optimization process, adapted from Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; E. Darian, O. Lopes, P.; Vorobyov, I; Mackerell, A. D., Jr. *J. Comput. Chem.* **2010**, 31, 671.

The $(R_{\text{minij}}/r_{ij})^{12}$ term describes the interatomic Pauli short range repulsion, while $(R_{\text{minij}}/r_{ij})^6$ describes the long range attraction. The electrostatic or Coulombic interactions is the other term of the nonbonding interactions, where D is the dielectric constant, q_i and q_j are the partial atomic charges of atoms i and j , respectively.

CGenFF Parametr Optimization Protocol

The aim of the CGenFF parameter optimization methodology is accuracy and simplicity so that new drug-like molecules can be optimized in reasonable computational time.⁶⁷ Therefore, although the CHARMM bimolecular force fields use experimental target data supplemented with QM data, CGenFF uses mainly QM results as target data with empirical scaling factors to make the QM data appropriate for bulk phases (Figure 2-1).

Identification of missing parameters in the molecule of interest and selection of a model compound is the first step in the parameter optimization process. The model compound must be in its equilibrium geometry. Next, intermolecular parameters are optimized. This involves optimization of partial atomic charges and Lennard-Jones, LJ (vdW) parameters. CGenff encompasses a wide range of diverse atom types with experimentally driven LJ parameters. Therefore, straight assignment of LJ parameters from already existing parameters is suitable. LJ parameters optimization is limited in cases where the partial atomic charge optimization could not be achieved within convergence criteria. Partial atomic charges are optimized to reproduce the interaction energies and distances of the model-water complexes determined from QM calculations at the HF/6-31G(d) level. This theory level is used by the CHARMM bimolecular force field when optimizing the charges of biological molecules like proteins and nucleic acids. Therefore, use of HF/6-31G(d) rather than a higher theory level in the parameter

optimization of drug-like molecules provides balanced nonbonding interactions between the ligand and its biological environment.⁶⁷ After optimization of the intermolecular parameters, the intramolecular parameters for bonds, angles and dihedrals are considered.

Bond and valence angle equilibrium parameters are optimized to reproduce the QM equilibrium geometry. The force constants for bonds, angles, UB distances and improper dihedral angles are optimized to reproduce the QM calculated vibrational spectra. Finally force constants, multiplicities and phases of the dihedral angles are optimized to reproduce the minima and barrier heights of the QM potential energy surface (PES). Because there is a strong correlation between the parameters, following optimization of the dihedrals, validation of the new parameters is performed by repeating the charge optimization calculations and re-optimizing the charges if necessary. If charges are re-optimized, then the optimization of the intramolecular parameters is repeated until the results converge.⁶⁷

Computational Details

The ParamChem web interface was used to assign atom types and to make initial guesses for charges and parameters.⁶⁷⁻⁷⁰ Molecular mechanics energy minimizations were performed using CHARMM version 35.⁵³ Energy minimization employed 200 steps of the conjugate gradient and Newton-Raphson algorithms to an RMS gradient of 10^{-5} kcal/(mol Å) and with infinite cut off distance.⁵³ Vibrational analyses were performed with the VIBRAN and MOLVIB modules in CHARMM.⁷¹ All quantum mechanics calculations were performed with Gaussian 09.⁷² The QM geometry optimization for the acylphosphate and *N*-phosphonosulfonimidoyl model compounds were performed using MP2/6-31+G(d) level of theory.^{73,74} Interactions with

water molecules were modeled using a series of monohydrate complexes with a water molecule (TIP3P geometry)⁷⁵ at each H-bond donor and acceptor group with the model compounds in their MP2/6-31G(d) geometries. The QM calculations were then performed on the monohydrate complexes using HF/6-31G(d) level of theory.^{76,77} The interaction energy was calculated as the difference between the total energy of the complex and the sum of the monomer energies. The QM vibrational frequencies were calculated using MP2/6-31G(d) and scaled by 0.943.⁷⁸ A potential energy scan (PES) on each selected dihedral was calculated using the scan keyword (Opt = ModRedundant) with 15° steps and optimization at each increment at the MP2/6-31+G(d) theory level. Single-point energy calculations using MP2/cc-pVTZ level was then employed for each optimized structure.

As a further validation of the new CGenFF parameters obtained for the acylphosphate and *N*-phosphonosulfonimidoyl model compounds, molecular dynamics (MD) simulations of these molecules in aqueous solution were performed. Each model compound (Figure 2-3) was solvated in an octahedral box (28 Å x 28 Å x 28 Å) of TIP3P water molecules.⁷⁵ The solvated models were energy minimized by steepest descent (SD) and adopted basis Newton-Raphson (ABNR) algorithms. Periodic boundaries were used in all MD simulations, with the particle mesh Ewald method used to obtain electrostatic energies.⁷⁹ The Leapfrog algorithm was used to integrate equations of motion with 1fs time step. The SHAKE algorithm⁸⁰ was used to constrain the length of covalent bonds to hydrogen atoms to their equilibrium values. After heating to 300 K (30 ps), each system was equilibrated for a further 40 ps in the canonical ensemble (NVT), before the “production” MD simulation CHARMM version 36b1 was performed in the

isobaric-isothermal ensemble(NPT) ensemble (16 ns for the acylphosphate model and \approx 17 ns for the *N*-phosphonosulfonimidoyl compound). The Hoover thermostat was used to maintain the temperature around the average.^{57,81}

Results and Discussion

Identification of Missing Parameters

The first step in parameterization process is the identification of the missing parameters and selection of a model compound to be parameterized. The ParamChem web site, based on CGenFF, was used to add atom types, assign charges and parameters to the missing parameters in the molecule by analogy to existing parameters in CGenFF.⁶⁷⁻⁶⁹ The ParamChem program produces high penalty scores to missing parameters, so that parts of the molecule to be parameterized can be identified. The structure of the selected compound subjected for parameter optimization should be in its equilibrium geometry. Both molecules; β -aspartyl-AMP and *N*-phosphonosulfonimidoyl inhibitor were uploaded as mol2 format to the ParamChem utility and missing parameters were identified and represented by model compounds (Figure 2-2) with each model compound at a global minimum on the potential energy surface. Different initial conformations of each molecule were optimized using MP2/6-31+G(d). The lowest energy structure from this conformational search was then subjected to frequency calculation to confirm that there were no imaginary frequencies, and thereby confirming that the lowest energy structure corresponds to the equilibrium geometry (Figure 2-3).

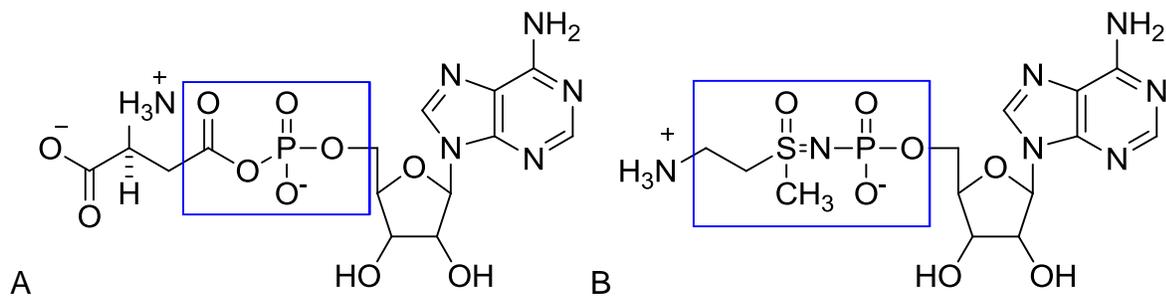


Figure 2-2. Parts of the molecules that require extensive optimization are enclosed in boxes. A) β -AspAMP. B) *N*-phosphonosulfonimidoyl inhibitor.

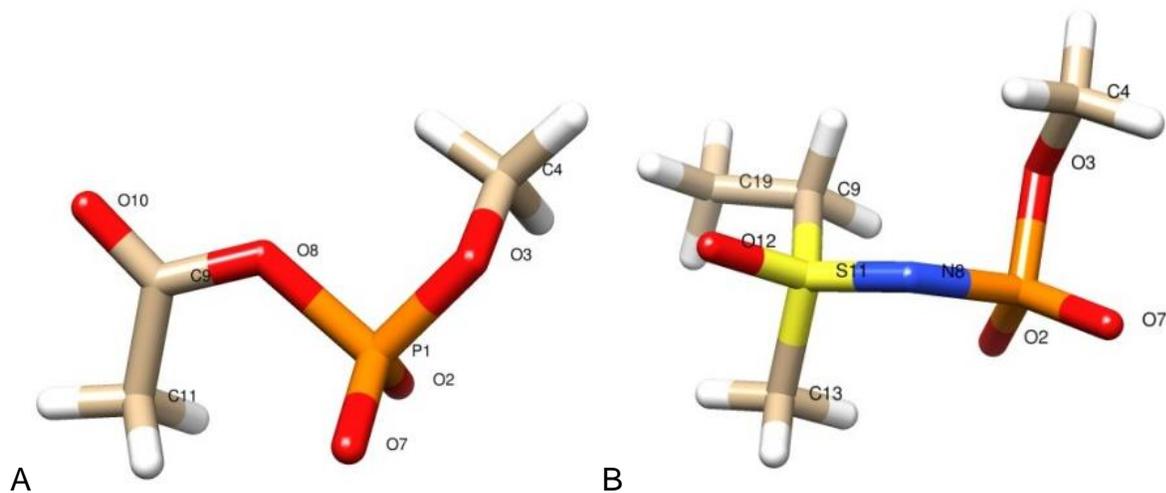
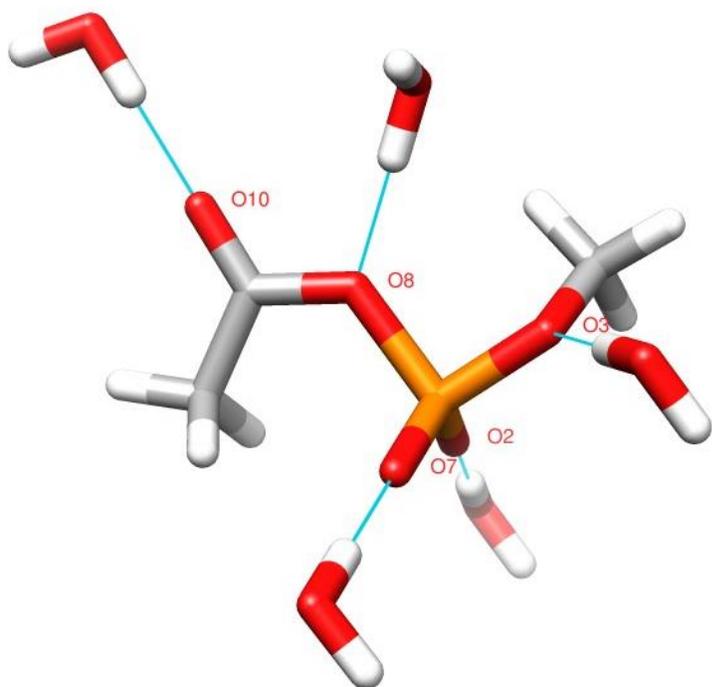


Figure 2-3. The equilibrium geometry and atomic labels used in the text and tables. A) Acylphosphate molecule. (B) *N*-phosphonosulfonimidoyl molecule.

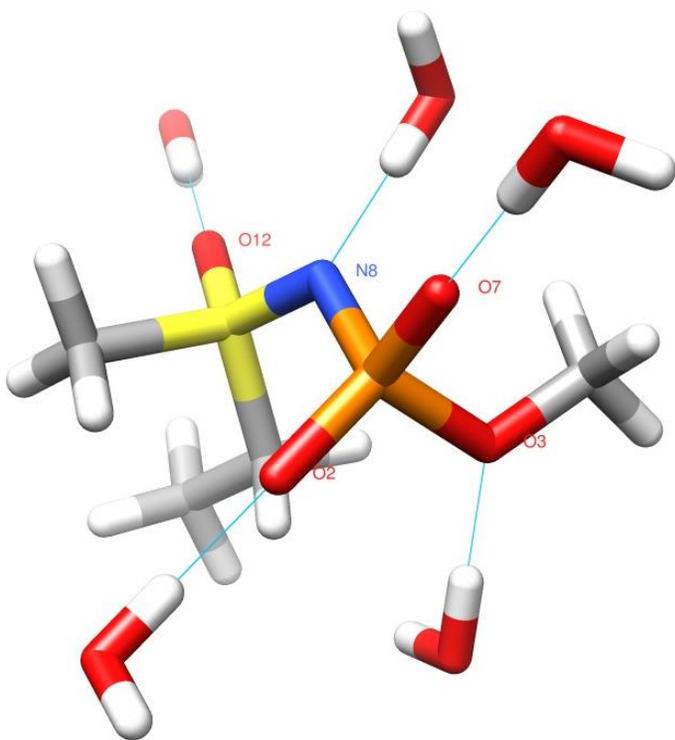
Charge Optimization

Partial atomic charges were then optimized to reproduce the QM interaction energies and distances of the model-single water complexes. Water molecules (TIP3P geometry)⁷⁵ are placed at each H-bond donor and acceptor group using the model compounds in their MP2/6-31G(d) geometries. Five monohydrates were generated for each model compound and each H-bond was optimized while fixing other degrees of freedom using Hf/6-31G(d) (Figure 2-4). The interaction energy was calculated as the difference between the total energy of the complex and the sum of the monomer energies. Although higher theory level will provide more accurate results, using the Hartree-Fock level of theory will be consistent with the CHARMM bimolecular force field and thereby providing well balanced non-bonded interactions.⁶⁷

The aliphatic hydrogen atoms were assigned a standard value of +0.09 with the methyl and methylene carbons given the same values as assigned by the ParamChem tool to allow transferability.⁶⁷⁻⁷⁰ In addition, overall charge of the molecule has to add -1, and O2 and O7 were assigned the same charge. Then MM charges were adjusted manually until the QM and MM interaction energies were within satisfactory agreement. MM interaction energies should be within 0.2 kcal/mol of the corresponding QM interaction energy and MM interaction distances should be 0.2 Å shorter than the respective QM interaction distances.



A



B

Figure 2-4. Water–model compound complexes. A) Acylphosphate molecule. B) *N*-phosphonosulfonimidoyl molecule with O3 in the trigonal position. A monohydrate complex was used in each calculation.

Table 2-1. Comparison of HF/6-31G(d) and CHARMM interaction energies and interaction distances before and after optimization for the acylphosphate molecule.

	Interaction Energies (Kcal/mole)			Interaction Distances(A°)		
	QM	MM	Optimized	QM	MM	Optimized
1) O2...HOH	-11.36	-12.17	-11.5	1.89	1.70	1.70
2) O7...HOH	-11.82	-11.42	-11.49	1.89	1.70	1.70
3) O3...HOH	-7.82	-11.65	-7.81	2.03	1.70	1.77
4) O8...HOH	-7.74	-5.46	-7.95	2.18	2.57	1.91
5) O10..HOH	-8.33	-8.83	-8.54	1.96	1.74	1.76
Average Difference				-0.05		
RMS Difference				0.21		
Average Absolute Deviation				0.18		

Table 2-2. Comparison of HF/6-31G(d) and CHARMM interaction energies and interaction distances before and after optimization for the *N*-phosphonosulfonimidoyl molecule.

	Interaction Energies (Kcal/mole)			Interaction Distances(A°)		
	QM	MM	Optimized	QM	MM	Optimized
1) O2...HOH	-13.0	-15.5	-12.8	1.83	1.64	1.71
2) O7...HOH	-10.0	-14.1	-11.3	1.87	1.70	1.73
3) O3...HOH(Lone pair)	-10.0	-12.4	-10.8	1.95	1.7	1.76
3) O3...HOH(Trigonal)	-9.95	-8.3	-8.15	2.02	1.74	1.81
4) N8...HOH	-10.0	-15.2	-10.1	2.07	1.84	1.99
5) O12..HOH	-7.8	-6.6	-8.0	1.98	1.82	1.77
Average Difference				-0.44		
RMS Difference				0.53		
Average Absolute Deviation				0.52		

Table 2-3. Optimized atomic partial charges for atoms in the acylphosphate and *N*-phosphonosulfonimidoyl molecules.

Acylphosphate model			<i>N</i> -phosphonosulfonimidoyl model		
Atom	Type	Charge	Atom	Type	Charge
P1	PG1	1.30	P1	PG1	0.20
O2	OG2P1	-0.71	O2	OG2P1	-0.46
O3	OG303	-0.46	O3	OG303	-0.28
C4	CG331	-0.17	C4	CG331	-0.17
H5	HGA3	0.09	H5	HGA3	0.09
H6	HGA3	0.09	H6	HGA3	0.09
O7	OG2P1	-0.71	O7	OG2P1	-0.46
O8	OG305	-0.38	N8	NG2D1	-0.38
C9	CG2O2	0.34	C9	CG321	0.02
O10	OG2D1	-0.48	H10	HGA3	0.09
C11	CG331	-0.27	S11	SG3O2	0.12
H12	HGA3	0.09	O12	OG2P1	-0.42
H13	HGA3	0.09	C13	CG331	0.11
H14	HGA3	0.09	H14	HGA3	0.09
H15	HGA3	0.09	H15	HGA3	0.09
			H16	HGA3	0.09
			H17	HGA2	0.09
			H18	HGA2	0.09
			C19	CG331	-0.27
			H20	HGA3	0.09
			H21	HGA3	0.09
			H22	HGA3	0.09

As shown in Table 2-1, the interaction energies and distances of the model-water complexes for the acylphosphate molecule are in excellent agreement with scaled QM data and satisfy the convergence criteria. The interaction distances are almost 0.2 Å smaller than QM interaction distances, which is excellent. For the *N*-phosphonosulfonimidoyl molecule, Table 2-2 shows that the interaction energies and distances for the optimized parameters compare well with the QM data, except the hydrogen bonds of O7 and O3 with water. The interaction energy of O7 with water is more favorable than the QM value by -1.3 kcal/mol. The QM interaction energy value of the hydrogen bond for the O7 water interaction is smaller than the interaction energy value of O2 water hydrogen bond. It was not possible to optimize to convergence criteria while satisfying the constraint of keeping the charges equal on O2 and O7. The *N*-phosphonosulfonimidoyl molecule is crowded with different functional groups which make charge optimization more difficult than in the acylphosphate molecule. However, the MM optimized charges are significantly improved compared to the ones initially assigned by analogy. The charge optimization for O3 was carried out using two conformations (the trigonal geometry and lone pair position) and a compromise between the two positions was considered during charge optimization. In general, the overall charge optimization for both the acylphosphate and *N*-phosphonosulfonimidoyl molecules yielded MM interaction energies and distances for the individual model-water complexes that were in good agreement with the QM data (Table 2-3).

Optimization of Equilibrium Values and Force Constants for Bond Lengths and Angles

Following optimization of the partial atomic charges, intramolecular or bonded parameters are optimized. Bonded parameters include equilibrium values and force

constants of bond lengths, valence angles and torsion angles. The equilibrium terms of bond lengths and angles for missing parameters were optimized to reproduce the ab initio equilibrium geometry. After optimization of the equilibrium terms, the MM optimized bond lengths and angles were in very close agreement with the QM data (Tables 2-4 and 2-5).

Force constants for bonds, angles, Urey-Bradley distances, improper dihedrals and dihedrals containing terminal hydrogen atoms were optimized based on the QM vibrational spectra. QM frequencies were calculated using MP2/6-31G(d) and were scaled by 0.943 to be appropriate for the bulk phase.⁷⁸ The vibrational spectra are composed of the frequencies and their assignments as well as their numerical values. MOLVIB module in CHARMM was used to generate the potential energy distribution (PED) for the QM and MM spectra and to assign the frequencies to the different contributions of the normal modes.^{71,82} Manual adjustment of the force constants was performed until agreement between the QM and MM spectra was obtained. In general, it is very difficult to achieve quantitative convergence between the QM and MM vibrational spectra because the normal modes contribute differently to a given frequency in the two methods. Generally, the MM vibrational frequencies should be within 5% of their corresponding QM frequencies, especially the lowest MM vibrational modes because these are the normal modes that involve conformational changes in the molecule. In addition, it is better for the MM vibrational frequencies to be lower than the QM frequencies rather than vice-versa to generate a more flexible molecule and to make conformational changes during MD simulations more flexible.⁶⁷

Table 2-4. Comparison of the CHARMM optimized geometry for missing parameters with the QM geometry for the acylphosphate molecule.

	MP2/6-31+G*	CHARMM	Difference
Bonds(Å)			
p1-o8	1.75	1.75	0.00
O8-C9	1.34	1.34	0.00
Angles(°)			
O3-P1-O8	94.12	94.3	-0.18
O2-P1-O8	105.99	107.6	-1.61
O7-P1-O8	108.07	107.6	0.47
C9-O8-P1	124.67	124.95	-0.28
O10-C9-O8	119.82	119.55	0.27
Dihedrals (°)			
P1-O8-C9-C11	2	2	0
P1-O8-C9-O10	-178	-178	0
O2-P1-O8-C9	68	66	-2
O7-P1-O8-C9	-68	-68	0
O3-P1-O8-C9	179	177	-2
O8-P1-O3-C4	-69	-71	-2
Improper Torsions (°)			
C9-C11-O10-O8	-0.1	0.0	0.1

Table 2-5. Comparison of the CHARMM optimized geometry for missing parameters with the QM geometry for the *N*-phosphonosulfonimidoyl molecule.

	MP2/6-31+G*	CHARMM	Difference
Bonds(Å)			
p1-n8	1.75	1.75	0.00
n8-s11	1.53	1.53	0.00
Angles(°)			
O3-P1-N8	99.79	99.79	0.00
O2-P1-N8	109.0	110.62	-1.62
O7-P1-N8	108.3	106.8	1.5
C9-S11-N8	112.17	111.8	0.37
P1-N8-S11	120	120	0
N8-S11-C13	111	111	0
N8-S11-O12	116	116	0
Dihedrals (°)			
N8-S11-C9-C19	-176	-169	7
O2-P1-N8-S11	-24	-23	1
O7-P1-N8-S11	-161	-155	6
O3-P1-N8-S11	85	90	5
P1-N8-S11-O12	-171	-171	0
P1-N8-S11-C9	-49	-48	1
P1-N8-S11-C13	67	67	0
N8-P1-O3-C4	71	74	3

Table 2-6. Vibrational spectra computed for the acylphosphate molecule at the scaled MP2 level and with optimized CGenFF^a parameters

MP2/6-31+G(d) scaled by a factor 0.943				CGenFF			
Freq ^b	Assign (%)	Assign (%)	Assign (%)	Freq ^b	Assign (%)	Assign (%)	Assign (%)
67.0	tdOPOC (98)			68.2	tdCCOP (65)	tdOPOC(29)	
81.8	tdCCOP (73)	tdCOPO (23)		79.8	tdCOPO (47)	tdOPOC (36)	scPO4' (7)
100.5	tdCOPO (55)	tdCCOP (33)	scPO4' (9)	106.5	tdOPOC (33)	tdCOPO (32)	tdCCOP (23)
140.2	tdPOCH (69)	dPOC9 (14)	scPO4' (11)	141.2	tdPOCH (57)	scPO4' (31)	dPOC9 (8)
180.2	dPOC9 (53)	tdPOCH (24)	tdCOPO (12)	207.9	tdPOCH (29)	dPOC9 (18)	rPO4 (16)
194.3	tdPOCH (89)			239.4	dPOC4 (33)	dPOC9 (19)	rPO4 (14)
219.0	dPOC4 (41)	twPO4 (18)	scPO4' (8)	268.0	tdPOCH (90)		
295.4	scPO4' (38)	twPO4 (20)	dCCO (12)	289.4	dPOC4 (31)	scPO4' (19)	twPO4 (9)
329.7	twPO4 (36)	dPOC4 (19)	saOP (12)	323.4	dCCO (28)	twPO4 (24)	dPOC9 (21)
343.8	dCCO (19)	ssOP (17)	saOP (16)	334.8	twPO4 (38)	wPO4 (28)	dCCO (9)
414.2	scPO4 (42)	wPO4 (20)	rPO4 (11)	417.2	ssOP (30)	scPO4 (24)	rPO4 (20)
460.0	wPO4 (32)	dCCO (24)	rC=O (15)	471.9	wPO4 (36)	dPOC4 (19)	twPO4 (14)
505.7	scPO4 (21)	wPO4 (16)	dPOC4 (15)	499.9	scPO4 (44)	dCCO (27)	dPOC9 (6)
515.2	rPO4 (54)	scPO4 (13)	ssOP (10)	537.0	rC=O (39)	saOP (12)	scPO4 (11)
564.4	tiOCOC (79)	rCH3-11 (14)		564.3	tiOCOC (90)		
685.5	sCC (34)	rC=O (25)	saOP (16)	671.9	saOP (47)	ssOP (11)	rPO4 (9)
723.7	saOP (38)	ssOP (32)	sO3C (9)	679.6	ssCC (34)	ssOP (18)	rC=O (11)
906.9	sCC (33)	sO8C (26)	rC=O (11)	952.4	sO8C (34)	ssPO (16)	r'CH3C11 (11)
993.8	r'CH3C11 (38)	ssPO (18)	rCH3C11 (13)	996.4	r'CH3C11 (35)	ssPO (34)	rCH3C11 (10)
1035.2	sO3C (59)	ssPO (25)	r'CH3C11 (10)	1023.3	sO3C (23)	ssPO (18)	r'CH3C11 (13)
1043.3	rCH3C11 (53)	r'CH3C11(20)	tiOCOC (20)	1040.1	sO3C (54)	ssPO (16)	ssOP (8)
1048.6	ssPO (46)	sO3C (27)	ssOP (9)	1046.7	rCH3C11 (64)	r'CH3C11 (19)	ad'CH3C11(10)
1135.9	rCH3C4 (72)	r'CH3C4 (25)		1139.0	rCH3C4 (40)	r'CH3C4 (37)	ad'CH3C4 (19)
1157.5	r'CH3C4 (66)	rCH3C4 (24)		1144.8	r'CH3C4 (39)	rCH3C4 (34)	adCH3C4 (24)
1243.9	saPO (93)			1187.6	saPO (92)	wPO4 (6)	

Table 2-6. Continued.

MP2/6-31+G(d) scaled by a factor 0.943				CGenFF			
Freq ^b	Assign (%)	Assign (%)	Assign (%)	Freq ^b	Assign (%)	Assign (%)	Assign (%)
1258.8	sO8C (45)	sCC (13)	rC=O (12)	1232.3	sO8C (33)	sCC (30)	rC=O (23)
1373.0	sdCH3C11 (86)	sCC (8)	sO8C (5)	1387.2	sdCH3C11 (98)		
1421.8	sdCH3C4 (99)			1429.7	ad'CH3C4 (58)	adCH3C4 (23)	r'CH3C4 (15)
1447.1	ad'CH3C4 (92)	rCH3C11 (6)		1434.3	adCH3C11 (91)		
1460.0	adCH3C11 (91)			1451.9	ad'CH3C11(88)	rCH3C11 (8)	
1465.0	ad'CH3C4 (92)	ad'CH3C4 (34)		1469.4	adCH3C4 (52)	ad'CH3C4 (21)	rCH3C4 (21)
1488.9	adCH3C4 (59)	adCH3C4 (36)		1615.8	sdCH3C4 (88)	sO3C (11)	
1673.5	sC=O (84)			1741.2	sC=O (88)	sCC (5)	
2914.5	ssCH3C4 (100)			2854.1	ssCH3C4 (100)		
2936.5	ssCH3C11(100)			2913.1	saCH3'C4 (71)	saCH3C4 (29)	
3001.6	saCH3C4 (99)			2915.8	ssCH3C11(100)		
3012.5	saCH3C4 (99)			2917.3	saCH3C4 (71)	saCH3'C4 (29)	
3023.1	saCH3C11 (76)	ssCH3C11 (24)		2973.4	ssCH3'C11 (75)	saCH3C11 (25)	
3041.8	ssCH3C11 (76)	saCH3C11 (24)		2975.9	saCH3C11 (75)	ssCH3'C11 (25)	

^a Optimized vibrational contributions from interactions for which parameters have been developed in this study are shown in bold font; s stands for bond stretching with the variations ss and sa for symmetric and asymmetric stretching, respectively; d means angle deformation with the variations sd and ad for symmetric and asymmetric deformation, respectively; td and ti stand for torsional and improper torsion deformation, respectively; sc stands for scissoring, r for rocking; w for wagging and tw for twisting. ^b Frequencies are expressed in units of cm⁻¹.

The optimized vibrational spectra for the acylphosphate molecule compare very well with the QM spectra (Table 2-6). Likewise, the optimized vibrational spectra for the *N*-phosphonosulfonimidoyl molecule are in good agreement with the QM vibrational spectra (Table A-1 in the appendix).

Optimization of Dihedral Terms

Dihedral parameters to be optimized include force constants, multiplicities and phases. Dihedrals are optimized to reproduce the QM adiabatic potential energy surface (PES) especially the minima and barrier heights. The phase is usually assigned 0° or 180° so that different stereoisomers have same dihedral energy with the same parameters. The selected dihedral was fixed at a specific value while constraining the other dihedrals at their equilibrium values, and the angle was increased by 15° with optimization after each increment until the complete surface was obtained. The QM adiabatic potential energy surface (PES) was calculated on the global minimum structure using MP2/6-31+G(d), followed by single point calculations using MP2/cc-pVTZ on the optimized geometries. The target for torsion optimization was the reproduction of the overall QM potential energy surface (PES), especially the minima and barrier heights. However, in some cases, it is impossible to reproduce the overall PES. In such situation, it is essential to reproduce the minima rather than barrier heights, because barriers over than 8kcal/mol will not be populated during MD simulations. The potential energy scans on the acylphosphate and *N*-phosphonosulfonimidoyl dihedrals are shown (Figures 2-5 and 2-6).

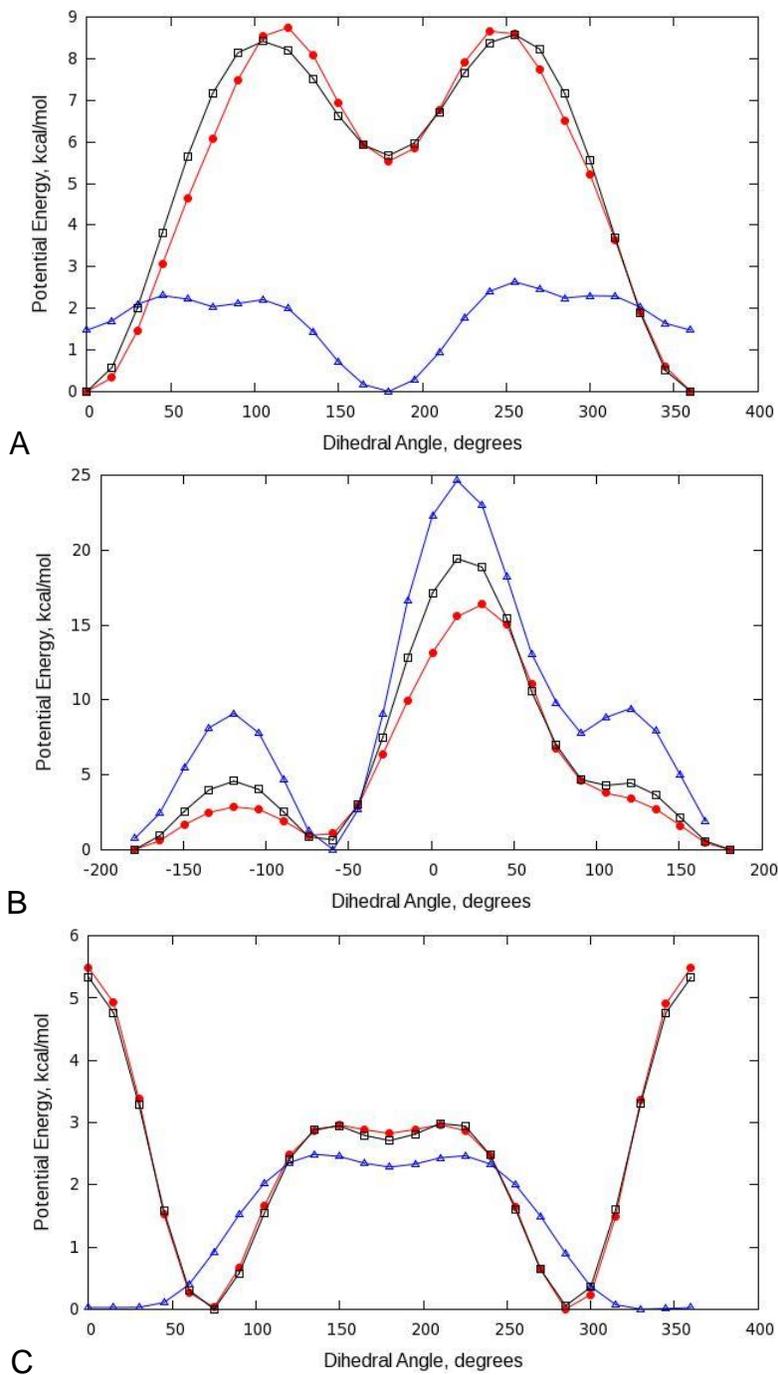
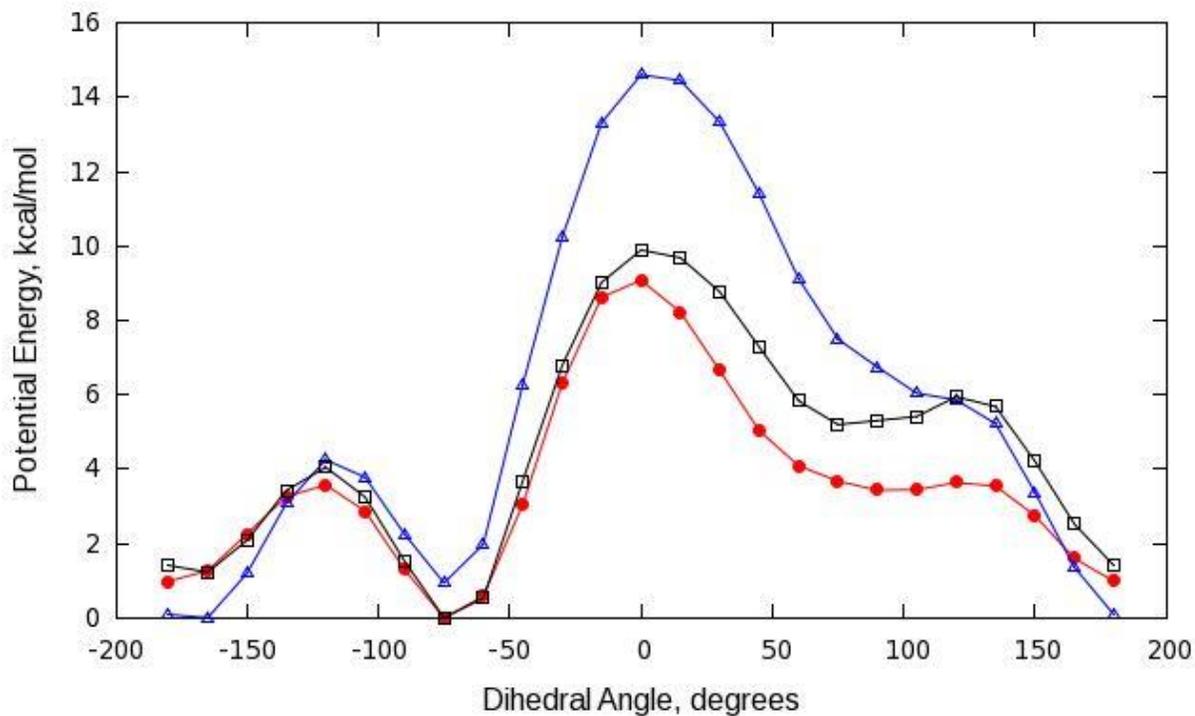
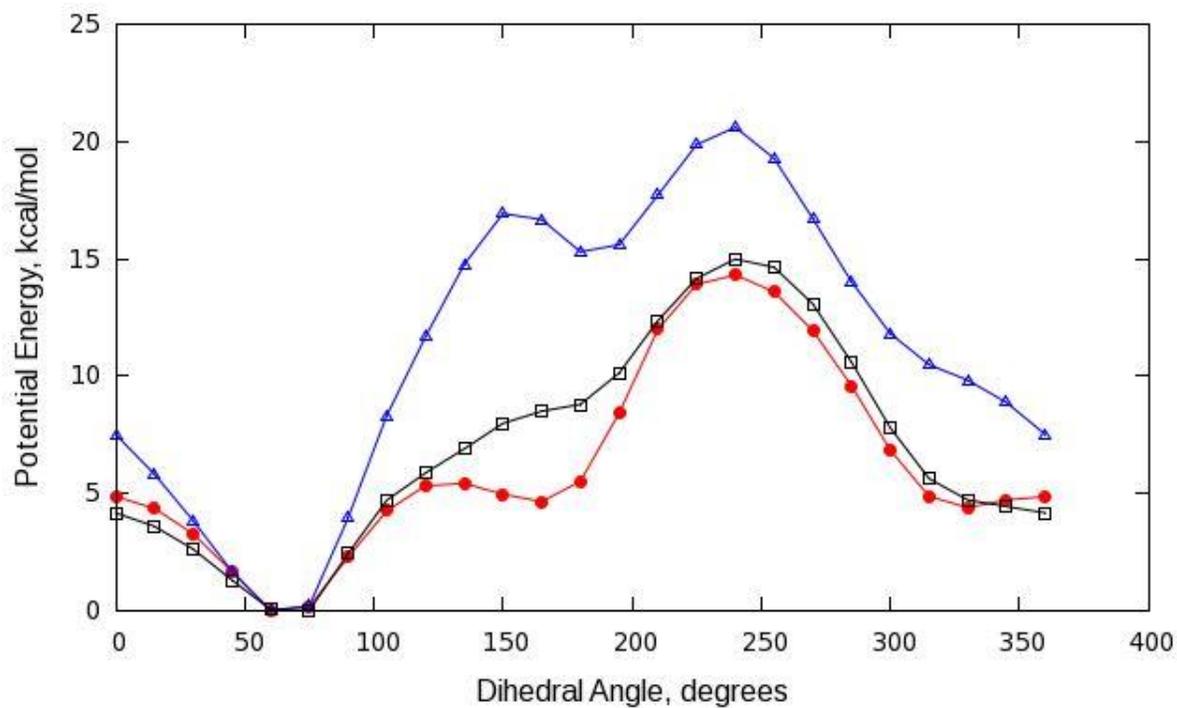


Figure 2-5. Potential energy scans on the acylphosphate molecule dihedrals. QM PES (red), optimized (black) and initial (blue) MM PES. A) C11-C9-O8-P1 PES. B) C9-O8-P1-O3 PES. C) O8-P1-O3-C4 PES. Interaction labels correspond to the atom numbers in Figure 2-3.

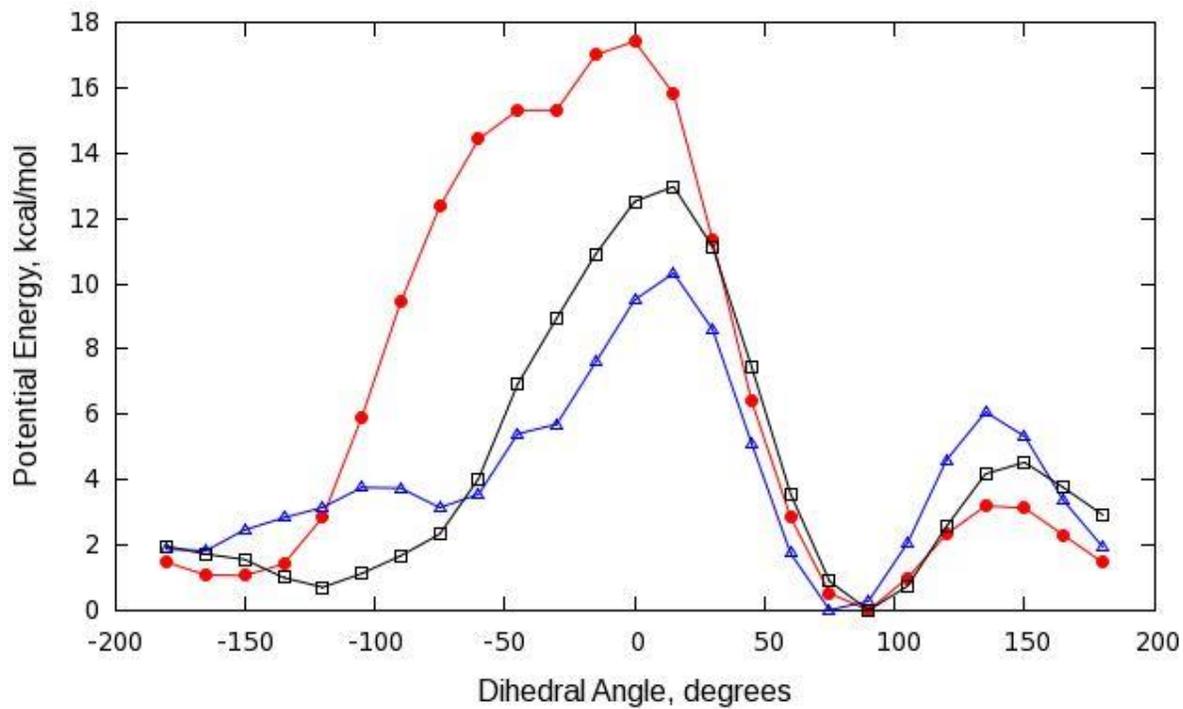


A

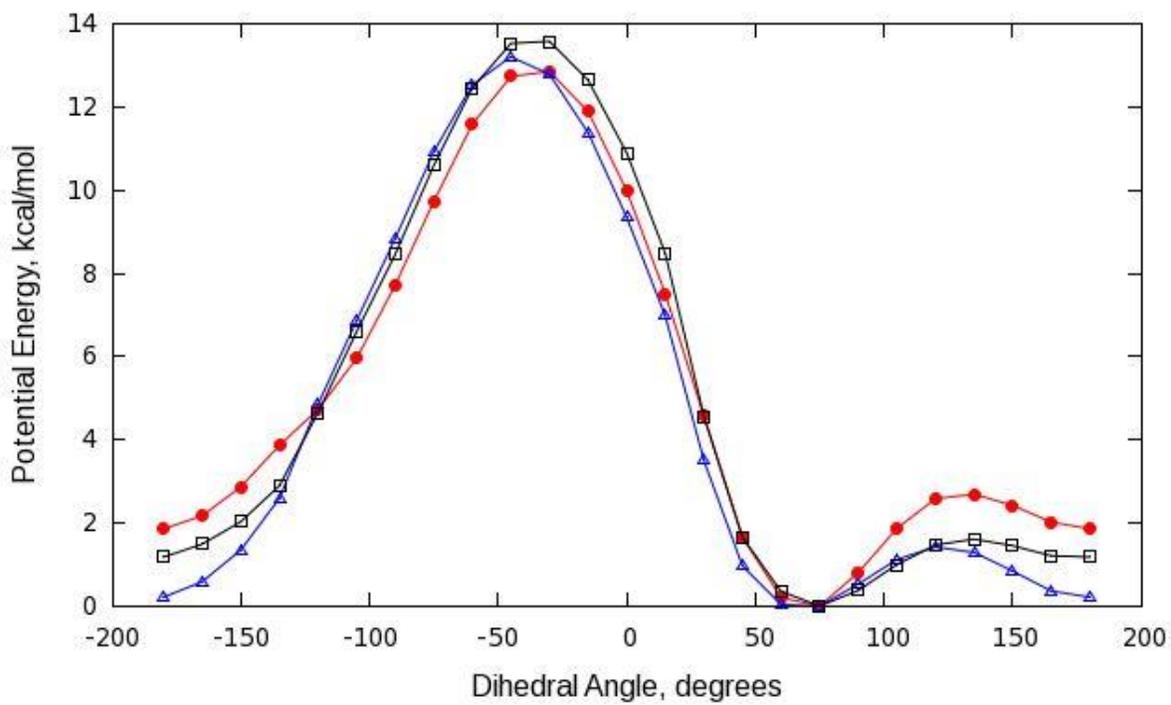


B

Figure 2-6. Potential energy scans on the *N*-phosphonosulfonimidoyl molecule dihedrals. QM PES (red), optimized (black) and initial (blue) MM PES. A) C19-C9-S11-N8 PES. B) C13-S11-N8-P1 PES. C) S11-N8-P1-O3 PES. D) N8-P1-O3-C4 PES. Interaction labels correspond to the atom numbers in Figure 2-3:



C



D

Figure 2-6. Continued.

It is very important when performing the potential energy surface scan that the conformation be the same in QM and MM calculations. For the acyl phosphate molecule, the MM geometry after minimization underwent conformational change that was different from the equilibrium QM geometry, making comparison of the QM and MM PES scans unreliable. Therefore, the MM geometry was constrained to the QM conformation during the CHARMM dihedral scan. The initial MM dihedral potential energy surfaces before optimization showed large deviations from the QM torsion PES, especially for the acylphosphate molecule. The O8-P1-O3-C4 dihedral for the acyl phosphate molecule (Figure 2-5) was not a missing parameter. However to improve the agreement of this dihedral with the QM one, a new atom type was assigned to O8, so that dihedral optimization could be performed. In general, it is preferable to assign a new atom type, rather than modifying already existing parameters to preserve the consistency of the force field and to allow transferability. This new atom type was assigned the same Lennard Jones parameters as O3. After optimization, the MM PES showed significant improvement. In summary, the newly developed parameters were able to reproduce the QM geometries, vibrational spectra, dihedral PES and interaction energies of the model compounds with water. The topology and parameter files for the acylphosphate and *N*-phosphonosulfonimidoyl models are provided (Appendix B).

Molecular Dynamics Simulations for Model Compounds in Aqueous Solution

As a further validation of the newly developed CGenFF parameters, molecular dynamics (MD) simulations of the acylphosphate and *N*-phosphonosulfonimidoyl model compounds in aqueous solution were performed. Rotation of the phosphate group was observed (Figures 2-7 and 2-8), and no major bond length or bond angle distortions occurred during the simulation.

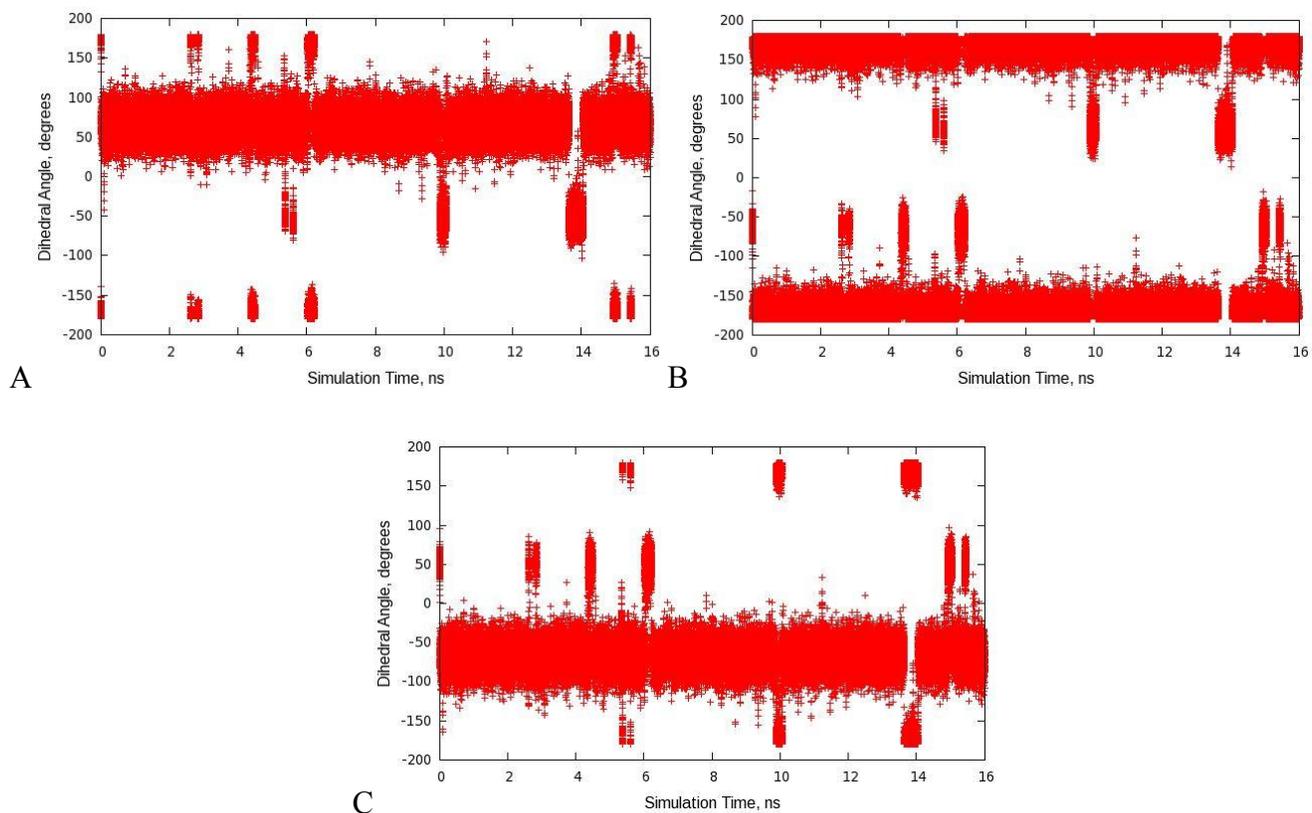


Figure 2-7. MD trajectory data showing that the phosphate moiety in the acylphosphate model undergoes rotation during the simulation. A) O2-P1-O8-C9. B) O3-P1-O8-C9, C) O7-P1-O8-C9. Dihedral angles are labeled with the atom numbers shown in Figure 2-3.

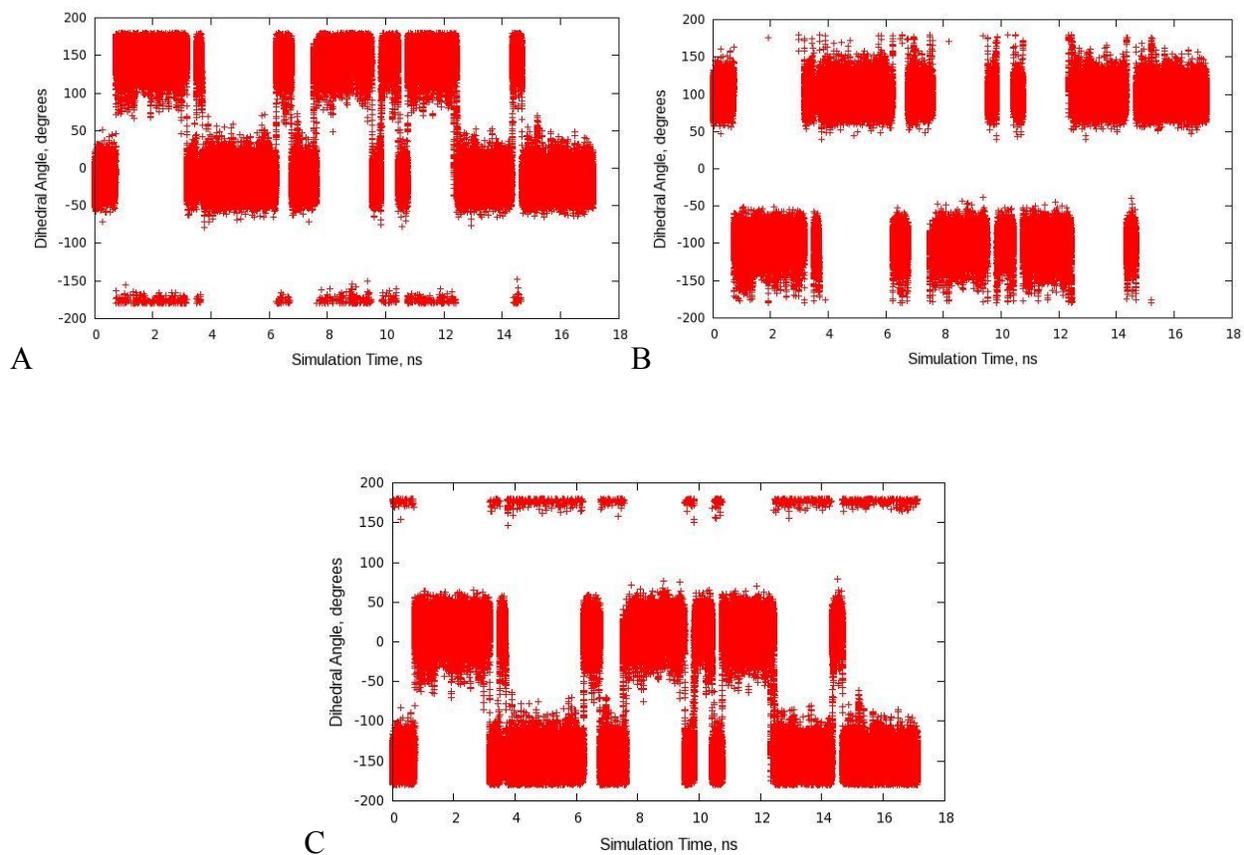


Figure 2-8. MD trajectory data showing that the phosphate moiety in the acylphosphate model undergoes rotation during the simulation. A) O2-P1-N8-C9. B) O3-P1-N8-C9. C) O7-P1-N8-C9. Dihedral angles are labeled with the atom numbers shown in Figure 2-3.

The temperature, pressure and total energy were monitored throughout the simulations with no large fluctuations were observed. In addition, the torsion angles for which new parameters had been developed fluctuated about values corresponding to minima on the potential energy surface. Therefore, these data suggests that these CGenFF parameters will be suitable for use in the simulated annealing, *in silico* docking, and free energy perturbation calculations that will be undertaken as part of the drug discovery efforts.

Summary

New force field parameters for the *N*-phosphonosulfonimidoyl and acylphosphate functional groups were developed. The parameterization process followed the CGenFF protocol with the ab initio results as the target data for the optimization. The charges were optimized to reproduce the QM interaction energies and distances for the model-water complexes. The results indicated that the new parameters were able to reproduce the QM geometries, vibrational spectra, potential energy surfaces for the dihedrals and the molecule-water interactions. The new parameters were further validated by molecular dynamics simulations of the *N*-phosphonosulfonimidoyl and acylphosphate models in aqueous solution. The newly developed parameters will enable computational studies using molecular mechanics approaches, such as molecular dynamics simulations and free energy calculations, on molecules containing the same functional groups using the CHARMM force field.

Table 2-7. New bonded interaction parameters assigned for the acylphosphate moiety.

Coordinate	Atom Types	Optimized Parameters			
Bonds		b_{eq}^a	K_b		
P1-O8	PG1-OG305	1.78	170		
C9-O10	CG2O2-OG305	1.34	230		
Bond Angles		θ_{eq}^b	K_θ	χ_{UB}	K_x
P1-O8-C9	PG1-OG305- CG2O2	121.5	70		
O2-P1-O8	OG2P1-PG1-OG305	103.0	60		
O3-P1-O8	OG303-PG1-OG305	90.8	60		
O7-P1-O8	OG2P1-PG1-OG305	103.0	60		
O8-C9-O10	OG305- CG2O2- OG2D1	118.0	70	2.26	160
O10-C9-C11	OG305- CG2O2- CG331	104.0	30	2.33	5
Dihedral Angles		K_ϕ^c	n	δ	
P1-O8-C9-O10	PG1-OG305-CG2O2-OG2D1	1.30	1	180	
P1-O8-C9-O10	PG1-OG305-CG2O2-OG2D1	2.60	2	180	
P1-O8-C9-C11	PG1-OG305-CG2O2-CG331	3.80	1	180	
P1-O8-C9-C11	PG1-OG305-CG2O2-CG331	1.60	2	180	
O2-P1-O8-C9	OG2P1-PG1-OG305-CG2O2	0.10	3	0	
O3-P1-O8-C9	OG303-PG1-OG305-CG2O2	0.10	2	180	
O3-P1-O8-C9	OG303-PG1-OG305-CG2O2	0.10	3	0	
C4-O3-P1-O8	CG331-OG303-PG1-OG305	1.47	2	0	
C4-O3-P1-O8	CG331-OG303-PG1-OG305	0.70	3	0	
Improper Torsion		K_ϕ^d	ϕ_o		
O8-O10-C11-C9	OG305-OG2D1-CG331-CG2O2	56	0		

^a b_{eq} , equilibrium bond distance (Å) and K_b , force constant (kcal/mol/Å²).

^b θ_{eq} , equilibrium bond angle (°) and K_θ , force constant (kcal/mol/rad²).

^c K_ϕ , torsional potential (kcal/mol), n and δ , periodicity and phase offset (°) of the torsion, respectively.

^d K_ϕ , improper dihedral potential (kcal/mol/rad²) and ϕ_o , equilibrium improper dihedral angle (°).

Table 2-8. New bonded interaction parameters assigned for the *N*-phosphonosulfonimidoyl moiety.

Coordinate	Atom Types	Optimized Parameters		
Bonds	Atom Types	b_{eq}^a	K_b	
P1-N8	PG1-NG2D1	1.72	100	
N8-S11	NG2D1-SG3O2	1.53	400	
Bond Angles	Atom Types	θ_{eq}^b	K_θ	
P1-N8-S11	PG1-NG2D1-SG3O2	113.0	30	
O2-P1-N8	OG2P1-PG1-NG2D1	106.0	50	
O3-P1-N8	NG2D1-PG1-OG3O3	98.8	94	
O7-P1-N8	OG2P1-PG1-NG2D1	106.0	50	
N8-S11-C9	NG2D1-SG3O2- CG321	114.3	65	
N8-S11-O12	NG2D1-SG3O2- OG2P1	119.0	65	
N8-S11-C13	NG2D1-SG3O2- CG321	114.0	79	
Dihedral Angles	Atom Types	K_ϕ^c	n	δ
P1-N8-S11-O2	PG1-NG2D1-SG3O2-OG2P1	2.50	1	180
P1-N8-S11-O2	PG1-NG2D1-SG3O2-OG2P1	1.00	2	0
P1-N8-S11-C9	PG1-NG2D1-SG3O2-CG321	1.00	2	0
P1-N8-S11-C13	PG1-NG2D1-SG3O2-CG331	1.00	1	0
P1-N8-S11-C13	PG1-NG2D1-SG3O2-CG331	0.60	2	0
O2-P1-N8-S11	OG2P1-PG1-NG2D1-SG3O2	0.50	4	0
O3-P1-N8-S11	OG3O3-PG1-NG2D1-SG3O2	1.80	1	0
O3-P1-N8-S11	OG3O3-PG1-NG2D1-SG3O2	3.00	2	0
N8-P1-O3-C4	NG2D1-PG1-OG3O3-CG331	0.40	1	0
N8-P1-O3-C4	NG2D1-PG1-OG3O3-CG331	0.80	2	0
N8-P1-O3-C4	NG2D1-PG1-OG3O3-CG331	0.35	3	0
N8-S11-C9-C19	NG2D1-SG3O2-CG321-CG331	1.40	1	180
N8-S11-C9-C19	NG2D1-SG3O2-CG321-CG331	0.001	3	0
N8-S11-C9-H17	NG2D1-SG3O2-CG321-HGA2	0.16	3	0
N8-S11-C13-H14	NG2D1-SG3O2-CG331-HGA3	0.18	3	0

^a b_{eq} , equilibrium bond distance (Å) and K_R , force constant (kcal/mol/Å²).

^b θ_{eq} , equilibrium bond angle (°) and K_θ , force constant (kcal/mol/rad²).

^c K_ϕ , torsional potential (kcal/mol/rad²); n and δ , periodicity and phase offset (°) of the torsion, respectively.

CHAPTER 3 MOLECULAR DYNAMICS SIMULATIONS OF UNLIGANDED ASPARAGINE SYNTHETASE AND ASNS/BASPAMP COMPLEX

Background

Proteins are not static structures; instead they are in constant motion. The energy landscape can be used to describe a protein's conformational space, and different protein conformations are populated according to their energies.^{83,84} Molecular dynamics (MD) simulation is a computational technique widely used to study biological macromolecules like proteins and to obtain valuable insights in to their dynamical behavior, including both the fast motions of the amino acid side chains and the slow conformational changes of active site loops.^{85,86} MD simulations also provide time average properties of biological systems, such as density, interaction energies and entropies.

The main goal of this research is the identification of potent lead inhibitors for human asparagine synthetase (hASNS) through virtual screening of large libraries of chemical compounds. This is very challenging because the protein receptor conformation used in the docking study is crucial to the outcome.⁸⁷⁻⁸⁹ For example; the conformations of an apo enzyme may not be adequate for docking new ligands, because of wrongly positioned side chains or loops that block ligand access to the binding site. Incorporating protein flexibility in the docking process through molecular dynamics (MD) simulations plays an important role in improving the docking calculations and thereby increasing the chances of predicting reliable lead inhibitors. Therefore, use of one representative protein structure of that best matches the most probable protein conformation observed after simulations, or docking the ligand into different protein conformations, can reduce the risks associated with docking to an inadequate protein

conformation.^{90,91} Therefore, it is important to prepare the structure of the target protein carefully before considering the docking process.

Currently, there exists only one crystal structure of glutamine-dependent asparagine synthetase, the *Escherichia coli* enzyme (AS-B) with a resolution of 2Å (Figure 1-2, Chapter 1).¹ Various attempts to obtain crystal structures for hASNS or AS-B complexed with the intermediate β -aspartyl-AMP (β AspAMP) have been unsuccessful.⁹² Therefore, there is incomplete information on the important catalytic residues in the synthetase active site, because the available crystal structure does not include the substrate or the intermediate and has only the byproduct adenosine monophosphate (AMP) bound in the synthetase active site. A computational model was developed previously in the Richards group where β -aspartyl-AMP intermediate, PPI and one Mg^{2+} ion were modeled in the synthetase active site using the very highly similar enzyme β -lactam synthetase complexed with the adenylated substrate (CEA), ATP analog (AMP-CPP) and one Mg^{2+} ion (PDB ID 1JGT) (Figure 3-2).^{92,94} The lack of many CHARMM parameters for the β AspAMP intermediate limited efforts towards the refinement of the structure of ASNS/ β AspAMP/PPI/ Mg^{2+} complex. However, after developing new CHARMM parameters for β AspAMP, as described in detail in Chapter 2, the ASNS/ β AspAMP/PPI/ Mg^{2+} model was refined using molecular dynamics simulations. The ligand-free ASNS was also simulated to detect related conformational changes as a result of intermediate binding and to identify key active site residues. The results of simulations were also compared to already available experimental data on mutagenesis and other experiments.

Molecular Dynamics Simulation

Molecular dynamics simulation is a widely used computational technique that explores microscopic information such as atomic positions and velocities. Statistical mechanics provides the mathematical equations that relate the microscopic information to macroscopic properties, such as pressure, energy and heat capacity. In statistical mechanics, averages of experimental observables are expressed as ensemble averages. An ensemble is a collection of molecular systems that have similar thermodynamic states, such as temperature, pressure, volume and number of particles. Each system is described by the phase space of the system which includes specific positions, r , and momenta, P , for all particles.⁹⁵

There exist different kinds of ensembles; for example, microcanonical ensembles, canonical ensembles and isobaric-isothermal ensembles. In microcanonical ensemble (NVE), thermodynamic state of the system is characterized by a fixed number of atoms, N , a fixed volume, V , and a fixed energy. In Canonical ensembles (NVT), the systems have fixed number of atoms, N , fixed volume, V , and a fixed temperature, T . Isobaric-isothermal ensembles (NPT) are composed of systems that have fixed number of atoms, N , fixed pressure, P , and fixed temperature, T .

In statistical mechanics, the phase space average of a property of interest, A , for a canonical ensemble, is the integral over all phase space.⁹⁶

$$\langle A \rangle = \iint A(r, p) P(r, p) dr dp \quad (3-1)$$

$P(r, p)$ is the probability of finding the system at every point in phase space and is given by

$$P(r, p) = \frac{e^{-\frac{E(r, p)}{k_B T}}}{Q} \quad (3-2)$$

where $E(r, p)$ is the total energy of the system, k_B is Boltzmann's constant. T is the temperature and Q is the canonical ensemble partition function

$$Q = \iint e^{-\frac{E(r, p)}{k_B T}} dr dp \quad (3-3)$$

Calculation of ensemble averages using Equation 3-3 is extremely difficult, especially taking in to account every possible state of the system. Alternatively, in molecular dynamics simulations, the average value for a property of interest is calculated as the time average attained when the motion of a single point through phase space is followed as a function of time.⁶⁶ The time average is expressed as

$$\langle A \rangle_t = \frac{1}{t} \int_0^t A(r_t, t_t) dt \quad (3-4)$$

According to the ergodic hypothesis; when the trajectory of the system of interest is allowed to evolve indefinitely, that system will cover all phase space, so both the time average and ensemble average will be equal (Equation 3-5).⁹⁵

$$\lim_{t \rightarrow \infty} \langle A(r, p) \rangle_t = \langle A(r, p) \rangle_Q \quad (3-5)$$

Therefore, if the system of interest is allowed to evolve in time until sufficient representative conformations are generated, experimentally relevant information can be calculated. Molecular dynamics simulations are based on solving Newton's second law or equation of motion

$$F_i = m_i a_i = m_i \frac{d^2 r_i}{dt^2} \quad (3-6)$$

where F_i is the force acting on particle i , m_i is the mass of the particle i , a_i is its acceleration given the second derivative of the particle's position, r_i , with respect to time, t . The force can also be obtained from the gradient (∇_i) of the potential energy function, $U(r)$ as

$$F_i = -\nabla_i U(r) \quad (3-7)$$

Combining Equation 3-6 and Equation 3-7 gives the acceleration:

$$-\nabla_i U(r) = m_i \frac{d^2 r_i}{dt^2} \quad (3-8)$$

The potential energy, $U(r)$ is a very complicated function of all atomic coordinates of the system. It ca not be solved analytically and must be evaluated numerically. There are numerous algorithms used in the integration of equations of motion: for example, the Verlet, leap-frog and velocity Verlet algorithms. These algorithms use Taylor expansion to approximate positions, velocities and acceleration.⁶⁶ The leap frog algorithm was used in the present study. In this algorithm, the velocities are calculated at the midpoints of the time steps and it can be written as

$$r(t + \Delta t) = r(t) + v\left(t + \frac{1}{2} \Delta t\right) \Delta t \quad (3-9)$$

$$v\left(t + \frac{1}{2} \Delta t\right) = v\left(t - \frac{1}{2} \Delta t\right) + a(t) \Delta t \quad (3-10)$$

Where Δt represents the time step. The size of the time step is an important factor in the integration quality. The smaller the Δt , the more accurate the integration. However, use

of very small Δt , makes the simulations very slow and does not allow sufficient sampling of phase space in a reasonable computational time. The proper time step should be 0.05 of the time involved in the fastest motion in the system. For example, in proteins the fastest motion is the stretching vibration of the bond of hydrogen with heavy atoms (X-H). This occurs at around 3000cm^{-1} with a period of about 10 fs thus the time step size should not exceed 0.5fs.⁶ This value of time step size is too small to sample enough of the ensemble phase space in a reasonable computer time. To overcome this obstacle, the SHAKE algorithm is used to constrain X-H bonds to their equilibrium values. Thus, freezing those vibrations saves simulation time by a factor of 2 ($\Delta t = 0.1$ fs) and hence offers more integration stability. Initial coordinates are assigned from experimentally determined x-ray crystallography and NMR experiments or are generated by computer modeling. The structure is then further submitted to minimization to relieve nonbonded interactions as well as bond lengths and angle distortions, before carrying out the molecular dynamics simulations. The initial velocities are randomly assigned from a Gaussian or Maxwell-Boltzmann distribution (Equation 3-11),⁶ which gives the probability that an atom, i , with mass, m_i , has a velocity, v , at specific temperature, T

$$P(v_i) = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{1}{2}} \exp \left[\frac{-m_i v_i^2}{2k_B T} \right] \quad (3-11)$$

It is advisable to start the simulations with a gradual heating phase starting from an initial low temperature and then to heat gradually until the desired temperature is reached. This gradual heating process avoids accidental assignment of high velocities to a group of atoms, thereby creating a “hot spot” leading to unstable simulations.⁶⁶

Following the heating phase, the system is equilibrated for several hundred picoseconds. The system properties, such as total energy, pressure and temperature are monitored during this phase to ensure that the system is free of unpredictable fluctuations. After equilibrating the system, the system is submitted for production simulations which are allowed to run for nanoseconds, depending on the system size and the computer time available, and the generated trajectory is stored for further analysis.

Computational Details

ASNS complexed with β AspAMP, PPI and one Mg^{2+} ion was computationally modeled previously.^{51,94} This model used the very highly similar enzyme, β -lactam synthetase (β -LS) complexed with its substrate, N_2 -(carboxyethyl)-L-arginine (CEA), α,β -methyleneadenosine 5'-triphosphate (AMP-CPP), an ATP analog, and one Mg^{2+} ion in the active site (PDBID 1JGT) to model β AspAMP, PPI and the Mg^{2+} ion in the active site of ASNS. However, several x-ray crystallographic snapshots of the β -LS catalytic cycle were resolved afterwards, and all showed two Mg^{2+} ions, instead of only one Mg^{2+} ion as in the previously resolved β -LS/CEA/AMP-CPP complex.⁴² Because the residues involved in binding of the ATP and Mg^{2+} ions are conserved between AS-B and β -LS (Figure 3-1), we decided to modify the original model and another Mg^{2+} ion was modeled on the opposite side of inorganic pyrophosphate (PPI) by superimposing the β -lactam synthetase (PDBID 1MBZ) and ASNS active sites (Figure 3-2). β -Lactam synthetase converts N_2 -(carboxyethyl)-L-arginine (CEA) to deoxyguanidinoproclavaminc acid (DGPC) in the presence of ATP and two Mg^{2+} ions.

```

β-LS 1 MGAPVLPAAF GFLASART-- --GG-GRAP- GPVFAT---- -RGSHTDIDT POGERSIAAT 49
AS-B 1 MC SIFGVFDIKT DAVELRKKAL ELSR-LMRER GPDWSGIYAS ----DNAILA 47

β-LS 50 LVHAPSVAPD RAV-ARSLTG APTTAVLAGE IYNRDELLSV LP---AGPAP EGDAELVLRÍ 105
AS-B 48 HERLSIVDVN AGAQPLYNQQ KTHVLAVNGE IYNHQALRAE YGDYQFQTG SDCE-VILAL 106

β-LS 106 LERYDLHAFR LVNGRFATVV RT--GDRVLL A'DHAGSVPL YTCVAP-GEV RASTEAKALA 162
AS-B 107 YQEKGPFLD DLQGMFAFAL YDSEKDAYLI GRDHLGIIPL YMGYDEHGQL YVASEMKALV 166

β-LS 163 AHRDPKGFPL ADARRVAGLT GVYQVPAGAV MDIDLGSFTA VTHRTWTP-- GLSRRIL-PE 219
AS-B 167 PVC-----R TIKEFPAGSY LWSQ----DG EIRSYHRDW FDYDAVKDNV 206

β-LS 220 GEAVAAVRAA LEKAVARVT PGDTPLVVLS GGDSSGVAA CAHRAA----- 265
AS-B 207 TDK-NELRQA LEDSVKSHLM SDVPYGVLLS GGLDSSIISA ITKKYARRV EDQEREAWWP 265

β-LS 266 -GELDTVSMG TDTSNEFREA RAVVDHLRTR HREITIPTE LLAQLPYAVW ASESVDPII 324
AS-B 266 Q--LHSFAVG LPGSPDLKAA QEVANHLGTV HHEIHFTVQE GLDAIRDVIY HIETYDVTTI 323

β-LS 325 EYLLPLTALY RALD-GPERR ILTGYGADIP LGMHREDRL PALDTV--LA HDMATFDGLN 381
AS-B 324 RASTPMYLMS RKIKAMGIKM VLSGEGSDEV FGGYLYFHKA PNAKELHEET VRKLLALHMY 383

β-LS 382 EMSP--VLST LAGHWTHPY WDREVLDLLV $LEAGLKRRH ---GRDKWVL RAAMADALPA 436
AS-B 384 DCARANKAMS AWGVEARVPF LDKKFLDVAM RINPQDKMC- GNGKMEKHIL RECFEAYLPA 442

β-LS 437 ETVNRPKL-G VHEGSGTTSS FSRLLLDHGV AEDRVHEAKR QVVRELFDLT VGGGRHPSEV 495
AS-B 443 SVAWRQEQ- -----FS DVGYSW--- -----ID 462

β-LS 496 DTDDVVRSVA D-RTARGAA 513
AS-B 463 TLKEVAAQOV SDQOLETARF RFPYNTPTSK EAYLYREIFE ELFPLPSAAE CVPGGPSVAC 523

AS-B 524 SSAKAIEWDE AFKKMDDPSG RAVGVHQSAY K 553

```

Figure 3-1. Structure based sequence alignment of β-LS and ASNS. © 2001, rights managed by nature publishing group.⁹³

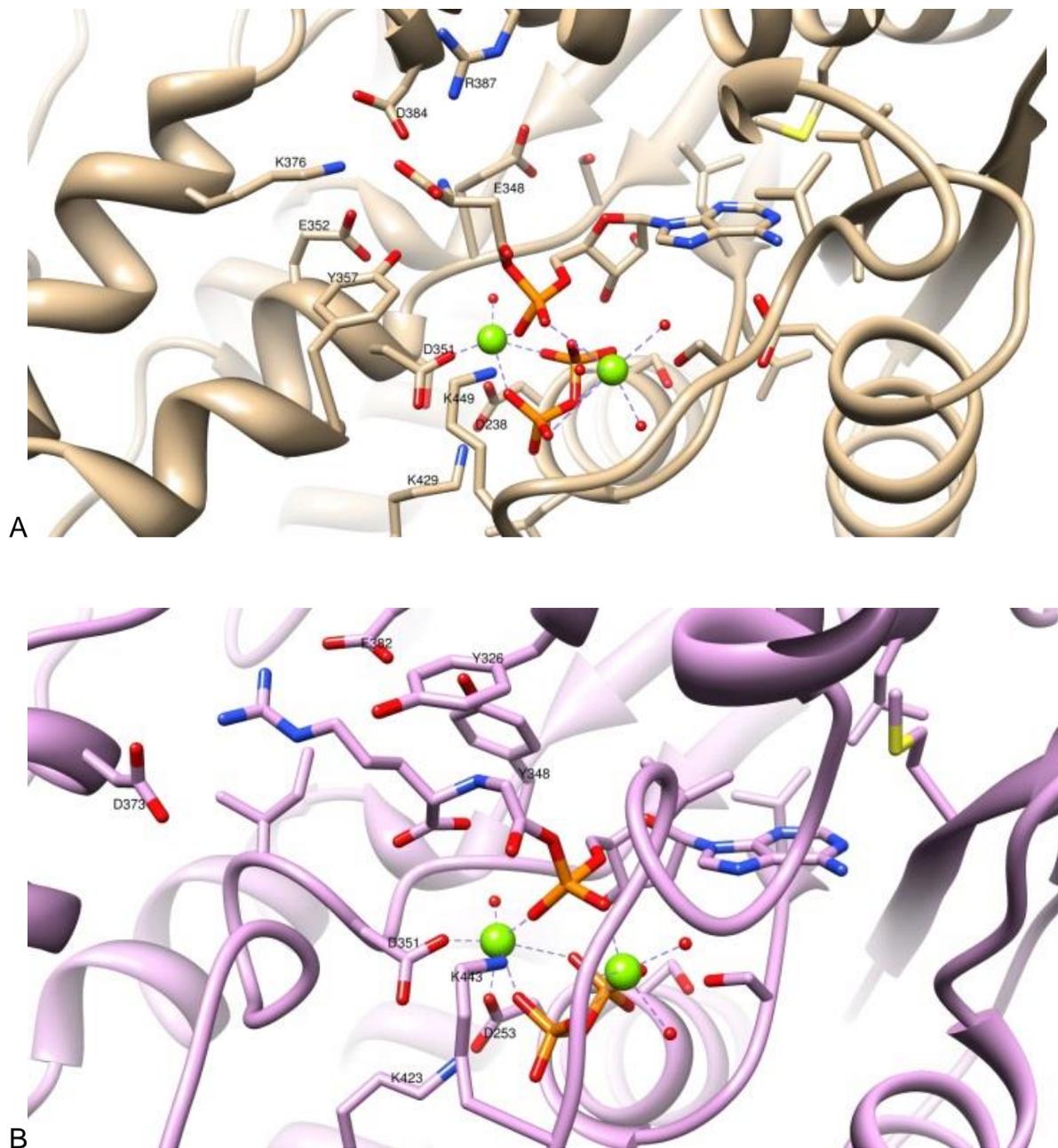


Figure 3-2. Stereoview of the AS-B and β -LS active sites A) The AS-B model active site with β AspAMP, PPI and two Mg^{2+} ions bound. B) The β -LS (PDBID 1MBZ) active site with CMA-AMP, PPI and two Mg^{2+} ions bound (PDBID 1MBZ). The intermediates and key residues are shown in stick representation. Mg^{2+} ions are shown as green spheres.

Both AS-B and β -LS facilitate intermediate formation by adenylation, but AS-B catalyzes intermolecular amide bond formation by an amino group, in contrast to β -LS, which catalyzes intramolecular amide bond formation to form a β -lactam ring (Figure 3-3). Two systems were designed for simulations: the unliganded ASNS and the ASNS/ β AspAMP/PPi/ Mg^{2+} complex. To sample adequate conformational space in reasonable computer time, each system was simulated as three independent trajectories as opposed to a single long trajectory. Each trajectory was assigned a different initial velocity by using a different seed for the random number generator.⁹⁷ The hydrogen atoms were then added to the resulting model using the HBUILD utility in CHARMM.⁵² The pKa for charged amino acid residues side chains were checked using PROPKA web interface.⁹⁸⁻¹⁰⁰ Charged amino acids side chains were modeled at their usual protonation states at physiological pH. The topology and parameter files for the enzyme used the CHARMM27 biological force field. The parameters for β AspAMP were optimized following a standard protocol (Chapter 2). The system was minimized first in vacuum for 50 steps using steepest descent (SD), followed by 50 steps of the adopted basis Newton-Raphson (ABNR) algorithm. Harmonic restraints were applied on the intermediates and protein backbone atoms with force constant of 50 kcal/mol/Å². The harmonic restraints on the protein atoms were then turned off while keeping the restraints on the Mg^{2+} ions and the coordinated water molecules. The system was then minimized again for 100 steps using the ABNR algorithm. The system was then solvated in an octahedral box (92 Å x 92 Å x 92 Å) of TIP3P water molecules.⁷⁵ The ions were added to give a system of zero net charge. Coulombic interactions were treated by the particle mesh Ewald method¹⁰¹ with a cut-off value of 12 Å.

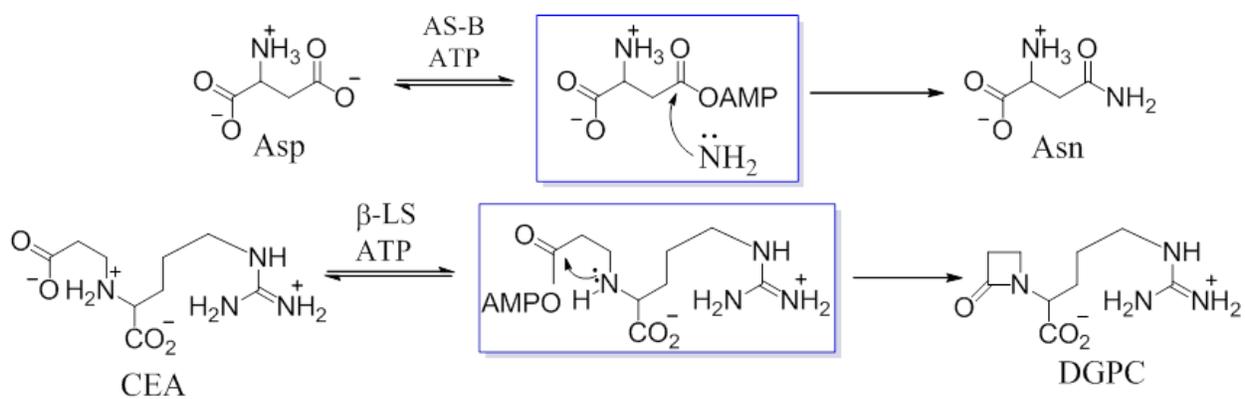


Figure 3-3. Comparison between Chemical reactions of AS-B and β -LS. A) AS-B. B) β -LS.

The Lennard-Jones (LJ) interactions were treated by a force switching function¹⁰³ over the range 10-14 Å. In the presence of periodic boundary conditions,⁷⁹ each system was minimized for 500 steps using the SD algorithm and 5000 steps using the ABNR algorithm. The minimized structure was then gradually heated for 30ps. The Leapfrog algorithm was used to integrate equations of motion with a 1fs time step. The SHAKE algorithm⁸⁰ was used to constrain the length of covalent bonds to hydrogen atoms to their equilibrium values.

Each system was then equilibrated at 300 K using a canonical ensemble (NVT) over 40ps. The restraints on the Mg²⁺ ions and coordinating water molecules were maintained during minimization in solvent, heating and the initial 40 ps of the equilibration period, after which the harmonic restraints were released. Final equilibration and production simulations of the system were performed using an isothermal-isobaric (NPT) ensemble. The Hoover thermostat was used to maintain the temperature around the average.^{57,81} The total simulation time for each system was 30 ns, including heating and equilibration time. The simulations were performed using CHARMM version 36b1.

CHARMM analysis tools were used to calculate root mean square deviations (RMSD) and root mean square fluctuations (RMSF) and to monitor hydrogen bonds and torsion angles. An equilibration time of 1ns was excluded from each trajectory during the analysis, resulting in a total production simulation time of 27 ns. The root mean square deviation (RMSD), which is a measure of the distance difference between two atoms averaged over all atoms, is a frequently used tool to assess the stability of the

simulations. The RMSD is defined by Equation 3-12, where δ_i is the distance difference between two equivalent atoms

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (\delta_i)^2}{N}} \quad (3-12)$$

Root mean square fluctuation measures the deviation of a specific atom, x_i , from a reference position of the same atom (average position in our case) averaged over time. The RMSF gives information about the structure mobility and is defined as following:

$$RMSF = \sqrt{\frac{1}{T} \left(\sum_{t=0}^T (x_i(t_j) - \bar{x}_i) \right)} \quad (3-12)$$

Hydrogen bond analysis was performed using default criteria with a donor-acceptor distance cutoff of 2.4 Å and no angle cut-off value. The conformations of the residues in the protein were monitored using χ_1 and χ_2 dihedral angles. The χ_1 corresponds to the dihedral angle of N-C α -C β -C γ in the amino acid residue of interest, while χ_2 corresponds to the dihedral angle C α -C β -C γ -C δ .

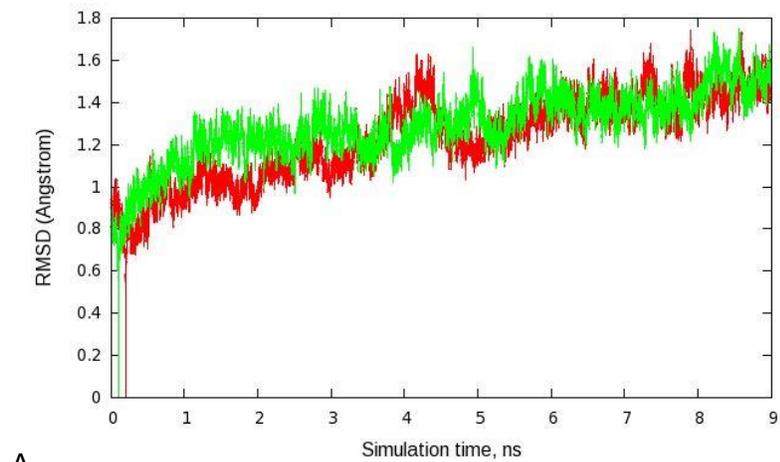
Results and Discussion

Root Mean Square Deviation (RMSD)

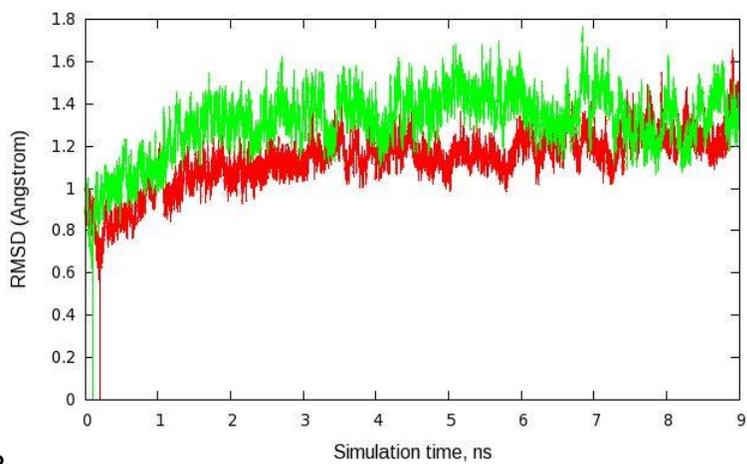
The RMSD of the backbone atoms for each independent trajectory was calculated as described above (Figure 3-4). The starting geometry of production simulations was used as the reference structure. RMSD did not exceed 2.0 Å thereby assessing the simulations' stability.

Root Mean Square Fluctuations (RMSF)

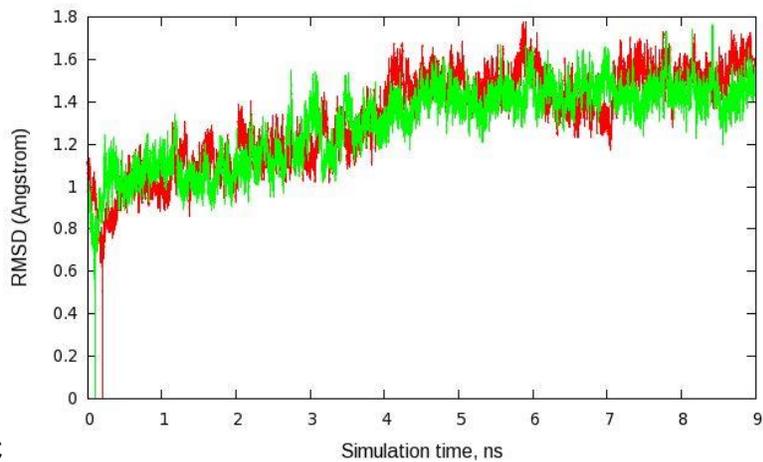
RMSF were calculated per residue backbone atoms for both the free and liganded enzyme independent trajectories (Figures 3-5 through 3-8).



A



B



C

Figure 3-4. Root mean square deviation (RMSD) of backbone atoms for unliganded (green) and ligand bound (red) asparagine synthetase enzyme. A) MD1. B) MD2. C) MD3.

The RMSF difference between the liganded and unliganded enzyme for each independent trajectory reveals changes in fluctuations in many backbone regions. Enzyme regions with RMSF differences greater than 0.5 Å are colored for clarity. Each trajectory exhibits different flexibility regions. Regions which are far from the active site will not be considered further. The loop extending in the region 442-452 exhibits reduced fluctuations in two simulations due to presence of the intermediate in the active site. This loop bears the catalytically important residue Lys449. The loop 442-452 moves towards the inside of the active site forming a more closed active site in the ASNS/ β Asp/PPi/ Mg^{2+} complex (Figure 3-8).

Mg^{2+} Coordination

The two Mg^{2+} ions displayed approximately octahedral coordination in all three simulations (Figure 3-9). The Mg^{2+} ion (Mg1) has direct contacts with the β - and γ -terminal oxygen atoms of the pyrophosphate (PPi), one oxygen atom from the phosphate group of β AspAMP, a one side chain carboxylate oxygen from each of the Asp-351 and Asp-238 residues, and one water molecule. The second Mg^{2+} ion (Mg2) is located on the opposite side of the pyrophosphate group and has direct contacts with oxygen atoms from each of the pyrophosphate α -terminal, β -bridging and γ -terminal groups and three water molecules. In the third independent MD simulations, Mg2 is observed to coordinate with the side chain carboxylate of Asp-279. Mg2 has indirect contact with the other oxygen of β AspAMP phosphate group through a bridging water molecule which is observed in all three simulations of the ASNS/ β AspAMP/PPi/ Mg^{2+} complex. Mg^{2+} ions stabilize the negative charge on PPi especially during the adenylation reaction. This behavior of Mg^{2+} coordination was also observed in β -LS/CMA-AMP/PPi/ Mg^{2+} complex (PDBID 1MBZ) (Figure 3-2).⁴²

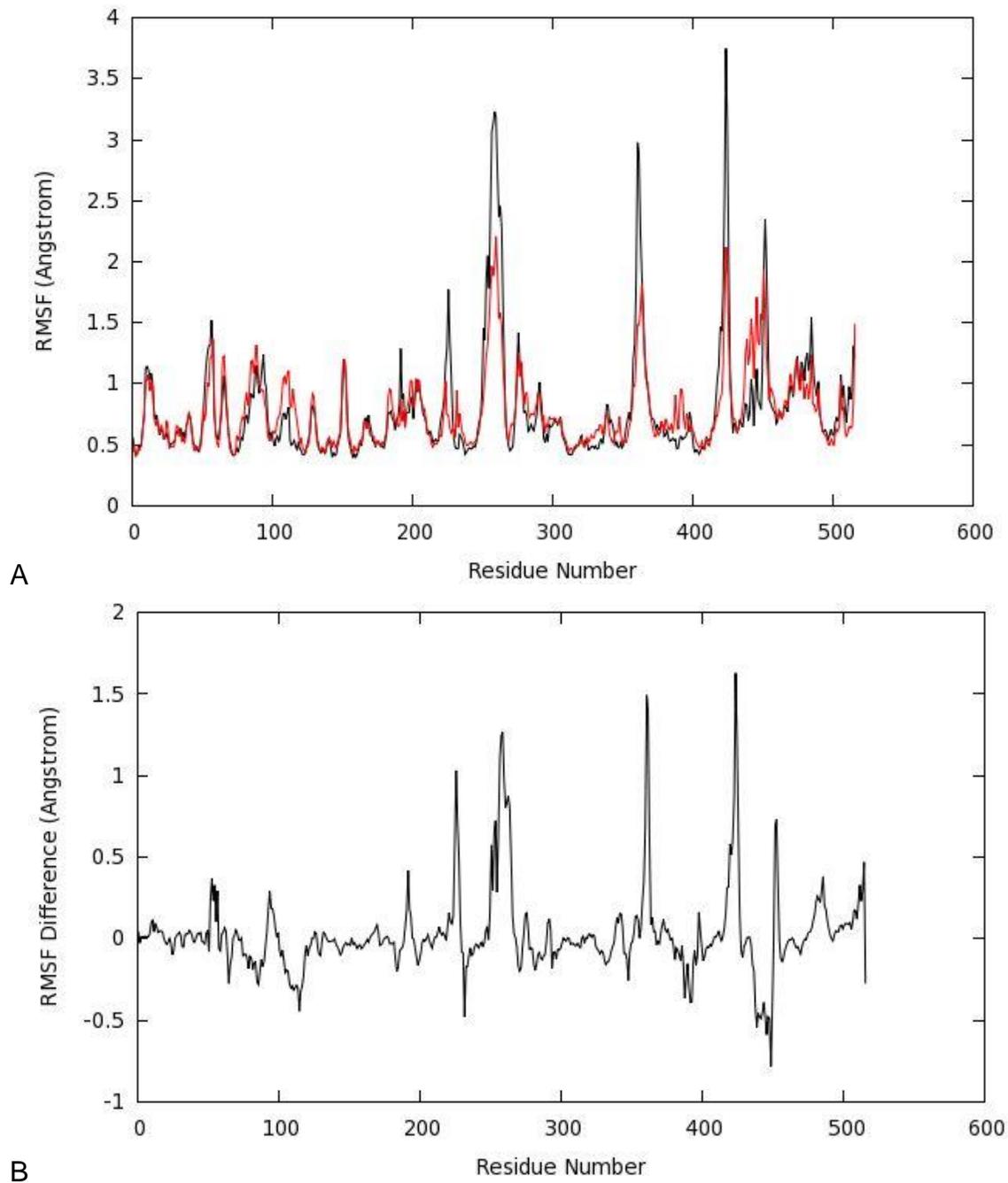


Figure 3-5. RMSF of backbone atoms for first independent trajectory simulations (MD1) (A) RMSF for unliganded (red) and ligand bound (black) enzyme. (B) RMSF difference between liganded and unliganded simulations.

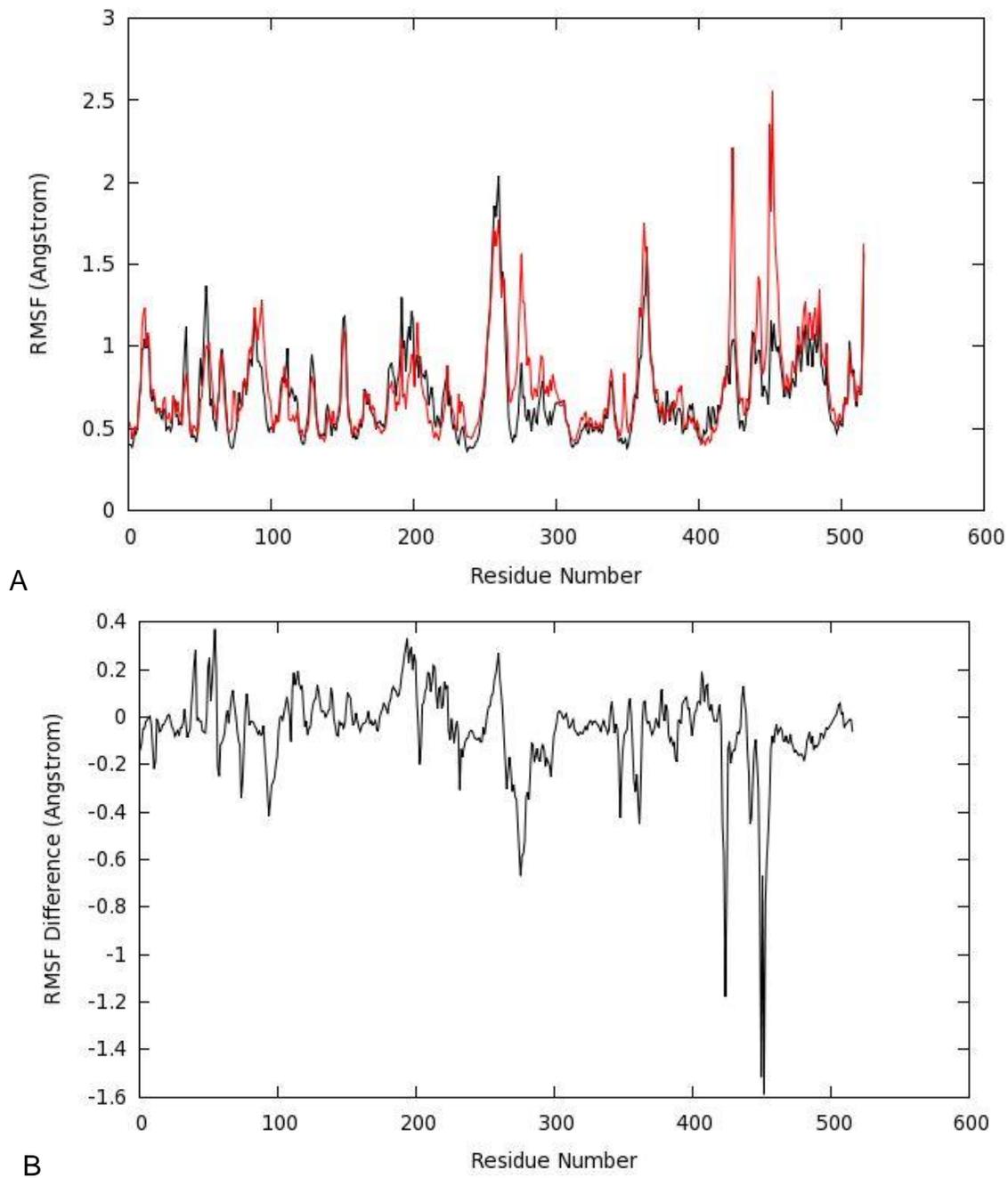
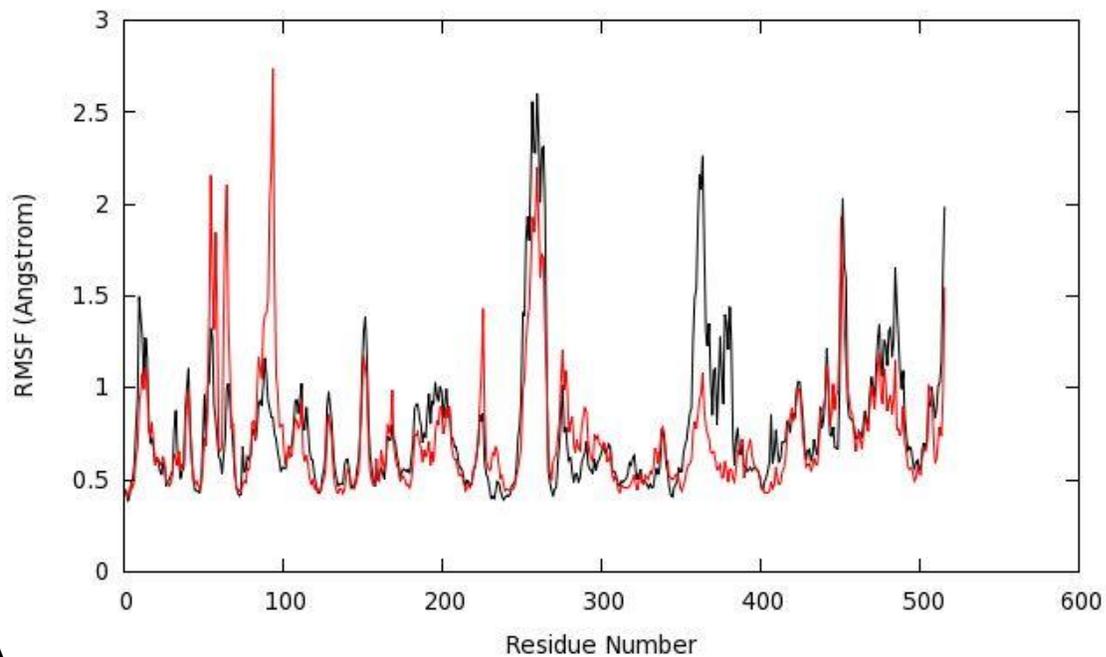
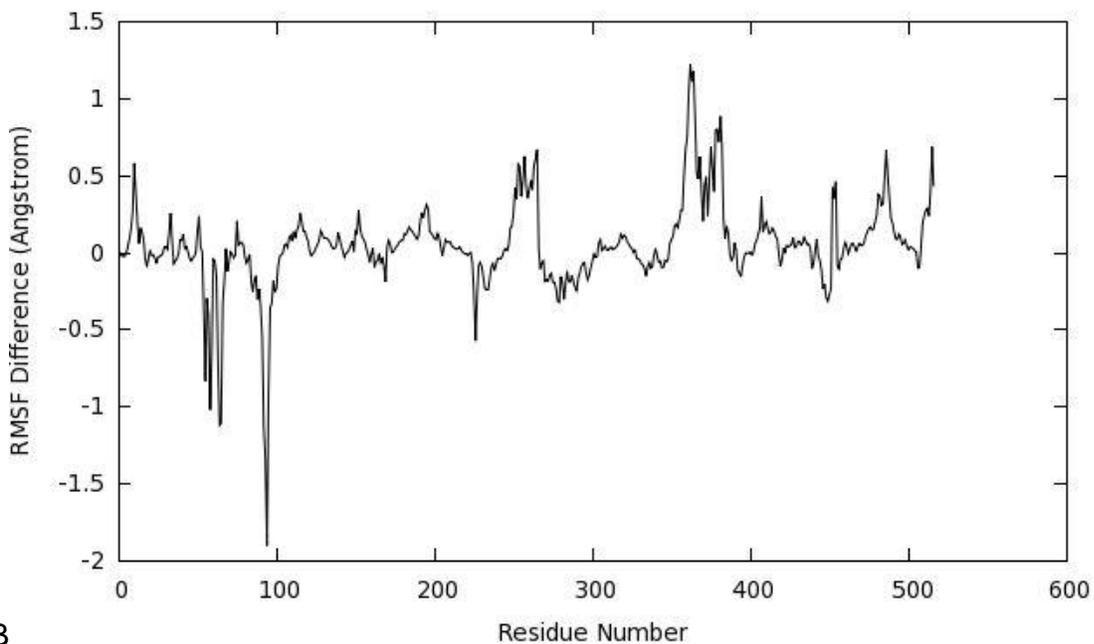


Figure 3-6. RMSF of backbone atoms for the second independent trajectory simulations (MD2). A) RMSF for unliganded (red) and ligand bound (black) enzyme. B) RMSF difference between liganded and unliganded simulations.



A



B

Figure 3-7. RMSF of backbone atoms for third independent trajectory simulations (MD3) (A) RMSF for unliganded (red) and ligand bound (black) enzyme. B) RMSF difference between liganded and unliganded simulations.

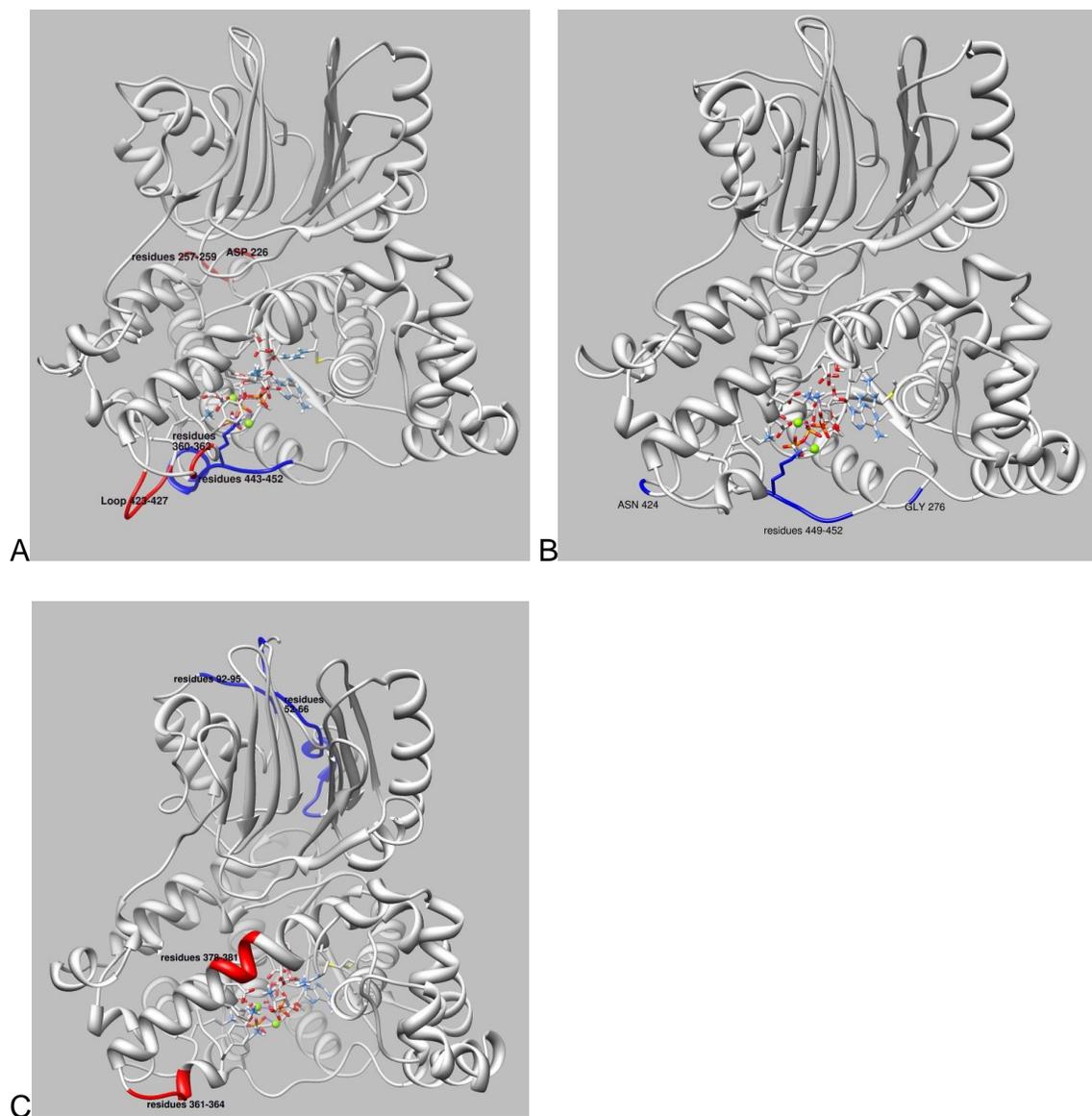


Figure 3-8. Enzyme regions with changes in RMSF of backbone atoms greater than 0.5 Å are colored. Red indicates increased fluctuations in the ASNS/BASP/PPi/Mg²⁺ complex and blue indicates reduced fluctuations in the ASNS/βAsp/PPi/Mg²⁺ complex. A) MD1. B) MD2. C) MD3.

Hydrogen Bonds

The stability of hydrogen bonding interactions over all simulations was also investigated (Tables 3-1 and 3-2). Table 3-1 shows the average occupancy and average lifetime for the hydrogen bonds between the intermediate β -aspartyl-AMP and active site residues in the three independent simulations for the ASNS/ β AspAMP/PPi/ Mg^{2+} complex. The amino group of β AspAMP forms three very stable hydrogen bonds with the carboxylate side chains of Glu-348, Glu-352 and Asp-384. The carboxylate group of β AspAMP is stabilized by a hydrogen bond to Lys-376. One of the phosphate oxygen atoms coordinates with one Mg^{2+} ion (Mg1). The other phosphate oxygen on the opposite side makes a hydrogen bond with a bridging water molecule that coordinates with the other Mg^{2+} ion (Mg2). This bridging water molecule is stable in the three simulations of ASNS/ β AspAMP/PPi/ Mg^{2+} complex.

In kinetic studies to investigate the inhibitory effect of transition state analogs, compounds **3** and **4**, on human ASNS showed that compound **3** has much less inhibitor activity than compound **1** while compound **4** did not show any inhibitory effect (Figures 1-4 and 1-5 in Chapter 1). These results suggested that the negatively charged phosphate group is essential to inhibit hASNS, and it was hypothesized that the negatively charged oxygen of the phosphate group makes a critical electrostatic interaction with the catalytically important Lys 449 in compounds **2** and **3**, producing a resonance-stabilized negative charge on the sulfamate.⁵¹ However, according to the MD simulation results, the negative charge of the phosphate group makes critical electrostatic interactions by coordinating to the Mg^{2+} ions and stabilizing the overall integrity of the active site for subsequent asparagine formation. Therefore, the negative resonance forms of the sulfamate groups in compounds **2** and **3** are able to mimic

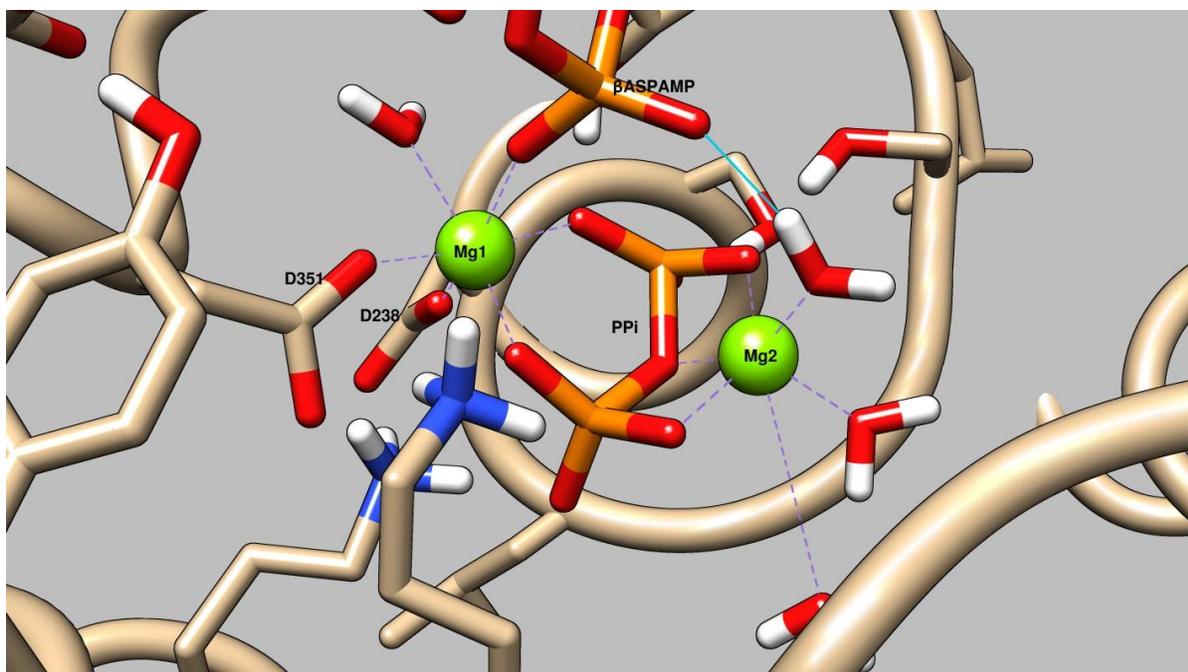


Figure 3-9. Focused view of the Mg^{2+} ions' coordinating residues. Coordinate bonds are indicated by purple pseudobonds.

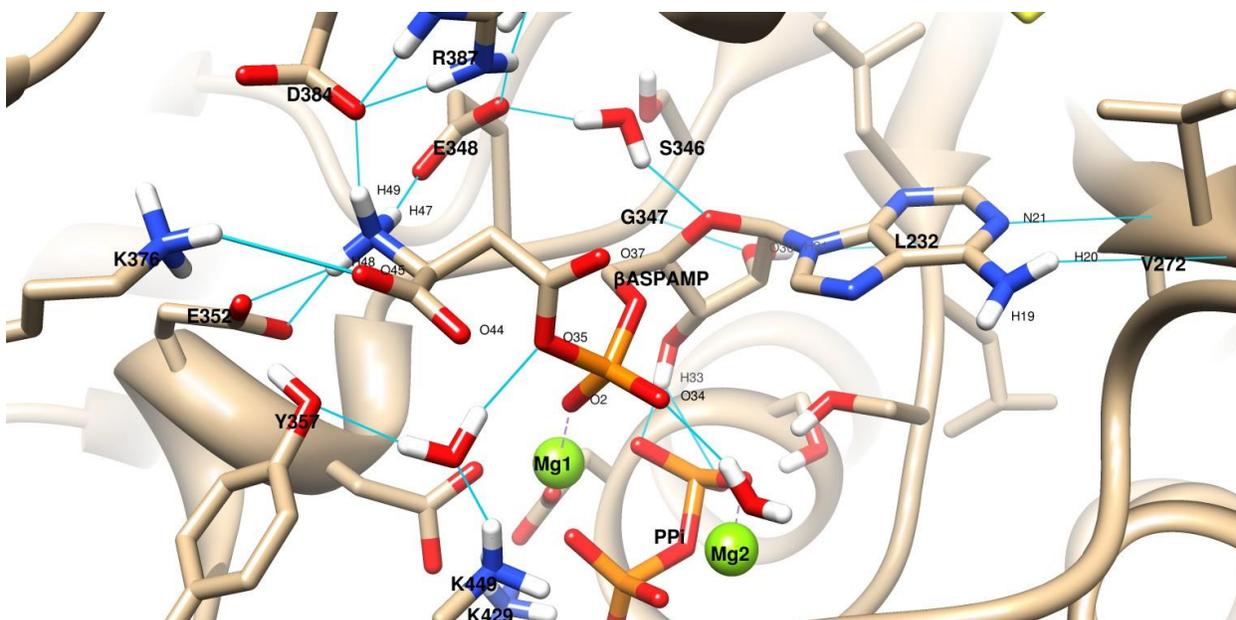


Figure 3-10. Hydrogen bonding interactions of β AspAMP with active site residues. Hydrogen bonds are indicated by blue lines.

the negative phosphate group in the β AspAMP intermediate and hence exhibit their inhibitory effects. Lys-449 is observed to interact with the acyl phosphate oxygen of β AspAMP through a bridging water molecule. One of the adenosine hydroxyls interacts with the backbone atoms of Gly-347 and Leu-232, and the other hydroxyl makes bifurcated hydrogen bonds with two pyrophosphate terminal β - oxygen atoms. The adenine nitrogen and amino group make hydrogen bonds to the backbone of Val-272.

The PPI binding pocket, called the PP-motif, is conserved in many enzymes which catalyze ATP hydrolysis from the α - β bond of ATP such as argininosuccinate synthetases, asparagine synthetases, and ATP sulfurylases and is characterized by the sequence SGGXDS.⁴⁴ From the hydrogen bonding analysis, PPI forms stable hydrogen bonds within the PP-motif with the side chain of Ser-234, backbone of Leu-237 and both backbone and side chain atoms of Ser-239 (Figure 3-11). The absence of ATP/PPI exchange, the very low rate of PIX (i.e. β AspAMP intermediate formation is irreversible) and the very stable interactions of PPI deep within the PP-motif all suggest that PPI is the last product to be released.¹⁰⁴ Similarly, in β -LS, PPI binds deeply in the active site cleft and kinetic studies suggested that PPI is the last product released.^{42,105}

Stability of hydrogen bonding interactions between amino acid residues within 7Å from β AspAMP were also investigated (Tables 3-3 and 3-4). There is no significant difference in the hydrogen bonding pattern in the ASNS active site in presence and absence of intermediate binding. That is due to the presence of a set of charged amino acids that make a network of electrostatic interactions and keep the active site residues positioned in an orientation ready for binding of the substrate.

Table 3-1. Hydrogen bonding interactions of β -aspartyl-AMP (BAA) with active site residues.

H-Bond	Average occupancy, % (Average life time, ps)		
	MD1	MD2	MD3
BAA-H47...OE1-E348	100 (4500)	100 (9000)	61 (2741)
BAA-O30...HN-G347	75 (12)	61 (7)	86 (19)
BAA-H31...O-L232	23 (6.8)	0.2 (4)	59 (12)
BAA-H31...OG-S346	58 (21)	88 (25)	24 (51)
BAA-H19...O-V272	17 (57)	19 (15)	93 (111)
BAA-H20...O-V272	80 (52)	69 (24)	5 (105)
BAA-N21...HN-V272	62 (7)	83 (15)	71 (9)
BAA-O44...HH-Y357	2 (3)	72 (150)	-
BAA-O45...HH-Y357	30 (9)	0.2 (3)	-
BAA-O45...NH3-K376	94 (356)	96 (222)	7 (65)
BAA-H48...OE1-E352	93 (77)	100 (1515)	36 (18)
BAA-H48...OE2-E352	64 (16)	57 (7)	44 (16)
BAA-H49...OD1-D384	99 (998)	100 (1135)	100 (1228)
*BAA-O34...NH3-K449	23 (14)	6 (10)	2 (8)
*BAA-O35...NH3-K449	17 (10)	23 (11)	33 (14)
*BAA-O35...OH-Y357	7 (4)	5 (3)	29 (4)
*BAA-O30...O-L232	34 (7)	33 (5)	6 (4)
*BAA-O30...HN-G347	18 (4)	19 (4)	1 (3)
*BAA-O32...O-G347	60 (6)	40 (5)	23 (7)
*BAA-O44...OD2-D351	10 (17)	2 (3)	-
*BAA-O44...OH-Y357	21 (5)	9 (3)	27 (4)
*BAA-O44...HH-Y359	12 (9)	8 (3)	
*BAA-O44...NH3-K449	52 (37)	60 (48)	22 (13)
*BAA-O45...OH-Y357	10 (4)	9 (3)	7 (3)
*BAA-O45...OOC-D384	48 (13)	41 (10)	100 (20)
*BAA-O45...HH12-R387	46 (6)	17 (5)	25 (5)

H-bonding interactions through bridged water molecules are indicated by an asterisk (*).

Table 3-2. Hydrogen-bonding interactions of the pyrophosphate group (PPi) with amino acid residues in the active site.

H-Bond	Average occupancy,% (Average life time, ps)		
	MD1	MD2	MD3
PPi-O1G...NH3-K449	62 (60)	3 (9)	1 (8)
PPi-O2G...NH3-K449	36 (14)	58 (30)	89 (35)
PPi-O2G...NH3-K429	29 (15)	30 (5)	1 (5)
PPi-O3A...HG1-S234	99 (1496)	93 (36)	100 (9000)
PPi-O3G...HN-L237	55 (7)	85 (18)	92 (32)
PPi-O3G...HZ3-K429	20 (415)	100 (4549)	100 (9000)
PPi-O2B...HN-S239	99 (640)	100 (4549)	100 (2260)
PPi-O2B...HG1-S239	100 (4499)	100 (9000)	100 (9000)
*PPi-O1B...O-G347	23 (3)	7 (3)	8 (3)
*PPi-O1G...NH3-K449	17 (16)	55 (50)	20 (15)
*PPi-O1G...HN-E450	19 (6)	83 (19)	68 (19)
*PPi-O3G...H22-R433	10 (10)	87 (25)	77 (15)
*PPi-O3G...OE2-E428	100 (78)	0.3 (6)	7 (5)
*PPi-O3G...NH3-K429	79 (945)	-	-

H-bonding interactions through bridged water molecules are indicated by an asterisk (*).

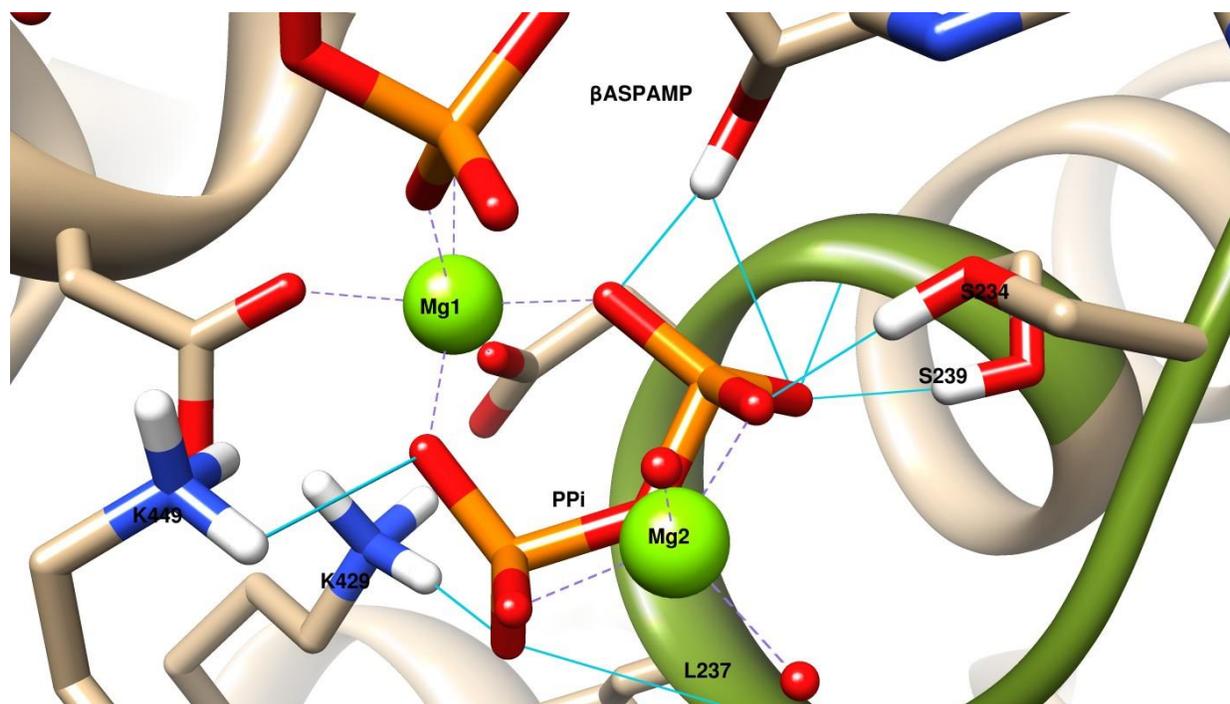


Figure 3-11. Interactions of PPi with active site residues. Hydrogen bonding interactions are indicated by blue lines The PP-motif is colored in green.

Table 3-3. Hydrogen bonding interactions between the residues in the active site in MD simulations of ASNS/ β AspAMP.

H-Bond	Average occupancy, % (Average life time, ps)		
	MD1	MD2	MD3
E348-NH...OG-S346	28 (4)	27 (3)	30 (4)
E348-OE1...HN-G349	89 (25)	91 (29)	63 (12)
E348-OE2...HH21-R387	97 (126)	28 (16)	60 (3)
E352-COO...HN-E352	90 (100)	99 (376)	45 (32)
E352-OE1...NH3-K376	27 (240)	98 (3000)	33 (120)
E352-OE2...NH3-K376	75 (759)	-	20 (85)
E352-COO...HH-Y357	71 (25)	8 (550)	90 (50)
D384-NH...OD2-D384	71 (10)	73 (16)	73 (11)
D384-O...HN-A388	89 (25)	36 (4)	92 (38)
D384-OD1...NH3-K376	40 (10)	87 (8)	-
D384-OD1...NH2-R387	100 (312)	70 (80)	99 (70)
D384-OD2...NH3-K376	69 (142)	80 (52)	-
K376-O...NH-A379	98 (17)	87 (21)	49 (12)
K376-NH...O-E372	82 (17)	89 (28)	94 (44)
K376-NH3...OE1-E352	99 (1000)	99 (3751)	50 (220)
K376-NH3...OOC-D384	76 (148)	90 (62)	-
L237-O...HN-I241	83 (16)	87 (21)	91 (30)
K429-NH...OOC-E428	95 (64)	99 (2700)	99 (1130)
K429-O...HN-L432	19 (4)	36 (4)	23 (3)
K429-O...HN-R433	75 (13)	88 (21)	52 (6)
K429-NH3...OOC-D238	99 (120)	99 (3000)	60 (10)
K429-NH3...OOC-D351	90 (195)	99 (4550)	99 (9000)
K449-NH...OOC-E428	73 (11)	90 (60)	89 (26)
K449-NH3...OOC-D351	92 (145)	99 (60)	60 (30)
S346-OG...HN-E348	28 (4)	27 (3)	30 (4)
S346-NH...O-R400	56 (7)	60 (6)	50 (6)
S234-HN...OG-S239	61 (7)	6 (4)	73 (10)
S234-HN...OG-S234	26 (4)	2	23 (4)
S234-OG...HN-G236	7 (5)	62 (7)	6 (5)
S239-OG...HN-S234	61 (7)	6 (4)	73 (10)
S239-OG...HN-G235	20 (3)	-	64 (7)
S239-O...HN-S243	92 (31)	98 (142)	98 (118)
S239-O...HG1-S243	60 (41)	99 (13)	2 (17)

Table 3-4. Hydrogen bonding interactions between the residues in the active site in MD simulations of unliganded ASNS.

H-Bond	Average occupancy, % (Average life time, ps)		
	MD1	MD2	MD3
E348-HN...OG-S346	31 (4)	34 (4)	3 (3)
E348-OE1...HN-G349	5 (10)	-	-
E348-OE2...NH2-R387	70 (40)	300 (18)	63 (45)
E348-OE1...NH2-R387	50 (35)	270 (16)	70 (40)
E352-OE2...HN-E352	97 (96)	80 (19)	99 (357)
E352-OE1...NH3-K376	100 (2998)	80 (1300)	99 (4500)
E352-OE2...NH3-K376	75 (759)	-	-
E352-COO...HH-Y357	99 (50)	92 (73)	34 (4)
D384-HN...OD2-D384	4 (3)	37 (8)	10 (3)
D384-O...HN-A388	79 (17)	20 (4)	86 (24)
D384-OD1...NH3-K376	2 (3)	125 (15)	9 (10)
D384-OD1...NH2-R387	100 (75)	95 (32)	99 (183)
D384-OD2...NH3-K376	83 (80)	28 (40)	98 (948)
K376-O...NH-A379	78 (13)	84 (18)	80 (13)
K376-NH...O-E372	93 (39)	89 (28)	90 (28)
K376-NH3...OOC-E352	100 (2998)	70 (1232)	99 (3700)
K376-NH3...OOC-D384	83 (80)	20 (30)	99 (1000)
L237-O...HN-I241	82 (19)	96 (79)	80 (15)
K429-NH...OE1-E428	80 (40)	99 (389)	99 (1500)
K429-O...HN-L432	30 (4)	21 (4)	40 (4)
K429-O...HN-R433	90 (28)	97 (89)	72 (10)
K429-NH3...OOC-D238	99 (132)	97 (111)	99 (9000)
K429-NH3...OOC-D351	99 (153)	99 (50)	99 (2000)
K449-NH...OE2-E428	51 (15)	-	-
K449-NH3...OOC-D351	57 (71)	-	99 (200)
K449-NH3...OOC-D279	16 (100)	90 (40)	-
S346-OG...HN-E348	31 (4)	33 (4)	3 (2.8)
S346-NH...O-R400	80 (15)	77 (13)	94 (44)
S346-HG1...OOC-E348	23 (40)	7 (9)	-
S234-HN...OG-S239	5 (3)	4 (3)	18 (4)
S234-HN...O-L232	41 (11)	55 (8)	-
S234-HN...OG-S234	11 (3)	9 (3)	16 (3)
S234-OG...HN-G236	18 (4)	19 (4)	18 (4)
S239-OG...HN-S234	5 (3)	4 (3)	18 (4)
S239-HG1...OG-S234	41 (14)	52 (15)	54 (8)
S239-OG...HN-G235	38 (6)	43 (6)	10 (3)
S239-OG...HN-G236	5 (3)	10 (3)	4 (3)
S239-O...HN-S243	98 (128)	98 (155)	97 (102)
S239-O...HG1-S243	82 (27)	64 (52)	86 (65)

The hydrogen bonding interactions together with conformational behavior for each residue will be discussed in detail in the following section

Behavior of Important Residues in the Active Site of Unliganded ASNS and the ASNS/ β AspAMP/PPi/Mg²⁺ Complex.

Glu-348

The χ_1 and χ_2 torsions of Glu-348, which is conserved among all ASNS enzymes,¹ exhibit different conformational behavior in the presence and absence of the intermediate (Figures 3-12 and 3-13). In the unliganded ASNS, the Glu-348 side chain samples different conformations with the carboxylate group forming hydrogen bonds with the guanidine group of Arg-387, and the backbone of Glu-348 making a hydrogen bond with the hydroxyl group of Ser-346. In the ASNS/ β AspAMP complex, the χ_1 and χ_2 torsions of Glu-348 change their conformation to orient Glu-348 into the active site with one of the carboxylate oxygens making a hydrogen bond with the amino group of β AspAMP and the backbone of Gly-349. The other carboxylate oxygen still hydrogen bonds with the guanidine group of Arg-387 and the hydroxyl group of Ser-346. Oxygen-18 isotope transfer experiments have shown that E348A and E348Q AS-B mutants were not able to form the β AspAMP intermediate although their glutaminase activity is retained as the wild type enzyme.¹⁰⁴ On the other hand, E348D AS-B mutant was able to form a β AspAMP intermediate, but the Glu:Asn ratio for the E348D increased to a value of 18:1 compared to a value of 7.2:1 in the wild-type enzyme. This indicates that the glutaminase and synthetase active sites became more uncoupled in the mutant enzyme. These results suggest that Glu-348 has an essential role in catalyzing the formation of the β AspAMP intermediate and may be playing a role, albeit weakly, on

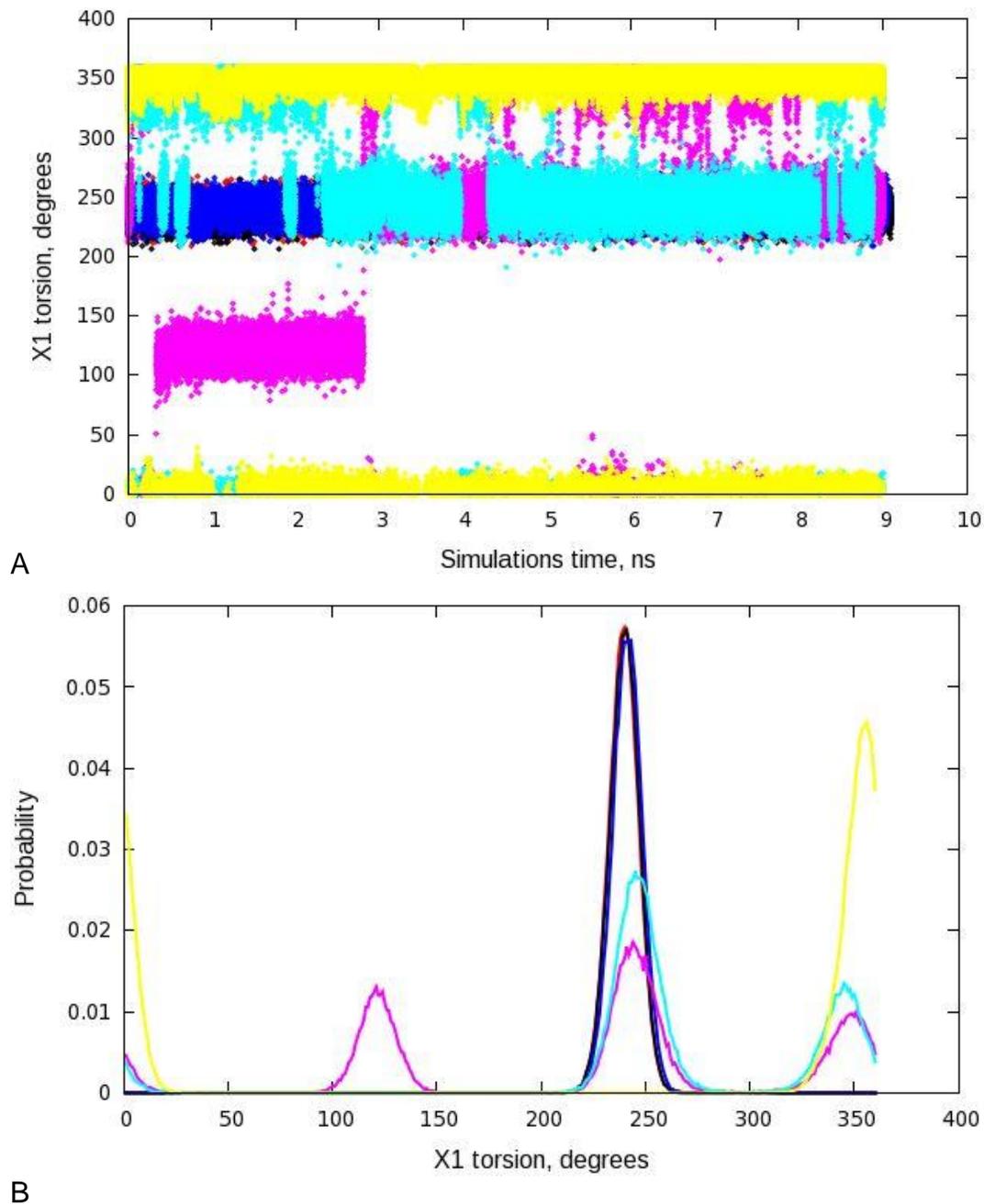
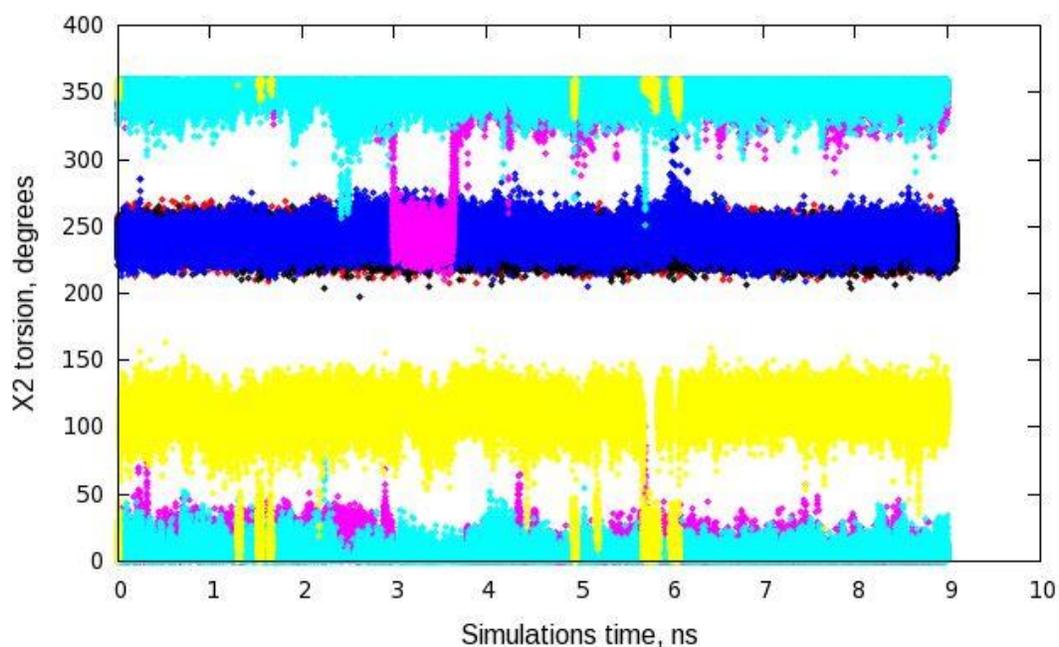
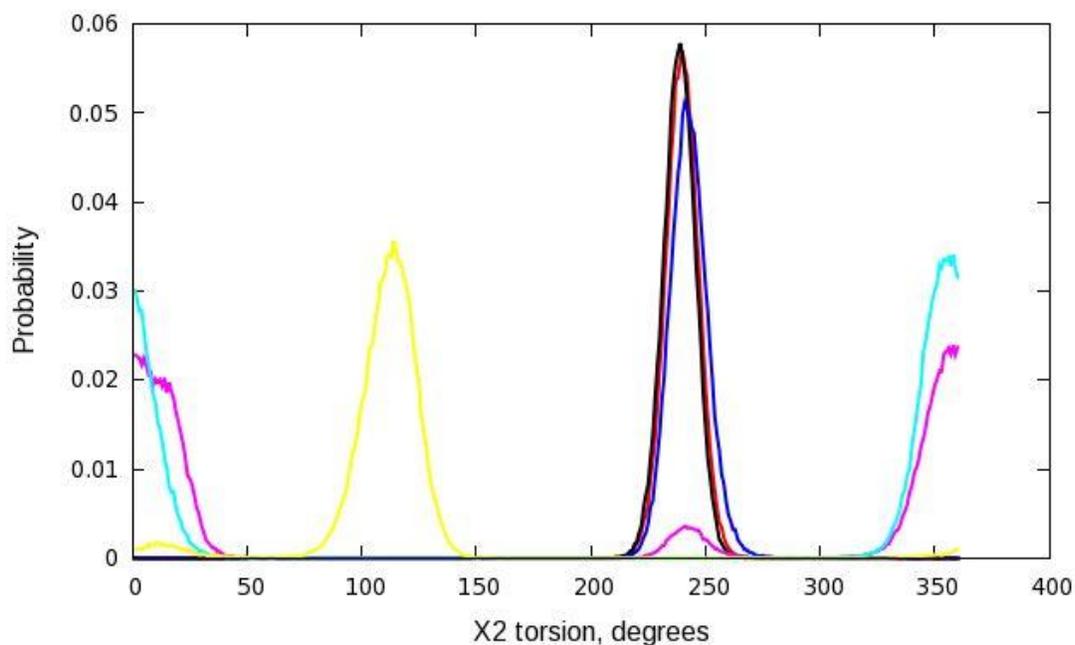


Figure 3-12. Glu-348 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue respectively); and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion. Note; χ_1 torsion for the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue respectively) have the same dihedral angle value.



A



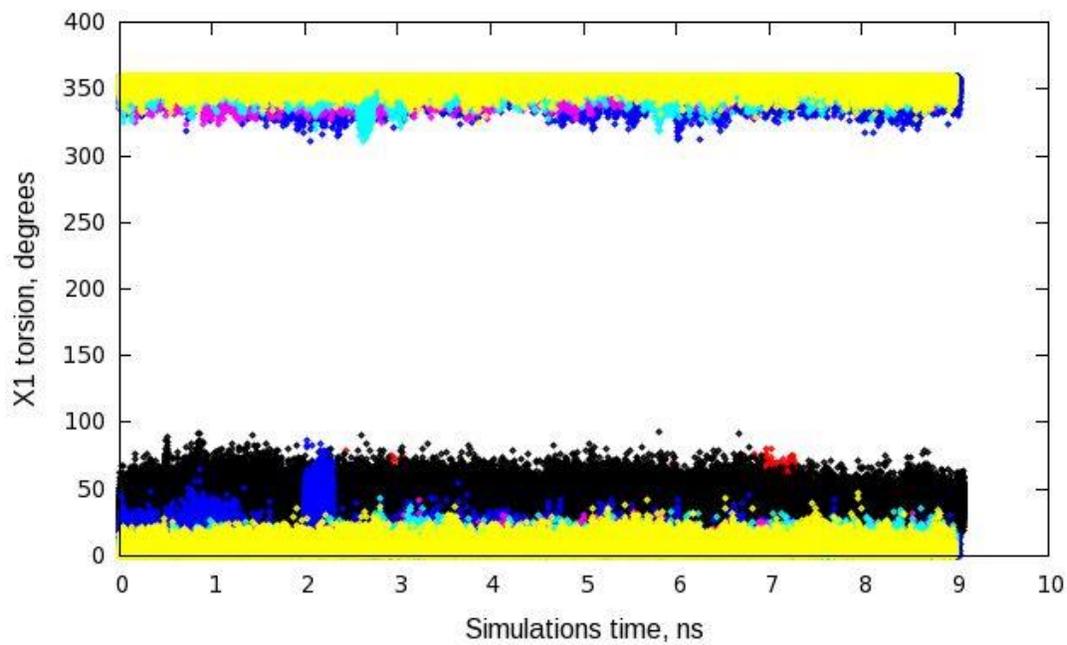
B

Figure 3-13. Glu-348 χ_2 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively); and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_2 torsion. B) Probability distribution of χ_2 torsion. Note; χ_1 torsion for the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue respectively) have the same dihedral angle value.

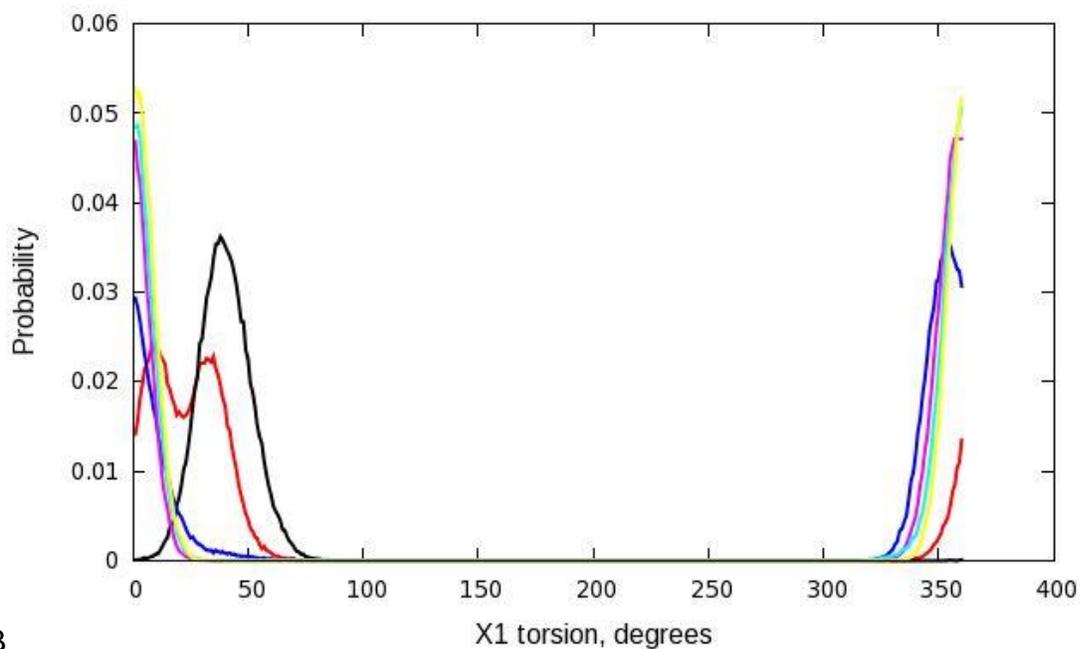
coordinating the glutaminase and synthetase active sites by acting as a general base to facilitate the attack of ammonia on the β AspAMP intermediate. Therefore, it is proposed that mutating Glu-348 to aspartate changes the length of the side chain and leads to a conformation that is not proper for ammonia translocation and proper positioning of the intermediate in the active site.¹⁰⁴ Based on these simulations, Glu-348 adopts a specific conformation in the ASNS/ β AspAMP simulations, and the side chain carboxylate forms a stable hydrogen bond with the amino group of the β AspAMP intermediate. This explains the role of Glu-348 in stabilizing the β AspAMP intermediate in the synthetase active site. The coordination of the Glu-348 to the two active sites could not be understood by these simulations. Further MD simulations are needed with glutamate bound in the glutaminase active site, in addition to β AspAMP and ammonia molecule present in the synthetase active site. Glu-348 in AS-B corresponds to Tyr-348 in the evolutionally related and structurally similar β -lactam synthetase enzyme (β -LS) (Figure 3-2). Both Glu-348 and Tyr-348 lie in the same positions in the active sites of AS-B and β -LS respectively. It has been shown that Tyr-348 together with Glu-382 in β -LS forms a catalytic dyad to deprotonate the secondary amine of the adenylated intermediate and catalyze intramolecular ring closure (Figure 3-3).¹⁰⁶

Glu-352

The Glu-352 χ_1 torsion adopts slightly different conformations in MD simulations 1 and 2 of the ASNS/ β AspAMP complex (Figures 3-14). The χ_2 torsion of the ASNS/ β AspAMP complex populates the same dihedral as the χ_2 torsion of the unliganded enzyme at 290° and also populates another dihedral angle at about 50° (Figures 3-15). The χ_2 torsion of MD1 of ASNS/ β AspAMP complex adopts only one conformation at about 50°.



A



B

Figure 3-14. Glu-352 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively); and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.

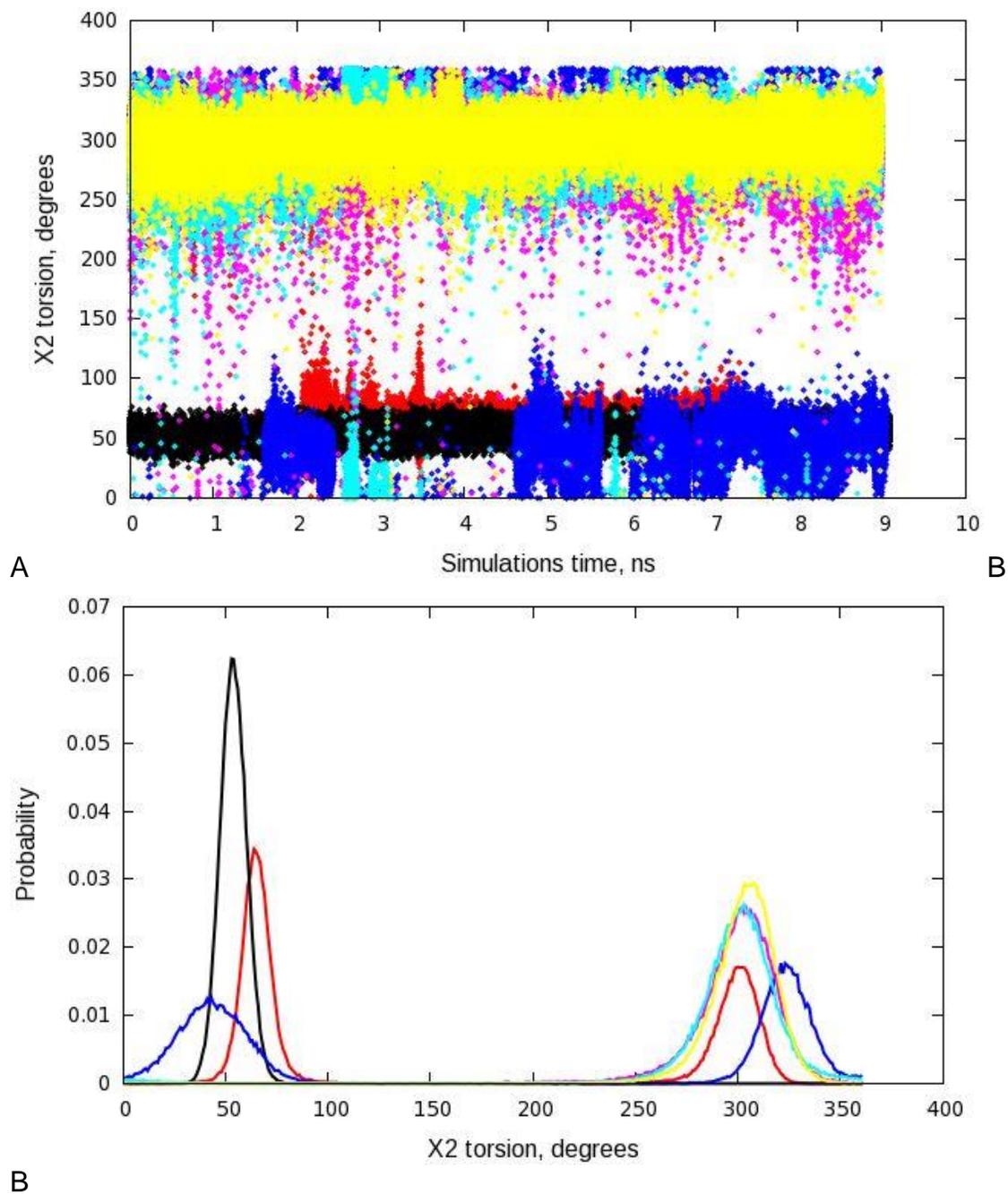


Figure 3-15. Glu-352 χ_2 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_2 torsion. B) Probability distribution of χ_2 torsion.

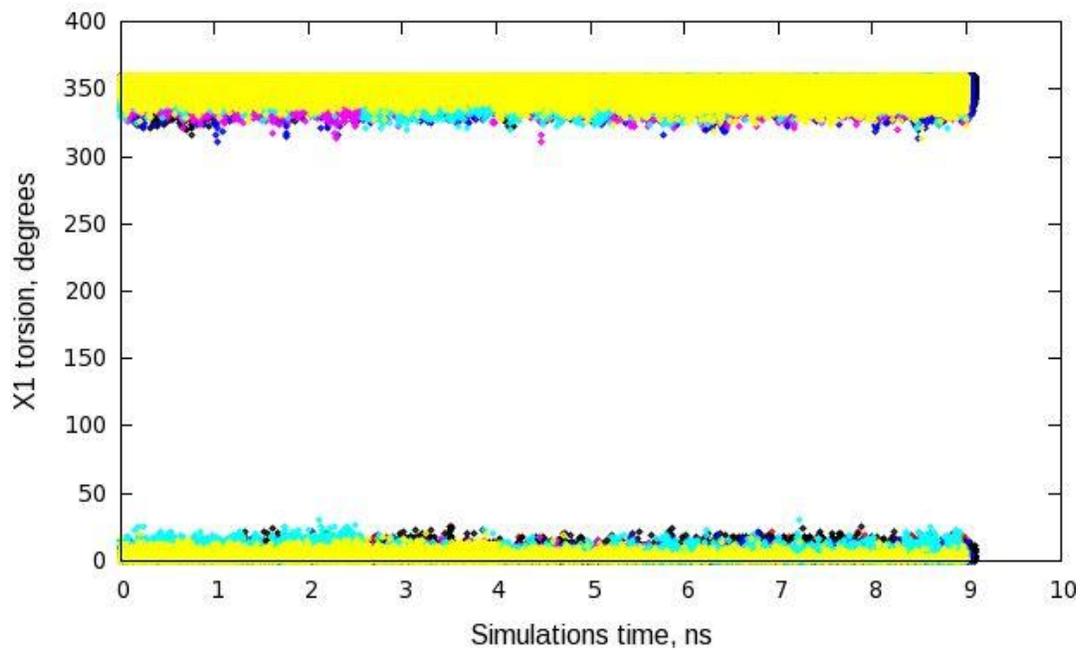
The side chain carboxylate of Glu-352 in the unliganded ASNS forms hydrogen bonds to its backbone atoms and to the side chains of Lys-376 and Tyr-357. The carboxylate group of Asp-352 preserves the same hydrogen bonding interactions but with less occupancy. It is also involved in a key hydrogen bonding interaction with the amino group in β AspAMP.

Asp-384

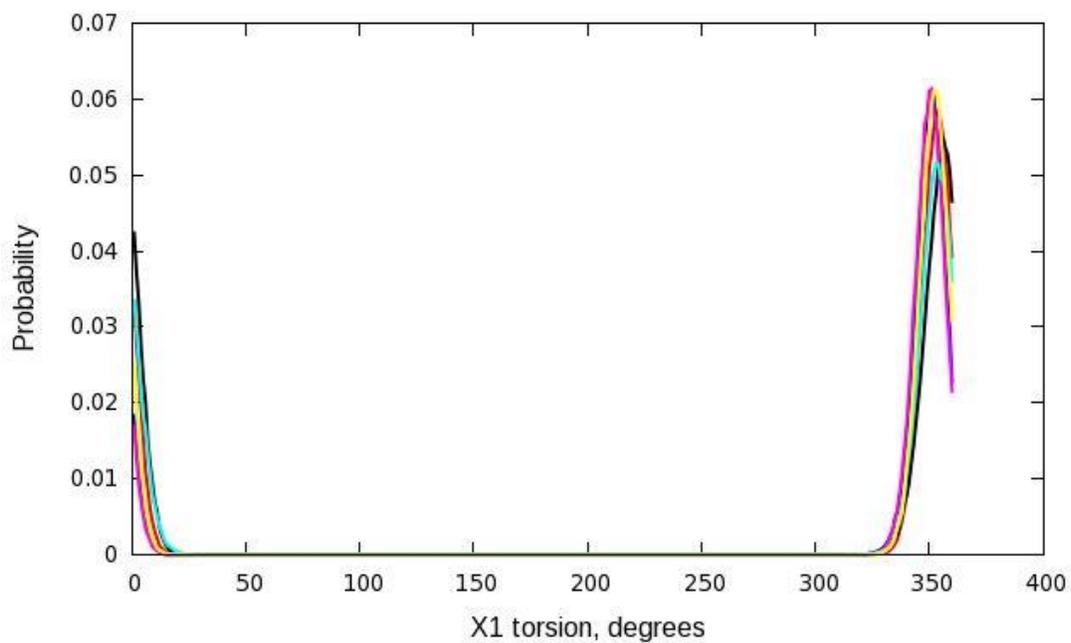
Asp-384 χ_1 torsion did not change its conformation in the presence of the intermediate. However, a small conformational change is observed for χ_2 torsion for the ASNS/ β AspAMP complex with the dihedral angle primarily populated at 70° . This dihedral angle is also sampled by the unliganded enzyme. The χ_2 torsion for Asp-384 of the unliganded enzyme is primarily populated at about 125° (Figures 3-16 and 3-17). This minor conformational change is due to presence of electrostatic interactions of both sides of the Asp-384 carboxylate with the amino group side chain of Lys376 and the guanidine group of Arg-387, thereby clamping Asp-384 in place and ready for interaction with the substrate.

Lys-376

Two close conformations for the χ_1 torsion are sampled by both the ASNS/ β AspAMP complex and ligand free ASNS. In the case of the χ_2 torsion, a dihedral angle of about 250° is primarily populated in ASNS/ β AspAMP simulations 1 and 3. However, MD2 of the ASNS/ β AspAMP complex, together with simulations 1 and 3 of the ligand-free ASNS, populates χ_2 torsion are populated at about 120° (Figures 3-18 and 3-19). Despite the minor change in populations of χ_2 torsion, the hydrogen bonding interactions between Lys-376 and the carboxylate groups of both Glu-352 and Asp-384 are retained. Together they form a network of hydrogen bonding interactions holding

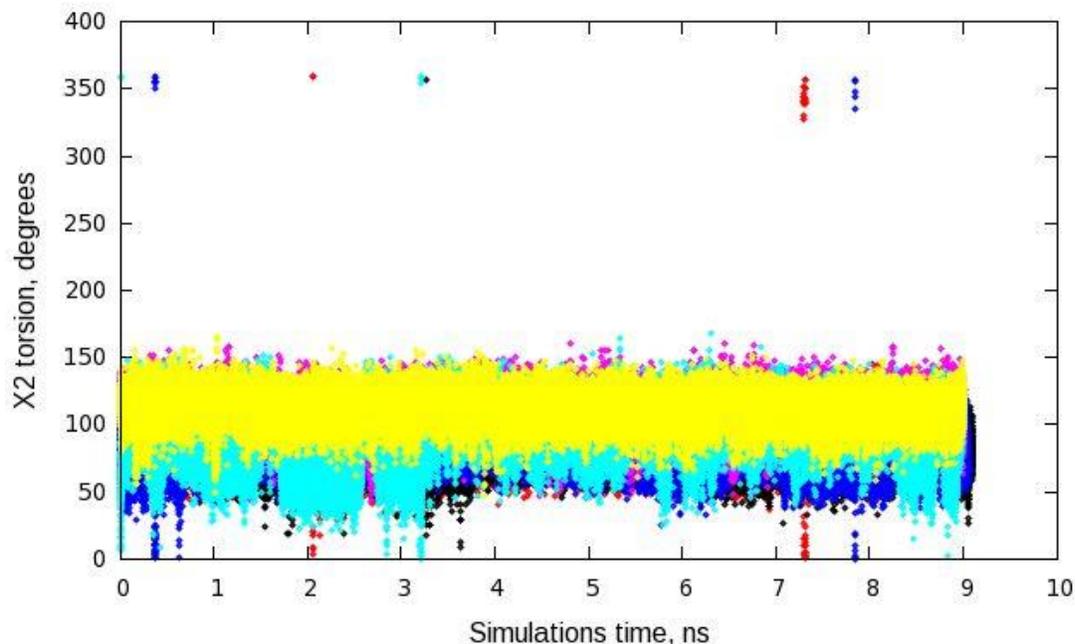


A

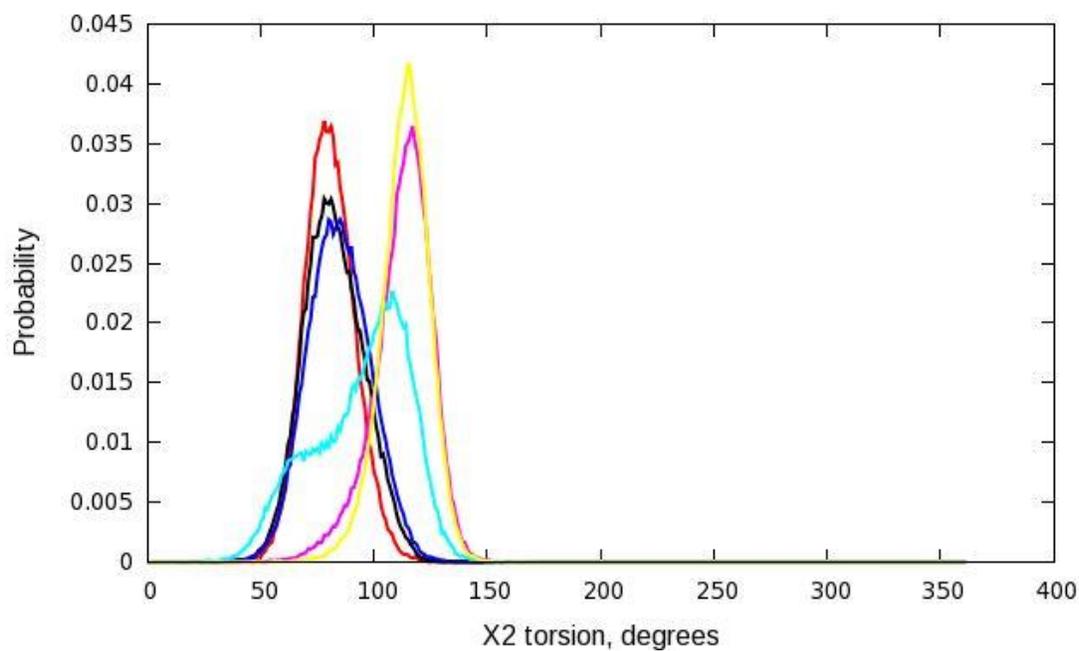


B

Figure 3-16. Asp-384 χ_1 torsion change during MD simulations of the ASNS- β ASPAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.

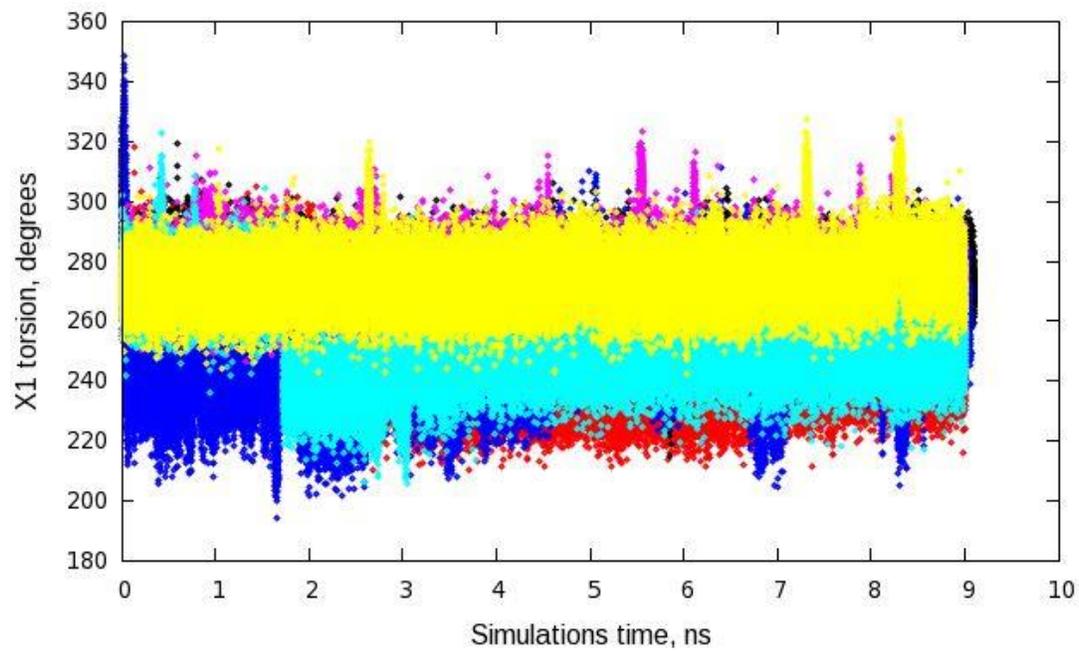


A

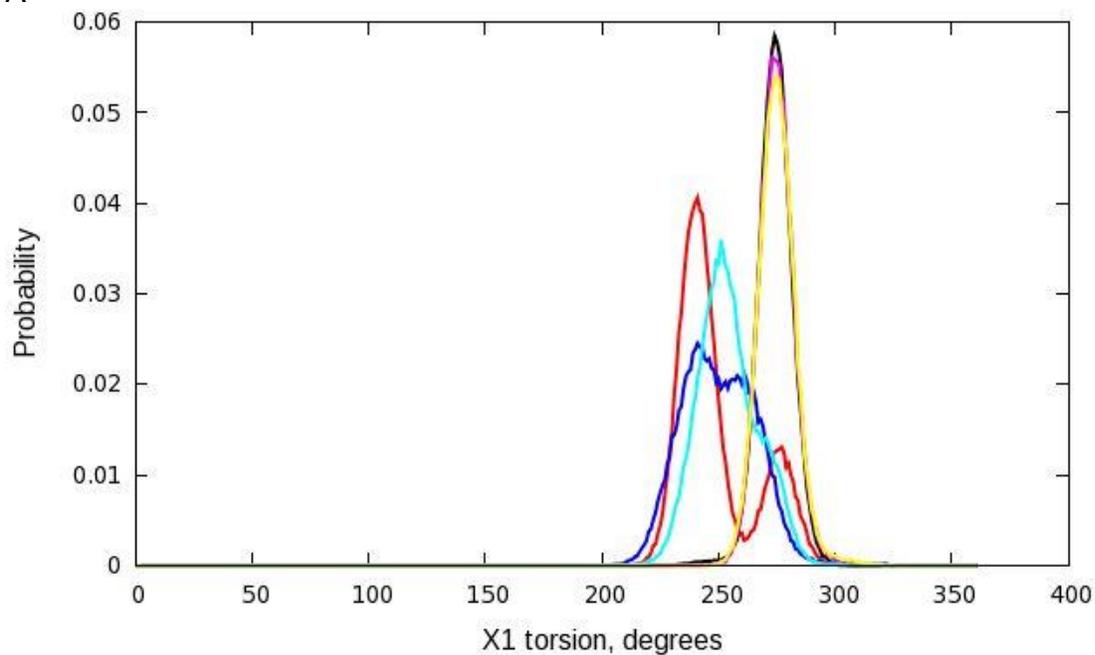


B

Figure 3-17. Asp-384 χ_2 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_2 torsion. B) Probability distribution of χ_2 torsion.

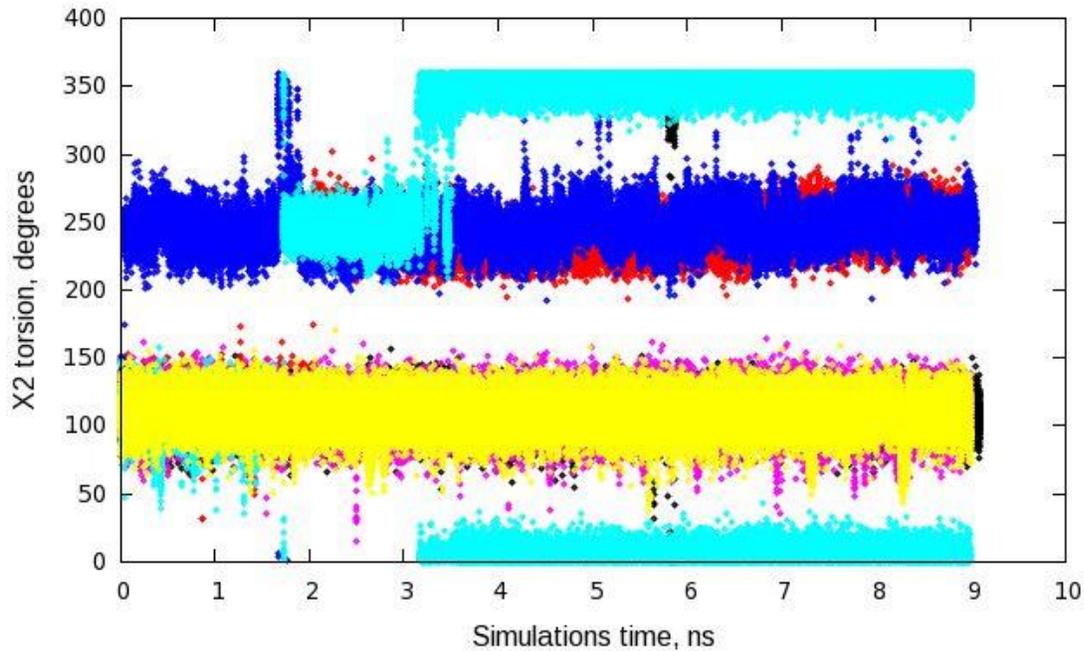


A

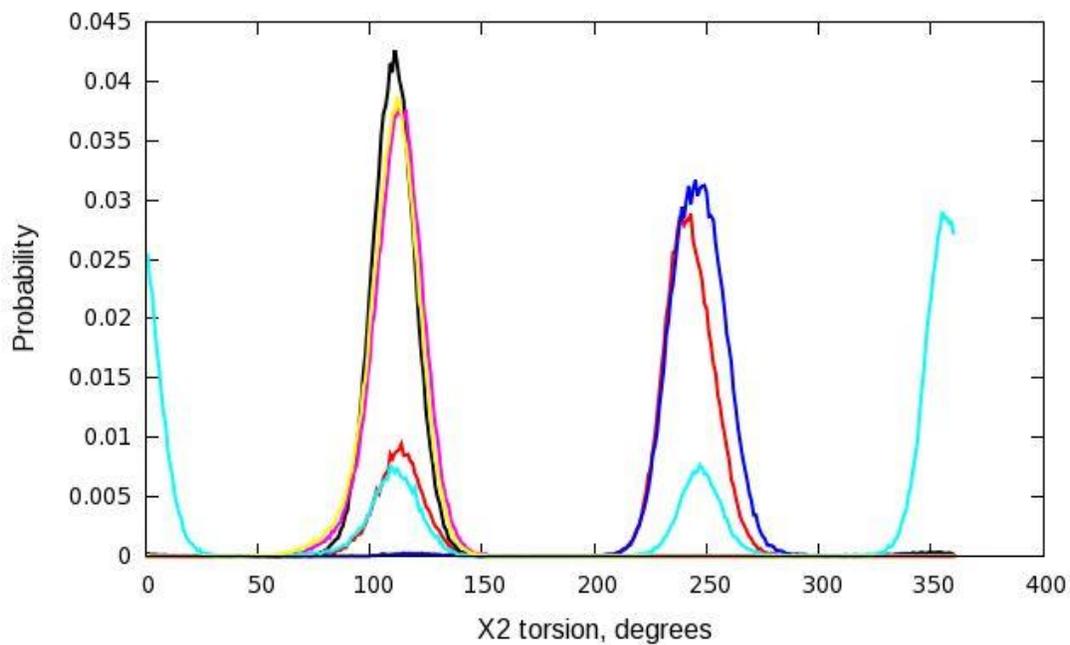


B

Figure 3-18. Lys 376 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.



A



B

Figure 3-19. Lys-376 χ_2 torsion change during MD simulations in ASNS- β AspAMP complex simulations; 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_2 torsion. B) Probability distribution of χ_2 torsion

them in a position set to interact with the intermediate β AspAMP, with Lys-376 interacting with the carboxylate group of the intermediate and the carboxylate groups of both Glu-352 and Asp-384 interacting with the amino group.

Lys-429

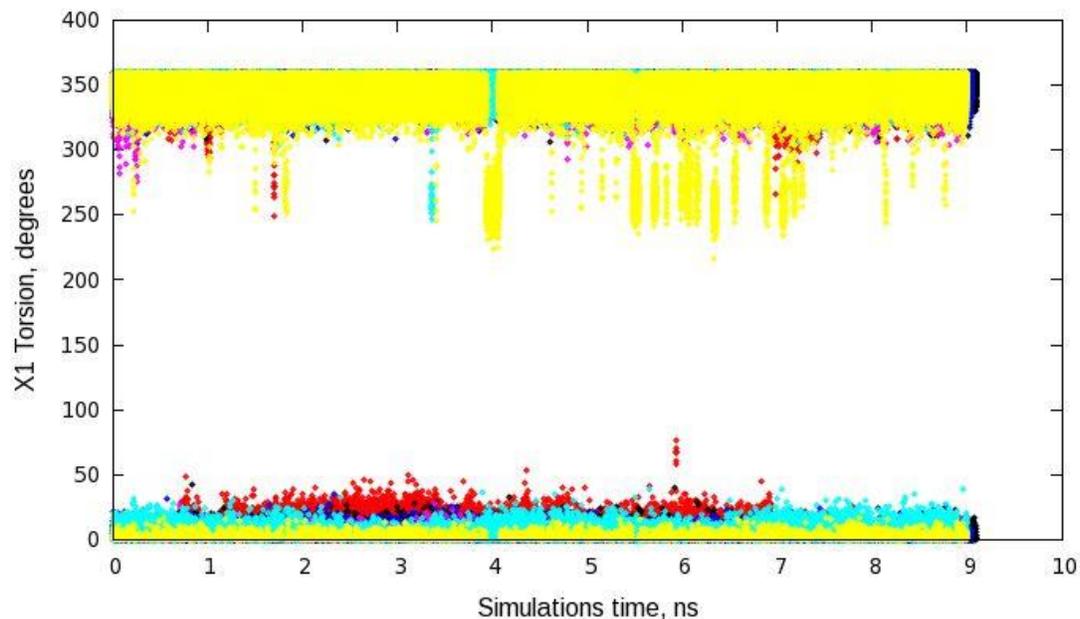
The χ_1 torsion did not change its conformation due to presence of the intermediate. The χ_2 torsion of Lys-429 in ASNS/ β AspAMP complex is populated primarily at 360° , while it exhibits different conformations in the unliganded ASNS simulations, indicating that it is sampling more conformations (Figures 3-20 and 3-21). Lys-429 shows the same hydrogen bonding interactions with neighboring residues in both ligand free ASNS and ASNS/ β AspAMP complex. The side chain of Lys-429 undergoes hydrogen bonding interactions with Asp-238, Asp-351 and inorganic pyrophosphate (PPi).

Ser-346

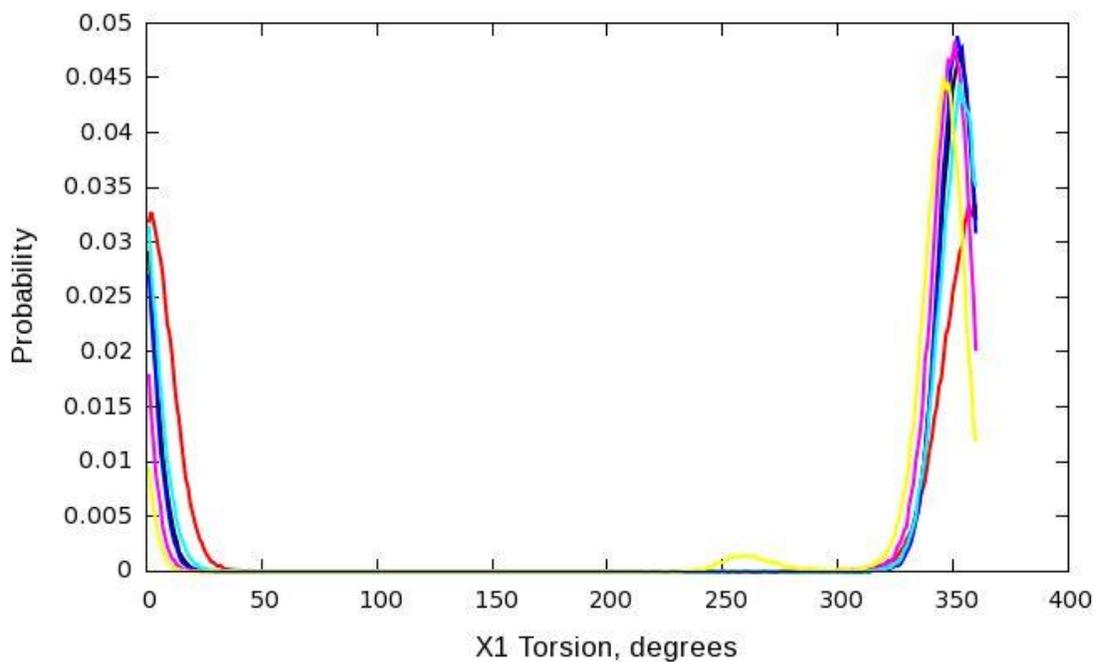
Ser-346 side chain hydroxyl group displays a hydrogen bonding interaction with the adenosine of β AspAMP, and it also interacts by hydrogen bonding with the backbone of Glu-348 and Arg-400 in the ASNS/ β AspAMP/PPi/Mg²⁺ and ligand-free ASNS simulations. No conformational change was observed for the χ_1 torsion of Ser-346 (Figure 3-22).

Ser-234 and Ser-239

Both Ser-234 and Ser-239 occur on the PP-motif (Figure 3-11). Both residues form hydrogen bonding interactions with pyrophosphate (PPi) thereby stabilizing PPi within the PP-motif in the ASNS/ β AspAMP/PPi/Mg²⁺ simulations. Ser-234 side chain forms also hydrogen bonding interaction with Gly-236 and Ser-234. Ser-239 side chain displays hydrogen bonding interactions with Gly-235, Gly-236 and Ser-234.

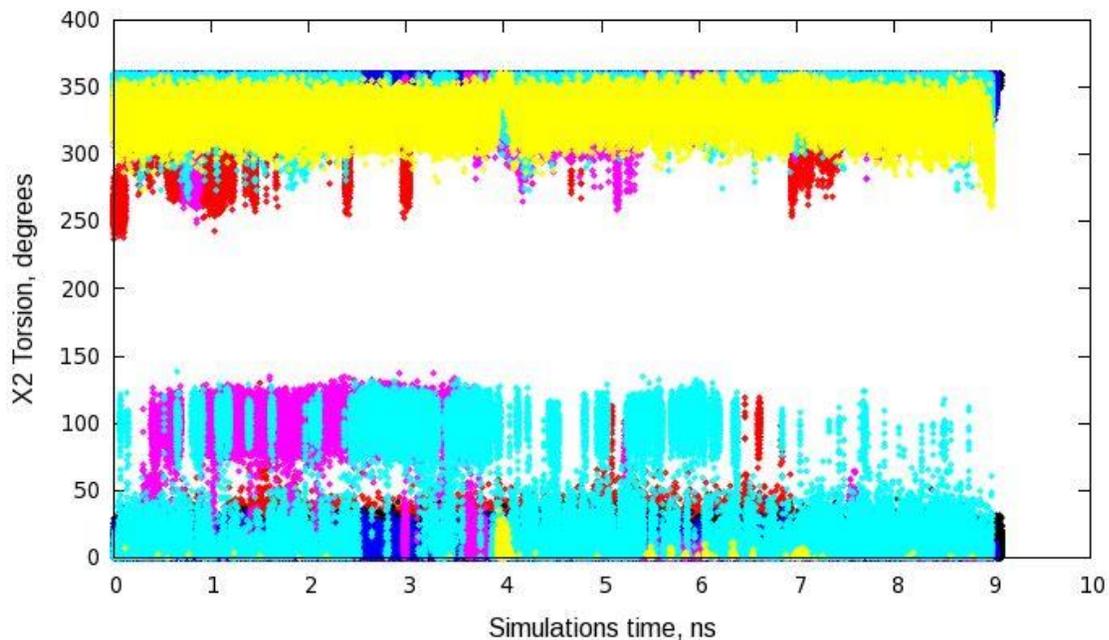


A

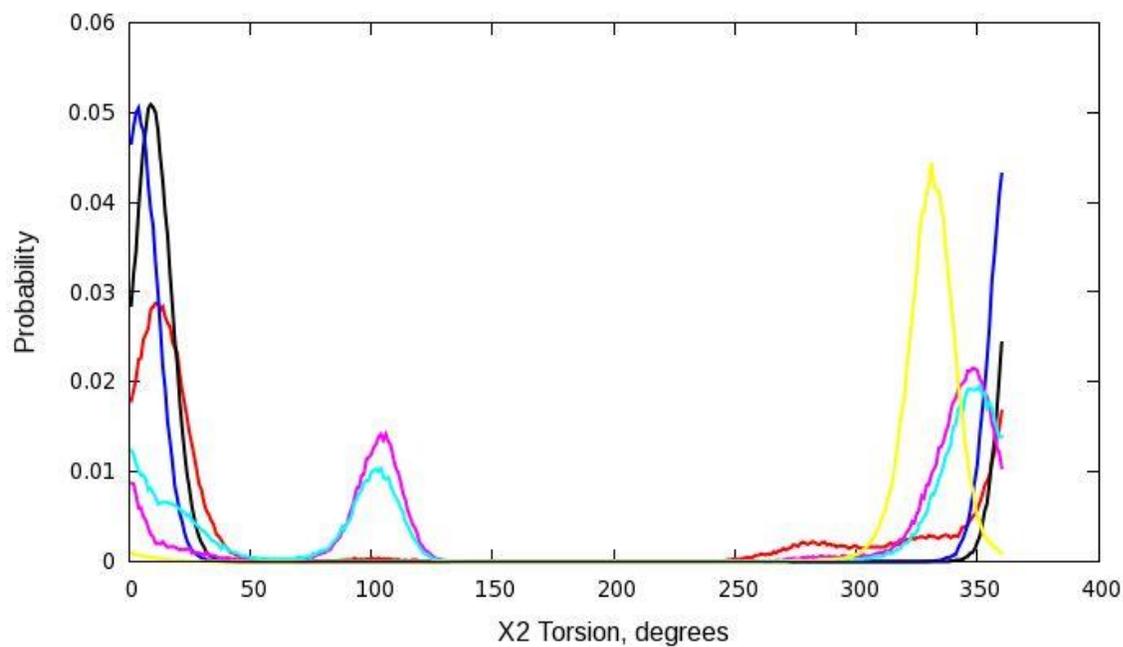


B

Figure 3-20. Lys-429 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations; 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.

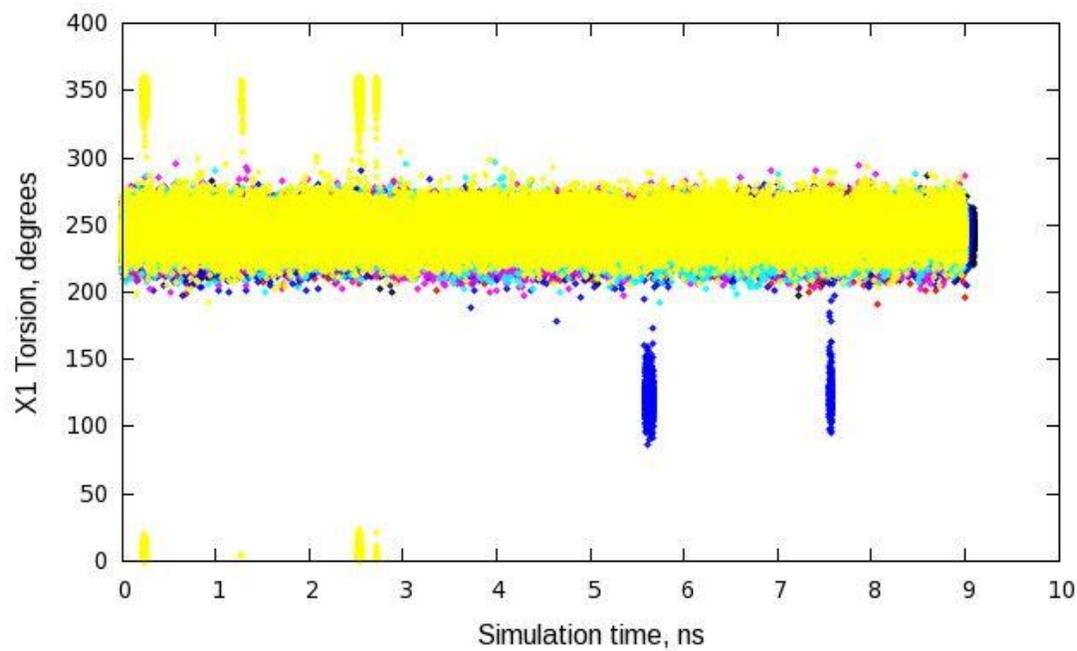


A

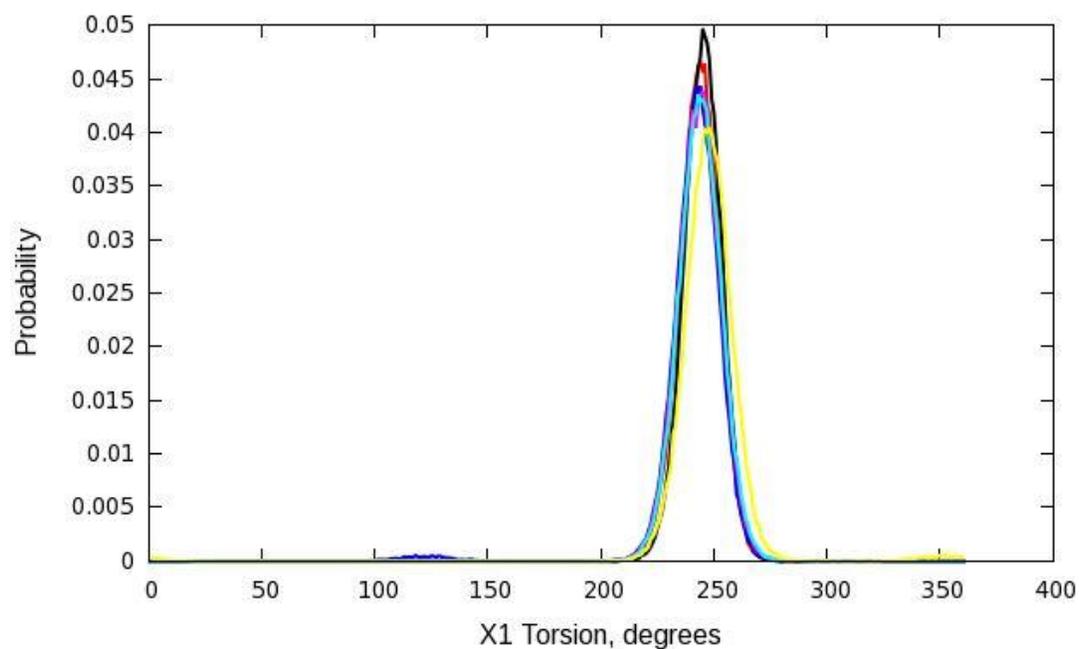


B

Figure 3-21. Lys-429 χ_2 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_2 torsion. B) Probability distribution of χ_2 torsion.

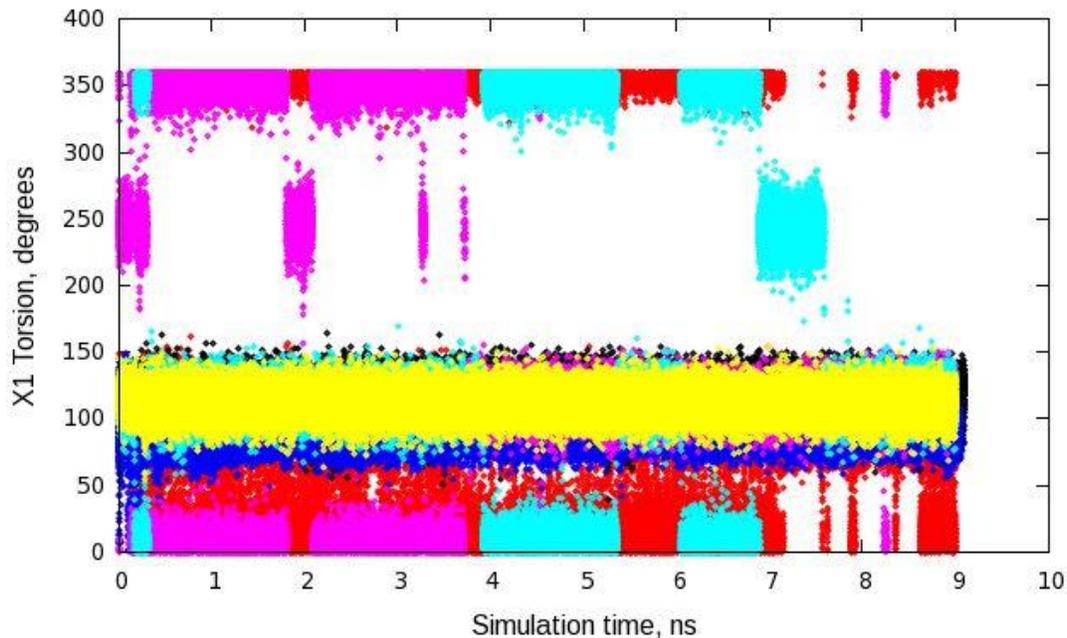


A

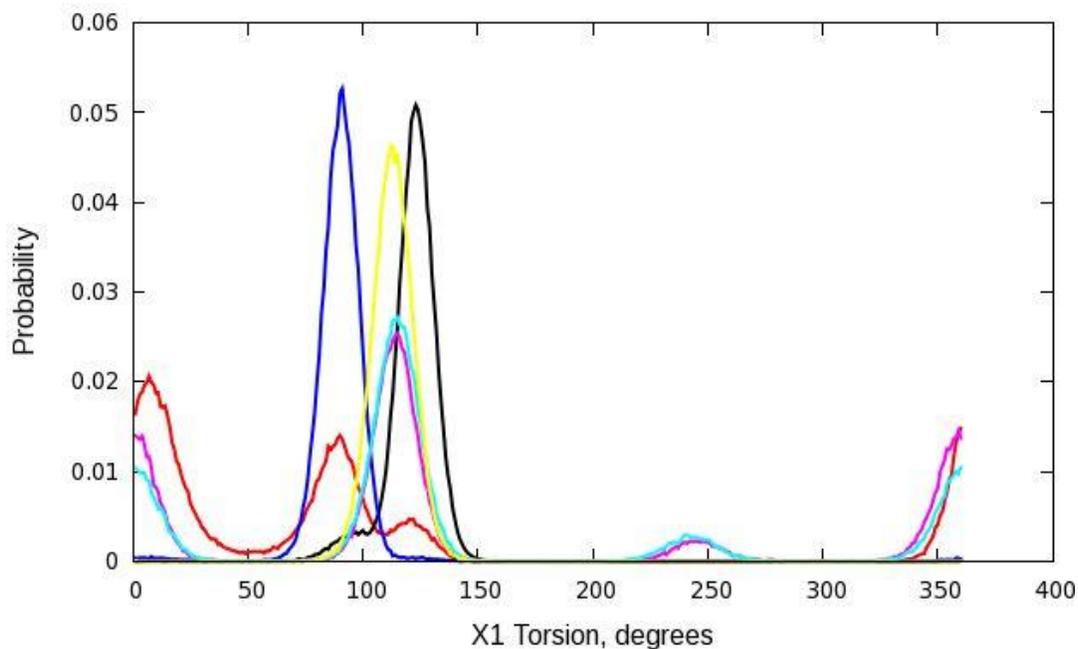


B

Figure 3-22. Ser346 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion

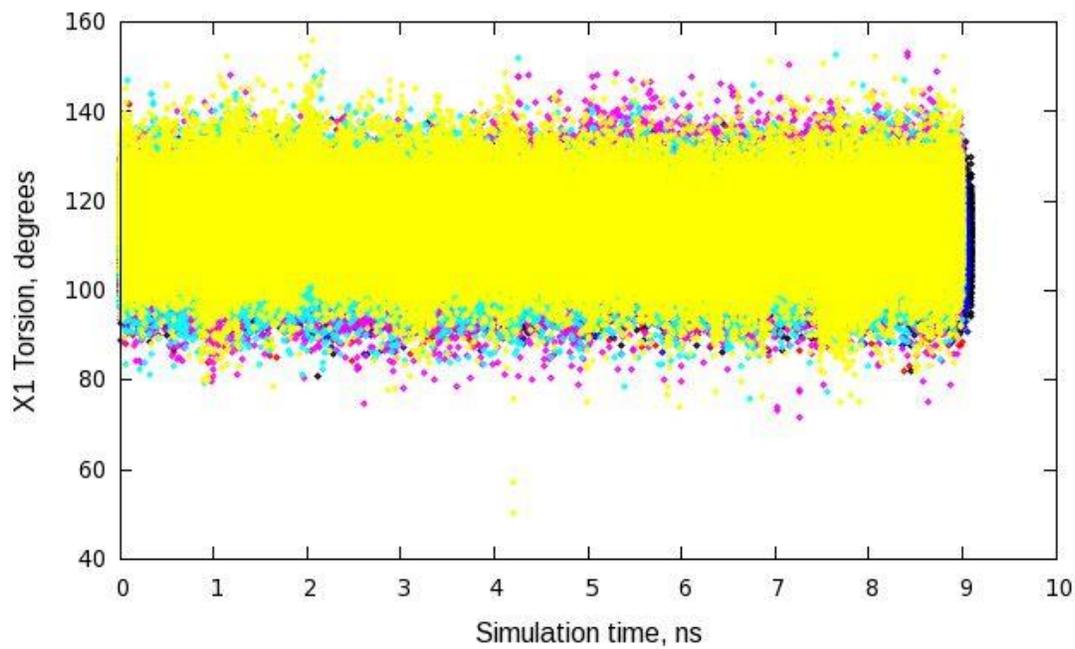


A

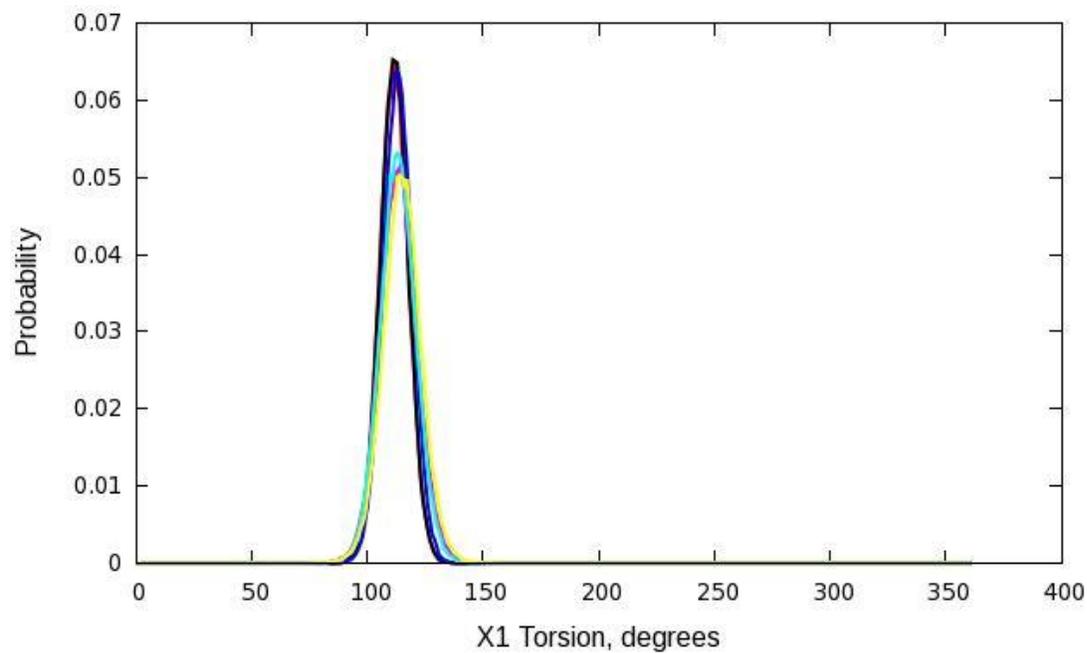


B

Figure 3-23. Ser234 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.



A



B

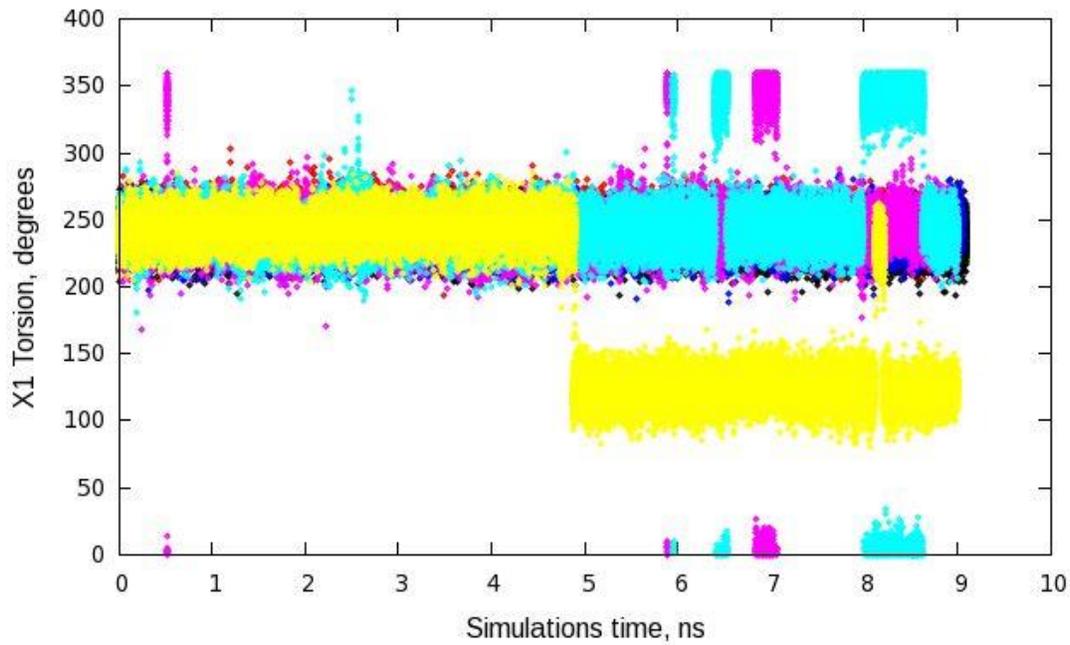
Figure 3-24. Ser239 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.

No distinct χ_1 conformational change for Ser-234 and Ser-239 in the absence and presence of intermediate in the ASNS active site (Figures 3-23 and 3-24).

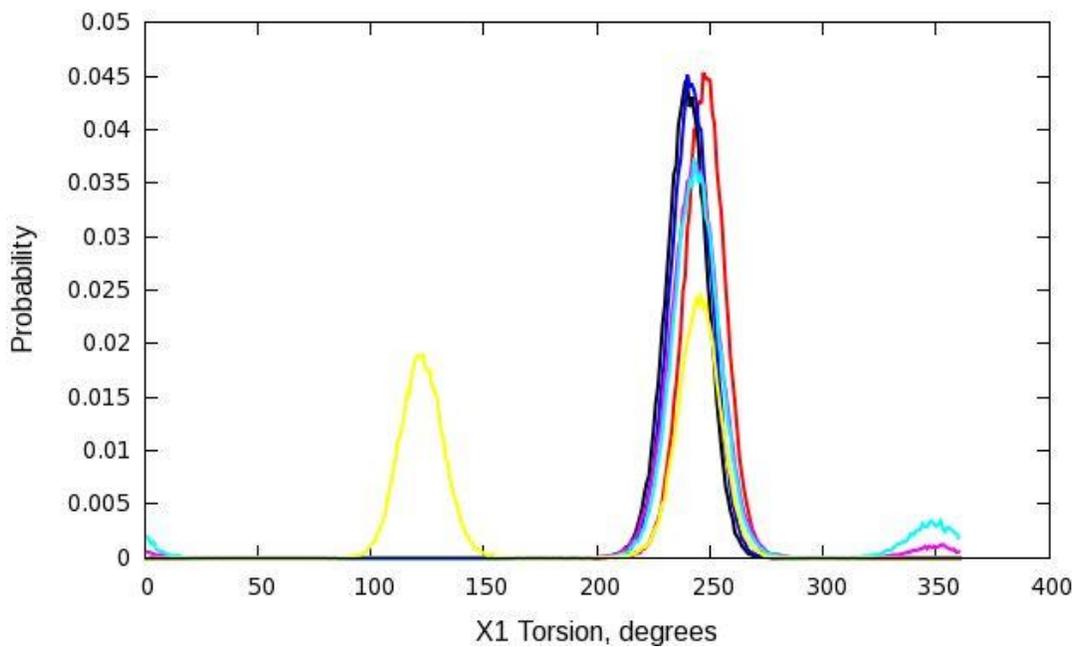
Lys-449

Lys-449 is conserved among all asparagine synthetase enzymes. Investigation of the χ_1 , χ_2 , χ_3 , χ_4 torsions of Lys-449 (Figures 3-25 - 3-28) shows that the χ_4 torsion in two independent simulations of the ASNS/ β AspAMP/PPi/Mg²⁺ complex is primarily populated at 250°. The χ_4 torsion of Lys-449 in the unliganded ASNS samples different conformations, indicating its ability to move freely. Lys-449 makes direct hydrogen bonding interactions with the inorganic pyrophosphate (PPi) and with phosphate oxygen of the β AspAMP intermediate through bridging water molecules (Tables 3-1 and 3-2). The side chain amino group is also involved in hydrogen bonding interactions with the side chain carboxylate of Asp-351. In the ligand-free ASNS, a hydrogen bonding interaction is observed between the side chain amino group of Lys-449 and the carboxylate side chain of Asp-279 (Table 3-4).

Mutation of Lys-449 in to arginine or alanine resulted in complete loss of synthetase activity, indicating the importance of Lys-449 for enzyme catalysis.⁹² Previous studies on a set of sulfonamide inhibitors have shown that the resonance stabilized negative charge of the sulfonamide group mimics the negative charge on the intermediate and is critical for inhibitor binding (Figures 1-4 and 1-5 in Chapter 1). It was hypothesized that the negatively charged phosphate group of β AspAMP is involved in a critical electrostatic interaction with Lys-449. However, according to MD simulation results, Lys-449 forms stable hydrogen bonding interactions primarily with PPi and the carboxylate side chain of Asp-351 in the ASNS/ β AspAMP complex.

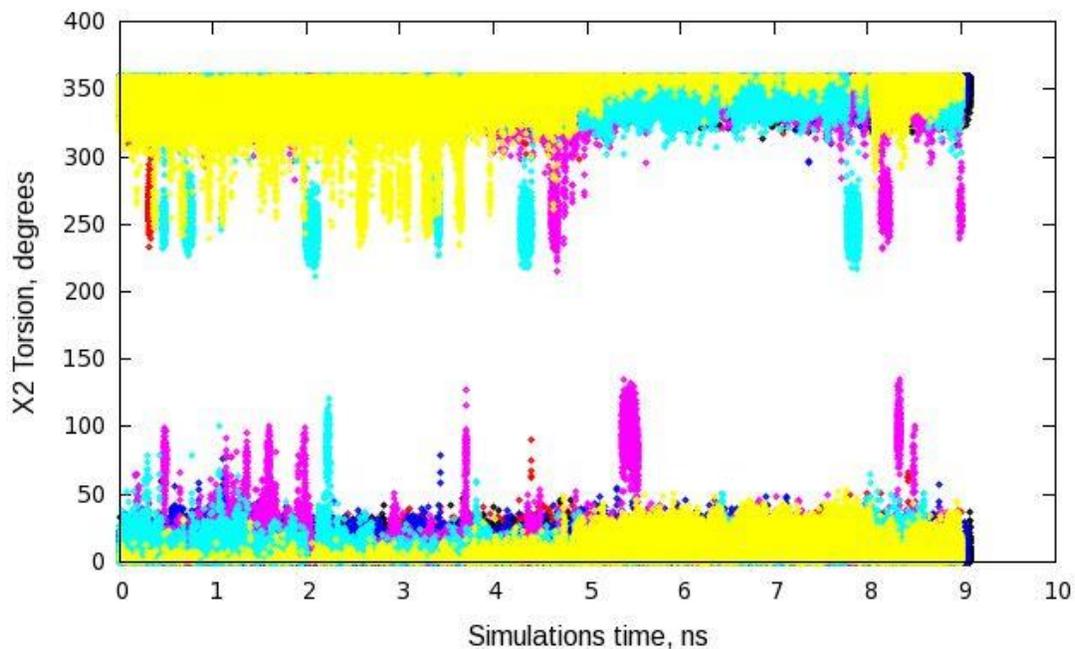


A

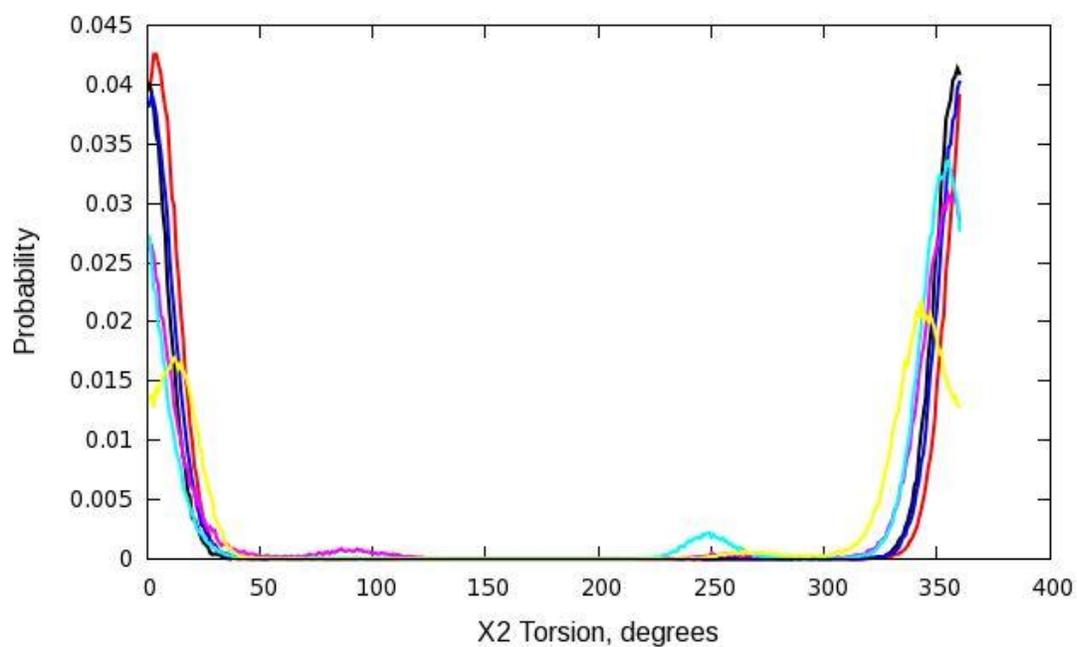


B

Figure 3-25. Lys-449 χ_1 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_1 torsion. B) Probability distribution of χ_1 torsion.

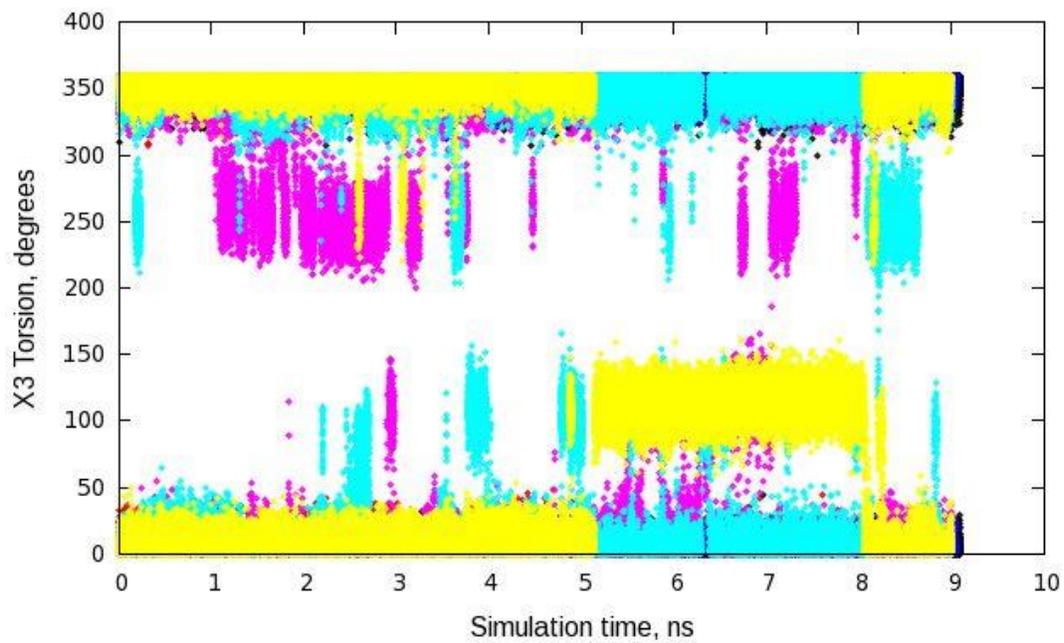


A

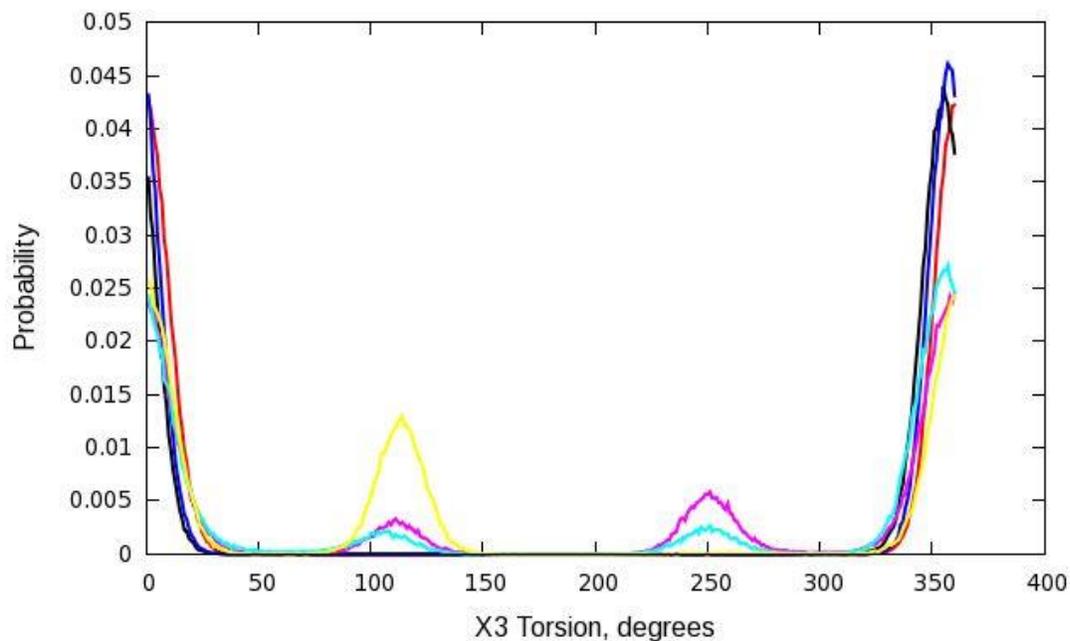


B

Figure 3-26. Lys-449 χ_2 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_2 torsion. B) Probability distribution of χ_2 torsion

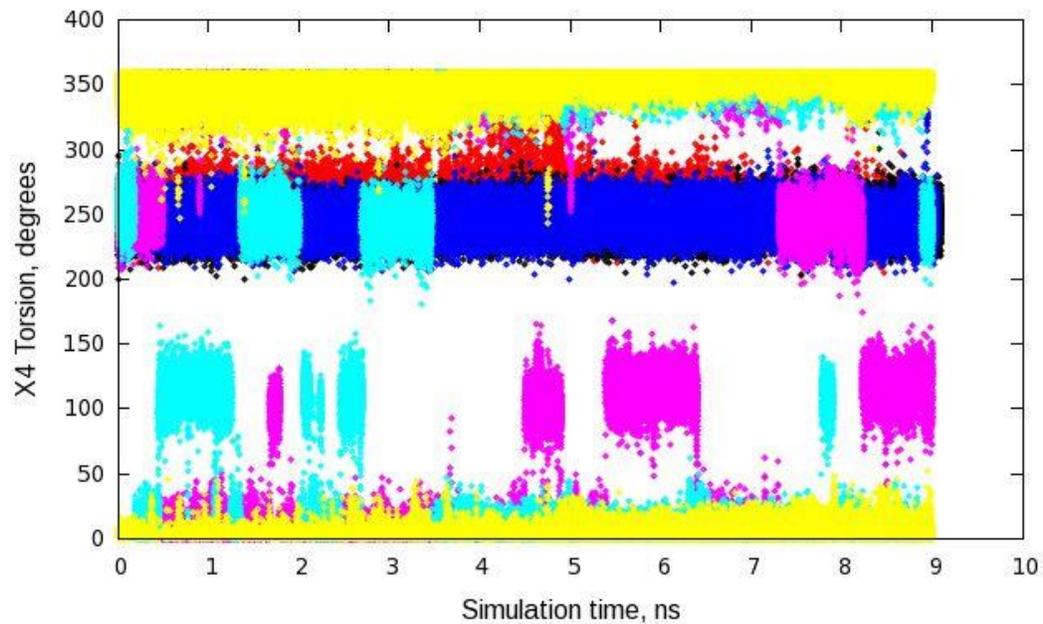


A

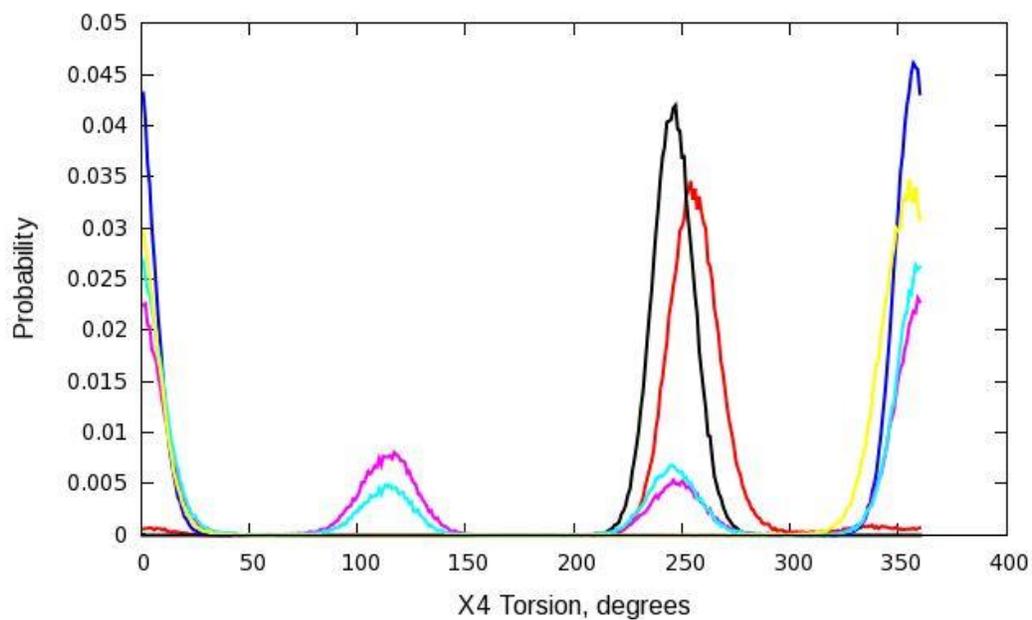


B

Figure 3-27. Lys-449 χ_3 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_3 torsion. B) Probability distribution of χ_3 torsion.



A



B

Figure 3-28. Lys-449 χ_4 torsion change during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Time series of χ_4 torsion. B) Probability distribution of χ_4 torsion

In addition, Lys-449 occurs on the loop 443-452, which is observed to have reduced fluctuations when the intermediate is bound in two simulations of the ASNS/ β AspAMP/PPi/Mg²⁺ complex and is observed to close upon ligand binding (Figure 3-29). Therefore, the catalytic importance of Lys-449 may be due to its role in retaining key hydrogen bonding interactions in the ASNS active site, in addition to its mediating the loop conformational change as discussed below.

Lys-449 corresponds to Lys-443 in β -LS, and it occurs on a similar loop comprising residues 444-453 in β -LS (Figure 3-2). This loop closes tightly on the adenylated intermediate in β -LS, and it has been shown that the conformational change of this loop is partially rate limiting.^{42,107} In addition, Lys-443 is observed to form hydrogen bonding interactions with the carbonyl of the intermediate CMA and the product DGPC (Figure 3-30).⁴² It was later shown by kinetic, mutagenesis, and density functional theory (DFT) studies that Lys-449 stabilizes the transition state of β -lactam formation.¹⁰⁸ However, this interaction is not observed in the ASNS/ β AspAMP simulated model, because the β AspAMP intermediate is shorter with its carboxyl group hydrogen bonding with Lys-376 and repelling the carbonyl from Lys-449 (Figure 3-10). Lys-376 is not present in the active site of β -LS. The absence of Lys-376 in β -LS and the additional electrostatic interaction of Lys-443 with the carbonyl of the CMA intermediate may be part of the evolutionary changes in the β -LS active site to accommodate the longer substrate (CEA) and formation of the β -lactam (Figure 3-2).

Loop 443-452 Mobility

Loop 443-452 is observed in two independent simulations to encompass reduced fluctuations when the intermediate is bound (Figure 3-8). This loop carries the catalytically important residue Lys-449. Movement of the loop extending from residue

443 to residue 452 was monitored by measuring the distance between the side chain nitrogen of Lys-449 and the side chain carboxylate of Asp-351 over all simulation trajectories (Figure 3-31). The distance in two simulations of the ASNS/ β AspAMP/PPi/Mg²⁺ system remains at 3 Å, where Lys-449 is clamped in position by hydrogen bonding interactions with the side chain carboxylate of Asp-351 and inorganic pyrophosphate oxygen atoms. This state represents the closed form of the loop and prevents the hydrolysis of the intermediates. The distance becomes slightly longer, approximately 4.5 Å, in the third MD simulation of ASNS/ β AspAMP/PPi/Mg²⁺. In the ligand-free ASNS simulations, the loop is observed to open widely to about 10 Å, and Lys-449 flips down to make a hydrogen bond with the side chain carboxylate of Asp-279, thereby exposing the active site to more solvent. The loop is observed in the third independent MD simulation of the ligand-free ASNS to be semi-closed around 4.0 Å, because the hydrogen bonding between Lys-449 and Asp-351 is maintained. This explains the similar RMSF of this loop in ASNS/ β AspAMP and ligand-free ASNS. Most probably if this trajectory is allowed to run for a longer time, this loop will experience open states as in the other two independent unliganded ASNS simulations. The behavior of loop 443-452 explains previous mutagenesis studies where mutating Lys-449 to arginine retained the glutaminase activity of the enzyme while the synthetase activity was lost.⁹² Mutating Lys-449 in to alanine makes the enzyme lose both its glutaminase and synthetase activity. According to MD simulation results, removal of Lys-449 by mutation may lead to loss of important hydrogen bonding interactions with the side chain carboxylate of Asp-351 and PPi. This would lead to a widely open

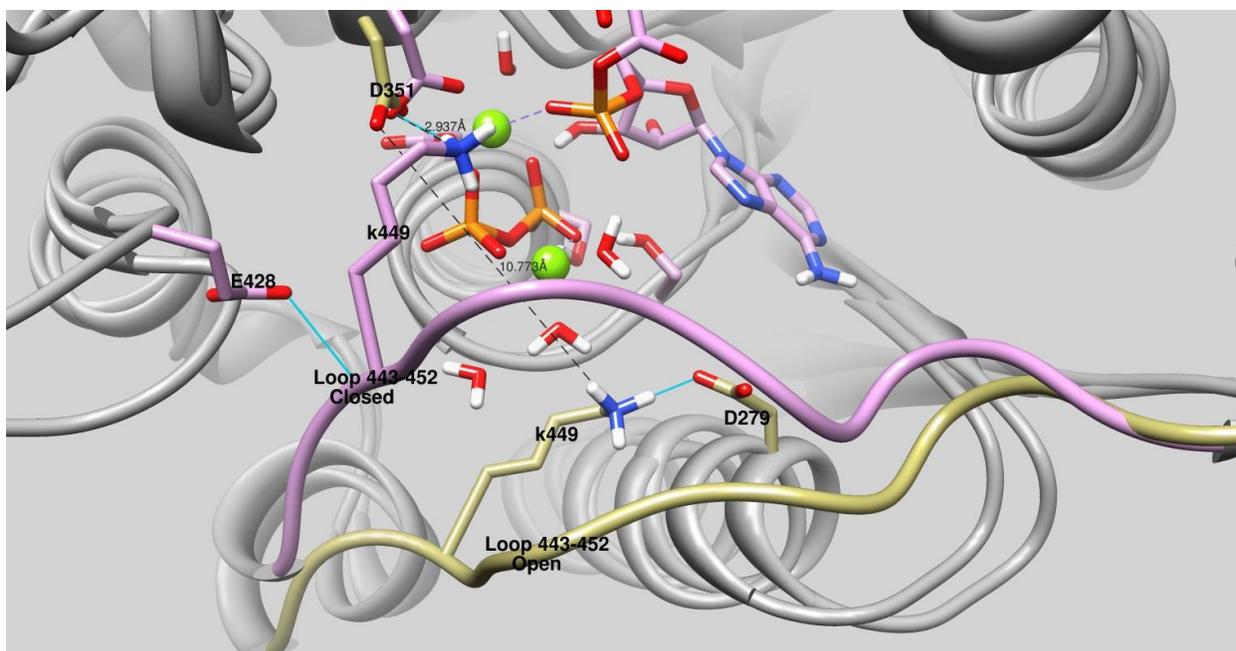


Figure 3-29. The closed (purple) and open (yellow) forms of loop 443-452 in ASNS.

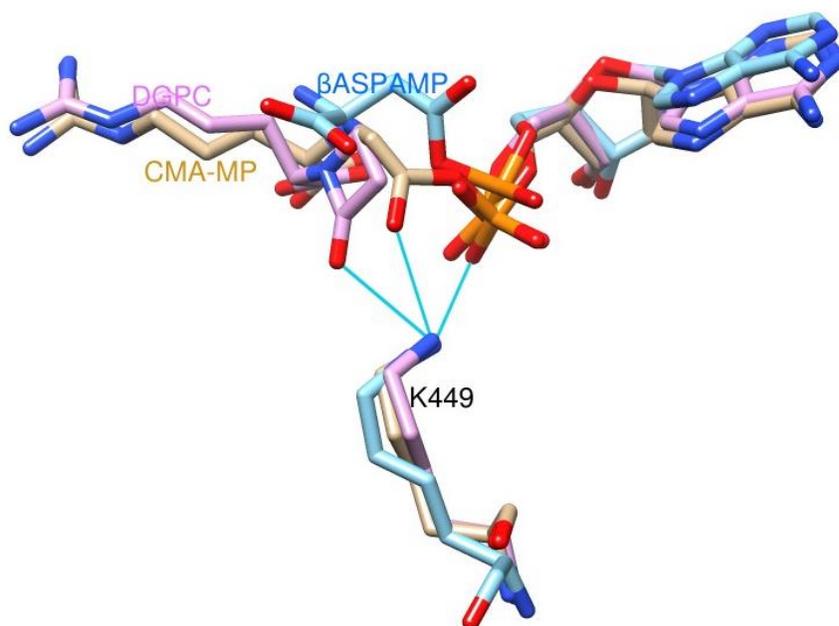
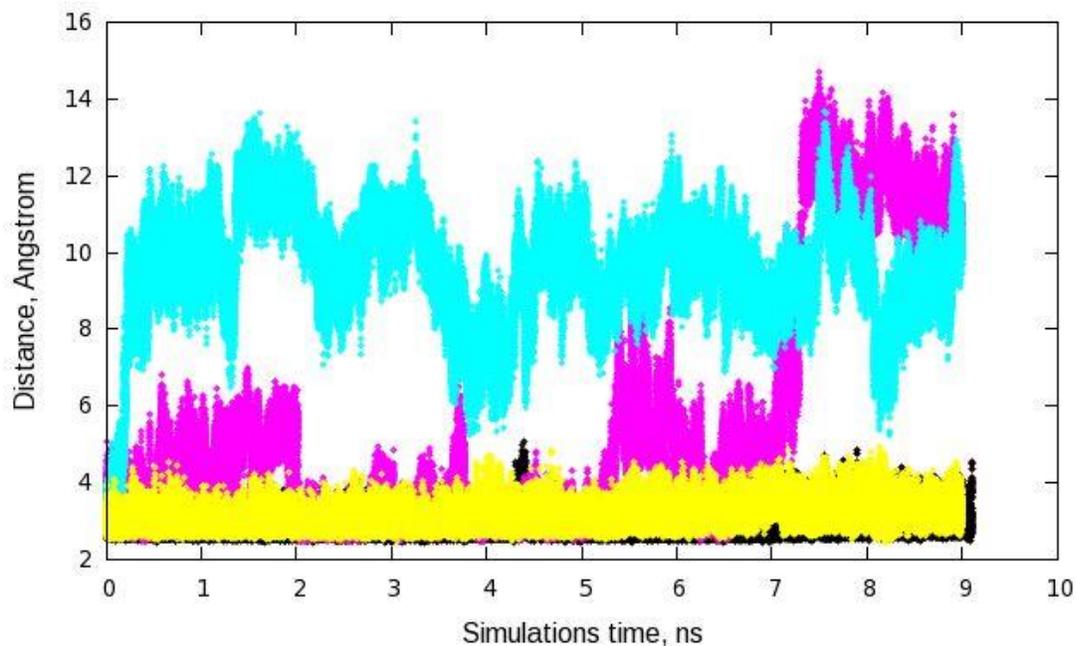
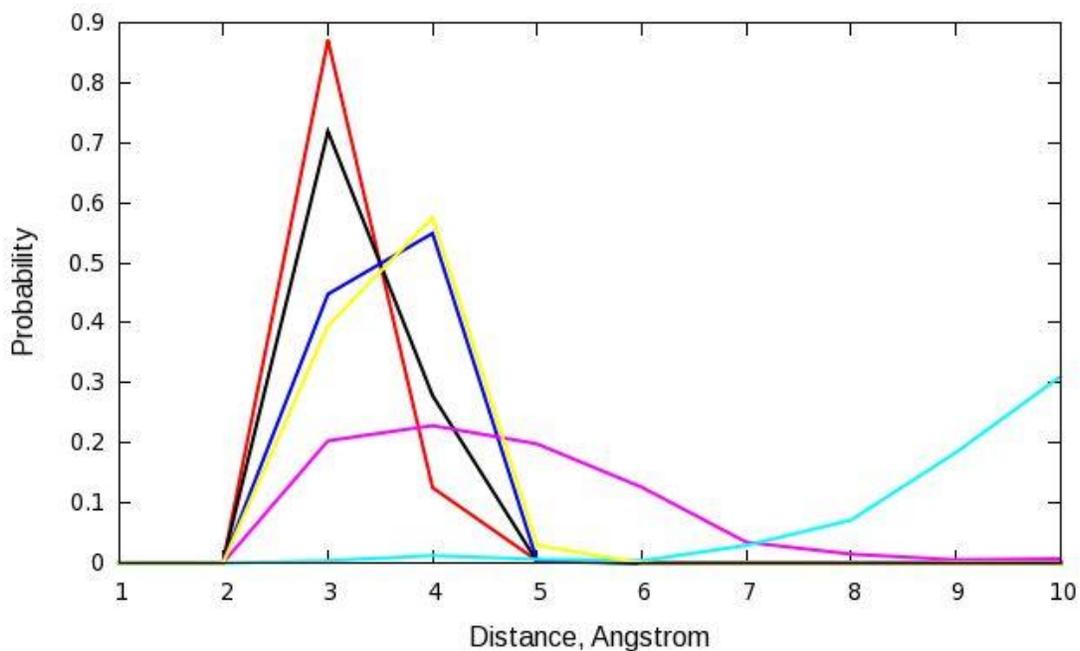


Figure 3-30. Structural overlay of the conserved lysine residue (K449 in AS-B, K443 in β -LS) and β AspAMP intermediate in AS-B (blue), CMA-AMP intermediate analog in β -LS (tan) and the product DGPC in β -LS (pink). Hydrogen bonds between K443 and the carbonyl and phosphate oxygens of the intermediate and product in β -LS are shown as blue lines. In contrast, K449 in ASNS does not hydrogen bond with the carbonyl and phosphate oxygen of the β AspAMP intermediate.



A



B

Figure 3-31. Distance between K449-N and D351-C α during MD simulations of the ASNS- β AspAMP complex; simulations 1, 2 and 3 (red, black and blue, respectively) and unliganded ASNS; simulations 1, 2 and 3 (purple, cyan and yellow, respectively). A) Distance time series. B) Probability distribution.

state of loop 443-452, with the active site becoming more exposed to solvent, possibly leading to hydrolysis of the intermediates and hence loss of activity. As mentioned earlier, a similar loop comprising residues 444-453 closes tightly on the adenylated intermediate in β -LS, and it has been shown that the conformational change of this loop is partially rate limiting.^{42,107}

Summary

Molecular dynamics simulations of the ASNS/ β AspAMP/PPi/Mg²⁺ complex and the unliganded ASNS have been performed. The MD simulation results revealed important interactions in the active site that mediate intermediate binding. The adenosine group binds in a well-defined ATP binding pocket, where the adenine undergoes hydrogen bonding interactions with the Val-272 backbone atoms. The adenosine hydroxyls forms hydrogen bonds with the side chain of Ser-346, the backbone of Gly-347 and inorganic pyrophosphate (PPi). The phosphate oxygen atoms experience critical electrostatic interactions by coordinating to the Mg²⁺ ions. The amino and carboxylate groups of the β AspAMP intermediate bind with the substrate (aspartate) binding residues, in which the amino group is stabilized by electrostatic interactions with the carboxylate side chains of Glu-348, Asp-384 and Glu-352. The carboxylate group of β AspAMP is stabilized by hydrogen bonding with Lys-376.

According to kinetic results, the absence of ATP/PPi exchange, and the very low PIX of Inorganic pyrophosphate (i.e. β ASPAMP intermediate formation is irreversible), PPi is proposed to be the last product released from the active site. Inorganic pyrophosphate lies in a deep binding pocket called the PP-motif and is stabilized by hydrogen bonds with Ser-234, Ser-239 and Leu-237 within the PP-motif as well as with Lys-429 and Lys-449. The Mg²⁺ ions exhibit approximately octahedral coordination in all

of the simulations, and are very important in balancing the negative charge of the PPI and facilitating asparagine formation and preventing reversible β AspAMP formation.

Based on structural similarity and sequence alignment, the β AspAMP intermediate, PPI and Mg^{2+} ions were modeled in the ASNS active site based on the β -LS/CMA-AMP/PPI/ Mg^{2+} crystal structure (1MBZ). Both enzymes catalyze intermediate formation by an adenylation reaction. AS-B catalyzes intermolecular amide bond formation by an amino group, in contrast to β -LS, which catalyzes intramolecular amide bond formation to form a β -lactam ring (Figure 3-3). Although the β -LS active site is elongated to accommodate the longer intermediate (Figure 3-2), the behavior of the simulated ASNS/ β AspAMP-/PPI/ Mg^{2+} model resembles that of the β -LS/CMA-AMP/PPI/ Mg^{2+} complex. The ATP binding residues, PP-motif, and Mg^{2+} coordination residues, in addition to the presence of a catalytic loop that carries a conserved lysine residue are similar in the active sites of both enzymes. However, MD simulations reveal changes in active site residues that bind to the substrate. For example, Asp-382 and Glu-348 in AS-B are replaced by Glu-382 and Tyr-348 in β -LS. Asp-382 and Glu-348 in AS-B bind strongly to the amino group of the β AspAMP intermediate. Also Glu-348 plays a role in coordinating the glutaminase and synthetase active sites, albeit weakly, by acting as a general base to facilitate the attack of ammonia on the β AspAMP intermediate. On the other hand, Glu-382 and Tyr-348 in β -LS are reported to act through acid/base catalysis to deprotonate the secondary amine of the intermediate and facilitate ring closure. Another difference revealed by the MD simulations is the behavior of a conserved lysine residue (Lys-449 in AS-B and Lys-443 in β -LS). This residue in both enzymes occurs on a catalytic loop that closes upon intermediate binding and

prevents hydrolysis. In β -LS, a stable hydrogen bond between the lysine side chain and the intermediate carbonyl was observed in the X-ray crystal structures snapshots, and it was shown to stabilize the transition state during β -lactam formation. However, based on the MD simulation results, electrostatic interaction between Lys-449 and the carbonyl of the β AspAMP intermediate is not observed and its catalytic importance is primarily to mediate loop closure by electrostatic interactions with Asp-251 and PPi when the intermediate is bound in the active site. That is due to the negative carboxylate of β AspAMP, which repels the carbonyl oxygen and binds with the side chain of Lys-376, which is not present in β -LS.

In summary, the MD simulations results reveal important information about active site residues. The behavior of ASNS/ β AspAMP active site was compared to known data about the β -LS active site. Finally, through molecular modeling and MD simulations, we were able to develop a reliable model of ASNS with β AspAMP intermediate bound in the active site after various unsuccessful efforts to obtain a crystal structure of this complex. A model representative of the average structure will be used further in virtual screening efforts to discover potent ASNS inhibitors.

CHAPTER 4
DISCOVERY OF A NEW INHIBITOR OF HUMAN ASPARAGINE SYNTHETASE
IDENTIFIED BY STRUCTURE-BASED VIRTUAL SCREENING

Background

The core of the drug discovery process is the identification of new inhibitors (leads) which show weak affinities against a particular biological target and the optimization of the leads into more potent and effective drugs.¹⁰⁹ High-throughput screening (HTS) is the experimental screening of a large number of drug like compounds for their biochemical activity against the biological target. In recent years, automated systems have emerged which can screen tens of thousands of compounds per day. However, despite the recent advances in HTS, certain problems are associated with its use in random screening. The high cost associated with assays and synthesis of lots of compounds, in addition to the low chance of identifying new hits, contribute to these problems.

As an alternative, virtual screening (VS) is the computational screening of very large chemical libraries through molecular docking. The screened compounds are ranked according to different criteria, such as their electrostatic and van der Waals complementarity with the target receptor, to identify new drug candidates (hits). The interplay between HTS and VS is very important, because VS generates a more focused library for subsequent experimental testing and biochemical assays. The advances in high performance computing centers have made the virtual screening more affordable and time efficient.¹⁰⁹

Identifying true hits and rejecting false ones through virtual screening of a large chemical library is very challenging. Towards achieving successful virtual screening, the

binding site flexibility of the target receptor, together with exploration of the ligand conformation, should be considered.¹¹⁰

Molecular dynamics (MD) simulations are widely used to study the dynamic behavior of biological macromolecules. MD simulations can provide valuable information on the flexibility of the target protein upon ligand binding. The three-dimensional structure of the target protein bound with the ligand is very important in docking studies. The ligand free receptor may not be adequate in virtual screening because of the incorrectly positioned amino acid side chains or inadequate loop conformation that can hinder ligand access and positioning in the active site.¹¹⁰ In a study where several ligands were docked into a series of x-ray crystal structure complexes for a particular target, the best results were obtained when a ligand was docked in the native crystal structure conformation, while poor results occur when the ligand was docked into a crystal structure of a different complex.¹¹¹ Therefore, it is essential to carefully prepare the target receptor before performing the docking calculations.

Inhibitors of hASNS have been shown to suppress the proliferation of resistant MOLT-4 leukemia cells.^{51,22,59} However, because the reported hASNS inhibitors are needed in high concentration, their clinical utility is limited. The virtual screening in the following study attempts to identify new leads with more potent and effective physiochemical properties that can inhibit human asparagine synthetase (hASNS). The target receptor used in this study is a representative of the average structure from MD simulations explained in Chapter 3. The NCI database was virtually screened using GLIDE (Grid-based Ligand Docking with Energetics) software.¹¹² Ten hits were

identified and subjected to kinetic assays to determine their activity against hASNS.

Compound NSC605322 inhibited hASNS and was confirmed as a new lead inhibitor of hASNS.

Structure-Based Virtual Screening

Virtual screening is the computational screening of large chemical libraries to identify new drug leads that can exhibit activity against a particular target. There are two main approaches used in virtual screening: structure-based virtual screening (SBVS) and ligand based virtual screening (LBVS). LBVS is designed to find compounds similar to known active compounds, while SBVS involves rapid docking of many chemical compounds into the active site of a biological target. Docking is the process of bringing two molecules together to form a stable complex; for example, docking a small molecule in the active site of an enzyme. The molecules are further scored and ranked according to scoring functions. Thus the quality of the scoring function, in terms of accuracy and efficacy to predict the binding affinity of each docked structure, is very important.¹¹³

Three different types of conformational searches are employed by docking software.¹¹⁴ Programs such as GLIDE,¹¹² AFFINITY, and ICM perform conformational search on each docked molecule in the binding pocket using stochastic search methods, such as Monte Carlo (MC) and simulated annealing methods. Other softwares, for instance FRED and SLIDE perform conformational search on each molecule before docking it in the active site. The different generated conformations for each molecule are then docked rigidly in the active site and the rotation and translation of the rigid conformers are considered. The third category of conformational search makes use of an incremental construction algorithm. In incremental construction, the

rotating bonds of the ligand are broken to produce fragments. The fragments are then docked rigidly at different favorable positions in the binding site with the larger fragment as the starting base fragment, and the entire ligand is reassembled. Softwares that perform incremental construction include DOCK and HOOK.¹¹⁴

Scoring functions are approximate methods to estimate the binding affinity between the docked ligand and the receptor based on ligand complementarity, such as electrostatic and van der Waals interactions with the active site. Because scoring functions tend to be sufficiently simple to screen large chemical library in reasonable computer time, they often suffer from poor accuracy. However, the scoring functions are in continuous development of to yield accurate binding without sacrificing accuracy. There are different types of scoring functions implemented in docking programs, such as force field-based, empirical-based, and knowledge-based scoring functions, as well as consensus scoring.¹¹⁴⁻¹¹⁶

Force field-based scoring functions sum up the electrostatic and vdW contributions between the ligand and receptor active site.¹¹⁷ The parameters for force field based scoring functions are obtained from both quantum mechanical calculations and experimental data. For example, the energy parameters for DOCK are obtained using the AMBER force field.¹¹⁸⁻¹²⁰ Solvation effects are usually represented using implicit solvent models, such as the Poisson–Boltzmann surface area (PB/SA)¹²¹⁻¹²³ model and the generalized-Born surface area (GB/SA) models.¹²⁴⁻¹²⁶ Empirical scoring functions estimate the binding free energy between receptor and ligand as a set of weighted energy components¹²⁷

$$\Delta G = \sum_i W_i \cdot \Delta G_i \quad (4-1)$$

where ΔG_i represents different energy components such as electrostatic, hydrogen bond, desolvation, entropy and vdW. The coefficients are determined by calibration against a set of experimentally known receptor-ligand structures and their binding affinities using regression analysis.¹²⁸⁻¹³⁴ The empirical energy function depends on the training set used in deriving the weighting coefficients. Well known docking programs use empirical energy functions such as FlexX and GLIDE. The Glide Score is a modified version of the ChemScore function as follows:¹³⁵

$$\begin{aligned}
 \Delta G_{bind} = & \\
 & C_{lipo-lipo} \sum f(r_{lr}) + \\
 & C_{hbond-neut-neut} \sum g(\Delta r)h(\Delta \alpha) + \\
 & C_{hbond-neut-charged} \sum g(\Delta r)h(\Delta \alpha) + \\
 & C_{hbond-charged-charged} \sum g(\Delta r)h(\Delta \alpha) + \\
 & C_{max-metal-ion} \sum f(r_{lm}) + C_{rot}H_{rot} + C_{polar-phob}V_{polar-phob} + \\
 & C_{coul}E_{coul} + C_{vdW}E_{vdW} + solvation - terms
 \end{aligned} \tag{4-2}$$

The ΔG coefficients are obtained by multiple linear regressions.¹³⁵ Distances or angles that lie within specific limits are given a full score by the functions f , g and h , and a partial score is given for distances and angles that lie outside those limits.¹³⁵ Δr and $\Delta \alpha$ are the deviations of the hydrogen bonds distances and angles from their ideal values, respectively.

Knowledge-based scoring function is another kind of scoring function and is also referred to as statistical-potential based scoring functions. A statistical potential is derived based on intermolecular distances observed in experimentally known

structures. Frequently observed nearby intermolecular interactions between different atom types are proposed to be energetically advantageous and contribute favorably to the binding affinity.¹³⁶⁻¹³⁸ Because each scoring function has its own imperfections and advantages, a recent trend is of using consensus scoring has been introduced.¹³⁹ Consensus scoring combines scores from different scoring functions so that increasing the probability of finding true hits. An example of consensus scoring is X-Cscore.^{139,140}

Computational Details

The first step in structure-based virtual screening is the preparation of the protein receptor structure. A representative of the average structure from the molecular dynamics simulations described in Chapter 3 was used as the target for this virtual screening study. Solvent was removed except for the water molecules coordinated with the Mg²⁺ ions. As was mentioned in Chapter 3, inorganic pyrophosphate (PPi) is stabilized by many hydrogen bonding interactions in a deep binding pocket called the PP-motif and covered with a catalytic loop. According to kinetic data and the absence of ATP/PPi exchange, it is believed that PPi is the last product released from the active site. In addition, absence of the PPi will lead to loss of hydrogen bonding interactions between the key residue Lys-449 and PPi which will result in a wide open state of the catalytic loop (residues 443-452) with the active site becoming more exposed to solvent, possibly leading to hydrolysis of the intermediates and hence loss of activity (Figure 3-29). Therefore, PPi and Mg²⁺ ions were kept in the active site during docking calculations.

Virtual Screening Using GLIDE

The Protein Preparation Wizard panel and Maestro's structure-editing capabilities were used to examine and adjust the formal charges and the atom types for the metal

ions and the cofactors. Following protein preparation, the receptor grid, which represents the position and size of the active site, was generated. GLIDE uses two boxes for receptor grid generation (both having the same center): the inner box (ligand diameter midpoint box) and the enclosing box (outer box). The ligand center, which is the midpoint between the most widely separated atoms, lies within the inner box during site point search. The ligand can move outside of the inner box, but must be enclosed within the enclosing box. The enclosing box should be large enough to contain the ligand and inner box lengths. The relationship between the lengths of the outer enclosing box, E, inner box, B and ligand size, L is: $E \geq B + L$

Following receptor grid generation, the native ligand β AspAMP was extracted from the ASNS- β AspAMP model and re-docked in the active site. The bond orders and formal charges for the β AspAMP ligand were adjusted using LigPrep module in Maestro. GLIDE produced a docking binding pose that closely resembled the binding pose of the MD simulated ASNS- β AspAMP complex and thereby verifying the docking approach. The docking score of the native ligand was -4.9Kcal/mol. The docking score of the previously identified inhibitor (compound **1**, Figure 1-4) was also -4.9Kcal/mol.

Three-dimensional structures of compounds from the NCI's chemical library were downloaded from the website (<http://cactus.nci.nih.gov/download/nci/>) and processed with LigPrep software (available from Schrödinger).¹³⁵ Flexible docking, in which the docked ligands undergo conformational search generating a set of conformations for each ligand, was then performed. A site point search was executed for each conformation to find best orientation and position of the ligand within the active site. Orientations that displayed steric clashes with active site atoms were excluded. Then

rotation of each orientation was followed, and hydrogen bonding interactions of the ligand conformation with the active site residues were scored with modified version of the ChemScore empirical scoring function.¹⁴¹ The best selected poses (200-400) were subjected to energy minimization using the energy function for the OPLS-AA force field, followed by a Monte Carlo sampling procedure, in which torsion angles of rotamer-groups were examined to improve sampling and scoring.¹⁴¹ Finally the minimized poses were ranked using Glide Score Standard Precision scoring function. A total of 500,000 compounds from NCI's chemical library were screened and ranked according to their scores. However, the best docked pose for each ligand was selected by using a model energy score called Emodel, which is a composite of GlideScore, the molecular mechanics interaction energy between the ligand and receptor, and the strain energy of the ligand. Glide Score is a modified version of the ChemScore function (Equation 4-2).¹³⁵

Virtual Screening Using DOCK

Virtual screening of two million drug-like compounds from the ZINC database³³ was performed using DOCK6.¹⁴³ The ligand was oriented within the active site using a negative surface image of the receptor binding site, prepared using the DMS tool in UCSF Chimera.²⁸ The SPHGEN tool within the DOCK package was used to create overlapping spheres to represent the binding site and describe the shape of the molecular surface.¹⁴⁴ The size and position of the active site grid was defined by the program SHOWBOX and the grid was generated by the program Grid tool in DOCK. Flexible docking was used, in which the conformational space for the docked molecules was sampled using the anchor and grow incremental algorithms.¹⁴⁵ The anchor, which undergoes rigid orientation within the active site, was first identified as the largest rigid

substructure of the ligand, followed by identification of the flexible parts. Anchor orientations were optimized using the grid scoring function and simplex minimizer. The native ligand β AspAMP was extracted from the ASNS- β AspAMP model and re-docked in the active site. DOCK generated a docking pose that closely resembled the native binding pose of β AspAMP in the active site, thereby verifying the docking approaches (Figure 4-1). The docking score of the native ligand was -98Kcal/mol. Drug-like compounds from the ZINC database were then docked in the ASNS active site and scored using the grid-based scoring function, which is based on the non-bonded terms of the molecular mechanics force field to approximate the interaction energies.

Results and Discussion

Three dimensional structures from the NCI database (500,000 compounds) were docked in the active site of asparagine synthetase (ASNS) using GLIDE. In addition, two million drug-like compounds from the ZINC database were virtually screened using DOCK. Molecules with docking scores higher than the score of the native ligand were inspected for their “drug likeness”. In 1997, Christopher A. Lipinski formulated a rule to evaluate drug likeness called rule of five (RO5) with respect to physicochemical factors linked with aqueous solubility and intestinal permeability.¹⁴⁶ The rule of five (RO5) defines four physicochemical characteristics of 90% of orally active drugs:(1) molecular weight less than 500; (2) H-bond donors fewer than 5; (3) H-bond acceptors fewer than 10; (4) Partition function (log P) less than five. Compounds that do not satisfy at least three of those four parameters are less likely to be oral active drugs and cannot pass phase II of clinical trials. Therefore, it is important to keep this rule in mind during drug discovery process. Compounds that are similar to other enzyme substrates or peptides

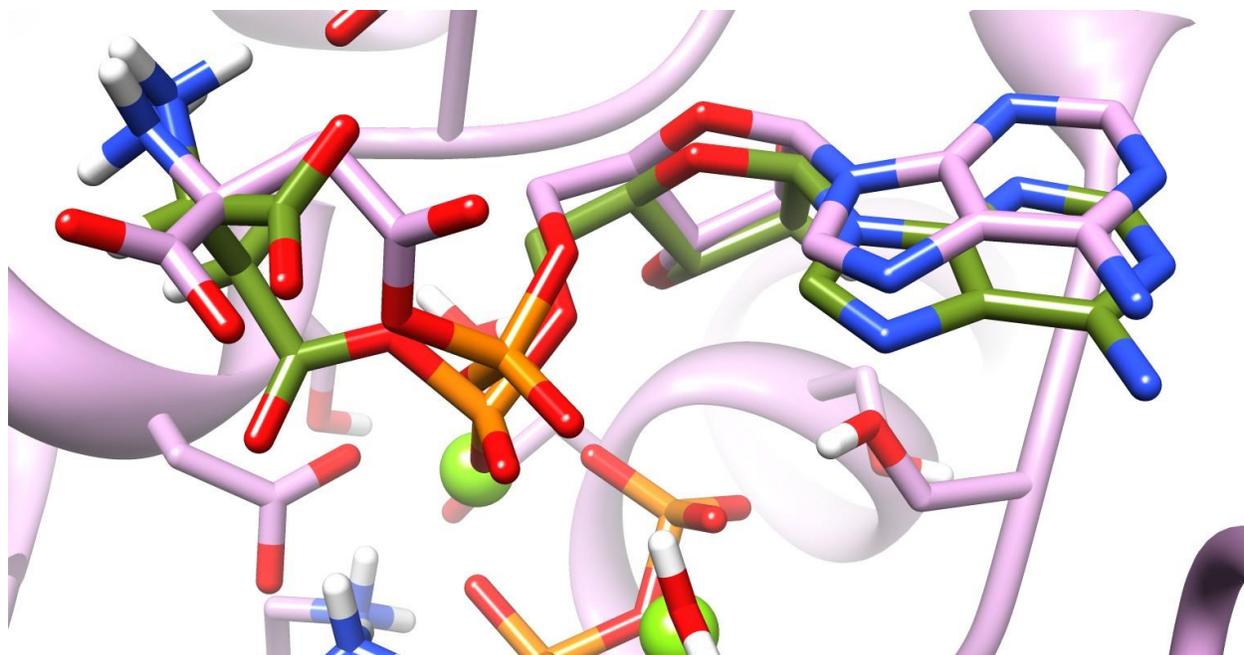


Figure 4-1. Overlay of the re-docked β AspAMP intermediate using DOCK (green) and the native binding pose of β AspAMP (pink)

Table 4-1. Docking score of the ten selected hits from GLIDE virtual screening, their Lipinski properties and their residual activity values.

Compound	Docking score (kcal/mol)	H-bond acceptors	H-bond donors	Log P	Molecular weight (g/mol)	Residual activity (%)
NSC605322	-9.7	6	2	4.34	474.96	65
NSC333462	-9.6	12	5	5.7	650.72	91
NSC8805	-9.5	7	4	1.74	386.38	71
NSC9219	-9.4	7	5	1.99	302.24	25
NSC18719	-9.2	7	6	0.19	302.71	89
NSC43663	-8.9	7	6	-0.5	268.27	93
NSC352274	-8.8	8	5	0.27	386.4	91
NSC36398	-8.2	7	5	-0.18	304.3	39
NSC1678	-8.3	8	6	-4.05	238.19	94
NSC369	-6.9	6	0	1.58	408.49	94

The residual activity test was performed by lab member Yongmo Ahn. Reaction mixtures for residual activity consisted of human ASNS (4 μ g), 0.5 mM ATP, 10 mM L-aspartate, either 25 mM L-glutamine or 100 mM NH_4Cl as nitrogen source, and the pyrophosphate reagent (350 μ L) dissolved in 100 mM EPPS buffer, pH 8.0, containing 100 μ M of each compound, 10 mM MgCl_2 and 10% DMSO (1 mL total volume).

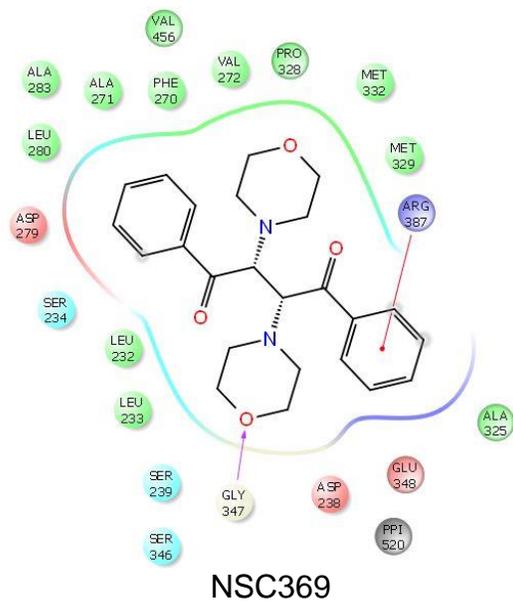
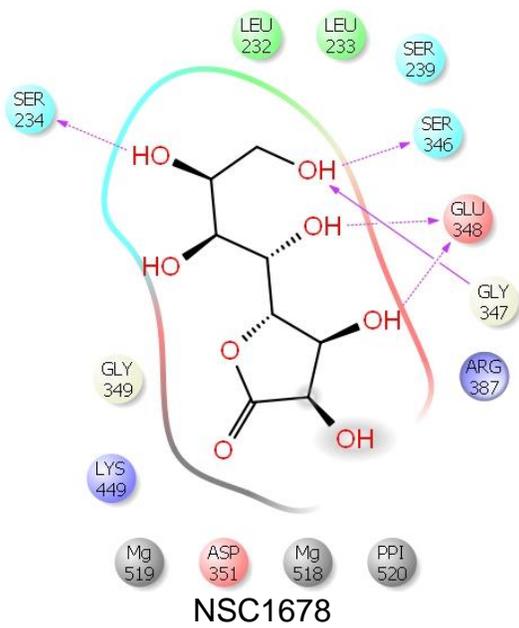
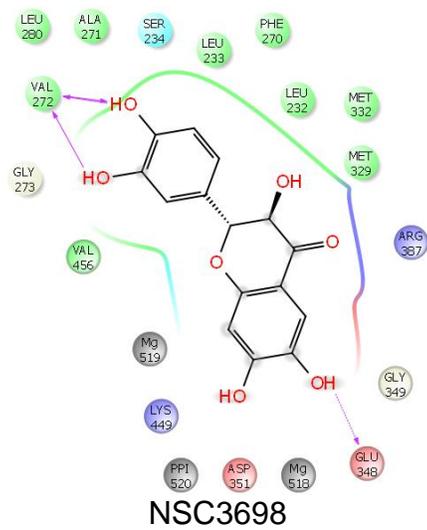
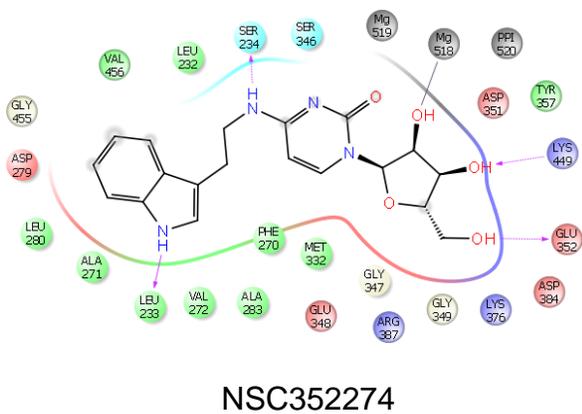


Figure 4-2. Continued.

or glycosides were excluded because of specificity concerns. It has been also observed that compounds with more than 10 rotatable bonds correlate with decreased rat oral bioavailability.¹⁴⁷ Therefore compounds with more than 10 rotatable bonds were excluded.

A final set of ten hits from GLIDE virtual screening was selected for experimental testing against hASNS inhibition (Figure 4-2, Table 4-1). All in vitro experiments were performed by lab member Yongmo Ahn. The ten selected hits were subjected to a residual activity test to approximate the effect of inhibitors on PPI production under steady state conditions in which PPI formation is coupled to NADH consumption (340 nm). The residual activity of hASNS(%) is defined as the quotient of the PPI production rate of the assay enzyme in the presence of 100 μ M inhibitor divided by the pyrophosphate production rate by the same enzyme in the absence of inhibitor. The lower the residual activity value, the more potent is the inhibitor. For example; the previous identified nanomolar inhibitor, compound **1** (Figure 1-4, Chapter 1) has a residual activity of 20 %, however compound **3** with micro-molar inhibition potency has a residual activity of 60%. A standard residual activity below 75% was set for further kinetic analysis experiments. Compounds NSC8805, NSC9219, NSC605322 and NSC36398 had residual activities less than 75% (Table 4-1), but only compound NSC605322 was subjected to further analysis, because compounds NSC8805, NSC9219 and NSC36398 were shown to inhibit the coupling enzymes by control experiments in the pyrophosphate reagent used for the continuous assay.

The inhibitory effect of compound NSC605322 on glutamine-dependent hASNS was assayed by measuring the rate of pyrophosphate (PPI) production.¹⁴⁸

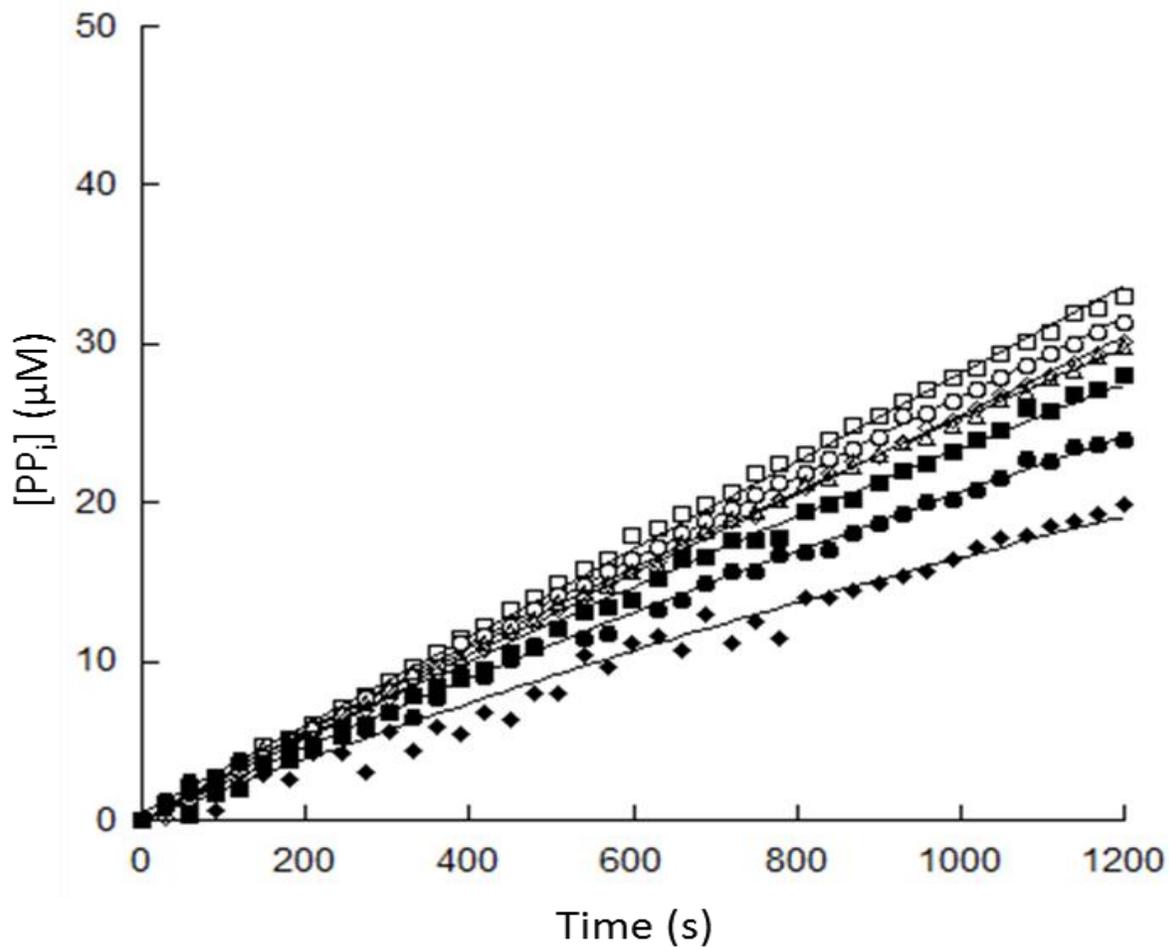


Figure 4-3. Glutamine-dependent production of PPI in the presence of the inhibitor NSC 605322, (0–100 μM). open squares, 0 μM ; open circles, 10 μM ; open diamonds, 25 μM ; open triangles, 40 μM ; filled squares, 60 μM ; filled circles, 80 μM ; filled diamonds 100 μM . Experiment performed by Yongmo Ahn

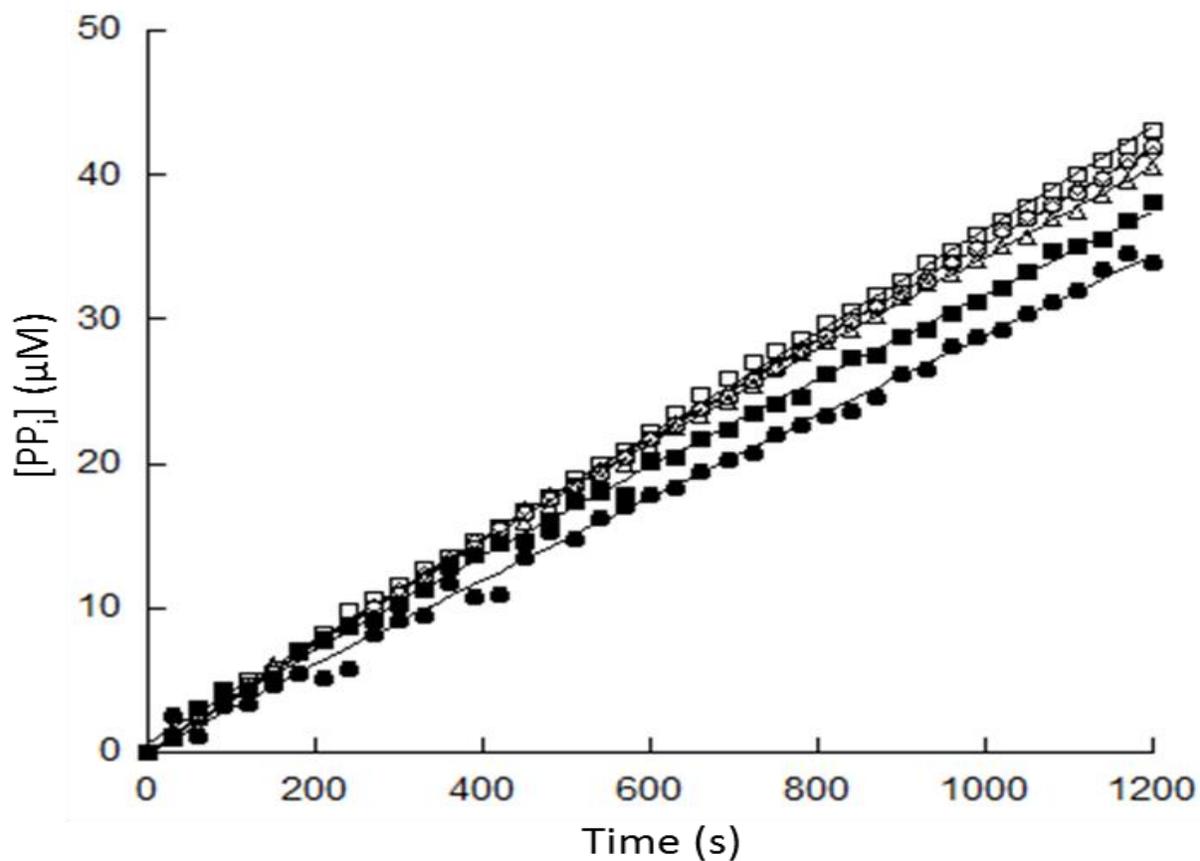


Figure 4-4. Ammonia-dependent production of PP_i in the presence of the inhibitor NSC 605322, (0–80 μM). open squares, 0 μM; open circles, 10 μM; open diamonds, 25 μM; open triangles, 40 μM; filled squares, 60 μM; filled circles, 80 μM. Experiment performed by Youngmo Ahn

Compound NSC605322 showed slow tight binding inhibition¹⁴⁹ (Figures 4-3 and 4-4) with an overall inhibition constant of 1.6 μM when ammonia was the nitrogen source and 2.2 μM when L-Gln was employed as the nitrogen source.

Although compound NSC605322 is much weaker inhibitor than compound **1**, NSC605322 provides key structural features and can be optimized further to a more potent lead inhibitor with improved cell permeability. The docking pose of compound NSC605322 in the active site of ASNS was investigated (Figure 4-5). Overlay between the ASNS/inhibitor complex and ASNS/ β AspAMP complex is shown (Figure 4-6). The predicted binding pose of compound NSC605322 employs a docking pose similar to that of the native intermediate β AspAMP. The acridine moiety is settled in the adenine binding pocket and makes hydrophobic contacts with hydrophobic amino acid side chains. Acridine-based pharmacophores were reported to show biological activities, such as intercalation into DNA, as well as anticancer and antimicrobial activities.^{150,151} Therefore, new drugs based on acridine derivatives are regularly designed and synthesized as antitumor and antibacterial therapeutic agents.^{152,153} The sulfonamide group occupies the same position as the phosphate group. The sulfonamide group of the inhibitor was predicted by Epik¹⁵⁴ (A program within the GLIDE package that predicts the pKa of the ionizable groups in ligands) during docking calculations to be deprotonated. The pKa values for the amide nitrogen of some sulfonamide derivatives are reported to be around 5.5-7.0 (Table 4-2). Therefore; the pKa of compound NSC605322 is likely in the range 6.0-7.0. Thus under the conditions used in the assay, pH 8.0, the amide nitrogen will be mainly deprotonated. The sulfonamide deprotonated form is also stabilized by resonance with the pyridine aromatic ring.

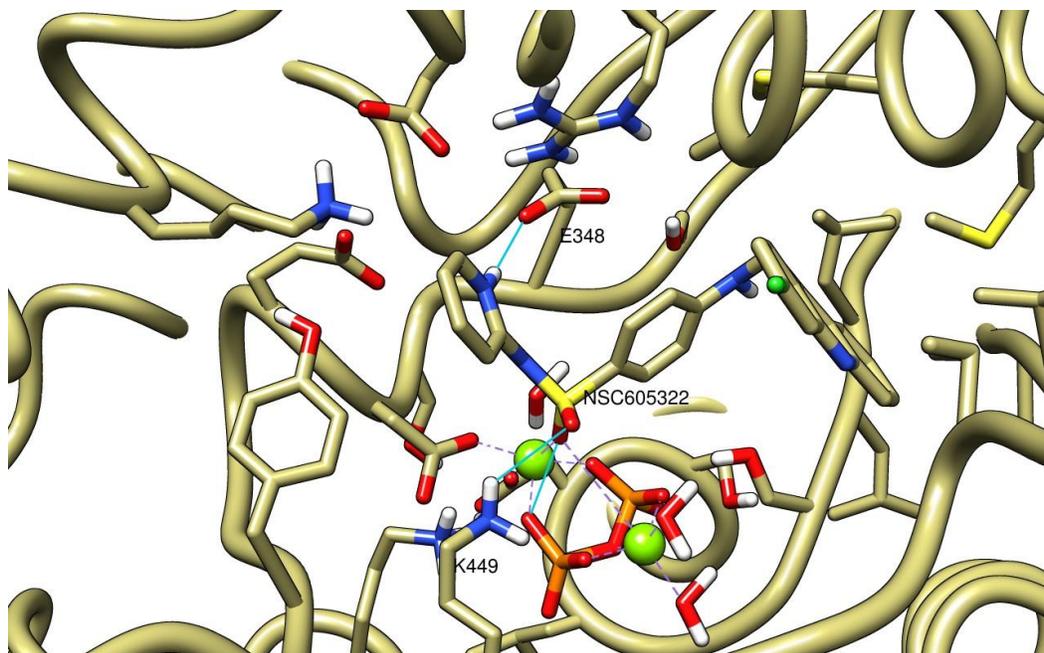


Figure 4-5. Putative interactions of compound NSC605322 with ASNS active site residues.

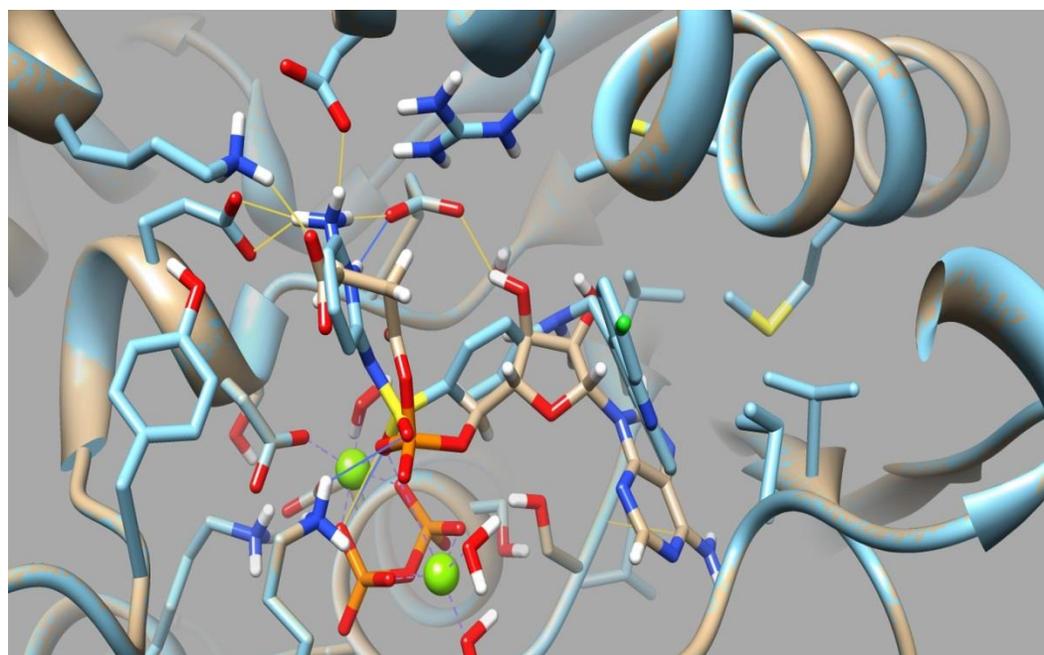


Figure 4-6. Overlay of the docking pose of ASNS-inhibitor complex (blue backbone atoms) and the ASNS-βAspAMP complex (brown backbone atoms).

Table 4-2. pKas of amide hydrogens of some sulfonamide derivatives in water.¹⁵⁵

R	pKa
	6.33
	6.71
	5.77
	5.19
	5.41
	6.22

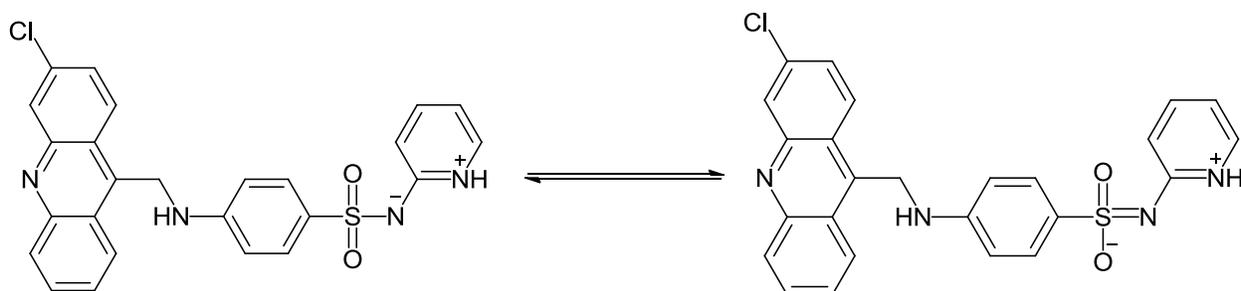


Figure 4-7. Conjugate base of ionized sulfonamide functional group of compound NSC605322.

Therefore, the sulfonamide group can bind as its conjugate base and coordinates with Mg^{2+} ions, thereby imitating the phosphate group of the native intermediate β AspAMP (Figure 4-7). This result is consistent with a previous study on hASNS inhibitors (Figures 1-4 and 1-5, Chapter 1), in which only compounds **2** and **3** were able to inhibit hASNS because they contain an ionized functional group that can mimic the charged phosphate group while compound **4** did not exert any inhibitory effect because it does not have an ionized functional group. In addition, the molecular dynamics simulations study performed on the ASNS/ β AspAMP/PPi/ Mg^{2+} complex have shown that Lys-449 occurs on a catalytic loop that closes tightly in the presence of the intermediate (Figure 3-29, Chapter 3). Loop closure is mediated mainly by hydrogen bonding interaction of Lys-449 side chain with PPi and the side chain carboxylate of Asp-351. Based on the docking results, the negatively charged sulfonamide group of compound NSC605322 in its ionized form makes electrostatic interactions with the side chain of Lys-449 (Figure 4-5). This electrostatic binding may increase the probability of loop closure on the active site and result in tighter binding of the inhibitor. Compound NSC605322 is the only hit from the experimentally tested set that has an ionized functional group that can mimic the charged phosphate group. Although no experiments have been performed to test the inhibition activity of selected hit compounds from DOCK, many of the identified hits contain a sulfonamide moiety, thereby emphasizing the importance of the presence of a functional group that can mimic the native intermediate phosphate group (Figure 4-8). Therefore, this functional group should be considered in further design and optimization of lead inhibitors for hASNS.

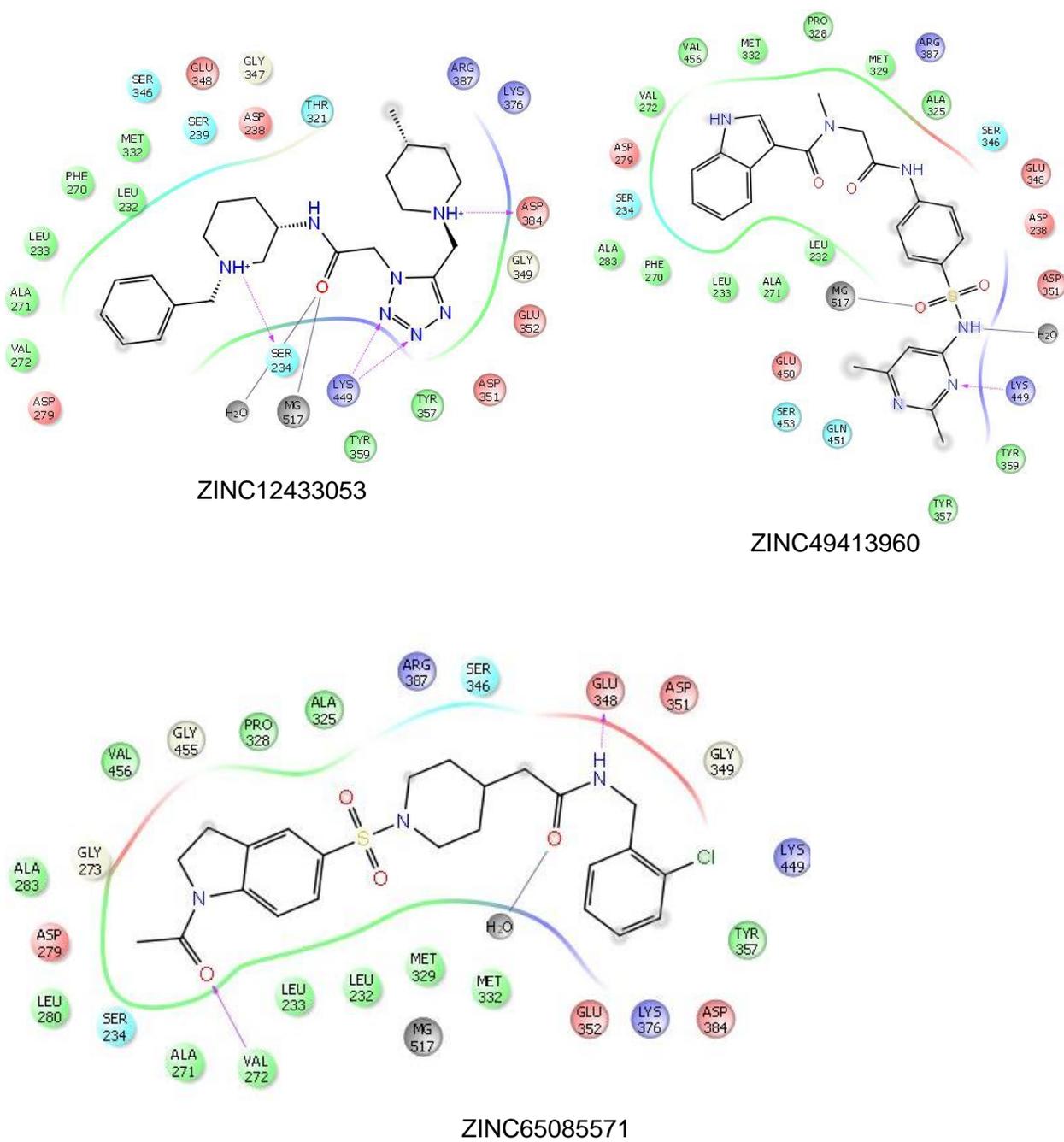


Figure 4-8. 2D predicted binding poses of some hits identified by virtual screening using DOCK.

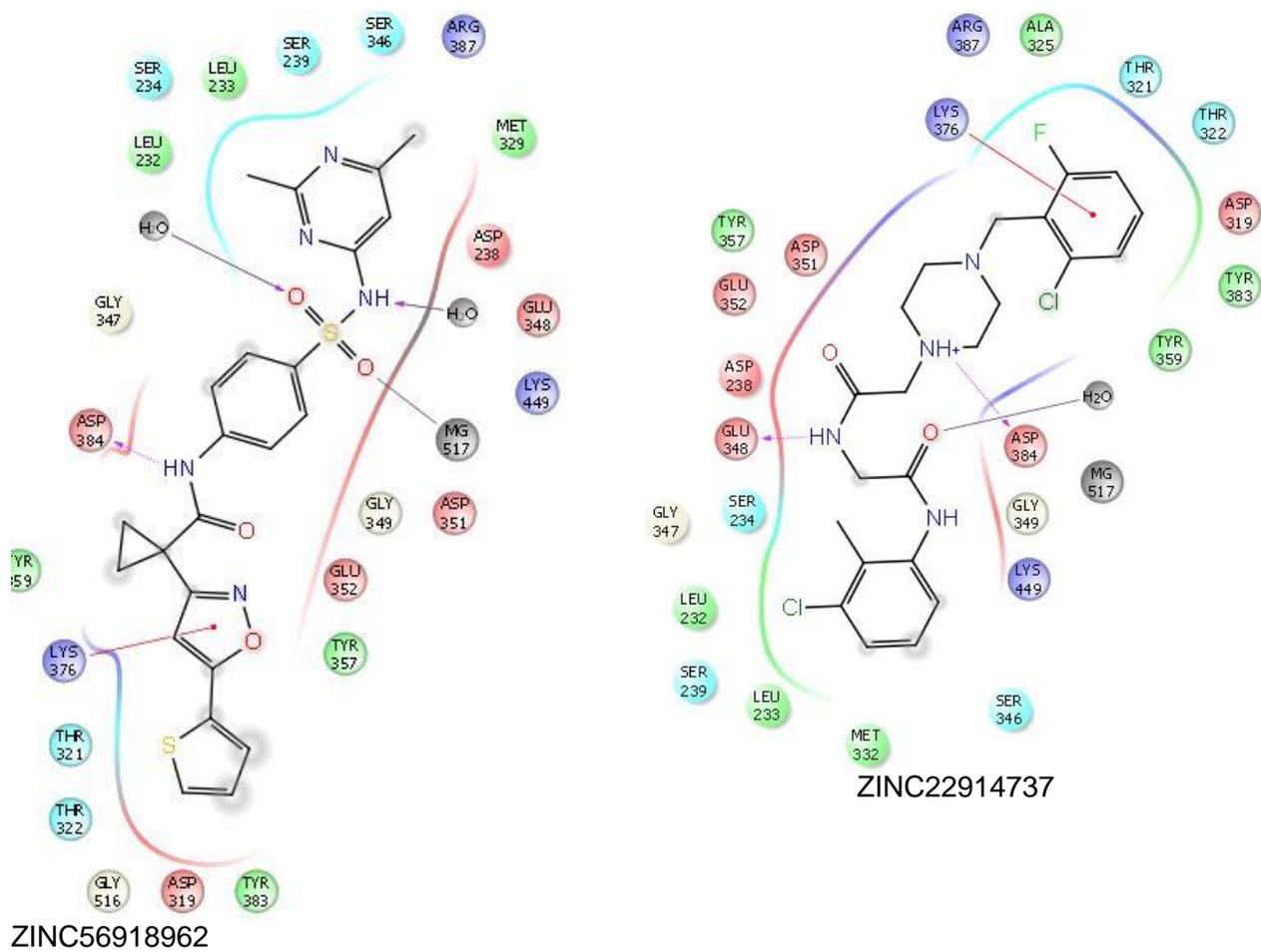


Figure 4-8.Continued.

Table 4-3. Docking score of some selected hits from DOCK virtual screening

Compound	Docking score (kcal/mol)
ZINC12433053	-85
ZINC49413960	-83
ZINC65085571	-82
ZINC56918962	-84
ZINC22914737	-84

In addition, the protonated nitrogen of the pyridine ring forms a hydrogen bond with the carboxylate group of Glu-348, which is a key residue in the active site and undergoes a stable hydrogen bonding interaction with the amino group of β AspAMP intermediate. Glu-348 was reported to be important for intermediate binding and for mediating coordination of the glutaminase and synthetase active sites.¹⁰⁴

Summary

Virtual screening to identify new inhibitors of human asparagine synthetase was performed. Two million drug-like molecules from the ZINC database were docked in the active site of hASNS using DOCK. In addition 500,000 compounds from the NCI chemical database were docked using GLIDE. Ten hits were selected from the NCI database, based on satisfying Lipinski's rule of 5 and the interaction with key active site residues. A new lead inhibitor of hASNS (NSC605322) was discovered. Kinetic experiments showed that compound NSC605322 inhibits hASNS with an overall inhibition constant of 1.6 μ M with ammonia as the nitrogen source and 2.2 μ M with L-Gln as the nitrogen source. Further experiments, such as cell based assays are required to test the effect of compound NSC605322 on the growth of cancer cell lines, for example, MOLt-4 leukemia cell lines. The new inhibitor has an ionized sulfonamide group that mimics the charged phosphate group in coordinating with Mg^{2+} ions. In addition the newly identified inhibitor makes electrostatic interactions with the key residues Lys-449 and Glu-348.

The discovery of a lead inhibitor is the first step in the drug discovery process. Focused libraries of novel compounds based on compound NSC605322 with acridine and sulfonamide moieties should be synthesized to gain better understanding of the structure-activity relationships (SAR) responsible for the inhibitory effect against

hASNS. For example, the pyridyl group may be replaced with other electron-withdrawing heterocyclic groups, such as thiazole, oxazole, thiadiazole and pyrazole. In addition the acridine ring can be modified to contain H-bond donor and acceptor groups that can undergo electrostatic interactions, such as the interaction of β AspAMP adenine with the backbone atoms of Val-272.

For further discovery of new hASNS inhibitors, kinetics experiments to test the inhibitory effect of selected hits from the DOCK virtual screening calculations can be performed. Also, structure-based virtual screening of new focused libraries that contain ionizable sulfonamide moieties and maintain electrostatic interactions with key residues should be considered. Another approach is to perform-ligand based virtual screening based on the identified hit to provide a library of structurally similar compounds that can be tested experimentally.

In summary, a new micromolar inhibitor of hASNS was discovered using structure-based virtual screening. This lead inhibitor has a new structural motif that should permit the development of compounds with improved cell permeability than those previously reported.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

Concluding Remarks

The ability of the identified hASNS *N*-phosphonosulfonimidoyl based-inhibitors (Figure 5-1) to suppress the proliferation of MOLT-4 leukemia cells with nanomolar potency proved that the hASNS inhibitors can be used in the treatment of acute lymphoblastic leukemia. However, these inhibitors are difficult to synthesize, and high concentrations are needed for cell internalization. Thus new hASNS inhibitors with more drug-like properties are needed. Structure-based virtual screening of large chemical libraries is a widely used technique to identify new lead inhibitors by docking many drug-like compounds to the biological target of interest. Incorporating protein flexibility in the docking process plays an important role in improving the docking calculations and thereby increasing the chances of predicting reliable lead inhibitors. To reduce the risks associated with docking to an inadequate protein conformation, one representative protein structure having the most probable protein conformation is used, or the ligands are docked to different protein conformations. The absence of good parameters that describe the acylphosphate and *N*-phosphonosulfonimidoyl functional groups precluded previous computational efforts to study ASNS bound with the intermediate or the inhibitor in the active site.

New parameters that describe the acylphosphate and *N*-phosphonosulfonimidoyl functional groups were optimized following the CGenFF protocol. The parameters are compatible with the CHARMM additive all-atom force field for biological molecules. The newly optimized charges were able to reproduce the interaction energies and distances for the model-water complexes calculated by QM calculations. In addition, QM

geometries, vibrational spectra and dihedral potential energy scans were reproduced well by the newly developed parameters. The newly developed parameters will enable computational studies using molecular mechanics approaches, such as molecular dynamics simulations and free energy calculations, on enzymes that catalyze acylphosphate formation and inhibitors with *N*-phosphonosulfonimidoyl functional groups.

After developing new CHARMM parameters for the acylphosphate functional group in the intermediate β AspAMP, the ASNS/ β AspAMP/PPi/Mg²⁺ complex was modeled and refined using molecular dynamics simulations. The ligand-free ASNS was also simulated to detect related conformational changes as a result of intermediate binding and to identify key active site residues. The MD simulation results revealed important interactions in the active site that mediate intermediate binding. The adenosine group binds in a well-defined ATP binding pocket. The phosphate oxygen atoms undergo critical electrostatic interactions by coordinating with the Mg²⁺ ions. The amino group is stabilized with electrostatic interactions with the carboxylate side chains of Glu-348, Asp-384 and Glu-352. The carboxylate group of β AspAMP is stabilized by hydrogen bonding with the side chain of Lys-376. Inorganic pyrophosphate lies in a deep binding pocket called the PP-motif, where it is stabilized by various hydrogen bonding interactions. According to kinetic results, the absence of ATP/PPi exchange, and the very low PIX of inorganic pyrophosphate (PPi), PPi was proposed to be the last product released from the active site. A catalytic loop that carries a conserved lysine residue (Lys-449) is observed to close upon the intermediate and prevent its hydrolysis. The behavior of the ASNS/ β AspAMP active site was compared to known data for the β -

LS active site. Finally, through molecular modeling and MD simulations, we were able to provide a reliable model of ASNS with the β AspAMP intermediate bound in the active site after various unsuccessful efforts to obtain a crystal structure of this complex.

A model of the ASNS/ β AspAMP/PPi/ Mg^{2+} complex representative of the average structure was further used in virtual screening calculations to identify potent ASNS inhibitors. A total of 500,000 compounds from the NCI chemical database were docked to the ASNS active site using GLIDE. Ten hits were selected based on agreement with Lipinski's rule of 5 and the presence of electrostatic interactions with key active site residues. From this group, a new lead inhibitor of hASNS (NSC605322) was identified. Compound NSC605322 inhibited hASNS with an overall inhibition constant of 1.6 μ M with ammonia as the nitrogen source and 2.2 μ M with L-Gln as the nitrogen source. The new inhibitor has an ionized sulfonamide group that mimics the charged phosphate group in coordinating with Mg^{2+} ions. In addition, the newly identified inhibitor interacts with the key residues Lys-449 and Glue-348.

The discovery of a lead inhibitor is the first step in the drug discovery process. Focused libraries of novel compounds based on compound NSC605322's acridine and sulfonamide moieties should be synthesized to gain better understanding of the structure-activity relationships responsible for the inhibitory effect against hASNS.

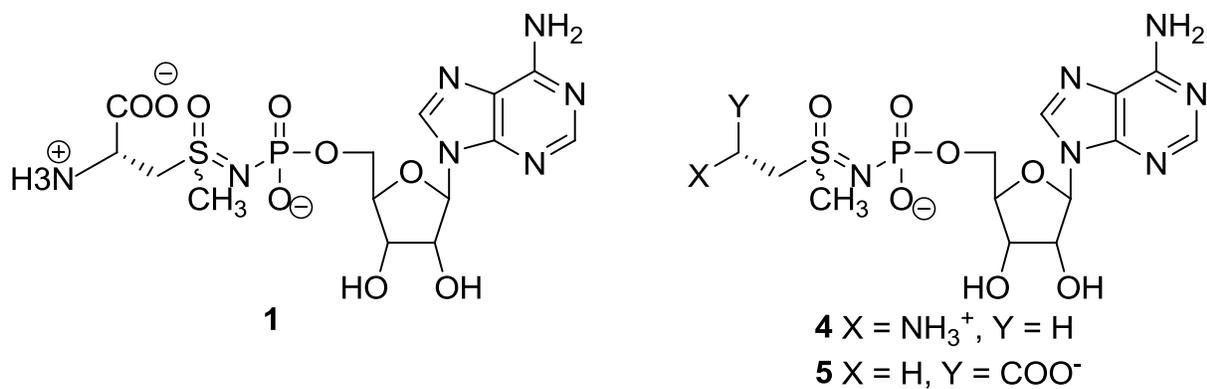


Figure 5-1. Chemical structures of hASNS *N*-phosphonosulfonimidoyl-based inhibitors.

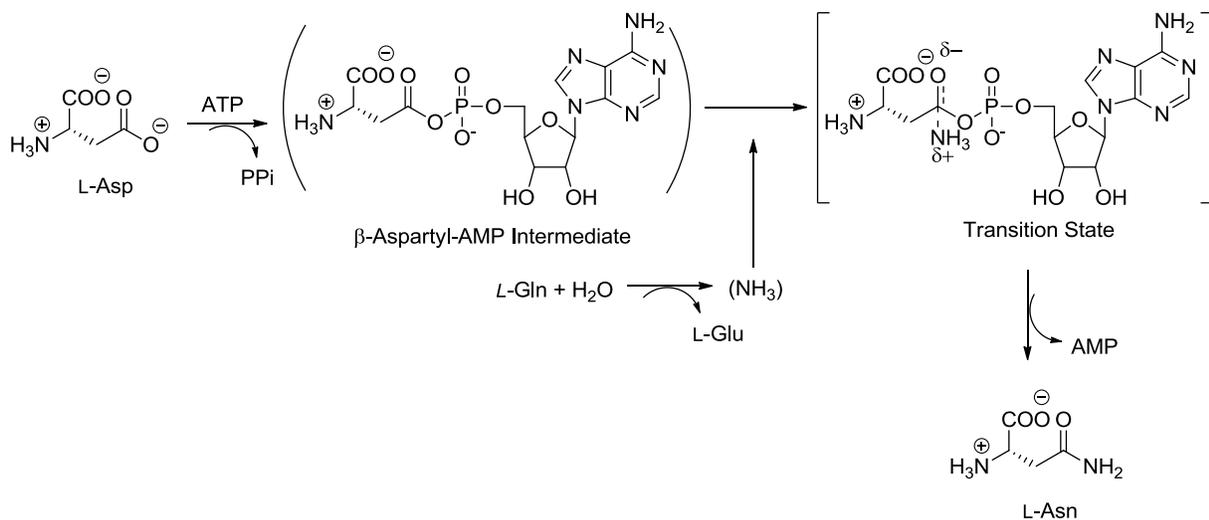


Figure 5-2. Overview of the reaction catalyzed by glutamine-dependent asparagine synthetase.

Future Work: Free Energy Calculations to Estimate the Difference in Binding Free Energy between (*R*)- and (*S*)- Isomers of an *N*-phosphonosulfonimidoyl Based Inhibitor

A sulfoximine-based inhibitor **1** of human asparagine synthetase (hASNS) is analogue of ASNS transition state and was able to inhibit hASNS with nanomolar potency (Figure 5-1). In addition, this inhibitor slowed the growth of ASNase-resistant leukemia cells when used in high concentrations and when L-asparagine was reduced in the external medium by ASNase.⁴⁹ It was hypothesized that compound **1** is needed in high concentrations to suppress the growth of ASNase-resistant leukemia cells, because the charged groups prevent its entry into the cell.²² In order to further investigate the relative contribution of the other charged groups to the inhibitory effect of compound **1**, the carboxyl and amino groups of compound **1** were removed and compounds **4** and **5** were synthesized and tested experimentally for the inhibition of hASNS.⁵⁹ Compound **4**, which has no net charge at cellular pH, inhibited hASNS in higher potency than compound **1**. In addition, Compound **4** slowed the growth of leukemia ASNase-resistant cells with $IC_{50} = 0.1$ mM and in the absence of ASNase in the external medium. This concentration is a factor of 10 lower than the IC_{50} of compound **1**.⁵⁹ Compound **4** has an overall inhibition constant (K_i^*) of 8 nM when L-glutamine was used as the nitrogen source. The K_i^* value for *N*-phosphonosulfonimidoyl **4** increased four-fold under the same conditions when ammonia was used instead of L-glutamine. Contrary to the higher inhibitory effect of amino *N*-phosphonosulfonimidoyl **4**, the carboxy *N*-phosphonosulfonimidoyl **5** was a relatively poor hASNS inhibitor with K_i^* values of 800 and 300 nM when the nitrogen source was glutamine or ammonia, respectively.

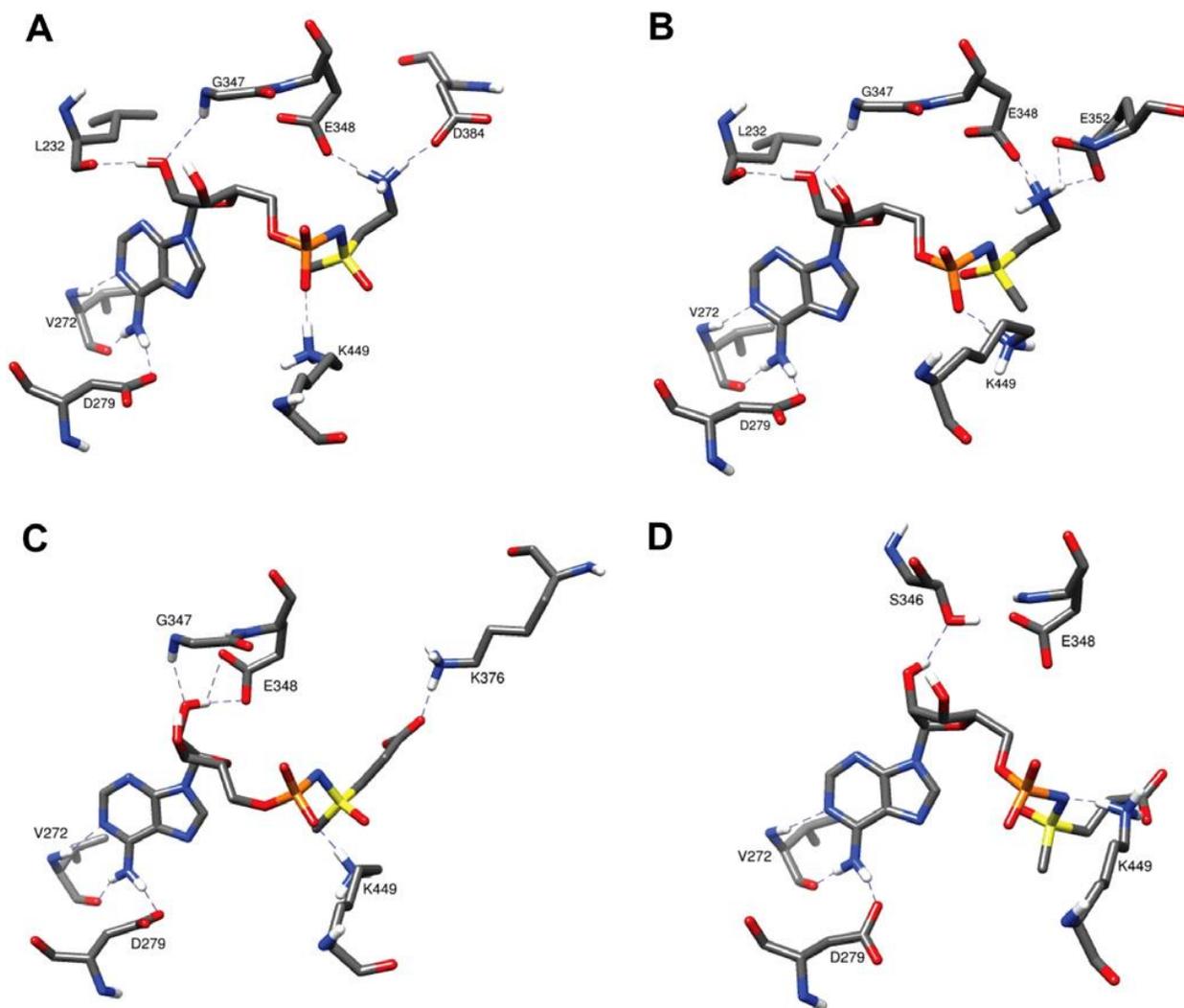


Figure 5-3. Putative interactions of the four diastereoisomers with key residues in the ASNS active site. (A and B) are the (*R*)- and (*S*)- diastereoisomers of compound **4**, respectively. (C and D) are the (*R*)- and (*S*)- diastereoisomers of compound **5**, respectively. Color scheme: C, grey; H, white; N, blue; O, red; P, orange; S, yellow. Reprinted with permission from Ikeuchi a, H.; Ahn, Y.; Otokawa, T.; Watanabe, B.; Hegazy, L.; Hiratake, J.; Richards, N. G. J. *Bioorg. Med. Chem.* **2012**, *20*, 5915.

Compounds **4** and **5** were employed as a mixture of diastereoisomers at the sulfur stereogenic center. Previous work showed that there is a strong dependence on the stereochemistry of tight binding sulfoximine inhibitors.^{156,157} For example, the (*R*)-configuration of a sulfoximine inhibitor of the enzyme γ -Glutamylcysteine Synthetase (γ -GCS) acted as a potent slow binding inhibitor, while the (*S*)-sulfoximine showed a simple reversible inhibition with much lower potency.¹⁵⁶ The molecular interactions of the diastereoisomers of compounds **4** and **5** within the synthetase active site were qualitatively investigated. Because of the structural resemblance of compounds **4** and **5** with the ASNS transition state, each isomer was placed in the synthetase active site in the same orientation of the transition state. Each model was solvated in a TIP3P octahedral box, followed by energy minimization using the CHARMM27 force field.⁵⁹

The diastereoisomer in which the methyl substituent on the stereogenic sulfur points towards the ammonia tunnel corresponds to the (*R*)-configuration of both compounds **4** and **5** (Figure 5-3). It is believed that the methyl group in the (*R*)-configuration mimics the ammonia molecule in its reaction with β -aspartyl-AMP (Figure 5-2.). Therefore, we believe that the (*R*)-configuration of both compounds **4** and **5** is the true inhibitor. The critical electrostatic interaction with Lys-449 was maintained in both diastereoisomers of compound **4**. The amino group of both isomers also makes a hydrogen bond with Glu-348, which is a key residue in the ASNS active site and is important in mediating the formation of the β -aspartyl-AMP intermediate.¹⁰⁴ However, the amino group in the stereoisomers of compound **4** undergoes electrostatic interactions with different carboxylate side chains.

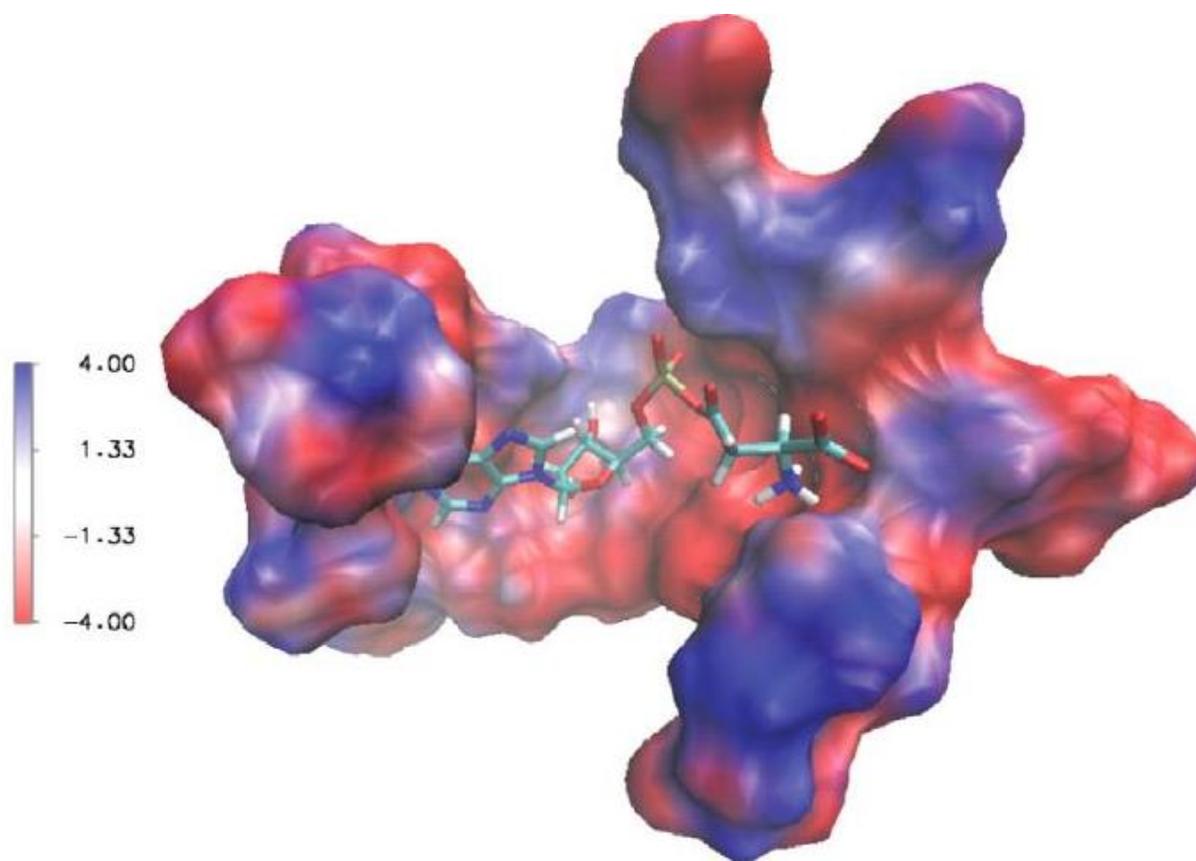


Figure 5-4. Electrostatic potential in the synthetase active site of the AS-B model rendered on the van der Waals surface of residues located within 0.4 nm of the β -aspartyl-AMP intermediate. Regions of positive and negative electrostatic potential are colored blue and red, respectively. The potential was computed using the Poisson-Boltzmann equation using all atoms in the uncomplexed protein, and was visualized using the VMD software package. Quantitative values at 300 K can be computed by multiplying those shown in the legend by 0.026 J/mol/C.). Reprinted with permission from Ikeuchi a, H.; Ahn, Y.; Otokawa, T.; Watanabe, B., Hegazy, L.; Hiratake, J.; Richards, N. G. J. *Bioorg. Med. Chem.* **2012**, 20, 5915.

In case of the carboxy *N*-phosphonosulfonimidoyl **5**, the key residue Glu-348 reorients and makes a hydrogen bond with other structures. The Lys-449 electrostatic interaction with the phosphate group is maintained only in the (*R*)-configuration.

Examination of the electrostatic potential map of the active site revealed that the active site has mainly negative electrostatic potential, with the exception of only two portions with positive potential stabilizing the negative charges of the phosphate and carboxylate groups (Figure 5-3). Therefore the higher affinity towards hASNS inhibition of compound **4** over compound **5** may be due to the overall charge and stability within the active site.

Experimental efforts to obtain compound **4** in an optically pure form were unsuccessful. Alternatively, computational methods will be used to estimate the free energy of binding. A free energy cycle is a widely used computational technique to calculate the difference in binding free energy of two different ligands to a given receptor. The thermodynamic cycle employed in this study is shown in Figure 5-5, which describes the binding of the (*R*)- and (*S*)- phosphonosulfonimidoyl inhibitor to ASNS. Free energy changes ΔG_3 and ΔG_4 describe the binding process and can be studied experimentally. However, they are very difficult to calculate computationally. Free energy changes ΔG_1 and ΔG_2 correspond to alchemical, nonphysical, mutation between the two isomers in the aqueous and enzyme-bound states. Since the free energy, ΔG , is a state function and therefore path independent, both the vertical and horizontal legs provide routes to the difference in the binding free energies, $\Delta\Delta G$, of the (*R*)- and (*S*)-*N*-phosphonosulfonimidoyl inhibitors to the enzyme (Equation 5-1).

$$\Delta\Delta G = \Delta G_1 - \Delta G_2 = \Delta G_3 - \Delta G_4 \quad (5-1)$$

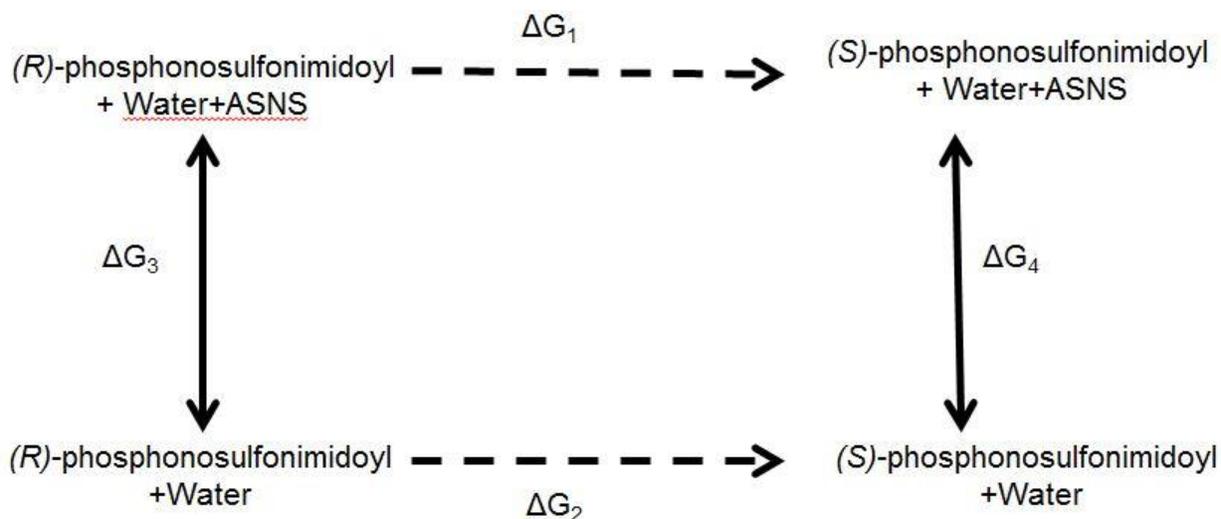


Figure 5-5. Thermodynamic cycle used for the calculation of relative free energy of binding of (*R*)- and (*S*)- diastereoisomers of the phosphonosulfonimidoyl inhibitor (compound **4**) to ASNS. ΔG_3 and ΔG_4 are the binding free energy changes of the two diastereoisomers to ASNS, while ΔG_1 and ΔG_2 correspond to the two hypothetical reactions that convert the (*R*)-phosphonosulfonimidoyl into (*S*)- phosphonosulfonimidoyl when bound with ASNS and in aqueous environment respectively. The direction of the alchemical transformation (dotted) arrows indicates that the change in free energy is calculated by taking the difference between the free energy of the state at the end and that at the beginning.

The free energy change for the alchemical transformation of the (*R*)- isomer into the (*S*)- isomer is calculated by the introduction of intermediate hybrid states coupled by an independent variable λ . The chimeric change is performed by the gradual vanishing of specific atoms in the initial state, (*R*)- configuration ($\lambda=0$), and the gradual appearance of the necessary atoms in the (*S*)- configuration through intermediate states, where the changes are assigned $\Delta\lambda$ until the final state is reached ($\lambda=1$). The free-energy change, ΔG_i , between any two successive states is given by:

$$\Delta G_i = k_B T \ln \left\langle \exp \left(\frac{H_{\lambda+\Delta\lambda} - H_{\lambda}}{k_B T} \right) \right\rangle \quad (5-2)$$

where the ensemble average is calculated as a time average over the Hamiltonian trajectory, $H(\lambda)$. The total free energy change, ΔG , of the initial state ($\lambda = 0$) to the final state ($\lambda = 1$) of transmutation is computed by summing the incremental free energy changes in each window encompassing λ from 0 to 1,

$$\Delta G = \sum_i (G(\lambda_i + \Delta\lambda) - G(\lambda_i)) \quad (5-3)$$

To avoid convergence errors, the initial and final states should be very similar and $\Delta\lambda$ should be kept small. Therefore, the alchemical pathway for the transformation of the (*R*)- isomer into the (*S*)- isomer will be performed on several steps (Figure 5-6): (1) In the first step (S1) of transformation, the methyl group of the (*R*)- inhibitor will be mutated into a hydroxyl group, (2) In S2, The hydroxyl group will be deprotonated and mutated into a negatively charged oxygen; (3) In S3, the negatively charged oxygen of the *N*-phosphonosulfonimidoyl inhibitor will be mutated into a hydroxyl group in the (*S*)- position of the stereospecific sulfur; (4) In the final step, S4, the hydroxyl group at the (*S*)- position, will be mutated into a methyl group at the (*S*)- position. The calculated free energies converge better when the atoms disappear instead of being added. Therefore,

the transmutation step (S4) will be computed in the reverse direction, i.e, the methyl group on the (S)- position will be mutated into a hydroxyl group on the (S)- position and the calculated free energy sign will be reversed when calculating the total transmutation free energy. The total transmutation free energy is the summing of the ΔG for each mutation step. We propose that the free energy of deprotonation of the hydroxyl derivatives of both diastereoisomers are equal in either aqueous or enzyme-bound states and their values cancel out when calculating the total transmutation free energies (steps S2 in the forward direction and S3 in the backward direction). Thus, only two transmutation step calculations will be required (steps S1 in the forward direction and S4 in the backward direction).

The free energy perturbation calculations will be carried out using the TSM module in CHARMM. Detailed parameter optimization for the *N*-phosphonosulfonimidoyl functional group for accurate representation of its chemical and physical behavior using computational calculations is explained in detail in Chapter 2. Also, both the (*R*)- and (*S*)- methyl derivatives of the *N*-phosphonosulfonimidoyl inhibitor (compounds **4** and **5**) have been modeled and equilibrated following the same equilibration procedure described in Chapter 3. However, the presence of the hydroxyl-substituent of the *N*-phosphonosulfonimidoyl inhibitor requires careful parameter validation and optimization. Parameters that represent the hydroxyl-derivative of the *N*-phosphonosulfonimidoyl functional group will be calculated following the same parameter optimization protocol described in detail in Chapter 2.

The free energy change in each mutation step (S1 and S4) between $\lambda=0$ and $\lambda=1$ will be calculated using Equation 5-3. A very small step increment ($\Delta\lambda = 0.05$) will be

used with equilibration and simulation of each chimeric state. The difference in binding free energy will be computed using Equation 5-1.

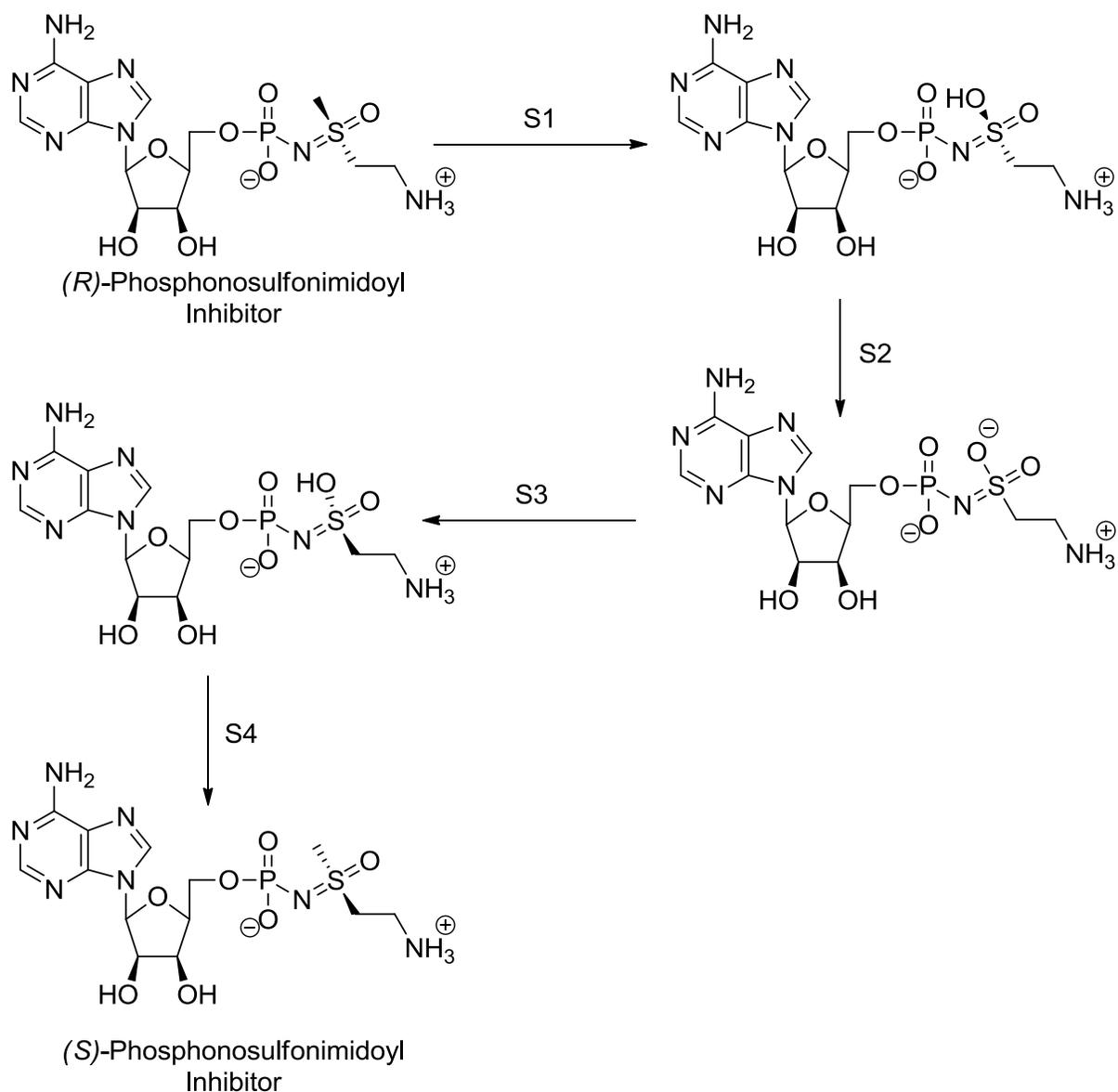


Figure 5-6. Pathway for the transmutation between the (*R*)- and (*S*)-phosphonosulfonimidoyl inhibitor. Free energy perturbation calculations for S1 will be performed in the forward direction. Free energy perturbation calculations for S4 will be performed in the backward direction. S2 and S3 free energies are proposed to cancel each other.

APPENDIX A

QM AND MM OPTIMIZED VIBRATIONAL SPECTRA FOR *N*-PHOSPHONOSULFONIMIDOYL MODEL COMPOUND.Table A-1. Vibrational spectra computed for the *N*-phosphonosulfonimidoyl molecule at the scaled MP2 level and with optimized CGenFF^a parameters

MP2/6-31+G(d) scaled by a factor 0.943				CGenFF			
Freq ^b	Assign (%)	Assign (%)	Assign (%)	Freq ^b	Assign (%)	Assign (%)	Assign (%)
62.6	tdOPNS (98)	tdNSC9H (7)		41.1	tdOPNS (89)	tdC4OPN (6)	
81.4	tdPNSC9 (53)	tdC4OPN (30)	scO3PN8 (11)	70.6	tdPNSC9 (59)	tdC4OPN (38)	
100.2	tdNSC9H (70)	tdPNSC9 (15)		85.7	tdNSC9H (50)	tdC4OPN (30)	tdPNSC9 (9)
114.9	tdC4OPN (52)	tdPNSC9 (32)		99.6	tdPNSC9 (38)	tdNSC9H (29)	dPNS (18)
151.4	dPNS (35)	tdPOC4H (20)	dC9SN (18)	154.9	dPNS (51)	tdPOC4H (26)	dC9SN (17)
168.4	tdPOC4H (67)	dPNS (16)	tdPNSC9 (7)	182.8	tdPOC4H(24)	dPNS (24)	dC13SN (14)
196.1	scSC9C19 (35)	dPNS (24)	dC9SN (17)	207.2	dC9SN (29)	twPO2O7 (23)	rPO2O7 (14)
218.1	tdSCC19H (65)	dPOC (7)	dC13SC9 (5)	219.1	rPO2O7 (16)	twPO2O7 (16)	scSC9C19 (16)
224.7	dPOC (35)	twPO2O7 (14)	tdSCC19H(11)	229.7	tdSCC19H(38)	tdNSC13H (28)	tdPOC4H (9)
250.6	dC13SN (36)	twPO2O7 (28)	dC13SC9 (9)	240.3	dC13SN (36)	tdNSC13H (12)	wPO2O7 (12)
259.7	tdNSC13H (84)			258.2	tdNSC13H(35)	tdSCC19H (30)	dPOC (10)
281.6	dC13SC9 (25)	twPO2O7 (14)	rSO' (11)	277.6	dPOC (52)	twPO2O7 (13)	tdNSC13H (10)
290.7	scSC9C19 (25)	scO3PN8 (10)	rSO' (8)	285.8	scSC9C19(37)	rSO (20)	sC9S (10)
319.2	dC13SC9 (25)	sN8P (21)	dC9SN (10)	335.8	dC13SC9 (62)	rSO (13)	wPO2O7 (4)
340.0	scO3PN8 (30)	rSO' (20)	rSO (14)	376.6	scO3PN8 (36)	dC13SN (20)	wPO2O7 (14)
387.2	scPO2O7 (16)	wPO2O7 (15)	dC9SN (11)	383.3	rSO' (54)	rSO (14)	
419.6	dC13SN (29)	rSO (24)	twPO2O7 (7)	389.7	dC13SN (24)	wPO2O7 (20)	scSC9C19 (14)
450.1	wPO2O7 (24)	rSO' (13)	rPO2O7 (11)	418.2	scO3PN8 (29)	rSO (18)	wPO2O7 (13)
491.1	scPO2O7 (37)	dPOC (18)	scO3PN8 (15)	455.3	scPO2O7 (31)	rPO2O7 (25)	dPOC (13)
515.6	wPO2O7 (36)	rPO2O7 (25)	scPO2O7 (14)	482.4	scPO2O7 (37)	dPOC (19)	twPO2O7 (14)
594.5	sC9S (34)	sC13S (33)	sN8P (12)	580.2	sN8P (27)	sC9S (17)	sC13S (13)
669.2	sO3P7 (67)	wPO2O7 (5)		645.3	sC9S (31)	sO3P7 (17)	sC13S (9)
701.5	sC13S (28)	sC9S (25)	rCH2C9 (11)	686.5	sO3P7 (48)	sN8P (16)	wPO2O7 (14)
720.7	sN8P (34)	sC13S (13)	rCH2C9 (11)	719.7	sC13S (58)	sC9S (17)	rCH2C9 (5)

Table A-1. Continued.

MP2/6-31+G(d) scaled by a factor 0.943				CGenFF			
Freq ^b	Assign (%)	Assign (%)	Assign (%)	Freq ^b	Assign (%)	Assign (%)	Assign (%)
781.4	rCH2C9 (34)	rCH3C19' (23)	sC13S (15)	821.5	rCH2C9 (70)	rCH3C19' (14)	tdNSC9H (5)
958.7	rCH3C13 (58)	rCH3C19 (11)	rCH3C13' (6)	941.0	rCH3C13' (81)	sN8S (9)	
964.6	rCH3C13' (53)	sC9C19 (21)	rCH3C19 (6)	973.7	rCH3C13 (87)	adCH3C13 (6)	
977.3	sC9C19 (19)	rCH3C13 (19)	rCH3C19 (15)	993.5	rCH3C19 (42)	sC9C19 (30)	wCH2C9 (9)
1003.4	ssPO (81)			1003.2	ssPO (66)	sO3C4 (8)	
1028.4	sO3C4 (59)	sN8S (14)	sS=O (12)	1027.7	rCH3C19' (45)	sN8S (14)	rCH3C13' (7)
958.7	rCH3C13 (58)	rCH3C19 (11)	rCH3C13' (6)	941.0	rCH3C13' (81)	sN8S (9)	
964.6	rCH3C13' (53)	sC9C19 (21)	rCH3C19 (6)	973.7	rCH3C13 (87)	adCH3C13 (6)	
977.3	sC9C19 (19)	rCH3C13 (19)	rCH3C19 (15)	993.5	rCH3C19 (42)	sC9C19 (30)	wCH2C9 (9)
1003.4	ssPO (81)			1003.2	ssPO (66)	sO3C4 (8)	
1028.4	sO3C4 (59)	sN8S (14)	sS=O (12)	1027.7	rCH3C19' (45)	sN8S (14)	rCH3C13' (7)
1037.4	sO3C4 (34)	sS=O (25)	sN8S (12)	1038.1	sO3C4 (67)	sdCH3C4 (9)	ssPO (8)
1048.5	sC9C19 (43)	rCH3C19 (34)	scSC9C19(7)	1043.8	sC9C19 (34)	rCH3C19' (31)	rCH3C19 (11)
1072.9	rCH2C9 (18)	rCH3C19' (17)	sN8S (17)	1052.7	sN8S (29)	sC9C19 (15)	rCH3C19 (12)
1135.9	rCH3C4 (81)	r'CH3C4 (15)		1142.7	rCH3C4 (74)	ad'CH3C4 (17)	adCH3C4 (10)
1154.2	r'CH3C4 (77)	rCH3C4 (14)		1154.0	r'CH3C4 (76)	adCH3C4 (17)	
1181.1	sS=O (42)	sN8S (35)		1175.2	saPO (92)	rPO2O7 (5)	
1212.8	saPO (82)	twCH2C9 (5)		1222.7	twCH2C9 (74)	sS=O (10)	
1246.0	twCH2C9 (60)	rCH3C19 (19)		1244.2	sS=O (66)	twCH2C9 (13)	
1287.3	wCH2C9 (56)	sdCH3C13(28)		1365.1	wCH2C9 (73)	scCH2C9 (11)	
1311.3	sdCH3C13(67)	wCH2C9 (22)		1400.4	sdCH3C13(80)	adCH3C13'(10)	
1374.5	sdCH3C19(94)			1416.4	sdCH3C19(79)	adCH3C19' (8)	
1418.7	sdCH3C4 (90)	scCH2C9 (6)		1424.4	adCH3C19'(49)	adCH3C19(26)	sdCH3C19(15)
1423.7	adCH3C13(82)	sdCH3C4 (5)	rCH3C13 (5)	1432.5	ad'CH3C4 (66)	adCH3C4 (17)	r'CH3C4 (12)
1432.0	scCH2C9 (93)	adCH3C13' (6)		1438.5	adCH3C13(37)	adCH3C13'(26)	adCH3C19'(15)
1454.5	adCH3C13'(86)	scCH2C9 (6)		1444.3	adCH3C19(33)	adCH3C19'(29)	adCH3C13'(21)

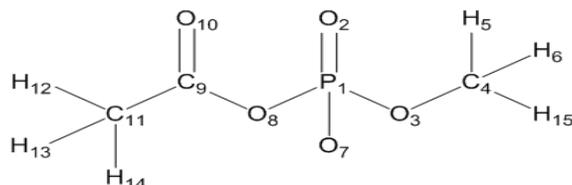
Table A-1. Continued.

MP2/6-31+G(d) scaled by a factor 0.943				CGenFF			
Freq ^b	Assign (%)	Assign (%)	Assign (%)	Freq ^b	Assign (%)	Assign (%)	Assign (%)
1462.6	adCH3C4 (80)	ad'CH3C4 (15)		1458.8	scCH2C9 (92)	wCH2C9 (8)	
1472.7	adCH3C19(70)	adCH3C19'(21)	rCH3C19'(6)	1462.4	adCH3C13(44)	adCH3C13'(36)	sdCH3C13 (8)
1480.4	adCH3C19'(71)	adCH3C19(21)	rCH3C19 (6)	1484.1	adCH3C4 (55)	rCH3C4 (21)	ad'CH3C4 (13)
1489.3	ad'CH3C4 (80)	adCH3C4 (14)		1615.4	sdCH3C4 (89)	sO3C4 (10)	
2894.2	ssCH3C4 (98)			2847.9	ssCH3C13(57)	saCH3C13'(41)	
2918.0	ssCH3C19 (94)			2854.0	ssCH3C4 (99)		
2924.6	ssCH3C13 (68)	saCH3C13'(31)		2855.9	ssCH3C9 (98)		
2945.5	ssCH3C9 (99)			2894.1	saCH3C9 (99)		
2974.9	saCH3C4 ' (87)	saCH3C4 (11)		2902.4	ssCH3C19 (98)		
2992.8	saCH3C19(83)	saCH3C19 ' (9)		2913.4	saCH3C4' (73)	saCH3C4 (26)	
2995.1	saCH3C4 (88)	saCH3C4 ' (12)		2914.9	saCH3C13(81)	saCH3C13'(20)	
3005.8	saCH3C9 (81)	saCH3C19'(13)	saCH3C19(9)	2918.5	saCH3C4 (74)	saCH3C4' (26)	
3023.4	saCH3C19'(76)	saCH3C9 (17)		2922.1	ssCH3C13(45)	saCH3C13'(41)	saCH3C13(18)
3038.8	saCH3C13(99)			2958.2	saCH3C19(98)		
3043.7	saCH3C13'(67)	ssCH3C13(27)		2960.8	saCH3C19'(98)		

^a Optimized vibrational contributions from interactions for which parameters have been developed in this study are shown in bold font; s stands for bond stretching with the variations ss and sa for symmetric and asymmetric stretching, respectively; d means angle deformation with the variations sd and ad for symmetric and asymmetric deformation, respectively; td and ti stand for torsional and improper torsion deformation, respectively; sc stands for scissoring; r for rocking; w for wagging and tw for twisting. ^b Frequencies are expressed in units of cm⁻¹.

APPENDIX B
CHARMM TOPOLOGY AND PARAMETER FILE FOR ACYLPHOSPHATE AND N-
PHOSPHONOSULFONIMIDOYL MODEL COMPOUNDS

CHARMM Topology and Parameter File for Acylphosphate Model



```
* Toppar stream file generated by
* CHARMM General Force Field (CGenFF) program version 0.9.6 beta
* For use with CGenFF version 2b7
*
* Topologies generated by
* CHARMM General Force Field (CGenFF) program version 0.9.6 beta
*
36 1
```

```
!optimized parameters for acetyl phosphate (Lamees Hegazy)
! OG305 should be assigned the same mass and Lennard Jones parameters
! by the user in the main CGenFF topology and parameter files
```

```
RESI ampa      -1.000
GROUP          ! CHARGE
ATOM P1        PG1      1.300 !
ATOM O2        OG2P1   -0.710 !
ATOM O3        OG303   -0.460 !
ATOM C4        CG331   -0.170 !
ATOM H5        HGA3     0.090 !
ATOM H6        HGA3     0.090 !
ATOM O7        OG2P1   -0.710 !
ATOM O8        OG305   -0.380 !New atom type,same VdW parameters as OG303
ATOM C9        CG2O2    0.340 !
ATOM O10       OG2D1   -0.480 !
ATOM C11       CG331   -0.270 !
ATOM H12       HGA3     0.090 !
ATOM H13       HGA3     0.090 !
ATOM H14       HGA3     0.090 !
ATOM H15       HGA3     0.090 !

BOND P1 O2
BOND P1 O3
BOND P1 O7
BOND P1 O8
BOND O3 C4
```

BOND C4 H5
BOND C4 H6
BOND C4 H15
BOND O8 C9
BOND C9 O10
BOND C9 C11
BOND C11 H12
BOND C11 H13
BOND C11 H14
IMPR C9 C11 O10 O8

END
read param card flex append

BONDS

CG202 OG305 230.00 1.3400 !
OG305 PG1 170.00 1.7800 !

ANGLES

CG331 CG202 OG305 30.00 104.00 05.00 2.32600 !
OG2D1 CG202 OG305 70.00 118.00 160.00 2.25760 !
CG202 OG305 PG1 70.00 121.50 !
OG2P1 PG1 OG305 60.00 103.000 !
OG303 PG1 OG305 60.00 90.80 !

DIHEDRALS

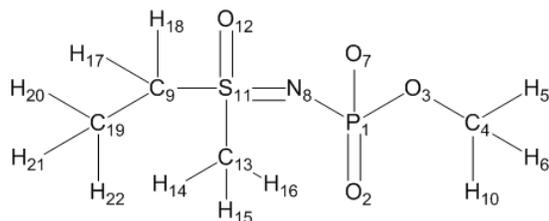
OG305 CG202 CG331 HGA3 0.0000 3 0.00 !
CG331 CG202 OG305 PG1 3.8000 1 180.00 !
CG331 CG202 OG305 PG1 1.6000 2 180.00 !
OG2D1 CG202 OG305 PG1 1.3000 1 180.00 !
OG2D1 CG202 OG305 PG1 2.6000 2 180.00 !
CG202 OG305 PG1 OG2P1 0.1000 3 0.00 !
CG202 OG305 PG1 OG303 0.1000 2 180.00 !
CG202 OG305 PG1 OG303 0.1000 3 0.00 !
CG331 OG303 PG1 OG305 1.4700 2 0.00 !
CG331 OG303 PG1 OG305 0.7000 3 0.00 !

IMPROPERS

CG202 CG331 OG2D1 OG305 56.0000 0 0.00 !

END
RETURN

CHARMM Topology and Parameter File for *N*-phosphonosulfonimidoyl Model



```

* Toppar stream file generated by
* CHARMM General Force Field (CGenFF) program version 0.9.6 beta
* For use with CGenFF version 2b7
*
* Topologies generated by
* CHARMM General Force Field (CGenFF) program version 0.9.6 beta
*
36 1
!optimized parameters for N-phosphonosulfonimidoyl functional group

```

```

RESI msxn -1.000 !
GROUP !
ATOM P1 PG1 0.200 !
ATOM O2 OG2P1 -0.460 !
ATOM O3 OG303 -0.280 !
ATOM C4 CG331 -0.170 !
ATOM H5 HGA3 0.090 !
ATOM H6 HGA3 0.090 !
ATOM O7 OG2P1 -0.460 !
ATOM N8 NG2D1 -0.380 !
ATOM C9 CG321 0.020 !
ATOM H10 HGA3 0.090 !
ATOM S11 SG302 0.120 ! !
ATOM O12 OG2P1 -0.420 !
ATOM C13 CG331 0.110 !
ATOM H14 HGA3 0.090 !
ATOM H15 HGA3 0.090 !
ATOM H16 HGA3 0.090 !
ATOM H17 HGA2 0.090 !
ATOM H18 HGA2 0.090 !
ATOM C19 CG331 -0.270 !
ATOM H20 HGA3 0.090 !
ATOM H21 HGA3 0.090 !
ATOM H22 HGA3 0.090 !

```

```

BOND P1 O2
BOND P1 O3
BOND P1 O7
BOND P1 N8
BOND O3 C4
BOND C4 H5
BOND C4 H6
BOND C4 H10

```

BOND N8 S11
 BOND C9 S11
 BOND C9 H17
 BOND C9 H18
 BOND C9 C19
 BOND S11 O12
 BOND S11 C13
 BOND C13 H14
 BOND C13 H15
 BOND C13 H16
 BOND C19 H20
 BOND C19 H21
 BOND C19 H22

END

read param card flex append

BONDS

NG2D1 PG1 100.00 1.7220 !
 NG2D1 SG3O2 400.00 1.5320 !

ANGLES

PG1 NG2D1 SG3O2 30.00 113.00 !
 NG2D1 PG1 OG2P1 50.00 106.00 !
 NG2D1 PG1 OG3O3 94.00 98.80 !
 CG321 SG3O2 NG2D1 65.00 114.30 !
 CG331 SG3O2 NG2D1 79.00 114.00 !
 NG2D1 SG3O2 OG2P1 65.00 119.00 !

DIHEDRALS

CG331 CG321 SG3O2 NG2D1 0.0010 3 0.00 !
 CG331 CG321 SG3O2 NG2D1 1.4000 1 180.00 !
 HGA2 CG321 SG3O2 NG2D1 0.1600 3 0.00 !
 HGA3 CG331 SG3O2 NG2D1 0.1800 3 0.00 !
 !SG3O2 NG2D1 PG1 OG2P1 0.5000 2 0.00 !
 SG3O2 NG2D1 PG1 OG2P1 0.5000 4 0.00 !
 SG3O2 NG2D1 PG1 OG3O3 3.0000 2 0.00 !
 SG3O2 NG2D1 PG1 OG3O3 1.8000 1 0.00 !
 CG321 SG3O2 NG2D1 PG1 1.00 2 0.0
 OG2P1 SG3O2 NG2D1 PG1 1.00 2 0.0
 OG2P1 SG3O2 NG2D1 PG1 2.50 1 180.0
 CG331 SG3O2 NG2D1 PG1 0.60 2 0.0
 CG331 SG3O2 NG2D1 PG1 1.00 1 0.0
 CG331 OG3O3 PG1 NG2D1 0.4000 1 0.00 !
 CG331 OG3O3 PG1 NG2D1 0.8000 2 0.00 !
 CG331 OG3O3 PG1 NG2D1 0.3500 3 0.00 !

IMPROPERS

END

RETURN

LIST OF REFERENCES

1. Larsen, T. M.; Boehlein, S. K.; Schuster, S. M.; Richards, N. G. J.; Thoden, J. B.; Holden, H. M.; Rayment, I. *Biochemistry* **1999**, *38*, 16146.
2. Zalkin, H. a.; Smith, J. L. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1998**, *72*, 87.
3. Badet, B.; Vermoote, P.; Haumont, P. Y.; Lederer, F.; Le Goffic, F. *Biochemistry* **1987**, *26*, 1940.
4. Tso, J. Y.; Zalkin, H.; Van Cleemput, M.; Yanofsky, C.; Smith, J. M. *J. Biol. Chem.* **1982**, *257*, 3525.
5. Van den Heuvel, R. H. H.; Curti, B.; Vanoni, M. A.; Mattevi, A. *Cell. Mol. Life Sci.* **2004**, *61*, 669.
6. Aslanian, A. M.; Fletcher, B. S.; Kilberg, M. S. *Biochem. J.* **2001**, *357*, 321.
7. Cooney, D. A.; Handschumacher, R. E. *Annu. Rev. Pharmacol.* **1970**, *10*, 421.
8. Chakrabarti, R.; Schuster, S. M. *Int. J. Pediatr. Hemotol./Oncol.* **1997**, *4*, 597.
9. Chakrabarti, R.; Wylie, D. E.; Schuster, S. M. *J. Biol. Chem.* **1989**, *264*, 15494.
10. Muller, H. J., BooS, J. *Crit. Rev. Oncol. Hematol.* **1998**, *28*, 97.
11. Pieters, R.; Hunger, S. P.; Boos, J.; Rizzari, C.; Silverman, L.; Baruchel, A. *Cancer* **2011**, *117*, 238.
12. Ueno, T.; Ohtawa, K.; Mitsui, K.; Kodera, Y.; Hiroto, M.; Matsushima, A.; Inada, Y.; Nishimura, H. *Leukemia* **1997**, *11*, 1858.
13. Richards, N. G.; Kilberg, M. S. *Annu. Rev. Biochem.* **2006**, *75*, 629.
14. Tallal, L.; Tan, C.; Oettgen, H.; Wollner, N.; McCarthy, M.; Helson, L.; Burchenal, J.; Karnofsky, D.; Murphy, M. L. *Cancer* **1970**, *25*, 306.
15. Barr, R.D.; DeVeber, L.L.; Pai, K.M.; Andrew, M.; Halton, J.; Cairney, A. E.; Whitton, A. C. *Am. J. Pediatr. Hematol.Oncol.* **1992**, *14*, 136.
16. Terebelo, H. R.; Anderson, K.; Wiernik, P. H.; Cuttner, J.; Cooper, R. M.; Faso, L.; Berenberg, J. L. *Am. J. Clin. Oncol.* **1986**, *9*, 411.
17. Lobel, J. S.; O'Brien, R. T.; McIntosh, S.; Aspnes, G. T.; Capizzi, R. L. *Cancer* **1979**, *43*, 1089.
18. Prager, M. D.; Bachynsky, N. *Biochem. Biophys. Res. Commun.* **1968**, *31*, 43.

19. Kiryama, Y.; Kubota, M.; Takimoto, T.; Kitoh, T.; Tanizawa, A.; Akiyama, Y.; Mikawa, H. *Leukemia* **1989**, *3*, 294.
20. Leslie, M.; Case, M. C.; Hall, A. G.; Coulthard, S. A. *Br. J. Haematol.* **2006**, *132*, 740.
21. Manning, J. M.; Moore, S.; Rowe, W. B.; Meister, A. *Biochemistry* **1969**, *8*, 2681.
22. Koizumi, M.; Hiratake, J.; Nakatsu, T.; Kato, H.; Oda, J. I. *J. Am. Chem. Soc.* **1999**, *121*, 5799.
23. Lorenzi, P. L.; Reinhold, W. C.; Rudelius, M.; Gunsior, M.; Shankavaram, U.; Bussey, K. J.; Scherf, U.; Eichler, G. S.; Martin, S. E.; Chin, K.; Gray, J. W.; Kohn, E. C.; Horak, I. D.; Von Hoff D. D.; Raffeld M.; Goldsmith P. K.; Caplen, N. J.; Weinstein, J. N. *Mol. Cancer Ther.* **2006**, *5*, 2613.
24. Lorenzi, P. L.; Llamas, J.; Gunsior, M.; Ozbun, L.; Reinhold, W. C.; Varma, S.; Ji, H.; Kim, H.; Hutchinson, A. A.; Kohn, E. C.; Goldsmith, P. K.; Birrer, M. J.; Weinstein, J. N. *Mol. Cancer Ther.* **2008**, *7*, 3123.
25. Sircar, K.; Huang, H.; Hu, L. M.; Cogdell, D.; Dhillon, J.; Tzelepim, V.; Efstathiou, E.; Koumakpayi, I. H.; Saad, F.; Luo, D. J.; Bismar, T. A.; Aparicio, A.; Troncoso, P.; Navone, N.; Zhang, W. *Am. J. Pathol.* **2012**, *180*, 893.
26. Oinonen, C.; Rouvinen, J. *Protein Sci.* **2000**, *9*, 2329.
27. Brannigan, J. A.; Dodson, G.; Duggleby, H. J.; Moody, P. C. E.; Smith, J. L.; *Nature* **1998**, *378*, 416.
28. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.
29. Lemke, C. T.; Howell, P. L. *Structure* **2001**, *9*, 1153.
30. Tesmer, J. J. G.; Klem, T. J.; Deras, M. L.; Davisson, V. J.; Smith, J. L. *Nat. Struct. Biol.* **1996**, *3*, 74.
31. Bachmann, B. O.; Li, R.; Townsend, C. A. *Proc. Natl. Acad. Sci. USA.* **1998**, *95*, 9082.
32. Miller, M. T.; Gerratana, B.; Stapon, A.; Townsend, C. A.; Rosenzweig, A. C. *J. Biol. Chem.* **2003**, *278*, 40996.
33. Horowitz, B.; Meister, A. *J. Biol. Chem.* **1972**, *247*, 6708.
34. Boehlein, S.K.; Schuster, S.M.; Richards, N.G.J. *Biochemistry* **1996**, *35*, 3031.
35. Boehlein, S. K.; Stewart, J. D.; Walworth, E. S.; Thirumoorthy, R.; Richards, N. G. J.; Schuster, S. M. *Biochemistry* **1998**, *37*, 13230.

36. Kim, J.H.; Krahn, J.M.; Tomchick, D.R.; Smith, J.L.; Zalkin, H. *J. Biol. Chem.* **1996**, *271*, 15549.
37. Van den Heuvel, R. H. H.; Ferrari, D.; Bossi, R. T.; Ravasio, S.; Curti, B.; Vanoni, M. A.; Florencio, F.J.; Mattevi, A.; *J. Biol. Chem.* **2002**, *277*, 24579.
38. Boehlein, S. K.; Stewart, J. D.; Walworth, E. S.; Thirumoorthy, R.; Richards, N. G. J.; Schuster, S. M. *Biochemistry* **1998**, *37*, 13230.
39. Hartman, S. C.; Buchanan, J. M. *J. Biol. Chem.* **1958**, *233*, 451.
40. Abrahms, R.; Bentley, M. *Arch. Biochem. Biophys.* **1959**, *79*, 91.
41. Moyed, H. S.; Magasinik, H. S. *J. Biol. Chem.* **1957**, *226*, 351.
42. Miller, M. T.; Bachmann, B. O.; Townsend, C. A.; Rosenzweig, A. C. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14752.
43. Fresquet, V.; Thoden, J. B.; Holden, H. M.; Raushel, F. M. *Bioorg. Chem.* **2004**, *32*, 63.
44. Bork, P.; Koonin, E. V. *Proteins Struct. Funct. Genet.* **1994**, *20*, 347.
45. Tesson, A. R.; Soper, T. S.; Ciustea, M.; Richards, N.G.J. *Arch. Biochem. Biophys.* **2003**, *413*, 23.
46. Boehlein, S. K.; Walworth, E. S.; Richards, N. G. J.; Schuster, S. M. *J. Biol. Chem.* **1994**, *269*, 26789.
47. Boehlein, S. K.; Nakatsu, T.; Hiratake, J.; Thirumoorthy, R.; Stewart, J.D.; Richards, N.G.J.; Schuster, S.M. *Biochemistry* **2001**, *40*, 11168.
48. Boehlein, S.K.; Walworth, E.S.; Richards, N.G.J.; Schuster, S.M. *J. Biol. Chem.* **1997**, *272*, 12384.
49. Gutierrez, J. A.; Pan, Y. X.; Koroniak, L.; Hiratake, J.; Kilberg, M. S.; Richards, N. G. J. *Chem. Biol.* **2006**, *13*, 1339.
50. Koroniak, L.; Ciustea, M.; Gutierrez, J. A.; Richards, N. G. *Org. Lett.* **2003**, *5*, 2033.
51. Ikeuchi, H.; Meyer, M. E.; Ding, Y.; Hiratake, J.; Richards, N. G. *Bioorg. Med. Chem.* **2009**, *17*, 6641.
52. Hegazy, L.; Richards, N. G. J. *J. Mol. Model.* **2013**, doi: 10.1007/s00894-013-1990-x

53. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cul, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
54. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M. Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. *J. Comput. Chem.* **2005**, *26*, 1668.
55. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
56. Kaminski, G.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**; *105*, 6474.
57. Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. *J. Comput. Chem.* **2001**, *22*, 1340.
58. Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1955.
59. Ikeuchi a, H.; Ahn, Y.; Otokawa, T.; Watanabe, B., Hegazy, L.; Hiratake, J.; Richards, N. G. J. *Bioorg. Med. Chem.* **2012**, *20*, 5915.
60. Agarwal. V.; Nair, S. K. *Med. Chem. Commun.* **2012**, *3*, 887.
61. LaRonde-LeBlanc, N.; Resto, M.; Gerratana, B. *Nat. Struct. Mol. Biol.* **2009**, *16*, 421.
62. Duckworth, B. P.; Nelson, K. M.; Aldrich, C. C. *Curr. Top. Med. Chem.* **2012**, *12*, 766.
63. Curnow, A. W.; Hong, K. W.; Yuan, R.; Kim, Sl.; Martins, O.; Winkler, W.; Henkin, T. M.; Soll, D. *Proc. Natl. Acad. Sci. USA* **1997**, *94*.11819.
64. Horiuchi, K. Y.; Harpel, M. R.; Shen, L.; Luo, Y.; Rogers, K. C.; Copeland, R. A. *Biochemistry* **2001**, *40*: 6450.
65. Oshikane, H.; Sheppard, K.; Fukai, S.; Nakamura, Y.; Ishitani, R.; Numata, T.; Sherrer, Rl.; Feng, L.; Schmitt, E.; Panvert, M.; Blanquet, S.; Mchulam, Y.; Soll, D.; Nureki, O. *Science* **2006**, *312*, 1950.
66. Becker, O. M., Mackerell, A. D., Roux, B.; Watanabe, M., Ed. *Computational Biochemistry and biophysics*; Marcel Dekker: New York, USA, 2001.

67. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; E. Darian, O. Lopes, P.; Vorobyov, I; Mackerell, A. D., Jr. *J. Comput. Chem.* **2010**, *31*, 671.
68. Vanommeslaeghe, K.; Mackerell, A. D., Jr. *J. Chem. Inf. Model.* **2012**, *52*, 3144.
69. Retrieved from <https://www.paramchem.org/>
70. Vanommeslaeghe, K.; Raman, E. P.; Mackerell, A. D., Jr. *J. Chem. Inf. Model.* **2012**, *52*, 3155.
71. Kuczera, K.; Wiorkiewicz, J.; Karplus, M. *CHARMM. Harvard University.* **1993**.
72. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb., *Gaussian, Inc., Wallingford CT.* **2009**.
73. Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D, J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654.
74. Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.
75. Jorgensen, w. I.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
76. Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
77. Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
78. Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.
79. Allen, M. P., Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
80. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
81. Shuichi, N.; Klein, M. L. *Mol. Phys.* **1983**, *50*, 1055.
82. Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E., *J. Am. Chem. Soc.* **1979**, *101*, 2550.
83. Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
84. Vinson, V. J. *Science* **2009**, *324*, 197.
85. Karplus, M.; McCammon, A. *Nat. Struct. Biol.* **2002**, *9*, 646.
86. Lee, G. M.; Craik, C. S. *Science* **2009**, *324*, 213.

87. Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, 34, 43.
88. McGovern, S. L.; Shoichet, B. K. *J. Med. Chem.* **2003**, 46, 2895.
89. Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. *J. Med. Chem.* **2004**, 47, 45.
90. Murray, C. W.; Baxter, C. A.; Frenkel, A. D. *J. Comput. Aided Mol. Des.*, **1999**, 13, 547.
91. Sugita, Y., Okamoto, Y. *Chem. Phys. Lett.*, **1999**, 314,141.
92. Meyer, M. E. Inhibition, Characterization, and Crystallization of Glutamine-Dependent Asparagine Synthetase, University of Florida, Gainesville, FL, 2010.
93. Miller, M. T.; Bachmann, B. O.; Townsend, C. A.; Rosenzweig, A. C. *Nature* **2001**, 8, 684.
94. Ding, Y. Experimental and Theoretical Characterization of Asparagine Synthetase Inhibitors. Ph.D. Dissertation, University of Florida, Gainesville, FL, 2002.
95. Mcquarrie, D. A. *Statistical mechanics*; University Science Books: Sausalito, CA, USA, 2000.
96. Cramer, C. J. *Essentials of computational chemistry*, 2nd ed.; John Wiley & Sons: England, U.K., 2004.
97. Hoover, W. G. *Phys. Rev. A* **1985**, 31, 1695.
98. Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, 61, 704.
99. Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins* **2008**, 73, 765.
100. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, 7, 525.
101. Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, 7, 2284.
102. Darden, T. A.; York, D.; Pedersen, L. G. *J. Chem. Phys.* **1993**, 98, 10089.
103. Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, 15, 667.
104. Meyer, M. E.; Gutierrez, J. A.; Raushel, F. M.; Richards, N. G. *J. Biochemistry* **2010**, 49, 9391.
105. Bachmann, B. O.; Townsend, C. A. *Biochemistry* **2000**, 39, 11187.

106. Raber, M. L.; Arnett, S. O.; Townsend, C. A.; *Biochemistry* **2009**, *48*, 4959.
107. Raber, M. L.; Arnett, S. O.; Townsend, C. A. *J. Biol. Chem.* **2009**, *284*, 217.
108. Raber, M. L.; Castillo, A.; Greer, A.; Townsend, C. A. *ChemBioChem.* **2009**, *10*, 2904.
109. Waszkowycz, B.; Perkins, T. D. J.; Sykes, R. A.; Li, J., *IBM Syst. J.* **2001**, *40*, 360.
110. Alonso, H.; Bliznyuk, A. A.; Gready, J. E., *Med. Res. Rev.* **2006**, *26*, 531.
111. Murray, C. W.; Baxter, C. A.; Frenkel, A. D., *J. Comput. Aided. Mol. Des.* **1999**, *13*, 547.
112. Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009.
113. Lyne, P. D. *Drug Discov. Today* **2002**, *7*, 1047.
114. Reddy, A. S.; Pati, S. P.; Kumar, P. P.; Pradeep, H. N.; Sastry, G. N., *Curr. Protein Pept. Sci.* **2007**, *8*, 329.
115. Stahl, M.; Rarey, M. *J. Med. Chem.* **2001**, *44*, 1035.
116. Perola, E.; Walters, W.P.; Charifson, P.S. *Proteins: Struct. Funct. Bioinf.* **2004**, *56*, 235.
117. Halgren, T.A. *J. Comput. Chem.* **1996**, *17*, 520.
118. Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. *J. Comput. Chem.* **1992**, *13*, 505.
119. Weiner, S. J.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1984**, *106*, 765.
120. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
121. Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128.
122. Grant, J. A.; Pickup, B. T.; Nicholls, A. *J. Comput. Chem.* **2001**, *22*, 608.
123. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. USA.* **2001**, *98*, 10037.
124. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
125. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122.

126. Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005.
127. Huang, S.; Grinter, S. Z.; Zou, X. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12899.
128. Böhm, H. J. *J. Comput. Aid. Mol. Des.* **1994**, *8*, 243.
129. Jain, A. N. *J. Comput. Aid. Mol. Des.* **1996**, *10*, 427.
130. Head, R. D.; Smythe, M. L.; Oprea, T. I.; Waller, C. L.; Green, S. M.; Marshall, G. R. *J. Am. Chem. Soc.* **1996**, *118*, 3959.
131. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G.V.; Mee, R. P.; *J. Comput. Aid. Mol. Des.* **1997**, *11*, 425.
132. Böhm, H. J. *J. Comput. Aid. Mol. Des.* **1998**, *12*, 309.
133. Wang, R.; Gao, Y.; Lai, L.; *J. Mol. Model.* **1998**, *4*, 379.
134. Wanga, R.; Laib, L.; Wanga, S. *J. Comput. Aid. Mol. Des.* **2002**, *16*, 11.
135. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739.
136. Muegge, I. *Perspect. Drug Discov. Des.* **2000**, *20*, 99.
137. Koppensteiner, W. A.; Sippl, M. J. *Biochemistry (Moscow)* **1998**, *63*, 247.
138. DeWitte, R. S.; Shakhnovich, E. I. *J. Am. Chem. Soc.* **1996**, *118*, 11733.
139. Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. *J. Med. Chem.* **1999**, *42*, 5100.
140. Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J.B. *J. Mol. Graph. Model.* **2002**, *20*, 281.
141. Eldridge, M. D.; Murray, C. W.; Auton., T. R.; Paolini, G. V.; Mee, R. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425.
142. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
143. Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA*. **2009**, *15*, 1219.
144. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E.; *J. Mol. Biol.* **1982**, *161*, 269.

145. Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. *J. Comput. Aided. Mol. Des.* **2001**, *15*, 411.
146. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3.
147. Veber, D. F.; Johnson, S. R.; Cheng, H.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615.
148. O'Brien, W. E. *Anal. Biochem.* **1976**, *76*, 423.
149. Morrison, J. F., and Walsh, C. T. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1988**, *61*, 201.
150. Albert, A. *The Acridines. Their preparation, physical, chemical and biological properties and uses*; Edward Arnold publishers Ltd: London, 1966.
151. Magidson, J.; Grigorowski, A. M. *Chem. Ber.* **1936**, *69*, 396.
152. Charmantary, F.; Demnynck, M.; Carrer, D. *J. Med. Chem.* **2003**, *46*, 967.
153. Patel, M. M.; Mali, M. D.; Patel, S. K. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6324.
154. Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. *J. Comp. Aided. Mol. Des.* **2007**, *21*, 681.
155. Sanli, S. S.; Altun, Y. K.; Sanli, N.; Alsancak, G.; Beltran, J. L. *J. Chem. Eng. Data* **2009**, *54*, 3014.
156. Tokutake, N.; Hiratake, J.; Kato, M.; Irie, T.; Kato, H.; Oda, J. *Bioorg. Med. Chem.* **1998**, *6*, 1935.
157. Hiratake, J. *Chem. Rec.* **2005**, *5*, 209.

BIOGRAPHICAL SKETCH

Lamees Hegazy was born in Dakahlia, Egypt in August of 1984. She received her Bachelor of Science degree in biochemistry with honors from Mansoura University in May of 2005. Lamees finished the graduate courses at Mansoura University in 2006 and then moved to the United States of America where she commenced her Ph.D. in computational biochemistry in January of 2009 under the supervision of Prof. Nigel Richards. During her course of study, Lamees has been focusing on modeling and discovery of lead inhibitors of asparagine synthetase. She was recognized with the certificate of outstanding achievement two times. Lamees graduated from University of Florida in December of 2013.