



eCOMMONS

Loyola University Chicago  
Loyola eCommons

---

Master's Theses

Theses and Dissertations

---

1997

## Molecular Dynamics of the Hammerhead Ribozyme

Vasant T. Gandhi  
*Loyola University Chicago*

Follow this and additional works at: [https://ecommons.luc.edu/luc\\_theses](https://ecommons.luc.edu/luc_theses)

 Part of the [Chemistry Commons](#)

---

### Recommended Citation

Gandhi, Vasant T., "Molecular Dynamics of the Hammerhead Ribozyme" (1997). *Master's Theses*. 4217.  
[https://ecommons.luc.edu/luc\\_theses/4217](https://ecommons.luc.edu/luc_theses/4217)

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).  
Copyright © 1997 Vasant T. Gandhi

LOYOLA UNIVERSITY CHICAGO

**MOLECULAR DYNAMICS OF THE HAMMERHEAD RIBOZYME**

A THESIS SUBMITTED TO  
THE FACULTY OF THE GRADUATE SCHOOL  
IN CANDIDACY FOR THE DEGREE OF  
MASTER OF SCIENCE  
DEPARTMENT OF CHEMISTRY

BY

**VASANT T. GANDHI**

CHICAGO, ILLINOIS

JANUARY 1997

Copyright by Vasant Gandhi, 1996  
All rights reserved.

## ACKNOWLEDGMENTS

I would like to thank all who gave their time, advice, and help to me in the completion of this thesis. I would like to give special thanks to Dr. Kenneth W. Olsen, Dr. David S. Crumrine, Dr. Willetta Greene-Johnson, Dr. Albert W. Herlinger, and Dr. Stephen F. Pavkovic who challenged, taught, inspired, and believed with their hearts and minds.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
LIST OF ILLUSTRATIONS	v
LIST OF TABLES	vii
LIST OF ABBREVIATIONS	viii
Chapter	
1. INTRODUCTION	1
2. THE HAMMERHEAD RIBOZYME	3
The Hammerhead Cleavage Reaction Mechanism	3
Structure Elucidation	4
3. FOLDING OF NUCLEIC ACIDS	13
4. METAL IONS AND RNA STRUCTURE	16
Counterionic Cloud	16
5. PROPERTIES OF NUCLEIC ACIDS	19
Conformational Statistics	19
Thermodynamics	20
Helix-Coil Transitions	20
The Molten Globule State	21
Molecular Stress	25
6. ATOMISTIC SIMULATIONS	31
Molecular Dynamics Theory	31
Full Newtonian Simulation or Deterministic Model of Atomic Motion	33
Stochastic Models of Atomic Motion	34
7. POTENTIAL ENERGY FUNCTIONS	42
8. SIMULATIONS	46
Molecular Complex Preparation	46
Energy Minimization	46

Solvent Representation	49
Dynamics	50
9. TRAJECTORY ANALYSIS	56
Comparison of Simulation Trajectories	59
Final Structures	60
In-Line Cleavage Mechanism	60
RMSD Evaluation	61
RMSD Calculations with respect to the Final Structure	63
Phosphate Backbone Changes	63
Base (Side-chain) Evolution	64
RMSD Calculations with respect to the Average Structure	65
Phosphate Backbone Flexibility	65
Base (Side-chain) Flexibility	66
10. CONCLUSION	95
APPENDIX A	97
VITA	119

## LIST OF ILLUSTRATIONS

Figure 1. (a) The hammerhead ribozyme catalytic pocket and location of the scissile bond (b) Proton abstraction in the hammerhead ribozyme cleavage mechanism (c) Ribozyme Engineering (d) Hammerhead ribozyme cleavage mechanism	7
Figure 2. (a) Pley structure (b) Tuschl structure	8
Figure 3. (a) Scott core structure (b) Tuschl core structure (c) Pley core structure	9
Figure 4. Scott structures showing nucleotide backbone	10
Figure 5. (a) <i>In vitro</i> pathway of protein folding illustrating possible interconversions between unfolded (U) and partially folded intermediates (I) which eventually fold to the compact intermediate (CI) state before reaching the native state (N) and a competing aggregation step (formation of occlusion bodies) for misfolded structures. (b) Pathway from unfolded (U) state to properly folded or native (N) state showing transition through intermediate states including compact intermediate and passage through a high energy state. (c) Model for structure of compact intermediates. (d) Energy landscape folding model	26
Figure 6. Minimized Structure showing Mg <sup>2+</sup> ions (small squares)	53
Figure 7. Backbone Side-Chain Correlation	69
Figure 8. Temperature versus Time	70
Figure 9. Potential Energy, Kinetic Energy, and Total Energy versus Time	72
Figure 10. Radius of Gyration versus Time	74
Figure 11. Hydrogen Bonds versus Time	76
Figure 12. Dipole Moment versus Time	78
Figure 13. Final Structures at 600 ps	80
Figure 14. O2'-P-O5' Bond Angle versus Time	81
Figure 15. Minimized ribozyme structure: (a) ribbon and (b) ladder views. Average Ribozyme structure over 550 ps dynamics: (c) ribbon and (d) ladder views	82
Figure 16. Average Structure of Ribozyme showing Mg <sup>2+</sup> ions	83
Figure 17. Ribbon and ladder views of ribozyme	84
Figure 18. RMSD with respect to Final	89
Figure 19. RMSD with respect to Average Structure	93

## LIST OF TABLES

Table 1. Comparison of Average Molecular Properties Between Newtonian and Langevin Dynamics Simulations at 300K and 500K	61
Table 2. Backbone (Phosphate Atom) Evolution towards Final Structure	65
Table 3. Side-Chain (Nucleotide Base) Evolution towards Final Structure	66
Table 4. Backbone (Phosphate Atom) Flexibility with respect to Average Structure	67
Table 5. Side-Chain (Base) Flexibility with respect to Average Structure	68



## LIST OF ABBREVIATIONS

<u>A</u>	<u>Adenosine</u>
<u>BPTI</u>	<u>Bovine Pancreatic Trypsin Inhibitor</u>
<u>C</u>	<u>Cytosine</u>
<u>CHARMm</u>	<u>Chemistry Harvard Macromolecular Mechanics</u>
<u>D</u>	<u>Daltons</u>
<u>DNA</u>	<u>Deoxyribonucleic Acid</u>
<u>FRET</u>	<u>Fluorescence Resonance Energy Transfer</u>
<u>G</u>	<u>Guanine</u>
<u>Mg<sup>2+</sup></u>	<u>Magnesium Ion</u>
<u>RMSD</u>	<u>Root-Mean-Square Deviation</u>
<u>RNA</u>	<u>Ribonucleic Acid</u>
<u>U</u>	<u>Uracil</u>

## CHAPTER 1

### INTRODUCTION

Conformations of biological macromolecules result from their interrelated structural, functional, and environmental attributes.<sup>1</sup> This thesis examines structural and energetic changes occurring in a specific type of biological macromolecule, nucleic acids, under simulated environmental (temperature, solvent) conditions. The computational biophysical approach through molecular dynamics simulations may help define knowledge essential to applications in molecular biology, pharmacology, medicine,<sup>2</sup> nanotechnology<sup>3</sup> and computer engineering.

Three candidate molecules for study were originally identified in the literature; each candidate molecule presented a different degree of nucleic acid structure: a synthetic oligoribonucleotide double helix, a transfer ribonucleic acid, and a catalytic ribonucleic acid, the hammerhead ribozyme. The hammerhead ribozyme is examined in this investigation.

The primary purpose of this thesis is to investigate and characterize the dynamics, structure, and thermodynamics of a catalytic ribonucleic acid, the hammerhead ribozyme, as the molecule searches conformational space at different temperatures and under different dynamical methods. A secondary purpose is to investigate the presence of a specific structural and energetic intermediate state, the “molten globule” state, in a nucleic acid’s search of conformational space during folding or unfolding. This study uses molecular

dynamics techniques to analyze the internal motions and thermodynamics of nucleic acids. These techniques present “an atomic picture of high resolution with respect to space, energy or time”<sup>4</sup> and are valuable “probes” of molecular structure and function.

As noted by one researcher, “[T]he characterization of a biomolecular system at the atomic level in terms of structure, mobility, dynamics, and energetics is incomplete \* \* \*

This incomplete molecular picture makes it difficult to establish the link between molecular structure, mobility, dynamics, and interactions on the one hand, and biological function on the other.”<sup>5</sup>

---

<sup>1</sup>For example, a change in pH will produce a change in the electrical charge associated with the side chain of an amino acid. This change in charge causes a change in the amino acid and overall protein conformation yielding a change in function.

<sup>2</sup>Ross, D.W. (1996). *Introduction to Molecular Medicine* (Second Edition). (New York: Springer-Verlag).

<sup>3</sup>Drexler, K.E. (1994). *Molecular Nanomachines: Physical Principles and Implementation Strategies*. 23 *Annu. Rev. Biophys. Biomol. Struct.* 377-405.

<sup>4</sup>van Gunsteren, W.F. and Mark, A.E. (1991). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 947-961.

<sup>5</sup>van Gunsteren, W.F., Luque, F.J., Timms, D. and Torda, A.E. (1994). *Molecular Mechanics in Biology: From Structure to Function, Taking Account of Solvation*. 23 *Annu. Rev. Biophys. Biomol. Struct.* 848. For example, molecular dynamics simulations were essential in showing the penetration of oxygen through folded myoglobin to reach the internal “buried” oxygen-binding site. Fosdick, L.D., Jessup, E.R., Schauble, C., and Domik, G. (1996). *An Introduction to High-Performance Scientific Computing*. (Cambridge: MIT Press) 534 citing Karplus, M. and McCammon, J. A. (1986). The dynamics of proteins. *Scientific American* 42-51.

## CHAPTER 2

### THE HAMMERHEAD RIBOZYME

The self-cleaving hammerhead ribozyme has been the subject of intensive recent examination. The hammerhead ribozyme, a catalytic RNA, is an interesting example of an RNA molecule which is rich in tertiary structure. The molecule's X-ray coordinates were received from the Protein Data Bank.<sup>1</sup> The ribozyme consists of three base-paired stems or helices and a core of non-complementary nucleotides.<sup>2</sup> This molecule "promotes" a magnesium ion-dependent site-specific cleavage of RNA by intermolecular or intramolecular reactions.<sup>3</sup> Upon the addition of hydrogens and magnesium ions, the molecular system in this study consists of 1078 fragments, 651 ring bonds, 386 rotatable bonds, and 1331 atoms and possesses a molecular weight of 13, 662.61 D. We characterize the temporal behavior of the system consisting of the molecule and an ion and solvent environment at different temperatures using various dynamical methods. The structural motif consists of three helical regions surrounding a conserved core.<sup>4</sup> When viewed in two-dimensions, the hammerhead ribozyme secondary structure resembles a "hammerhead."<sup>5</sup>

#### The Hammerhead Cleavage Reaction Mechanism

The self-cleavage or auto-cleavage reaction results from the nucleophilic attack by a core nucleotide's 2'-hydroxyl upon the adjacent phosphodiester bond.<sup>6</sup> Figure 1(a) shows the general architecture of the hammerhead ribozyme motif and the locations of the scissile bond and catalytic pocket. The reaction mechanism yields a 2',3'-cyclic phosphate terminal

and a 5'-hydroxyl terminal.<sup>7</sup> Significant inquiry is directed currently at the cleavage mechanism generally and the role of magnesium ions in proton abstraction (Figure 1(b)).<sup>8</sup> Studies have also been conducted on the *in vitro* activity of the hammerhead ribozyme.<sup>9</sup> Through so-called "ribozyme engineering", ribozymes may be used to "edit" or repair messenger RNA molecules (Figure 1(c)).<sup>10</sup> For example, a ribozyme may be used as an "antisense" nucleotide designed to bind with and subsequently cleave deleterious RNA and DNA sequences. Figure 1(d) presents ribozyme engineering and the mechanism of hammerhead ribozyme cleavage.

The hammerhead ribozyme is one of several structurally distinct classes of catalytic RNA differing in terms of their reaction products and nucleophile. Most are metalloribozymes requiring  $Mg^{2+}$  or  $Pb^{2+}$  as electrophiles for their cleavage reactions. These ribozymes may be further classified as either splicing or self-cleaving.<sup>11</sup>

### Structure Elucidation

Recent efforts by researchers at several institutions have produced structural models for the hammerhead ribozyme. The research groups adopted differing approaches to structure solution in terms of the biophysical technique utilized or the molecular system investigated. Pley, *et al.*, have produced a structure based on X-ray crystallography of a hammerhead RNA-DNA ribozyme-inhibitor complex to a resolution of 2.6Å.<sup>12</sup> Tuschl, *et al.*, have utilized fluorescence resonance energy transfer of the hammerhead ribozyme in solution to produce a three-dimensional model.<sup>13</sup> More recently, Scott, *et al.*, have

elucidated the X-ray crystal structure of an all-RNA hammerhead ribozyme.<sup>14</sup> Figure 2 presents the Pley , Tuschl, and Scott structures.

Comparisons of the determined structures of the hammerhead ribozyme yield useful insight into the efficacy of the method used and provide a basis for critical study of the biophysical methodology. We examined the structures obtained by FRET (Tuschl) and X-ray crystallography (Scott, Pley) to determine whether model building with the judicious application of distance constraints based upon physical measurements is sufficient to define the structure of small RNA molecules such as the hammerhead ribozyme.<sup>15</sup>

Figure 2(d) presents the Scott structure with two ribozyme molecules within the asymmetric unit. The molecular coordinate file (revised) as deposited in the Brookhaven Protein Databank was modified to produce separate coordinate files for each molecule in the Scott structure. The Pley structure contained three molecules within the asymmetric unit; one molecule was excised from the coordinate file. The Scott structure and the Pley structure were compared with the Tuschl structure.

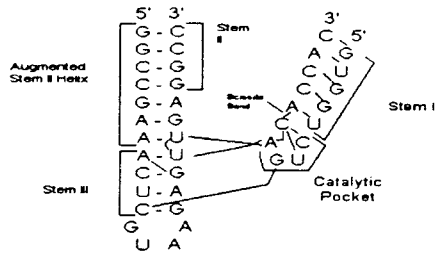
In order to correctly compare the methods used to generate each structure, a meaningful basis for comparison must be established, i.e. alignment of similar sequence residues. Conveniently, for our purposes, each hammerhead ribozyme molecule possesses a conserved core of nucleotides articulated into two structural domains. Because of the different hammerhead ribozyme complexes studied, a frame of reference is critical.

We define the common core in the following manner: the nucleotide sequence C U G A (the first domain) followed by U followed by the nucleotide sequence G A G C G A A

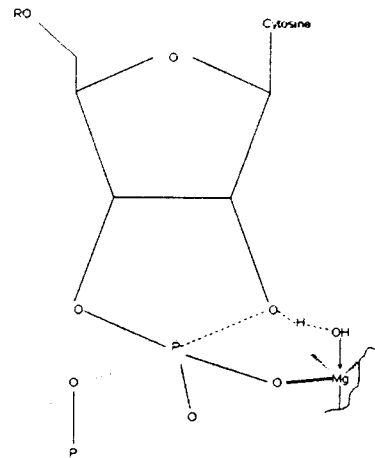
A C (the second domain). The core consists of 14 nucleotides or groups, 302 atoms, and 307 bonds. Schematically, the consensus structure of each molecule is shown in Figure 3 as a backbone representation.

The Pley and Tuschl structures exhibited a RMSD of 8.500 Å. The Scott and Tuschl structures exhibited a RMSD of 8.531 Å and the Scott and Pley structures showed a RMSD of 0.904 Å. In each case, the calculations encompassed 4 fragments, 380 ring bonds, and 160 rotatable bonds within 2 groups. 302 matches were conducted as part of the calculation. These results indicate that the FRET structure deviates strongly from either of the X-ray structures. The X-ray structures were in close agreement with each other. It is reasonable to conclude that FRET does not yet provide sufficient structural information in the case of small macromolecules to be a viable substitute for X-ray crystallography. Additionally presented are differing representations of the Scott structure, including backbone, ladder, and ball-and-stick depictions. Such representations are useful in assessing the degree of structure present in the hammerhead ribozyme.

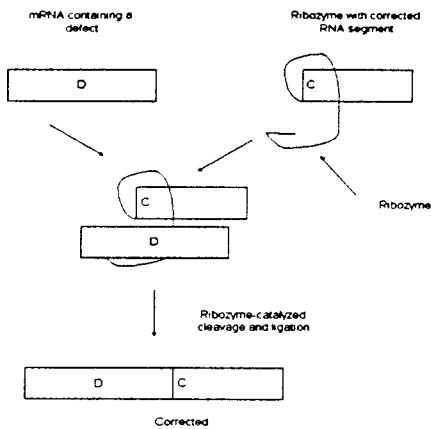
(a)



(b)



(c)



(d)

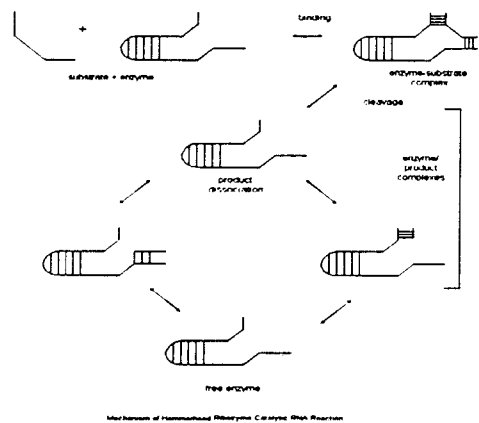


Figure 1. (a) The hammerhead ribozyme catalytic pocket and location of the scissile bond (b) Proton abstraction in the hammerhead ribozyme cleavage mechanism (c) Ribozyme Engineering (d) Hammerhead ribozyme cleavage mechanism.



(a)



(b)



(c)



(d)



Figure 2. (a) Pley structure (b) Tuschl structure (c) Scott structure (d) Scott structure with two ribozyme molecules within asymmetric unit.

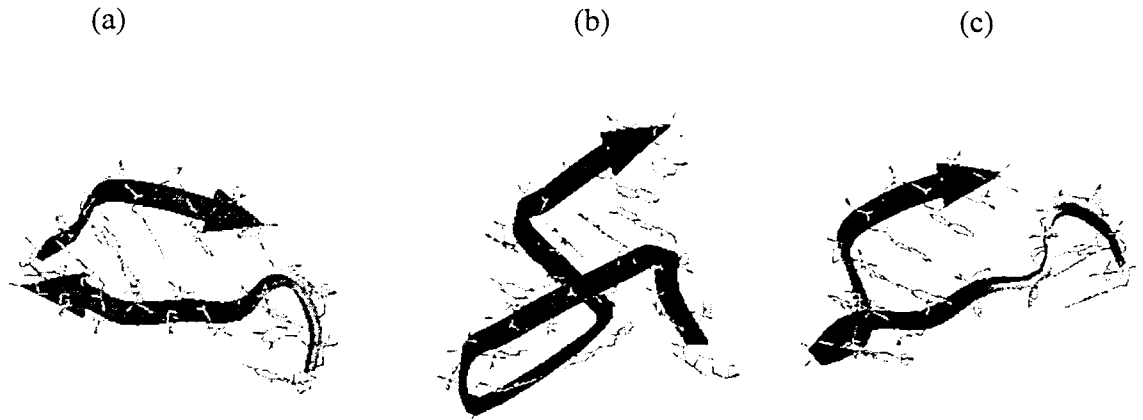


Figure 3. (a) Scott core structure (b) Tuschl core structure  
(c) Pley core structure.

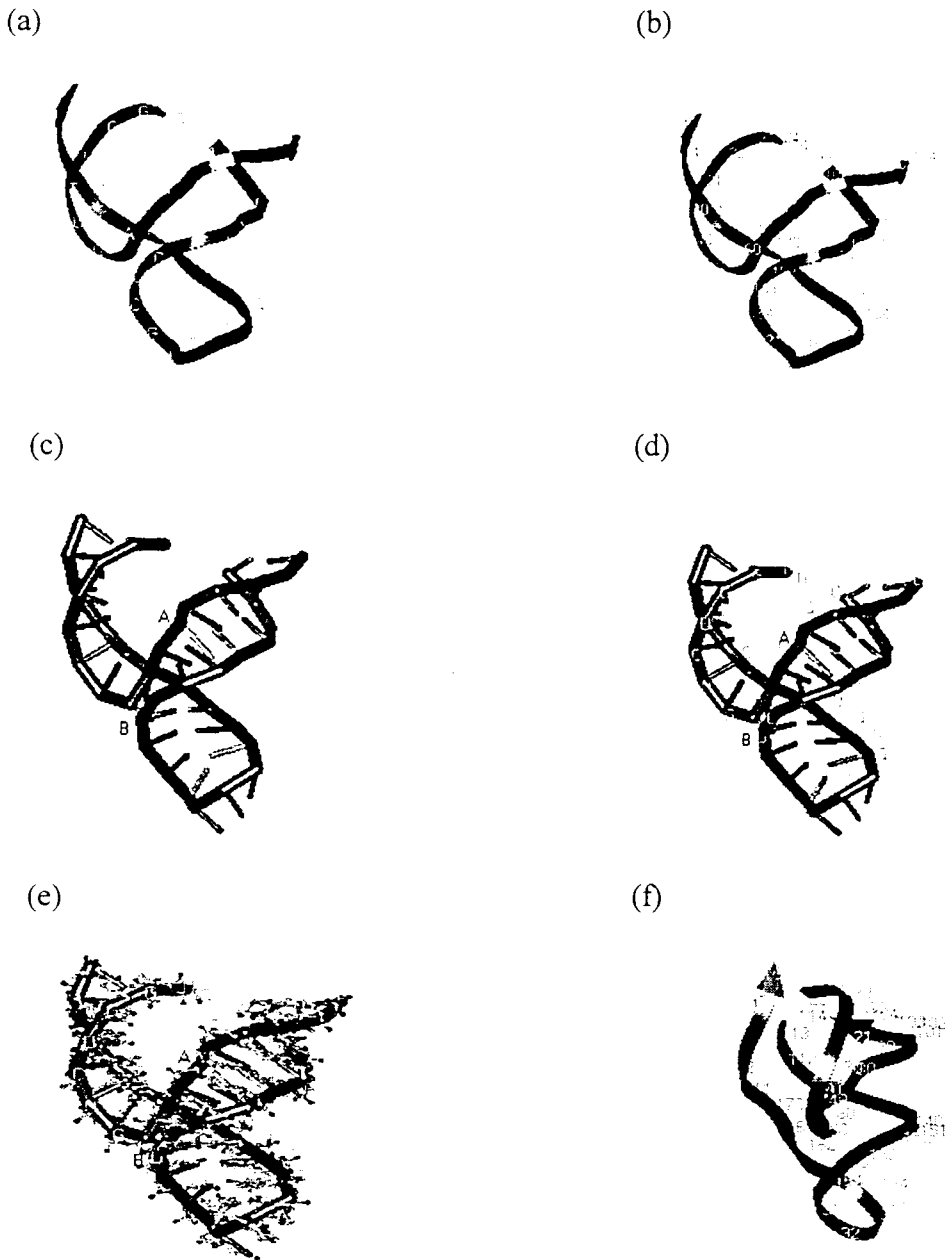


Figure 4. Scott structures showing nucleotide backbone

- 
- <sup>1</sup>Scott, W.G., Finch, J.T., Klug, A. (1995). The Crystal Structure of an All-RNA Hammerhead Ribozyme: A Proposed Mechanism for RNA Catalytic Cleavage. 81 Cell 991-1002. See also Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. 372 Nature 68-74. Tuschl, T., Gohlke, C., Jovin, T.M., Westhof, E., and Eckstein, F. (1995). A Three-Dimensional Model for the Hammerhead Ribozyme Based on Fluorescence Measurements. 266 Science 785.
- <sup>2</sup>A nucleotide (a phosphate ester of a nucleoside) consists of a sugar (ribose or deoxyribose), a nitrogen heterocyclic purine or pyrimidine base, and a phosphate group. The bases are connected to the sugar-phosphate backbone. The helical backbone is comprised of a linear series of sugars which are linked through bonds from the 3' OH of the sugar to a phosphate group to the 5' OH of the next sugar; the nitrogenous base is attached to the backbone by its N to the C1 of the sugar. Base pairing between adenine and uracil consists of 2 hydrogen bonds; base pairing between cytosine and guanine consists of 3 hydrogen bonds. The orientation of the bases with respect to each other is not perfectly coplanar. The actual twisting which occurs is often termed "propeller twist." Stryer at 650.
- <sup>3</sup>Tuschl, T., Gohlke, C., Jovin, T.M., Westhof, E., and Eckstein, F. (1995). A Three-Dimensional Model for the Hammerhead Ribozyme Based on Fluorescence Measurements. 266 Science 785.
- <sup>4</sup>Doudna, J.A. (1995). Hammerhead ribozyme structure: U-turn for RNA structural biology. 3 Structure 747-750.
- <sup>5</sup>Yarus, M. (1993). How many catalytic RNAs? Ions and the Cheshire cat conjecture. 7 The FASEB Journal 31 (the three helices coalesce into a T shape).
- <sup>6</sup>Doudna, J.A. (1995). Hammerhead ribozyme structure: U-turn for RNA structural biology. 3 Structure 747.
- <sup>7</sup>Doudna, J.A. (1995). Hammerhead ribozyme structure: U-turn for RNA structural biology. 3 Structure 747.
- <sup>8</sup>Scott, W.G., Finch, J.T., Klug, A. (1995). The Crystal Structure of an All-RNA Hammerhead Ribozyme: A Proposed Mechanism for RNA Catalytic Cleavage. 81 Cell 991-1002. See also Smith, D. (1995). Magnesium as the Catalytic Center of RNA Enzymes in Cowan, J.A., ed. The Biological Chemistry of Magnesium. (New York: VCH) 85-108.
- <sup>9</sup>Beck, J. and Nassal, M. (1995). Efficient hammerhead ribozyme-mediated cleavage of the structured hepatitis B virus encapsidation signal *in vitro* and in cell extracts, but not in intact cells. 23 Nucleic Acids Research 4954-4962. Hendry, P. and McCall, M.J. (1995). A comparison of the *in vitro* activity of DNA-armed and all-RNA hammerhead

---

ribozymes. 23 *Nucleic Acids Research* 3928-3936. Lustig, B., Lin, N.H., Smith, S., Jernigan, R.L., and Jeang, K. (1995). A small modified hammerhead ribozyme and its conformational characteristics determined by mutagenesis and lattice calculation. 23 *Nucleic Acids Research*. 3531-3538. Werner, M. and Uhlenbeck, O.C. (1995). The effect of base mismatches in the substrate recognition helices of hammerhead ribozymes on binding and catalysis. 23 *Nucleic Acids Research* 12:2092-2096. Beaudry, D., Bussiere, F., Laureau, F., Lessard, C. and Perreault, J. (1995). The RNA of both polarities of the peach latent mosaic viroid self-cleaves *in vitro* solely by single hammerhead structures. 23 *Nucleic Acids Research* 5: 745-752.

<sup>10</sup>Rawls, R. (1996). Splicing ribozyme can 'edit' mammalian RNA. *Chemical and Engineering News*. June 3, 1996 at 7. See also Cech, T.R. (1992). Ribozyme Engineering. 2 *Current Opinion in Structural Biology* 605-609.

<sup>11</sup>Cech, T.R. (1987). The Chemistry of Self-Splicing RNA and RNA Enzymes. 236 *Science* 1532-1539. See also Cech, T.R. (1993). Structure and Mechanism of the Large Catalytic RNAs: Group I and Group II Introns and Ribonuclease P. in Gesteland, R.F. and Atkins, J.F., eds. (1993). *The RNA World. The Nature of Modern RNA Suggests a Prebiotic RNA World.* (New York: Cold Spring Harbor Laboratory Press) 239-269.

<sup>12</sup>Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. 372 *Nature* 68-74.

<sup>13</sup>Tuschl, T. Gohkle, C., Jovin, T., Westhof, E., Eckstein, F. (1994). A Three-Dimensional Model for the Hammerhead Ribozyme Based on Fluorescence Measurements. 266 *Science* 785-788.

<sup>14</sup>Scott, W.G., Finch, J.T., Klug, A. (1995). The Crystal Structure of an All-RNA Hammerhead Ribozyme: A Proposed Mechanism for RNA Catalytic Cleavage. 81 *Cell* 991-1002.

<sup>15</sup>A comparison of the Scott and Pley X-ray structures found a strong similarity between the two structures. "[B]oth crystal structures must be a close approximation to the true solution structure of an unaltered hammerhead ribozyme." Scott, W.G., Finch, J.T., Klug, A. (1995). The Crystal Structure of an All-RNA Hammerhead Ribozyme: A Proposed Mechanism for RNA Catalytic Cleavage. 81 *Cell* 991-992.

## CHAPTER 3

### FOLDING OF NUCLEIC ACIDS

The architecture of folding patterns in nucleic acids<sup>1</sup> is currently the subject of extensive research particularly as the discovery of additional structures frees nucleic acids from the historically confining designation as “carrier[s] of genetic information”.<sup>2</sup> One central problem is that of defining the “arrangement of RNA structural elements in three-dimensional space” in a system analogous to that developed for proteins.<sup>3</sup>

In proteins, it is commonplace to describe polypeptide folding patterns in terms of structural motifs such as helices, sheets, barrels, etc.<sup>4</sup> The units of RNA folding<sup>5</sup> are generally viewed as the base-paired double helix secondary structure, pseudoknots, loops which cap helices, loops within helices, RNA mispairing regions, nucleoside triple interactions, quadruplexes, and U-turns<sup>6</sup> as well as multiplexes, junctions, and hairpins.<sup>7</sup> Tertiary structures form, in a manner analogous to protein folding, by the “condensation” of “individual blocks of secondary structure.”<sup>8</sup> The strength of the analogy is limited, however, by the view that “RNA secondary structural elements are very stable and capable of forming independently of tertiary structure. This has led to the concept that RNA secondary structure forms rapidly and precedes the packaging of RNA into tertiary structure, i.e. that base pairing interactions occur prior to helix or domain formation.”<sup>9,10</sup>

Nonetheless, the foundation for analyzing polynucleotides and polypeptides in similar fashion is a sound one. As noted by one author,

“From the energy landscape perspective it is natural to suggest that the considerations that lead to the theoretical developments of protein folding should also apply to RNA folding. In general terms, the requirements for RNA folding are analogous to those of protein folding. As is true for polypeptides, the number of conformations in the fully denatured state (the Levinthal limit) is large. For RNA sequences, the kinetic problem consists of forming the correct secondary structure, that is, Watson-Crick base pairs between complementary sequences, and achieving the correct three-dimensional organization of the structural elements.”<sup>11</sup>

A comparison of the kinetic folding pathways of proteins and RNA<sup>12</sup> shows clear differences in the folding pathway. The theoretical time scale for RNA unfolding is significantly large so as to exceed the boundaries of the present investigation. However, one of the purposes of this project, to characterize the short-term dynamics of the molecule, is not diminished by the realization that significant large-scale structural transitions will not likely occur before 1 millisecond.

---

<sup>1</sup>See generally Saenger, W. (1984). Principles of Nucleic Acid Structure. (New York: Springer-Verlag).

<sup>2</sup>Kochoyan, M. and Leroy, J. (1995). Hydration and solution structure of nucleic acids. 5 Current Opinion in Structural Biology 329-333.

<sup>3</sup>Pyle, A.M. and Green, Justin B. (1995). RNA folding. 5 Current Opinion in Structural Biology 303.

<sup>4</sup>See generally Branden, C. and Tooze, J. (1991). Introduction to Protein Structure. New York: Garland.

<sup>5</sup>See Zuker, M. (1989). On Finding All Suboptimal Foldings of an RNA Molecule. 244 Science 48-52.

<sup>6</sup>Pyle, A.M. and Green, Justin B. (1995) RNA folding. 5 Current Opinion in Structural Biology 303. For example, “[p]seudoknots are interlocked regions of coaxially stacked helices that are commonly involved in RNA binding and folding. Pseudoknotting is a motif with a high capacity for molecular recognition \* \* \*”. Pyle, A.M. and Green, Justin B. (1995) RNA folding. 5 Current Opinion in Structural Biology 303. See also Abrahams,

---

J., van den Berg, M., van Batenburg, E., and Pleij, C. (1990). Prediction of RNA secondary structure, including pseudoknotting, by computer simulation. 18 *Nucleic Acids Research* 10: 3035-3044; Fontana, W., Stadler, P. Tarazona, P. Weinberger, E. and Schuster, L. (1993). RNA folding and combinatorial landscapes. 47 *Physical Review (E)* 3:2086; Westhof, E and Patel, D. (1995). *Nucleic Acids: Diversity, folding, and stability of nucleic acid structures.* 5 *Current Opinion in Structural Biology* 279-281.

<sup>7</sup>See generally Lilley, D.M.J., Clegg, R.M., Diekmann, S., Seeman, N.C., von Kitzing, E., and Hagerman, P.J. (1995). A nomenclature of junctions and branchpoints in nucleic acids. 23 *Nucleic Acids Research* 17: 3363-3364.

<sup>8</sup>Pyle, A.M. and Green, J. B. (1995) at 304.

<sup>9</sup>Pyle, A.M. and Green, J. B. (1995) at 307.

<sup>10</sup>"As opposed to the protein case, the secondary structure of RNA sequences is well-defined; it provides the major set of distance constraints that guide the formation of tertiary structure, and covers the dominant energy contribution to the 3D structure." Fontana, W., Stadler, P. Tarazona, P. Weinberger, E. and Schuster, L. (1993). RNA folding and combinatorial landscapes. 47 *Physical Review (E)* 3:2086. See also Thirumalai, D. and Woodson, S.A. (1996). Kinetics of Folding of Proteins and RNA. 29 *Acc. Chem. Res.* 433-439.

<sup>11</sup>Thirumalai, D. and Woodson, S.A. (1996). Kinetics of Folding of Proteins and RNA. 29 *Acc. Chem. Res.* 433.

<sup>12</sup>Draper, D. (1996). Parallel worlds. 3 *Nature Structural Biology* 5:397-400. See also Draper, D. (1996). Strategies for RNA folding. 21 *TIBS* 145-149.



## CHAPTER 4

### METAL IONS AND RNA STRUCTURE

The presence of metal ions contributes greatly to the folding of RNA into various structures.<sup>1</sup> The structural and dynamic effects of the binding of metal ions to RNA may be specific or non-specific. Thermal denaturation experiments have shown that specific (site) binding of metal ions, as with tertiary RNA structures, produces a linear relationship between the reciprocal of the RNA melting temperature and magnesium ion concentration. Non-specific RNA binding by metal ions, as in the case of charge screening and duplex formation, produces a sigmoidal plot.<sup>2</sup>

It has been concluded that metal ions perform three distinct functions in RNA folding: non-specific binding, specific binding or coordination using the  $Mg^{2+}$  ion, and “binding to high-affinity sites through charge density or outer-sphere coordination rather than through the formation of direct metal contacts.”<sup>3</sup> An important concept is the counterionic cloud within the context of the Debye-Huckel theory.<sup>4</sup> Although there is evidence for the  $Mg^{2+}$  - independent folding of the hammerhead ribozyme<sup>5</sup>, one author questions the experimental conditions. Specifically the 2'-deoxyribose substitution of the cleavage site to prevent cleavage of substrate may have disrupted ion binding.<sup>6</sup>

#### Counterionic Cloud

The presence of ions produces, as an average property, a “counterionic cloud” which moves with the molecule.<sup>7</sup> The molecule's size increases due to the cloud's presence

which alters the molecule's electrical and hydrodynamic properties.<sup>8</sup> The Debye-Huckel theory explains the charge distribution of the cloud and the resultant effect upon a charged macromolecule's electrical potential  $V$  at a given location  $r$ .<sup>9</sup> Calculation of counterion density, i.e. position-dependent concentration, requires the following conditions: (1) non-constant concentration of ions due to thermal fluctuations, (2) equal concentrations of positive and negative counterions at remote distances from the molecule, and (3) a greater concentration of oppositely charged counterions than same charge as the molecule at local distances from the molecule.<sup>10</sup>

Additionally, the Boltzman distribution is assumed to govern charge distribution.<sup>11</sup> The counterions assume positions based upon the Boltzmann distribution of their respective potential energies;<sup>12</sup> the net charge density  $\rho$  equals the difference between the concentration of positive and negative counterions.<sup>13</sup> The relationship between net charge density and potential as a function of position is governed by the Poisson equation,  $d^2V/dx^2 + d^2V/dy^2 + d^2V/dz^2 = -4\pi\rho/D$ , where  $\rho$  is the net charge density and  $D$  is the dielectric constant of the medium (solvent).<sup>14</sup> The presence of a counterionic cloud decreases the potential energy of the molecular system in comparison to systems without a counterionic cloud (ionic atmosphere).<sup>15</sup>

---

<sup>1</sup>Sharp, K. and Honig, B. (1995). Salt effects on nucleic acids. 5 *Current Opinion in Structural Biology* at 323-328.

<sup>2</sup>Pyle, A.M. and Green, Justin B. (1995) at 305 citing Laing, L.G., Gluick, T.C., and Draper, D.E. (1994). Stabilization of RNA structure by Mg ions. 237 *J Mol Biol* 577-587.

- 
- <sup>3</sup>Pyle, A.M. and Green, Justin B. (1995). at 305 citing Laing, L.G., Gluick, T.C., and Draper, D.E. (1994). Stabilization of RNA structure by Mg ions. 237 J Mol Biol 577-587.
- <sup>4</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics. 221-246.
- <sup>5</sup>Heus, H. and Pardi, A. (1991). Nuclear magnetic resonance studies of the hammerhead ribozyme domain. Secondary structure formation and magnesium ion dependence. 217 J. Mol. Biol. 113-124.
- <sup>6</sup>Smith, D. (1995). Magnesium as the Catalytic Center of RNA Enzymes. in Cowan, J.A. , ed. The Biological Chemistry of Magnesium. (New York: VCH) 113. See, generally, Porschke, D. (1995). Modes and Dynamics of Mg<sup>2+</sup>-Polynucleotide interactions in Cowan, J.A. , ed. The Biological Chemistry of Magnesium. (New York: VCH) 85-108.
- <sup>7</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 222.
- <sup>8</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 223.
- <sup>9</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 223.
- <sup>10</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 223.
- <sup>11</sup>The concentrations of negative and positive counterions close to the molecule, respectively, is given by  $e^{-U/k_B T}$  where U is the electrical potential energy of the counterion at a given distance from the molecule,  $k_B$  is the Boltzmann constant, and T is the absolute temperature. Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 224.
- <sup>12</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 224.
- <sup>13</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics. 221-246.
- <sup>14</sup>Given that  $W_{ab} = D_{ab}q$  and that there is zero potential at infinity, the potential energy, U, of a given counterion is  $U = W = Vq$  where V equals the potential at the counterion's final location and q equals the counterion's charge. Ion concentration, C, is equal to  $C_0 e^{-qV/k_B T}$  where  $C_0$  is the bulk concentration of ions. Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 225.
- <sup>15</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics at 246.

CHAPTER 5  
PROPERTIES OF NUCLEIC ACIDS

Conformational Statistics

Polymeric chains, such as polypeptides and polynucleotides, may be viewed as linear systems made of independent elements or statistical segments.<sup>1</sup> In the case of polynucleotides, the relevant unit is the nucleotide. The flexibility of biopolymers, due to rotation around single bonds, produces a very large set or ensemble of possible conformations. Analysis of the potential energy of the biopolymer chain and its functional dependence upon the multitude of possible internal angles of rotation underlies the conformational statistical approach.

Internal motions of biopolymers may be functionally classified as local motions (atomic fluctuations, sidechain motions, loop motions), rigid-body motions (helix motions, domains or hinge-bending motions, and subunit), and larger-scale motions (helix-coil transitions, dissociation/association and coupled structural changes, opening and distortional fluctuations, and folding and unfolding transitions).<sup>2</sup>

The rotations  $x$  about the glycosidic bond are of two types: syn where  $x = 0$  and anti where  $x = 210$ .<sup>3</sup> The preferred state is a function of the degree of sugar “puckering” present.<sup>4</sup> Puckering refers to the relative spatial position of the 5 atoms constituting the ribose ring.<sup>5</sup> Four atoms are in the same plane; the fifth, C2' or C3', is above or below the plane.<sup>6</sup> An endo conformation exists when the displacement occurs on the same side as

C5'; exo occurs when the displacement occurs on the opposite side.<sup>7</sup> There are six rotational angles for each backbone unit.<sup>8</sup>

By considering factors such as steric (electronic) hindrance or bad (forbidden) contacts, the possible conformations may be limited by first-order interactions (between atoms whose interatomic distance is dependent upon a single rotation angle.)<sup>9</sup> Second-order interactions are those interactions whose separation distance is simultaneously dependent upon two adjacent angles.<sup>10</sup>

### Thermodynamics

Using molecular dynamics techniques, thermodynamic properties, such as entropy, may be ascertained from the system's thermally accessible structural conformations.<sup>11</sup> The most stable conformation of a biological macromolecule is found at its lowest energy structure.

### Helix-Coil Transitions

The helical structure of protein or polypeptide chains is produced by intermolecular hydrogen bonding. In the case of polynucleotides, however, intramolecular hydrogen bonding between base pairs provides helical structure.<sup>12</sup> The polyelectrolytic nature of nucleic acids due, in part, to the presence of negatively charged phosphate groups, plays a role as well in the structure of these biopolymers.<sup>13</sup>

It is well known that polynucleotides and polypeptides undergo helix-coil transitions under conditions of increasing temperature (thermal denaturation) and changing pH of solution.<sup>14</sup> Development of a partition function or free energy expression for a

polynucleotide system is essential to an understanding of the helix-coil transition. An example of such a relationship is

Equation 1

$$\Delta G(T) = \Delta H(T) - T\Delta S(T) = \Delta H(T_R) + \Delta C_p(T-T_R) - T\Delta S(T_R) + \Delta C_p \ln(T/T_R)$$

where  $\Delta G(T)$  represents the difference between the molar Gibbs free energy function of any state and a reference state,  $\Delta H(T)$  represents the state's relative enthalpy,  $\Delta S(T)$  represents the entropy in a two-state transition and equals  $\Delta H_{TM}/T_m$ ;  $\Delta C_p$  represents the heat capacity at constant pressure and equals  $d\Delta H/dT$ ; and  $T_R$  represents an appropriate reference temperature.<sup>15</sup>

### The Molten Globule State

The existence of the molten globule state as an intermediate in protein folding or unfolding pathways has been defined and discussed by several researchers.<sup>16</sup> A question arises whether such a state exists in the folding or unfolding pathways of other biological molecules such as nucleic acids.

Molten globules, for proteins, have been characterized as compact, mobile structures with large amounts of secondary structure, but diminished tertiary contacts relative to the native state.<sup>17</sup> The molten globule state may be a common early intermediate during folding and may be an equilibrium intermediate under a variety of unfolding conditions such as high temperature, extreme pH, presence of organic solvents, or removal of stabilizing ions.<sup>18</sup> Researchers have also found the molten globule state to be under kinetic control.<sup>19</sup>

Daggett and Levitt have performed extensive molecular dynamics simulations of temperature-induced protein unfolding using bovine pancreatic trypsin inhibitor (BPTI).<sup>20</sup> Their experimental approach began with the X-ray crystallographic structure of BPTI. All atoms were explicitly present during simulations. The protein and fragments were immersed in a box of water molecules and counterions were present to yield an electrically neutral system.<sup>21</sup> Daggett *et al.* observed the molten globule state through a structural analysis of protein parameters. The researchers started with BPTI in its native, folded state and observed the first transition from native state to molten globule. A second transition was observed from the molten globule to the unfolded state.

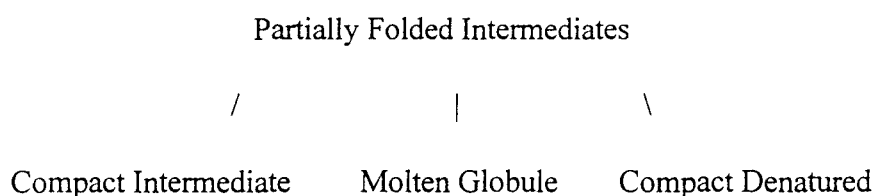
The results of phase 1 from the native state to the molten globule state were presented in terms of structural level from global to semi-local to local levels. The researchers followed time-dependencies of properties which show characteristic differences between the X-ray structure and molten globule states. The parameters and characteristics examined included global structural deviations from crystal structure (conformational sampling, size, packing interactions, hydrophobic core, and presence of secondary structural motifs) and semi-local and local structural deviations (secondary structure, turn formation, correlated motion).

Although the researchers found it “difficult to estimate numerically the deviation that one would expect for a molten globule or other partially unfolded intermediates; the values obtained “seem[ed] reasonable for a molten globule, which is native-like in many respects, and is not expected to deviate considerably from the native state.”<sup>22</sup>

Additional studies have observed the molten globule state in cytochrome *c* at low pH by examining changes in the molecule's thermodynamic properties.<sup>23</sup> Figure 5 presents a useful diagram illustrating the pathways.<sup>24</sup> Methods of identifying compact nonnative structures from the native globular protein structures have also been investigated.<sup>25</sup> The molten globule has been termed the third thermodynamical state of protein molecules.<sup>26</sup>

The transitions from the native folded state (N) to the molten globule state (MG) and from MG to the denatured unfolded state (U) have been determined to follow the “all or none” mechanism due to the absence of equilibrium intermediates between the states.<sup>27</sup>

It has been concluded that “all proteins, under the appropriate conditions, will form [compact intermediates]”<sup>28</sup> As recognized by several researchers, it is valuable to distinguish between the varying types of partially folded intermediates.<sup>29</sup> The following diagram attempts to show the relevant distinctions:



At the outset, it is important to recognize that “[t]he term compact intermediates encompasses a broad range of conformations and degrees of folding and compactness: compact intermediates have no single, unique conformation, but rather a whole plethora of structures that range from being very similar to the native state to being substantially expanded and significantly unfolded.”<sup>30</sup>



The compact intermediate generally possesses the following qualities: [a] substantial secondary structure with slight tertiary structure, [b] compact collapsed state with respect to the unfolded state as measured in terms of hydrodynamic radius, [c] exposed hydrophobic surfaces, [d] heat capacity similar to the unfolded state<sup>31</sup>, [e] absence of functional properties, and [f] less cooperative unfolding transition.<sup>32</sup>

With respect to compact denatured states of proteins in general: (a) An individual protein may possess several stable compact intermediate states depending on the individual protein character and observation conditions; (b) structurally, two types of compact intermediates predominate: “native state-like regions of secondary structure connected by disordered regions of polypeptide, but still retaining a relatively native state-like topology, or a core of native-like structure”;<sup>33</sup> and (c) it is difficult to isolate compact intermediates from compact substrates as “the energetic difference between relatively expanded compact intermediates and compact unfolded states resides mostly in the entropy and not [in] *sic* the enthalpy.”<sup>34</sup> The compact intermediate may also be an ensemble of states before passing through a high energy state, and bypassing a competing aggregation step, to the native state.<sup>35</sup> An energy landscape folding model is a valuable method of visualizing the folding pathway of biological macromolecules (Figure 5(d)).<sup>36</sup> The concern has been raised that the molten globule and other non-native states may in fact be incorrectly refolded states produced as a result of the observation conditions, which may be inappropriate for folding, instead of true folding intermediates: “Does [the protein or nucleic acid] fold and unfold through an intermediate state that is stable under some conditions and does not represent an

artifact of folding, i.e. is not a misfolded state trapped in a potential well?"<sup>37</sup> Such a state would represent "topological frustration" for the molecule.<sup>38</sup>

### Molecular Stress

For purposes of stress analysis on the molecular level, the oligonucleotide chain may be viewed as a basic unit. The molecules behave as entropic springs. Assume a polymer, e.g. a polynucleotide, of length  $L$  is subjected to a tensile force  $F$ . Macroscopic thermodynamic theory holds that

$$\text{Equation 2}$$

$$F = \left. \frac{\partial A}{\partial L} \right|_T$$

where  $A(L,T)$  is the Helmholtz free energy at absolute temperature  $T$ .

$$\text{Equation 3}$$

$$A(L,T) = U(L,T) - TS(L,T)$$

where  $U$  is the internal energy and  $S$  is the entropy of the system.

$$\text{Equation 4}$$

$$F = \left. \frac{\partial U}{\partial L} \right|_T - T \left. \frac{\partial S}{\partial L} \right|_T$$

The force necessary for polymer extension produces changes, to varying degrees, in both internal energy and entropy.<sup>39</sup> In the case of constrained macromolecules, the behavior is that of a molecular entropic spring in tension.<sup>40</sup> The concept of intrinsic monomer stress (IMS) has been developed to explain stress in polymeric systems at the atomic level.<sup>41</sup>

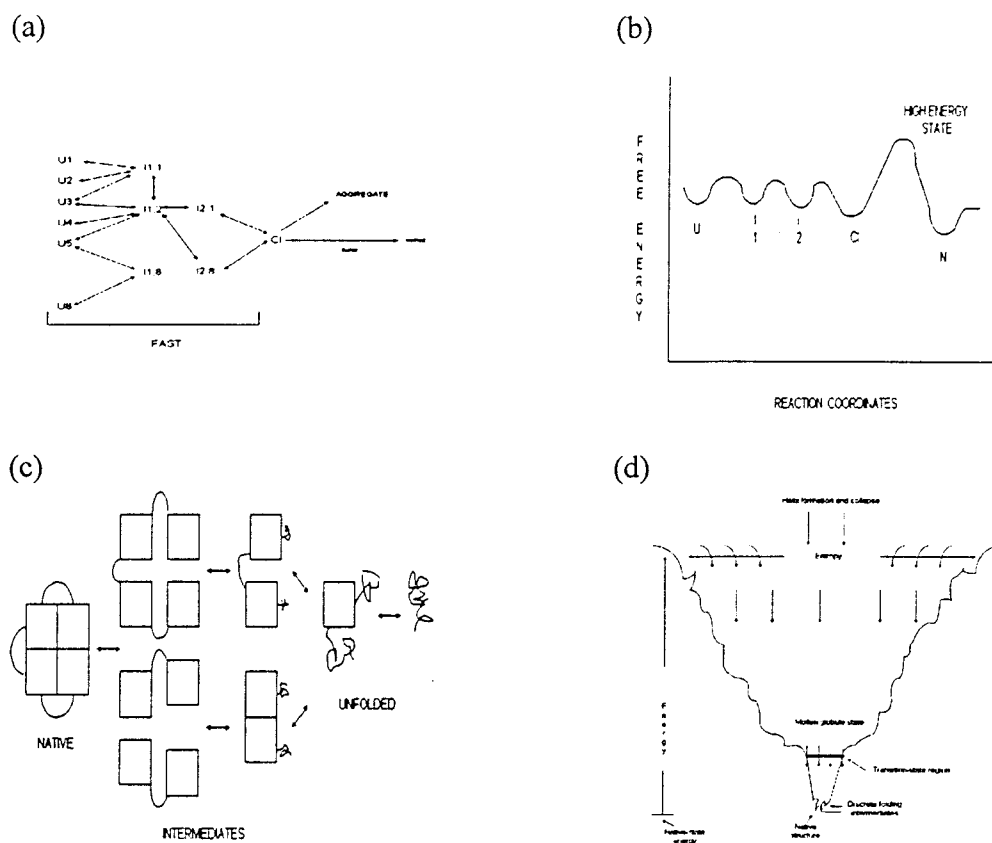


Figure 5. (a) *In vitro* pathway of protein folding illustrating possible interconversions between unfolded (U) and partially folded intermediates (I) which eventually fold to the compact intermediate (CI) state before reaching the native state (N) and a competing aggregation step (formation of occlusion bodies) for misfolded structures. (b) Pathway from unfolded (U) state to properly folded or native (N) state showing transition through intermediate states including compact intermediate and passage through a high energy state. (c) Model for structure of compact intermediates. (d) Energy landscape folding model.

- 
- <sup>1</sup>Ptitsyn, O.B. (1966). Conformations of Macromolecules. 1.
- <sup>2</sup>Brooks, C.L., Karplus, M., and Pettitt, B.M. (1988). Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. (Advances in Chemical Physics Volume LXXI) (John Wiley and Sons: New York) 20-21.
- <sup>3</sup>Cantor and Schimmel (1980) at 313.
- <sup>4</sup>Cantor and Schimmel (1980) at 313.
- <sup>5</sup>Cantor and Schimmel (1980) at 314.
- <sup>6</sup>Cantor and Schimmel (1980) at 314.
- <sup>7</sup>Cantor and Schimmel (1980) at 314.
- <sup>8</sup>Cantor and Schimmel (1980) at 314.
- <sup>9</sup>Cantor and Schimmel (1980) at 314.
- <sup>10</sup>Cantor and Schimmel (1980) at 314.
- <sup>11</sup>McCammon and Harvey at 68-73.
- <sup>12</sup>Ptitsyn, O.B. (1966). Conformations of Macromolecules. 297.
- <sup>13</sup>Ptitsyn, O.B. (1966). Conformations of Macromolecules. 297.
- <sup>14</sup>Cantor, C.R. and Schimmel, P.R. (1980). Biophysical Chemistry, Part III: The behavior of biological macromolecules 1041-1073.
- <sup>15</sup>Haynie, D. and Freire, E. (1993). Structural Energetics of the Molten Globule State. 12 PROTEINS: Structure, Function, and Genetics 122.
- <sup>16</sup>See, for example, Skolnick, J., Koliniski, A., and Godzik, A. (1993). From independent modules to molten globules: Observations on the nature of protein folding intermediates. 90 Proc. Natl. Acad. Sci. USA 2099-2100. See, also, Udgaonkar, J. and Baldwin, R. (1995). Nature of the Early Folding Intermediate of Ribonuclease A. 34 Biochemistry 12: 4088-4096; Vidugiris, G., Markley, J., Royer, C. (1995). Evidence for a Molten Globule-like Transition State in Protein Folding from Determination of Activation Volumes. 34 Biochemistry 15: 4909-4912. On the subject of protein folding, see generally Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. (1995). Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. 21 PROTEINS: Structure, Function, and Genetics 167-195; Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D., and Chan, H.S. (1995). Principles of protein folding-A perspective from simple exact models. 4 Protein Science 561-602.

- 
- <sup>17</sup>Daggett, V. and Levitt, M. (1993). Protein Unfolding Pathways Explored Through Molecular Dynamics Simulations. 232 *J. Mol. Biol.* 600.
- <sup>18</sup>Ptitsyn, O.B.I. (1987). Protein folding: hypotheses and experiments. 6 *J. Protein Chem.* 273-293 and references therein; Kuwajima, K. (1989). The molten globules state as a clue for understanding the folding and cooperativity of globular protein structure. 6 *Proteins* 87-103 and references therein. See also Luthey-Schulten, Z., Ramirez, B.E., and Wolynes, P.G. (1995). Helix-Coil, Liquid Crystal, and Spin Glass Transitions of a Collapsed Heteropolymer. 99 *J. Phys. Chem.* 2177-2185.
- <sup>19</sup>Baker, D., Sohl, J.L. and Agard, D. (1992). A protein folding reaction under kinetic control. 356 *Nature* 263-265.
- <sup>20</sup>Daggett, V. and Levitt, M. (1993). Realistic Simulations of Native-Protein Dynamics in Solution and Beyond. 22 *Annu. Rev. Biophys. Biomol. Struct.* 353-80.
- <sup>21</sup>For details of the parameters, water model and the potential function utilized, see Daggett, V. and Levitt, M. (1992). A model of the molten globule state from molecular dynamics simulations. 89 *Proc. Natl. Acad. Sci, USA* 5142-5146; Daggett, V. and Levitt, M. (1993). Realistic Simulations of Native-Protein Dynamics in Solution and Beyond. 22 *Annu. Rev. Biophys. Biomol. Struct.* 353-80.
- <sup>22</sup>Daggett, V. and Levitt, M. (1992). A model of the molten globule state from molecular dynamics simulations. 89 *Proc. Natl. Acad. Sci, USA* 5142-5146.
- <sup>23</sup>Kuroda, Y., Kidokoro, S., Wada, A. (1992). Thermodynamic Characterization of Cytochrome *c* at Low pH: Observation of the Molten Globule State and of the Cold Denaturation Process. 223 *J. Mol. Biol.* 1139-1153.
- <sup>24</sup>Fink, A.L. (1995). Compact Intermediate States in Protein folding. 24 *Annu. Rev. Biophys. Biomol. Struct.* 495-522.
- <sup>25</sup>Wang, Y., Zhang, H., Li, W., and Scott, R. (1995). Discriminating compact nonnative structures from the native structure of globular proteins. 92 *Proc. Natl. Acad. Sci, USA* 709-713.
- <sup>26</sup>Ptitsyn, O.B., and Uversky, V.N. (1994). The molten globule is a third thermodynamical state of protein molecules. 341 *FEBS Letters* 15-18. See especially Haynie, D. and Freire, E. (1993). Structural Energetics of the Molten Globule State. 16 *PROTEINS: Structure, Function, and Genetics* 115-140; Haynie, D. and Freire, E. (1994). Thermodynamics Strategies for Stabilizing Intermediate States of Proteins. 34 *Biopolymers* 261-271 (presentation of theorems concerning the molten globule's thermodynamic properties). Ptitsyn, O. (1996). How molten is the molten globule? 3 *Nature Structural Biology* 6: 488-490.

- 
- <sup>27</sup>The distribution function of molecular dimensions is bimodal during the transitions. Uversky, V., Semisotnov, G., Pain, R. and Ptitsyn, O.B. (1992). 314 FEBS Letters 1:89-92.
- <sup>28</sup>Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 Annual Rev. Biophys. Biomol. Struct. 495-522.
- <sup>29</sup>Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 Annual Rev. Biophys. Biomol. Struct. 495-522. See also Haynie, D. and Freire, E. (1993). Structural Energetics of the Molten Globule State. 12 PROTEINS: Structure, Function, and Genetics 122.
- <sup>30</sup>Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 Annual Rev. Biophys. Biomol. Struct. 496.
- <sup>31</sup>However, see Haynie, D. and Freire, E. (1993). Structural Energetics of the Molten Globule State. 12 PROTEINS: Structure, Function, and Genetics 115-140. (detailing recent protein studies showing differences in enthalpy and heat capacity between the molten globule state and the unfolded state).
- <sup>32</sup>Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 Annual Rev. Biophys. Biomol. Struct. 495-497.
- <sup>33</sup>Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 Annual Rev. Biophys. Biomol. Struct. 497.
- <sup>34</sup>Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 Annual Rev. Biophys. Biomol. Struct. 497.
- <sup>35</sup>Landry, S.J. and Gierasch, L.M. (1994). Polypeptide Interactions with Molecular Chaperones and their Relationship to In Vivo Protein Folding. 23 Annu. Rev. Biophys. Biomol. Struct. 647-651.
- <sup>36</sup>Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. (1995). Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. 21 PROTEINS: Structure, Function, and Genetics 167-195.
- <sup>37</sup>Privalov, P.L. (1996). Intermediate States in Protein Folding. 259 J. Mol. Biol. 707-725.
- <sup>38</sup>See Thirumalai, D. and Woodson, S.A. (1996). Kinetics of Folding of Proteins and RNA. 29 Acc. Chem. Res. 433-439.
- <sup>39</sup>Gao, J. and Weiner, J.H. (1994). Nature of Stress on the Atomic Level in Dense Polymer Systems. 266 Science 748.
- <sup>40</sup>Gao and Weiner (1994) at 748 citing Guth, E. and Mark, H. (1934). 65 Monatsh. Chem. 93.

---

<sup>41</sup>Gao, J. and Weiner, J.H. (1994). Nature of Stress on the Atomic Level in Dense Polymer Systems. 266 *Science* 748-752.

## CHAPTER 6

### ATOMISTIC SIMULATIONS

#### Molecular Dynamics Theory

Molecular dynamics simulations are used to determine individual molecular motions within solids, liquids, or gases.<sup>1</sup> Molecular dynamics simulations investigate temporally-related sequences in the positions of molecules, or trajectories, through the numerical solution of the equations of motion and may be thus viewed as a purely deterministic method.<sup>2</sup>

It should be noted that the complexity of biomolecular systems prohibits solution of Schrodinger's equation due to the large number of atoms and interactions involved.<sup>3</sup> The use of classical mechanics theory and semi-empirical or effective interaction functions is a useful substitute for the exact quantum mechanical solutions.<sup>4</sup> This approach substitutes an analysis of the electronic degrees of freedom for atomic degrees of freedom.<sup>5</sup> Molecular mechanics theory treats electrons implicitly rather than explicitly.<sup>6</sup> The uncertainty principle requires the use of advanced quantum mechanical techniques because momentum and position cannot both be simultaneously known with certainty with respect to an atom.

There are three common stages to a molecular dynamics simulation: [a] model construction, [b] trajectory calculations, and [c] analysis of trajectories.<sup>7</sup> The model consists of the molecule under investigation, based upon defined atomic coordinates derived, for example, through X-ray analysis<sup>8</sup>, and, generally, a fixed environment or



volume of gas, solid, or liquid. The molecule and its environment comprise a system for investigation. The system is typically at thermodynamic equilibrium. Under equilibrium conditions, the physical parameters specifying a given state remain the same, e.g. pressure, volume, or temperature.

The thermodynamic state of the system is a function of the number of molecules ( $N$ ), the volume ( $V$ ), and the total energy ( $E$ ). Successive integrations of Newton's and Langevin's motion equations yield sequential time-related atomic or molecular positions.<sup>9</sup> The integrations produce trajectories or molecular configurations as a function of time.<sup>10</sup> These time-dependent position vectors are combined with time-dependent momentum vectors, formed in response to atomic interactions, to produce a  $6N$ -dimensional hyperspace or phase space. Phase space represents the combination of  $3N$ -dimensional configuration space and  $3N$ -dimensional momentum space.<sup>11</sup>

Three general concerns are relevant to the reliability of a molecular model: [a] explicit representation of the electronic or atomic degrees of freedom "essential" to the representation of the phenomenon of inquiry; [b] the interaction function; [c] a choice of the appropriate set of equations for the molecular motions based upon the available degrees of freedom. Equation sets commonly used in simulations include Schrodinger (quantum mechanical), Newton or Lagrange (classical mechanical) and Langevin (stochastic).<sup>12</sup>

Molecular dynamics of proteins and nucleic acids has been extensively investigated.<sup>13</sup> Specific internal motions of interest in nucleic acids include the relative

vibration of bonded atoms, longitudinal motions of bases in double helices, lateral motions of bases in double helices, global stretching and twisting, elastic vibration of globular region, sugar repuckering, torsional libration of buried groups, relative motion of different globular regions (hinge bending), global bending, allosteric transitions,<sup>14</sup> and local denaturation.<sup>15</sup> The parameters of interest associated with each of these motions include spatial extent and amplitude.<sup>16</sup> Molecular dynamics simulations of helix denaturation have been extensively examined for proteins.<sup>17</sup>

The method of molecular dynamics is also relevant. The commonly used methods include adiabatic dynamics,<sup>18</sup> used in the current study, as well as annealed dynamics,<sup>19</sup> canonical (TVN) dynamics,<sup>20</sup> and temperature-damped dynamics.<sup>21</sup> Under the simulated annealing or “slow cooling” approach, the atoms lose their thermal mobility.<sup>22</sup> These methods may be modified further by the use of “impulse” dynamics where the initial directional velocities of certain atoms are preset to overcome translation barriers and permitting relaxation or “quenched dynamics” where the molecular structure is minimized periodically and saved, following dynamic stages, allowing a search for low energy structures.<sup>23</sup>

### Full Newtonian Simulation or Deterministic Model of Atomic Motion

In a full Newtonian simulation, the complete equations of motion (Newton’s second law)

Equation 5

$$\ddot{\mathbf{r}}_i = \mathbf{F}_i^{\text{int ra}} / m$$

are integrated with respect to time for all atoms including hydrogens and, if applicable, solvent atoms. The trajectory  $\mathbf{r}_i(t)$  of every atom is thus determined for a given set of initial conditions  $\mathbf{r}_i(0)$ . The force exerted on each atom is the negative gradient of the potential energy function:

Equation 6

$$\mathbf{F}_i^{\text{int ra}} = -\partial V / \partial \mathbf{r}_i .$$

In integrating the equations of motion, a time step which is sufficiently short with respect to the smallest time scale of the atomic motions must be employed. The time scale of the atomic motion decreases as the atomic mass is reduced. Therefore, the time step of the simulation and, hence, the overall computational cost is determined by the motion of the lighter (hydrogen) atoms. This makes the computational cost of a full Newtonian simulation prohibitive. In the present study, in order to reduce the computational cost, the CHARMM-based SHAKE algorithm was employed to constrain the lengths of the hydrogen bonds during both the Newtonian and Langevin simulations. The harmonic motions represented by the fluctuation of the hydrogen bonds are weakly coupled to other atomic motions and, hence, may be constrained to within a certain interatomic distance.

#### Stochastic Models of Atomic Motion

Besides the motion due to the interatomic forces, another important component of the atomic motion is the harmonic vibration which corresponds to the thermodynamic

equilibrium of individual atoms with a temperature bath. The time scale of this motion is much shorter than the time scale of the motion which corresponds to the interatomic forces. Therefore, any attempt to accurately follow this motion during a full Newtonian simulation would make the computational cost of the simulation prohibitive. The representation of this vibration with a stochastic (random) motion offers an economical alternative.

The transition from the original deterministic system to the stochastic one may be better understood by introducing the concept of the probability of recurrence, i.e. the probability that a particular state (configuration) will reoccur after a finite time has elapsed. A physical system which follows Newton's laws of motion is deterministic. Therefore, every realizable state of the system can be expected to occur after a finite time with a probability equal to 1. If the expected time of recurrence is much longer than the time scale of the physical phenomena under investigation, then the system can be assumed to be random.

The state of a physical system which is stochastic in nature may be completely described by a probability density function defined in the appropriate phase space. The modeling of a physical system is made substantially easier by introducing a model system which does not obey the same deterministic laws as the original system under investigation but which is described by the same probability density function. For example, it can be shown that the probability density function which describes a model stochastic system which is governed by Brownian kinematics changes in the same way as the probability

density function of a physical system which is governed by gradient diffusion. The two systems are then said to be *statistically equivalent*.

Statistical equivalence is the foundation from which stochastic models of otherwise deterministic physical phenomena (such as Brownian or Langevin models) are derived. In essence, the trajectory  $r(t)$  of an atom is now assumed to be a superposition of a deterministic and a random motion

Equation 7

$$r_i(t) = r_i^{\text{det}}(t) + r_i^{\text{ran}}(t).$$

The deterministic component of the atomic motion  $r_i^{\text{det}}(t)$  is computed in the same manner as in the full Newtonian simulation, i.e. based on Newton's second law. The interatomic force is again computed as the gradient of a potential energy field. However, only the motions due to the interatomic forces are being followed while the harmonic motion which corresponds to the thermodynamic equilibrium with the temperature bath is represented by the random term  $r_i^{\text{ran}}(t)$ . The choice of this random term (closure assumption) is dictated by the requirements of statistical equivalence, i.e. that the effect that this random motion has on the probability density distribution function of the position of the atoms is the same as the effect of the actual harmonic vibration.

The simplest choice for the random component of the atomic motion is a Brownian (random walk) model. In the Brownian model of motion, the state variable (in the present case, position,  $r_i$ ) is assumed to change according to

Equation 8

$$dr_i^{\text{ran}} = dW_i$$

where  $dW_i$  is a random displacement, e.g. a random walk. The key assumption about the statistical properties of  $dW_i$  is that they depend only on the present state ( $r_i$ ) of the system, i.e.  $dW_i$  is independent of any previous states. Naturally, this assumption of uncorrelated changes of state is not consistent with the deterministic character of the underlying physical laws. However, this assumption is acceptable when the correlation time of these changes is much shorter than the time scale of the physical processes which are being modeled.

When the correlation time of these changes of state cannot reasonably be neglected, a Langevin model may be adopted instead. In the Langevin formulation, the assumption of negligible correlation time is shifted from position changes to velocity changes. The random velocity

Equation 9

$$\dot{r}_i^{\text{ran}} = dr_i^{\text{ran}}/dt$$

of the system is assumed to change according to

Equation 10

$$d\dot{r}_i^{\text{ran}} = A(\dot{r}_i^{\text{ran}}, r_i^{\text{ran}}, t)dt + dW_i$$

where  $A(\dot{r}_i^{\text{ran}}, r_i^{\text{ran}}, t)$  is a deterministic acceleration and  $dW_i$  is a random velocity change for which the same assumption of negligible correlation time is maintained. The function  $A(\dot{r}_i^{\text{ran}}, r_i^{\text{ran}}, t)$  determines the correlation profile of the displacements  $dr_i^{\text{ran}}$  which were

assumed to be uncorrelated in the Brownian model. It can be shown that, at the asymptotic limit of zero correlation time, the Langevin model reduces to the Brownian model.

In the present study, a Langevin model of the form

Equation 11

$$d\dot{r}_i^{\text{ran}} = -\frac{\zeta \dot{r}_i}{m} dt + \frac{F_i^{\text{inter}}}{m} dt$$

was adopted. The feedback function  $A(\dot{r}_i^{\text{ran}}, r_i^{\text{ran}}, t)$  has been chosen so that it has the same effect as a viscous damping force,  $-\zeta \dot{r}_i$ , where  $\zeta = m b$  and  $b$  is the friction coefficient of the particular atom. The random acceleration,  $dW_i$ , is the effect of a random interatomic force  $F_i^{\text{inter}}$  with the following statistical properties

Equation 12

$$\langle F_i^{\text{intra}}(t) \rangle = 0$$

$$\langle F_i^{\text{intra}}(t) F_i^{\text{intra}}(0) \rangle = 2k_b T_0 b m \delta(t)$$

where  $k_b$  is Boltzmann's constant and  $T_0$  is the absolute temperature of the temperature bath. It can be shown that the present choice of the feedback function  $A(\dot{r}_i^{\text{ran}}, r_i^{\text{ran}}, t)$  corresponds to an exponential temporal correlation of the random atomic displacements with an autocorrelation time  $\tau_c = m/\zeta$ .

The final form of the equation describing atomic motion in the case of the stochastic simulation (compare with Equation 6) is

## Equation 13

$$d\dot{r}_i = -\frac{\zeta \dot{r}_i}{m} dt + \frac{F_i^{\text{inter}}}{m} dt + \frac{F_i^{\text{intra}}}{m} dt$$

---

<sup>1</sup>See generally Haile, J. (1992). *Molecular Dynamics Simulation: Elementary Methods*. (John Wiley); van Gunsteren, W.F. and Berendsen, H.J.C. (1990). *Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry*. 29 *Agnew. Chem. Int. Ed. Engl.* 992-1021; Caspar, D.L.D. (1995). Problems in simulating macromolecular movements. 3 *Structure* 327-329. (insufficient sampling of conformational sub-states); Fernandez, A. (1993). Simulating an exploration of RNA conformational space with an appropriate parallel-updating strategy. 48 *Physical Review E* 4: 3107-3111.

<sup>2</sup>In a purely stochastic methodology, such as Metropolis Monte Carlo, a molecular configuration is dependent solely upon the immediately preceding configuration, i.e. a Markov chain. The spectrum between purely stochastic and purely deterministic simulation methods is composed of hybrid methods with elements, to varying degrees, of both methods: Force-Biased Monte Carlo, Brownian Dynamics, and General Langevin Dynamics in order of increasing deterministic properties. Haile, J. (1992) at 13.

<sup>3</sup>van Gunsteren, W. and Mark, A. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 947-948.

<sup>4</sup>van Gunsteren, W. and Mark, A. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 948. See also Roitberg, A., Gerber, R., Elber, R., Ratner, M. (1995). Anharmonic Wave Functions of Proteins: Quantum Self-Consistent Field Calculations of BPTI. 268 *Science* 1319. Comba, P. (1996). *Inorganic Molecular Mechanics*. 73 *Journal of Chemical Education* 2: 108. (pictorial description of parameters used in a molecular mechanics force field). Force field equations simulate various interactions that describe the potential energy surface of a molecule. See further Iachello, F. and Levine, R.D. (1995). *Algebraic theory of molecules*. (New York: Oxford University Press) 156-189.

<sup>5</sup>van Gunsteren, W. and Mark, A. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 948. Ab initio models include Hartree-Fock models and correlated models. The Hartree-Fock models include the Born Oppenheimer approximation involving separation of nuclear and electron motions, the Hartree-Fock approximation involving separation of electron motions, and the LCAO (linear combination of atomic orbitals) approximation. *Chemistry with computation: An Introduction to SPARTAN*. 1:1. (1995).



- 
- <sup>6</sup>Lipkowitz, K. (1995). Abuses of Molecular Mechanics. 72 *Journal of Chemical Education* 12: 1070-1075.
- <sup>7</sup>Haile, J. (1992) at 13.
- <sup>8</sup>Hendrickson, W.A. (1995). X Rays in Molecular Biophysics. 48 *Physics Today* 11:42-43.
- <sup>9</sup> $F_i(t) = m\dot{r}_i(t) = -\partial\mu(rN)/\partial r_i$ . Haile, J. (1992) at 15. Newtonian dynamics, which is used in the current study, should be contrasted with Hamiltonian dynamics, based upon Newton's second law, wherein a time-independent position and velocity function, the Hamiltonian, is used. Haile, J. (1992) at 40-42.
- <sup>10</sup>van Gunsteren, W. and Mark, A. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 948.
- <sup>11</sup>Haile, J. (1992) at 43. In  $3N$ -dimensional configuration space, the coordinate axes are components of position vectors,  $r_i(t)$ ; the coordinate axes are components of momentum vectors,  $p_i(t)$  in  $3N$ -dimensional momentum space. A single point in space represents the positions and momenta of the entire  $N$ -atom system at a single point in time.
- <sup>12</sup>van Gunsteren, W.F., Luque, F.J., Timms, D. and Torda, A.E. (1994). *Molecular Mechanics in Biology: From Structure to Function, Taking Account of Solvation*. 23 *Annu. Rev. Biophys. Biomol. Struct.* 849.
- <sup>13</sup>See generally McCammon, J. and Harvey, S. (1989). *Dynamics of proteins and nucleic acids*. (Cambridge: Cambridge University Press); Brooks, C.L., Karplus, M., and Pettitt, B.M. (1988). *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. (Advances in Chemical Physics Volume LXXI) (John Wiley and Sons: New York).
- <sup>14</sup>The term allosteric effects refers to conformational changes in a biomolecule's structure which alter affinity for substrates, e.g. oxygen-binding and hemoglobin. Sybesma, C. (1977). *Biophysics: An Introduction*. (The Netherlands: Kluwer Academic Press). 118-119.
- <sup>15</sup>McCammon (1989) at 29.
- <sup>16</sup>McCammon (1989) at 29.
- <sup>17</sup>Daggett, V. and Levitt, M. (1992). Molecular Dynamics Simulations of Helix Denaturation. 2223 *J. Mol. Biol.* 1121.
- <sup>18</sup>In this method, the molecular structure's temperature is maintained within a specified range by periodic scaling of the atomic velocities to constrain the kinetic energy. Molecular Simulations, Inc. (6/93). CERIUS Minimizer/Dynamics Document Update.18-7.

- 
- <sup>19</sup>The molecular structure's energy is gradually minimized to avoid local minimum "trapping". The temperature is incrementally increased, as a function of time, to a designated temperature and back to the starting temperature. The lowest energy structure within each temperature cycle is minimized. Molecular Simulations, Inc. (6/93). CERIOUS Minimizer/Dynamics Document Update.18-7.
- <sup>20</sup>In this isothermal method, temperature, volume, and number of atoms is held constant in isothermal fashion by exchanging energy with a heat bath. The degree of thermal coupling between the system and the heat bath may be regulated by modification of the relaxation time. The heat bath may provide energy needed to overcome rotational barriers to conformations. Molecular Simulations, Inc. (6/93). CERIOUS Minimizer/Dynamics Document Update.18-8.
- <sup>21</sup>This adiabatic method maintains a constant temperature through a "weak" coupling scheme of sequentially scaling atomic velocities. Molecular Simulations, Inc. (6/93). CERIOUS Minimizer/Dynamics Document Update.18-8.
- <sup>22</sup>Press, W.H., Flannery, B.P., Teukolsky, S. A., and Vetterling, W.T. (1986). Numerical Recipes. The Art of Scientific Computing. (Cambridge University Press). 327
- <sup>23</sup>Molecular Simulations, Inc. (6/93). CERIOUS Minimizer/Dynamics Document Update.18-9.

## CHAPTER 7

### POTENTIAL ENERGY FUNCTIONS

As noted by one author, “[t]he form of the potential function and parameters [force constants] are generally chosen with the goal of mimicking the physics of interatomic interactions.”<sup>1</sup> The correct choice of a reliable potential energy function is based upon considerations of intramolecular energy, charge distributions, hydrogen bonding, dispersion coefficients and other factors.<sup>2</sup> The essential components of a force field include atom types, auto-typing rules, functions for energy terms, parameters, and an unspecified parameters generator.<sup>3</sup> The potential energy function or force field equation expresses the functional dependence of potential energy upon individual atomic position.<sup>4</sup>

Three general interaction models are used in molecular simulations. The Hooke’s Law model treats the atoms or molecules as connected to their neighbors by springs. The hard sphere model treats particle interactions as “billiard ball” interactions with an interaction, i.e. bouncing off, taking place only when the particles are within a certain distance of each other. The Lennard-Jones model is characterized by “forces that are strongly repulsive at very short interparticle distances, attractive at larger distances, and extremely weak attractive at very large distances.”<sup>5</sup>

Typically, the model for the intermolecular potential,  $\mu(r/N)$ , is considered pairwise additive and accordingly represents the interaction energy among  $N$  atoms as the sum of

isolated two-body contributions:  $\mu(rN) = \sum\sum u(r_{ij})$  where  $i < j$ . The Lennard-Jones model for soft-sphere pair potential is generally modified, to varying degrees, in producing force fields to encompass the energy terms required for different classes of molecules, including biomolecules.<sup>6</sup> Molecular mechanics, or optimum geometries, and molecular or motional dynamics are evaluated from the forces calculated for two-, three-, and four-body interactions.<sup>7</sup> Using truncated potentials, minimum image criterion, and neighbor lists as timesaving measures often optimizes calculations of atomic forces.<sup>8</sup>

The potential function generally includes terms for bonded interactions (e.g. harmonic restoring forces between bonded nearest neighbors, penalties for angle deformations, and dihedral torsional potentials for hindered rotation of groups about bonds) and nonbonded interactions between separated atoms (including repulsive van der Waals forces, dispersion attraction, and partial charge electrostatic interactions).<sup>9</sup> Parameter data is obtained experimentally.

The governing relationship between potential energy, atomic positions and atomic velocities is given by

Equation 14

$$-\frac{dV}{dx} = F = ma = m\frac{d^2x}{dt^2}$$

where  $V$  equals the potential energy,  $x$  represents atomic positions,  $t$  represents time,  $m$  equals mass, and  $F$  equals force. The following is an example of a typical biomolecular force field or effective atomic interaction system:

Equation 15

$$\begin{aligned}
V(r_1, r_2, \dots, r_N) = & \sum \frac{1}{2} K_b [b - b_o]^2 + \sum \frac{1}{2} K_\theta [\theta - \theta_o]^2 + \sum \frac{1}{2} K_\xi [\xi - \xi_o]^2 \\
& \text{all bonds} \qquad \qquad \qquad \text{all bond angles} \qquad \qquad \qquad \text{improper dihedrals} \\
& + \sum K_\phi [1 + \cos(n\phi - \delta)] + \sum \left[ \frac{C_{12}(ij)}{r_{ij}^{12}} - \frac{C_6(ij)}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_o \epsilon_r r_{ij}} \right] \\
& \text{dihedrals} \qquad \qquad \qquad \text{all pairs (i, j)}
\end{aligned}$$

The first term represents the covalent bond-stretching interaction, a harmonic potential, along bond  $b$ . The bond lengths and force constants,  $K_b$ , depend upon the bond type. The second term represents the three-body interaction or bond-angle bending. The third term represents dihedral-angle or four-body interactions and the fourth term represents the effective non-bonded interactions as a sum over all pairs of atoms (van der Waals and Coulomb interactions).<sup>10</sup>

It should be noted, however, that “the actual behavior of molecules is governed by the nature of the free energy hyperspace, not that for the potential energy.”<sup>11</sup> The calculation of the relative free energy of systems has traditionally been computationally prohibitive. Modern free energy perturbation methods such as compositional or conformational (potential of mean force) calculations and increasingly powerful computational methods<sup>12</sup> will result in more accurate simulations.

---

<sup>1</sup>Malhotra, A., Tan, R.K-Z., and Harvey, S.C. (1994). Modeling Large RNAs and Ribonucleoprotein Particles Using Molecular Mechanics Techniques. 66 Biophysical Journal 1777, 1778.

- 
- <sup>2</sup>Rullmann, J.A.C. and van Duijnen, P.Th. (1990). Potential Energy Models of Biological Macromolecules: A Case for Ab Initio Quantum Chemistry. 1 Reports in Molecular Theory 1-21.
- <sup>3</sup>Molecular Simulations, Inc. (6/93). CERIOUS Open Force Field Document Update. 7-6.
- <sup>4</sup>Daggett, V. and Levitt, M. (1993). Realistic Simulations of Native-Protein Dynamics in Solution and Beyond. 22 Annu. Rev. Biophys. Biomol. Struct. 355.
- <sup>5</sup>Fosdick, L.D., Jessup, E.R., Schauble, C., and Domik, G. (1996). An Introduction to High-Performance Scientific Computing. (Cambridge: MIT Press). 535-547.
- <sup>6</sup>Haile, J. (1992) at 188-197.
- <sup>7</sup>Molecular Simulations, Inc. (6/93). CERIOUS Open Force Field Document Update. 7-4.
- <sup>8</sup>Haile, J. (1992) at 190-197.
- <sup>9</sup>Brooks, C.L., Karplus, M., and Pettit, B.M. (1988). Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. (Advances in Chemical Physics Volume LXXI) (John Wiley and Sons: New York) at 25.
- <sup>10</sup>van Gunsteren, W. and Mark, A. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 Eur. J. Biochem. 948-949.
- <sup>11</sup>Pearlman, D.A. and Kollman, P.A. (1991). Evaluating the Assumptions Underlying Force Field Development and Application Using Free Energy Conformational Maps for Nucleosides. 113 J. Am. Chem. Soc. 7177. See also Pearlman, D. and Kollman, P. (1990). Are Free Energy Calculations Necessary? A Comparison of DNA Modeling Studies in Beveridge, D. and Lavery, R., eds. (1990). Theoretical Biochemistry & Molecular Biophysics. (New York: Adenine Press). See generally Mezey, P.G. (1987). Potential Energy Hypersurfaces. Studies in physical and theoretical chemistry, v. 53. (New York: Elsevier).
- <sup>12</sup>Plimpton, S. and Hendrickson, B. (1996). A New Parallel Method for Molecular Dynamics Simulation of Macromolecular Systems. 17 Journal of Computational Chemistry 3:326-337.

## CHAPTER 8

### SIMULATIONS

#### Molecular Complex Preparation

The molecule was prepared for dynamics in several steps. Initially, the X-ray crystallographic coordinates of the hammerhead ribozyme, as deposited in the Brookhaven Protein Databank, were imported into the QUANTA<sup>1</sup> molecular modeling program. Hydrogen atoms were added to the ribozyme and a sufficient number of magnesium ions were added to the set of coordinates to create an electrically neutral molecule-ion complex. The placement of magnesium ions in each complex was essentially random although an attempt was made to place the ions in close proximity to the phosphate groups on each molecule under the expectation that the positively charged magnesium ions would move towards the negatively charged phosphate groups during minimization and dynamics. The complex was “solvated” through the use of a distance-dependent dielectric constant.

#### Energy Minimization

Minimization of the complex was accomplished through the CHARMM<sup>2</sup> program for molecular mechanics and dynamics within the QUANTA environment. CHARMM-based nucleic acid topology and potential function parameter files<sup>3</sup> were employed in the minimization and dynamics simulations.

Minimization is used to reduce a macromolecule’s potential energy to its lowest possible value before running molecular dynamics. The molecule must be “relaxed”

because “the crystallographic structure is itself a model, generated by balancing the desire for atomic positions that reproduce the observed diffraction pattern against the desire for a model with traditional values for bond lengths, bond angles, [etc].”<sup>4</sup> Theoretically, the molecule will have zero kinetic energy at its potential energy minimum. The potential energy gradient, or first derivative, is brought to zero through atomic positioning and repositioning through a variety of algorithms.<sup>5</sup> Minimization algorithms differentiate the potential energy function with respect to each of the molecule’s x, y, and z atomic coordinates to achieve a gradient very close to zero. The various minimization methods traditionally used in simulations may differ in their use of first and second derivatives in optimizing a molecule’s structure.<sup>6</sup>

Steepest descents is used to quickly descend toward the global minimum potential energy. Convergence takes place under the steepest descents method by successively adjusting the molecular coordinates in the negative direction of the first derivative of the potential energy function.

The size of each successive step is determined by the immediately prior change in potential energy. A reduction in potential energy dictates a larger step size; an increase in potential energy dictates a smaller step size. Convergence is difficult to achieve using steepest descents minimization alone and is regularly achieved through a combination of methods. In general, steepest descents will “rapidly improve a very poor conformation.”<sup>7</sup> Additionally, a large number of iterations are generally required.<sup>8</sup>



The method of conjugate gradients uses both prior and current gradients to determine the size of successive steps. Mathematically, convergence is achieved in  $N$  steps for a quadratic energy surface where  $N$  is the number of degrees of freedom in the energy.

The adopted basis Newton-Raphson minimization method iteratively solves the Newton-Raphson minimization equations by computing a derivative of the gradient using a second derivative matrix diagonalization to find an optimum step size along eigenvectors. Small or negative eigenvalues are subjected to additional energy and gradient determinations to set necessary step size and direction. Rapid convergence is often achieved through this algorithm. The adopted basis method avoids saddle points on the energy surface by numerically constructing the second derivative matrix from gradient vector changes.<sup>9</sup>

Minimization procedures also allow normal mode analysis. The equation for a molecular potential surface of  $n$  atoms can be used to examine the vibrations around a particular minimum by transforming the potential energy expressions from Cartesian coordinate representations to normal coordinates to calculate the vibrational frequency of a normal mode. Modes with frequencies corresponding to intramolecular vibrations, molecular translation, and molecular rotation as well as harmonic normal modes can be used to describe molecular properties such as average thermal atomic motion.<sup>10</sup>

The potential energy gradient of the complex was brought to 1.53 kcal per mole through the use of steepest descents, conjugate gradients, and adopted-basis Newton-

Raphson minimization techniques. The minimized structure of the hammerhead ribozyme complex is shown in Figure 6.

### Solvent Representation

The use of a suitable solvation model is very important in macromolecular simulations as longtime dynamics of such molecules in solution are “governed by diffusion.”<sup>11</sup> The relevant interaction is between solute atoms and solvent atoms—the free energy of hydration—which produces conformation and thermodynamic properties.<sup>12</sup> As stated by one researcher, “the enzyme plus medium must be regarded as a unit.”<sup>13</sup>

Among the roles of solvent molecules in simulations are [a] improvement in macromolecular packing; [b] reduction of “sizeable” cavities through minimization of free surface area due to surface tension effects; [c] satisfying unmatched hydrogen-bond relationships through hydrogen-bonding capacity, [d] exertion of shielding effect on electrical interaction between charges or dipoles in the protein or nucleic acid yielding solvent-polarity dependent changes in structure and stability and [e] viscosity of solvent.<sup>14</sup>

Several methods of solvation were contemplated for the solvation of the molecule-ion complex during the simulation: full-scale immersion into a box of water molecules with implementation of periodic boundary conditions resulting in a “unit” of tiled water boxes or cells<sup>15</sup> and the use of a distance-dependent dielectric constant. The use of periodic boundary conditions in a solvent simulation permits the incorporation of fewer water molecules in a simulation with no change in results from using bulk water. The Rahman method treats a rectangular box of water molecules as a period system—a constant number of neighboring

water molecules is maintained both inside and outside the cell. The departure of one water molecule from one exit is followed by the entry of the same water from the opposite direction.<sup>16</sup>

Both methods were viewed with an eye toward reducing the computational cost of the simulations. The large number of water molecules and amount of computational time required in a periodic boundary condition water box or bulk solvent approach prompted the use of a distance-dependent dielectric constant.<sup>17</sup>

### Dynamics

After minimization, two sets of dynamics simulations were initiated by Newtonian and Langevin dynamics, using the Verlet<sup>18</sup> algorithm, to perform initialization of atomic positions and velocities, heating and equilibration for the simulations. In the Newtonian case, heating to the desired temperatures (300K and 500K) as well as the subsequent equilibration, is achieved by scaling the initial velocities. The heating process begins at 0 K and the temperature was rapidly increased incrementally over several picoseconds. In the Langevin case, heating was accomplished through the use of a temperature bath at 300K or 500K. All molecular complex atoms came within the influence of the heatbath. The permissible temperature deviation from the final temperature for all phases of the simulations was plus or minus 10 degrees. Violation of these boundaries triggered atomic velocity scaling protocols.

Equilibration, after heating to the desired temperature, was accomplished over time in the simulations.<sup>19</sup> The complex at the different temperatures, and under the different

methods, was subjected to a 600 picosecond molecular dynamics simulation with 2 femtosecond intervals between steps and a history file written every ten steps.

Initialization of the molecule-ion complex is accomplished by setting the initial positions,  $r_i(0)$ , of the atoms to their respective X-ray, minimized, or previous molecular dynamics, for the ions, coordinates with time  $t = 0$ . The initial velocities are assigned through random number generation.<sup>20</sup> The simulations were run on Silicon Graphics Indigo2 Unix Workstations under the IRIX 5.3 operating system. A total of four simulations of the molecule-ion complex were conducted: Newtonian and Langevin at 300 and 500 degrees Kelvin, respectively.

A distance-dependent dielectric constant was incorporated into the simulation. The steps of production dynamics were run and trajectory data, in the form of conformation ensembles, was collected for analysis. The analysis items include conformation-dependent (internal motions)<sup>21</sup> and thermodynamic properties. The conformation-dependent properties include the radius of gyration, hydrogen bonding and dipole moment. The thermodynamic properties include temperature, potential energy, kinetic energy, and total energy. The properties are plotted as a function of time over the 600 picosecond simulations and the trajectories at each temperature are analyzed in terms of local and large-scale structural and global energy changes. Additional analysis is conducted with respect to the root-mean-square displacement of backbone (phosphate) and side-chain (base) atoms. Actual global relaxation is expected to occur on a time scale longer than that of the present simulations.

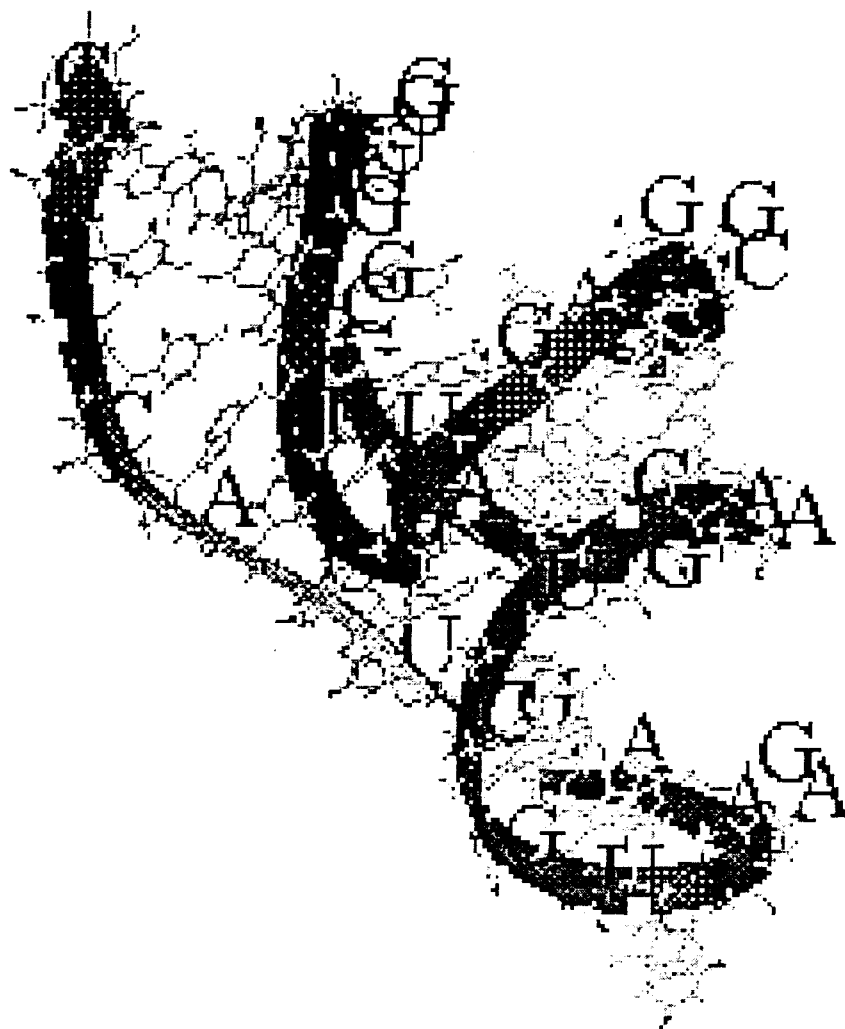


Figure 6. Minimized Structure showing  $Mg^{2+}$  ions (small squares).

---

<sup>1</sup>Molecular Simulations, Inc.

<sup>2</sup>CHARMm is an acronym for Chemistry at Harvard-Macromolecular Mechanics. See Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983). CHARMm: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. 4 *Journal of Computational Chemistry* 187. Representation of objects in Cartesian space (x,y,z coordinates) is substituted for a CHARMm internal coordinate system in which CHARMm utilizes a Cartesian coordinate representation for the first 3 coordinates and bond angle, bond distance/length, and rotational (torsional) angle for the fourth coordinate. The potential energy function ("force field") is manipulated in terms of the internal coordinates.

<sup>3</sup>MacKerell, A. and Wiorkiewicz, J., All-Hydrogen Nucleic Acid Parameter File (Developmental), Version 5.0 (June 1992); MacKerell, A. and Wiorkiewicz-Kucera, J., All-Hydrogen Nucleic Acid Topology File (Developmental), Version 5.0 (June 1992).

<sup>4</sup>McCammon, J. and Harvey, S. (1987) at 187.

<sup>5</sup>Daggett, V. and Levitt, M. (1993). Realistic Simulations of Native-Protein Dynamics in Solution and Beyond. 22 *Annu. Rev. Biophys. Biomol. Struct.* 357.

<sup>6</sup>CHARMm Course Outline (1990). Lecture 7.

<sup>7</sup>CHARMm documentation, MINIMIZ.DOC (1991) at 4.

<sup>8</sup>Many degrees of freedom in macromolecules may produce many local minima. See generally Warshel, A. (1991). *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. (New York: John Wiley and Sons). 113-117.

<sup>9</sup>CHARMm documentation, MINIMIZ.DOC (1991) at 5.

<sup>10</sup>Warshel, A. (1991). *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. (New York: John Wiley and Sons). 117-118. Lattice dynamics is used to examine vibratory motions of molecules in solids. Haile, J. (1992). *Molecular Dynamics Simulations. Elementary Methods*. (John Wiley) 1.

<sup>11</sup>Perico, A., Guenza, M., Mormino, M., and Fioravanti, R. (1995). Protein Dynamics: Rotational Diffusion of Rigid and Fluctuating Three Dimensional Structures. 35 *Biopolymers* 47; Beveridge, D.L. and DiCapua, F.M. (1989). Free Energy via Molecular Simulation: Applications to Chemical and Biomolecular Systems. 18 *Annu. Rev. Biophys. Biophys. Chem.* 431, 461. See also Still, C., Tempczyk, A., Hawley, R.C., and Hendrickson, T., (1990). Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. 112 *J. Am. Chem. Soc.* 6127-6129. (solvent treated as a statistical continuum); van Gunsteren, W.F., Luque, F.J., Timms, D. and Torda, A.E. (1994). *Molecular Mechanics in Biology: From Structure to Function, Taking Account of*

- 
- Solvation. 23 *Annu. Rev. Biophys. Biomol. Struct.* 847-863; Saenger, W. (1987). Structure and Dynamics of Water Surrounding Biomolecules. 16 *Ann. Rev. Biophys. Chem.* 93-114; Ben-Naim, A., Ting, K.-L., Jernigan, R.O.L. (1989). I. Separation of the Volume and Surface Interactions with Estimates for Proteins. 28 *Biopolymers* 1309-1325.
- <sup>12</sup>Kang, Y.K., Gibson, K.D., Nemethy, G., and Scheraga, H.A. (1987). Free Energies of Hydration of Solute Molecules. 4. Revised Treatment of the Hydration Shell Model. 92 *Journal of Physical Chemistry* 16: 4739-4742.
- <sup>13</sup>Welch, G.R. (1986). *The Fluctuating Enzyme*. (New York: John Wiley and Sons) ix.
- <sup>14</sup>van Gunsteren *et al.* (1994) at 852. See also discussion of boundaries, long-range electrostatic effects and the approximate treatment of solvent effects (853-856).
- <sup>15</sup>Daggett (1993) at 359.
- <sup>16</sup>Daggett, V and Levitt, M. (1993). Realistic Simulations of Native-Protein Dynamics in Solution and Beyond. 22 *Annu. Rev. Biophys. Biomol. Struct.* at 358. Such an approach may also be necessary to avoid “edge” effects in the case of a small number of molecules where the “fraction near the boundary of the system is far greater than a real system would have.” Winn, J. (1995). *Physical Chemistry* 651.
- <sup>17</sup>See generally Banks, J., Brower, R., Ma, J. (1995). Effective Water Model for Monte Carlo Simulations of Proteins. 35 *Biopolymers* 331-341.
- <sup>18</sup>The Verlet method is a finite-difference, third-order Störmer algorithm, used to determine positions and velocities. The addition of two Taylor expansion series of position from time  $t$  forward to time  $t + \Delta t$  and from time  $t$  backward to time  $t - \Delta t$ , respectively, produces the Verlet algorithm for positions.

Forward expansion series:

$$x(t + \Delta t) = x(t) + (dx(t)/dt) (\Delta t) + \frac{1}{2} (d^2x(t)/dt^2) \Delta t^2 + \frac{1}{3!} (d^3x(t)/dt^3) \Delta t^3 + O(\Delta t^4)$$

Backward expansion series:

$$x(t - \Delta t) = x(t) - (dx(t)/dt) (\Delta t) + \frac{1}{2} (d^2x(t)/dt^2) \Delta t^2 - \frac{1}{3!} (d^3x(t)/dt^3) \Delta t^3 + O(\Delta t^4)$$

Verlet algorithm for positions:

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + (d^2x(t)/dt^2) \Delta t^2 + O(\Delta t^4)$$

Estimation of velocities is accomplished through the use of a first-order central difference estimator,  $v(t) \approx [x(t + \Delta t) - x(t - \Delta t)]/2\Delta t$ . Acceleration is derived from Newton's second law and intermolecular forces. Haile, J. (1992) at 158-159 and references cited therein, including Verlet, L. (1967). *Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*. 159 *Phys. Rev.* 98.; Beeman,

---

D. (1976). Some Multistep Methods for Use in Molecular Dynamics Calculations. 20 J. Comput. Phys. 130. Additional Taylor expansion series-based finite-difference methods routinely used in molecular dynamics simulations include Euler's method, the multi-order Runge-Kutta methods, and the general and Gear predictor-corrector methods. The algorithms differ in their stability, i.e. amplification of truncation and round-off errors through sequential steps. Haile, J. (1992) 148-166.

<sup>19</sup>These procedures "decrease[s] [sic] the probability that localized fluctuations in the energy (e.g. "hot spots") will persist throughout the simulation." Brooks, C.L., Karplus, M., and Pettitt, B.M. (1988). *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. (Advances in Chemical Physics Volume LXXI) (John Wiley and Sons: New York) at 34. Equilibration occurs when the system's energy "settles into a reasonable approximation of an oscillation about a mean." Beveridge, D.L. and DiCapua, F.M. (1989). Free Energy via Molecular Simulation: Applications to Chemical and Biomolecular Systems. 18 *Annu. Rev. Biophys. Biophys. Chem.* 435.

<sup>20</sup>If  $x_x$ ,  $x_y$ , and  $x_z$  are random numbers distributed uniformly over the interval from -1 to +1, the corresponding initial velocity coordinates are given by  $v_{ix}(0) = dx_x/x$ ,  $v_{iy}(0) = dx_y/x$ , and  $v_{iz}(0) = dx_z/x$ .  $x = [dx_x^2 + dx_y^2 + dx_z^2]^{1/2}$  Haile, J. (1992) at 202.

<sup>21</sup>See generally Cantor and Schimmel. (1980). *Biophysical Chemistry, Part III: The behavior of biological macromolecules*. 980-1018. (theory of calculation of conformation-dependent properties of polymer chains).



## CHAPTER 9

### TRAJECTORY ANALYSIS

The analysis of the trajectory data generated in the simulations must be carefully considered.<sup>1</sup> Thermodynamic (temperature, kinetic energy, potential energy, total energy) and conformational (radius of gyration, hydrogen bonds, dipole moment) variable data were collected for analysis. Each trajectory was arbitrarily separated into two phases. The first phase, Phase I, consisted of the time period from 0 to 300 ps; Phase II covered the time period from 301 to 600 ps.

The selected properties were chosen for their value in contributing to an overall characterization of the biomolecular system. The temperature of a system is directly related to the kinetic energy by

Equation 16

$$\sum_{i=1}^{N_a} \frac{m_i v_i^2}{2} = \frac{N_f k_B T}{2}$$

where  $N_a$  is the number of atoms,  $m_i$  is the mass of atom  $i$ ,  $v_i$  is the velocity of atom  $i$ , and  $N_f$  is the number of degrees of freedom. The radius of gyration reflects the root mean square displacement of the atoms from center of mass of the molecule. The number of hydrogen bonds formed through intramolecular base pairing is an indication of the relative degree of structure of the molecule. In the case of a group of distributed charges, a dipole moment exists which may be defined by

Equation 17

$$\mu = \sum_i q_i \mathbf{r}_i$$

or in the case of a continuous distribution of charge,

Equation 18

$$\mu = \int \rho(\mathbf{r}) \mathbf{r} dV$$

where  $\rho(\mathbf{r})$  is the charge density at a given position  $\mathbf{r}$ . “Any system of charges (even if the system has a zero net charge) whose center of positive charge does not coincide with the center of negative charge will have a dipole moment.”<sup>2</sup> The dipole moment is essentially an indication of the alignment of the molecule’s electrical field. The dielectric constant is a measure of how much a particular substance will reduce the electric field.

Two major time frames are relevant to the analysis of the dynamics: (a) the pre-equilibration phase where heating is accomplished and equilibration occurs and (b) the post-equilibration “true” dynamics period where meaningful conclusions may be drawn regarding the molecule’s search of conformational space.

The graphs on the following pages show the evolution of the properties over Phase I and Phase II. The first 50 ps of each simulation represent the results of a combined heating and equilibration protocol within QUANTA. For temperature, equilibration appears to be reached in all cases (Figure 8(a)-(f)). For energy, in the Newtonian simulations, equilibration is reached by 50 ps (Figure 9). However, in the case of the Langevin method, equilibration does not appear to have occurred until 200 ps (Figure 9). By Phase II, energy

equilibration has been reached in all simulations. The results for the radius of gyration are similar to those of the energy. In the case of hydrogen bonds, Langevin simulations (Figure 11) equilibrated more quickly than Newtonian (Figure 11) but both are fully equilibrated by 50 ps at 300K and 500K in Phase II. The dipole moment equilibrated rapidly in the Newtonian simulations at both 300K and 500K (Figure 12), but showed at least two major transitions in the Langevin simulations before equilibrium positions were reached (Figure 12).

The decrease in potential energy and contemporaneous increase in hydrogen bonding (Figure 11) during the early phases of dynamics indicate that molecular dynamics functions as a very effective energy minimization routine. This may be due to the greater conformational space being explored in the dynamics simulation as compared to the energy minimization procedure. Overall, the properties equilibrated quickly and remained relatively constant during the simulations. These observations were maintained regardless of the dynamics method or the temperature utilized. In particular, the radius of gyration and dipole moment showed the most change between dynamics methods. The radius of gyration took a longer time to equilibrate or “settle down” and achieved higher values in the Langevin simulations than in the Newtonian simulations. The dipole moment required a longer time to reach a “steady state” value in the Langevin dynamics overall and showed an increase in value in Phase II dynamics as compared to a decreasing dipole moment returned in Newtonian Phase II dynamics. Additionally, over the long run, Langevin dynamics appears to have produced conformations with somewhat less structure than Newtonian

dynamics as evidenced by the formation of fewer hydrogen bonds and possibly greater radius of gyration in the Langevin case. No major conformational changes were observed during any of the simulations; however, detailed structural analysis (RMSD), in the case of the 500K Newtonian dynamics simulation, showed variation in flexibility for different parts of the molecule.

### Comparison of Simulation Trajectories

Table 1 presents the results for Phase I and Phase II dynamics. In general, the Langevin simulation maintained a slightly higher average potential energy in both phases, a correspondingly slightly higher total energy in both phases, and a significantly higher dipole moment in both phases at both temperatures. As required, the kinetic energy increased in the 500K simulations compared to the 300K ones, but there was a decrease in the potential energies at the higher temperatures. This appears to be due to the breaking of some non-covalent interactions. For example, in both the Newtonian and Langevin simulations, the number of hydrogen bonds observed decreased at the higher temperature. For the Newtonian, but not the Langevin, simulations, the radius of gyration increased at the higher temperature, possibly indicating fewer van der Waals interactions since the molecular mass would be distributed over a larger volume.

Table 1. Comparison of Average Molecular Properties Between Newtonian and Langevin Dynamics Simulations at 300K and 500K

SIMULATION	PE	KE	TE	TEMP	HB	RG	DM
ND 300(I)	-6105.35	1081.14	-5024.20	306.34	51.85	13.45	167.36
LD 300 (I)	-5425.18	1052.74	-4372.44	297.78	53.18	14.77	448.91
ND 300 (II)	-6188.09	1095.95	-5092.14	310.53	53.48	13.21	150.01
LD 300 (II)	-5886.77	1053.28	-4833.49	297.94	57.31	14.37	364.89
ND 500 (I)	-5454.14	1793.24	-3660.89	508.11	44.56	14.71	300.01
LD 500 (I)	-4927.57	1754.03	-3173.54	496.16	44.25	14.34	425.55
ND 500 (II)	-5538.52	1848.36	-3690.15	523.73	48.59	14.22	300.58
LD 500 (II)	-5403.34	1754.45	-3648.88	496.26	47.12	13.43	342.81

Symbols: ND (Newtonian Dynamics), LD (Langevin Dynamics), PE (Potential Energy, kcal/mole), KE (Kinetic Energy, kcal/mole), TE (Total Energy, kcal/mole), TEMP (Temperature K), HB (Hydrogen Bonds), RG (Radius of Gyration, A), DM (Dipole Moment, Debyes)

For higher level analysis, including RMSD analysis, the 500K Newtonian dynamics simulation was chosen because the Newtonian simulations appear to reach equilibrium sooner than the Langevin ones, allowing more points for analysis in the production phase, and because the 500K simulation would explore more conformational space than the 300K ones.

### Final Structures

The pictures in Figure 13 present the final structures of the hammerhead ribozyme complex at 600 ps.

### In-Line Cleavage Mechanism

The hammerhead ribozyme cleavage reaction is believed to proceed via an “in-line” mechanism whereby the O5' and P atoms of the adenine residue of the scissile bond and the

O2' atom of the cytosine residue on the other side of the scissile bond align themselves into a linear or 180 degree pattern.<sup>3</sup> The 500K Newtonian dynamics simulation was analyzed to determine whether the atoms had so aligned themselves during the simulation. Figure 14 shows the achieved angle as a function of time for the Phase I and Phase II dynamics, respectively.

As shown, the angular value stays far from the 180 degree value necessitated by the “in-line” mechanism. This may be an artifact of the substitution of the 2'-OH of the cytosine residue with a 2'O-methyl group to prevent cleavage during crystallization. Although the methyl group was invisible during resolution, it may be that a “bad” angle was adopted upon substituting the hydroxyl group for the final structure. The Mg<sup>2+</sup> ions may be in different places in the simulation than would be required for the transition state for the proposed “in-line” mechanism. Finally, it may be that the simulation needs to be much larger to observe this intermediate.

#### RMSD Evaluation

As observed in one article, “[t]he most direct way to assess the stability of a [molecular dynamics] simulation over the course of time is the evaluation of the difference between the initial experimental coordinates and the generated structures, measured by the [root-mean-square difference] RMSD.”<sup>4</sup> RMSD analysis provides a meaningful method of assessing molecular conformational changes which arise during the course of a simulation. In the context of biological macromolecules, two major structural domains are generally

examined: the main chain or backbone and the side chains. With respect to nucleic acids, the main chain comprises the sugar-phosphate backbone and the side chains comprise the bases.

Two sets of calculations were performed using CHARMM; examples of the Fortran programs appear in Appendix A. The first set of calculations used the final structure of the ribozyme complex at 600 ps (Newtonian 500K simulation) as a reference structure with respect to which RMSD calculations for the 550 ps dynamics trajectory were produced. The trajectory comprised a concatenation of every 100th structure generated during the dynamics. These calculations permit an analysis of the structural changes which took place in the molecular system as it evolved to the final structure.

The second set of calculations used the average structure of the ribozyme complex during the 550 ps dynamics period as a reference structure with respect to which RMSD calculations for the 550 ps dynamics trajectory were produced. These calculations permit an assessment of the overall flexibility of the molecule during dynamics. Both calculations focused on the two major structural elements of biological macromolecules: the backbone (phosphate atoms) and the side-chains (using a single average value for the nucleoside or base atoms).

Figure 13 presents the minimized and average ribozyme structures and the structures of the ribozyme at various points during the dynamics, including the final

structure at 600 ps. The base-pairing representations are particularly useful in representing the degree of hydrogen bonding and overall structural disorder which occurred during the simulations.

## RMSD Calculations with respect to the Final Structure

### Phosphate Backbone Changes

The phosphate backbone is examined in Figure 18 which presents three-dimensional plots of phosphate atom versus RMSD with respect to the final structure versus time in the form of rotated representations to more clearly visualize the dynamical changes. Additionally presented are two-dimensional plots of the average RMSD of the phosphate atoms over time versus phosphate atom position which collectively show general trends in the conformational changes of the ribozyme backbone.

The relatively higher magnitude changes in backbone atomic positions are summarized in Table 2. The average RMSD values were 0.727 for Phase I and 0.547 for Phase II. Larger motions are associated with the terminal or near-terminal, in the case of the Guanine 18 phosphate, atoms, especially during the Phase I dynamics and with respect to the enzyme strand. As time progresses, these movements are reduced and are replaced by larger movements of the Stem 3 phosphate atoms, including some of the conserved core phosphate atoms.



Table 2. Backbone (Phosphate Atom) Evolution towards Final Structure

Residue	Chain	Atom	Phase I RMSD	Phase II RMSD	Location
G13	A	P	1.04	1.46	Core
G14	A	P	0.91	1.28	Near-Terminus
C16	A	P	1.37	0.69	Terminus
G18	B	P	2.65	1.30	Near-Terminus
A24	B	P	1.03	1.32	Core
G32	B	P	1.47	0.58	Stem 3
A33	B	P	1.64	0.88	Stem 3

#### Base (Side-chain) Evolution

Figure 18 presents the RMSD of residue atoms in two general formats. First, the RMSD of the residues are presented as a function of residue position and time. Second, the average RMSD of all the atoms comprising an individual residue is plotted versus residue position. One particular residue, the guanine 17 of the substrate strand, shows the most movement in comparison to the other residues during Phase I dynamics; this behavior, however, is smoothed out in the Phase II dynamics suggesting perhaps that the initial larger movements may be due to full equilibration not having occurred as rapidly as originally thought. Not surprisingly, the generated RMSD values for the residues are greater than those of the phosphate backbone atoms. The average RMSD values were 0.940 for Phase I and 0.795 for Phase II.

In general, the guanine and cytosine residues showed larger deviations than the adenosine and cytosine residues. Of the 10 base pairs in the original structure, 7 were guanine-cytosine base pairs. These base pair residues are located in general at the terminii

of the molecule and may be more largely affected due to these positions. Further analysis of the breakdown in hydrogen bonding structure is necessary.

The following table summarizes the higher observed values for RMSD of base atoms. An average value for the atoms comprising each base was used for purposes of calculation and presentation. The movement of the bases seem to follow the same pattern as the backbone phosphate atoms, i.e. greater motions at the termini in Phase I and greater motions within the core region in Phase II.

Table 3. Side-Chain (Nucleotide Base) Evolution towards Final Structure

<b>Residue</b>	<b>Chain</b>	<b>Phase I RMSD</b>	<b>Phase II RMSD</b>	<b>Location</b>
C6	A	2.06	1.87	Near-Terminus
C16	A	2.69	2.38	Terminus
G17	B	3.16	2.45	Terminus
G18	B	1.73	0.85	Near-Terminus
A24	B	1.09	1.41	Core
C25	B	0.99	1.30	Core
U26	B	1.02	1.35	Core
C41	B	2.74	0.70	Terminus

#### RMSD Calculations with respect to the Average Structure

##### Phosphate Backbone Flexibility

The phosphate backbone is examined in Figure 19 which presents three-dimensional plots of phosphate atom versus RMSD with respect to the average structure versus time, including rotated representations to more clearly visualize the dynamical changes. Additionally presented are two-dimensional plots of the average RMSD of the

phosphate atoms over time versus phosphate atoms position which show general trends in the evolution of the ribozyme backbone towards the final structure. The values associated with the most predominant peaks in the graphs are summarized in Table 4. The average RMSD values were 0.393 for Phase I and 0.405 for Phase II. During the designated Phase I dynamics, greater movement or flexibility was associated with the phosphate atoms of the substrate strand (Chain B) as compared with the phosphate atoms of the enzyme strand (Chain A). The near-terminus atoms in both strands showed the most flexibility in Phase I. This behavior decreased with respect to these atoms in Phase II and was replaced, generally, with higher motions associated with the core structure and Stem 3 atoms. The phosphate atom associated with one of the nucleotides of the scissile bond showed significant movement during Phase I and Phase II dynamics as well.

Table 4. Backbone (Phosphate Atom) Flexibility with respect to Average Structure

<b>Residue</b>	<b>Chain</b>	<b>Atom</b>	<b>Phase I RMSD</b>	<b>Phase II RMSD</b>	<b>Location</b>
G14	A	P	0.74	0.69	Near-Terminus
G18	B	P	1.28	1.07	Near-Terminus
C20	B	P	0.48	0.36	Near-Terminus
G21	B	P	0.42	0.56	Near-Terminus
A24	B	P	0.41	0.46	Core
C25	B	P	0.44	0.56	Core
G32	B	P	0.56	0.53	Stem 3
A33	B	P	0.42	0.58	Stem 3
A37	B	P	0.48	0.46	Scissile Bond

#### Base (Side-chain) Flexibility

Figure 19 shows changes in flexibility of the ribozyme structure with respect to the average structure over the 550 ps production dynamics. The results are summarized for the

predominant peaks in Table 5. The average RMSD values were 0.381 for Phase I and 0.560 for Phase II. The bases showed significant flexibility as compared to the average structure during Phase I and Phase II. The greater motions are localized to residues located in the near-terminus, terminus, core, and Stem 3 areas. The values for these motions increased through Phase I and Phase II. This observation is in contrast to the flexibility of the phosphate atoms which generally evidence a decrease in flexibility as time progresses.

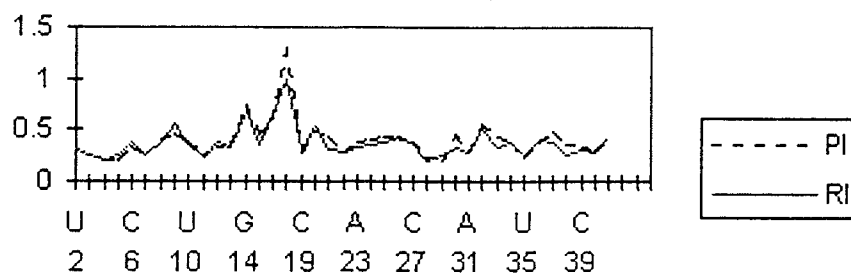
A strong correlation exists with respect to the average RMSD values for the backbone and base atoms in both Phases I and II (Figure 7). This suggests that the molecule is moving through conformational space as a concerted unit.

Table 5. Side-Chain (Base) Flexibility with respect to Average Structure

<b>Residue</b>	<b>Chain</b>	<b>Phase I RMSD</b>	<b>Phase II RMSD</b>	<b>Location</b>
C6	A	0.39	0.67	Near-Terminus
A9	A	0.55	0.80	
G14	A	0.68	0.80	Near-Terminus
C16	A	0.65	0.82	Terminus
G17	B	0.91	1.16	Terminus
G18	B	0.99	1.27	Near-Terminus
C20	B	0.52	0.70	Near-Terminus
A24	B	0.37	0.70	Core
C25	B	0.39	0.66	Core
G32	B	0.52	0.74	Stem 3
A33	B	0.33	0.61	Stem 3
C41	B	0.43	0.54	Terminus

(a)

### Backbone Side-Chain RMSD Correlation (Phase I)



(b)

### Backbone Side-Chain RMSD Correlation (Phase II)

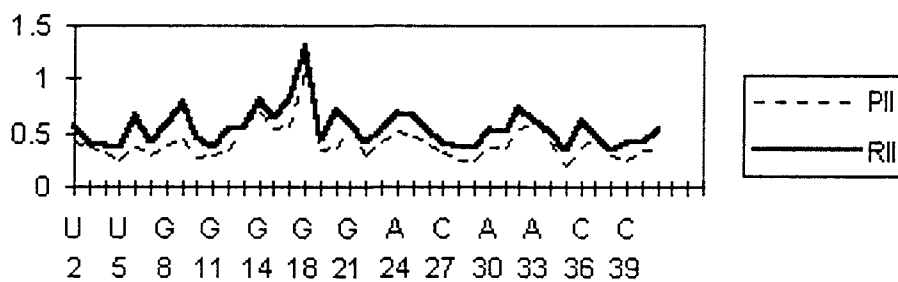
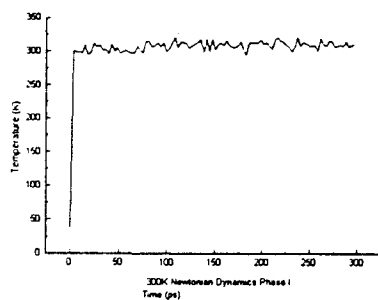


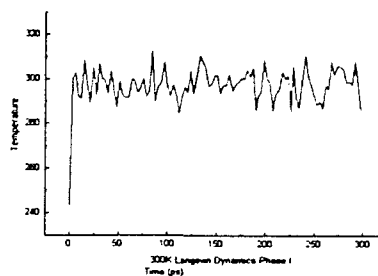
Figure 7. Backbone Side-Chain Correlation

- Figure 8. Temperature versus Time
- (a) 300K Newtonian Dynamics Phase I
  - (b) 300K Langevin Dynamics Phase I
  - (c) 500K Newtonian Dynamics Phase I
  - (d) 500K Langevin Dynamics Phase I
  - (e) 300K Newtonian Dynamics Phase II
  - (f) 300K Langevin Dynamics Phase II
  - (g) 500K Newtonian Dynamics Phase II
  - (h) 500K Langevin Dynamics Phase II

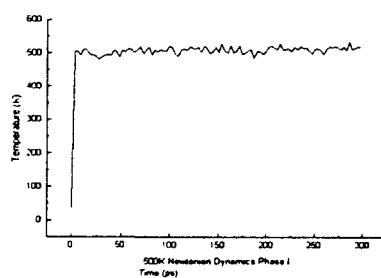
a)



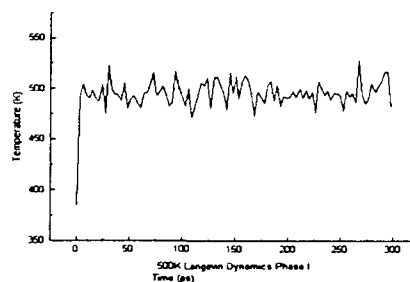
b)



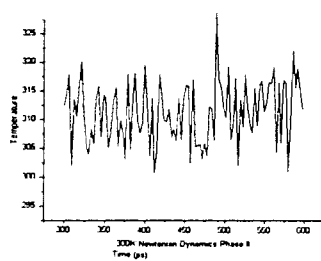
c)



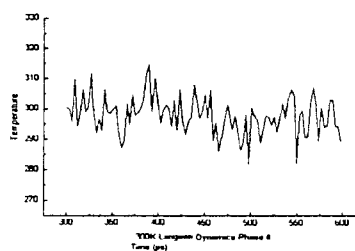
d)



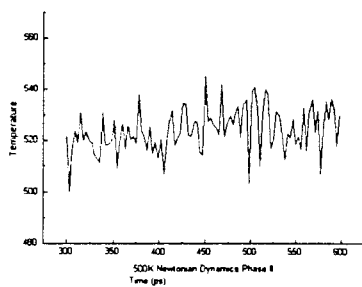
e)



f)



g)



h)

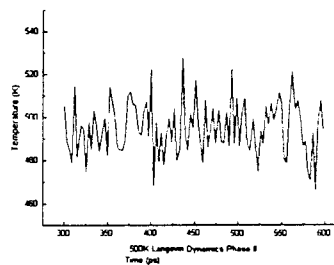
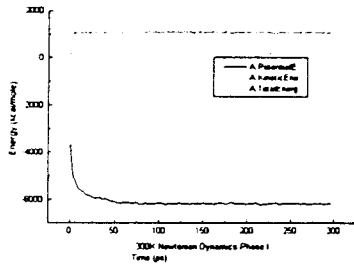


Figure 9. Potential Energy, Kinetic Energy, and Total Energy versus Time.

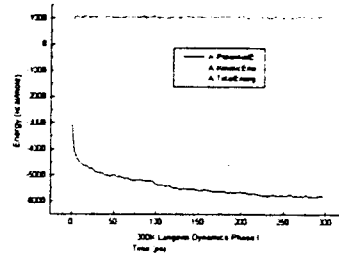
- (a) 300K Newtonian Dynamics Phase I
- (b) 300K Langevin Dynamics Phase I
- (c) 500K Newtonian Dynamics Phase I
- (d) 500K Langevin Dynamics Phase I
- (e) 300K Newtonian Dynamics Phase II
- (f) 300K Langevin Dynamics Phase II
- (g) 500K Newtonian Dynamics Phase II
- (h) 500K Langevin Dynamics Phase II



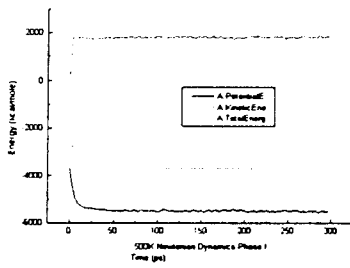
(a)



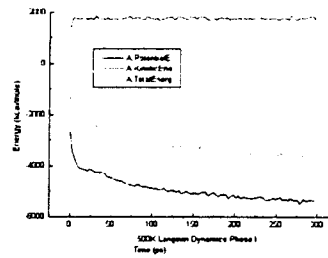
(b)



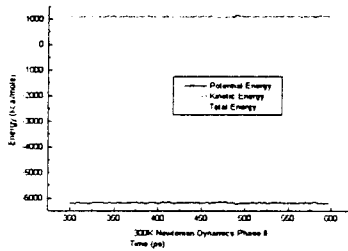
(c)



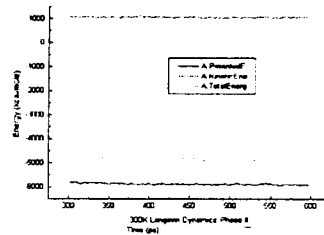
(d)



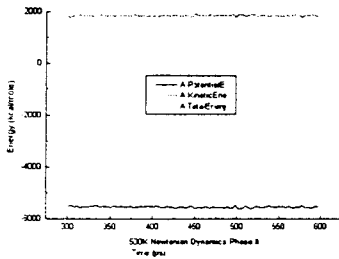
(e)



(f)



(g)



(h)

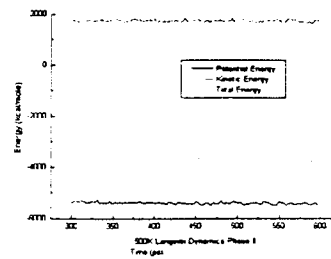
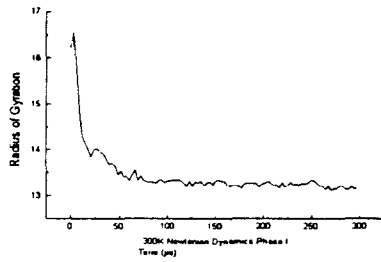


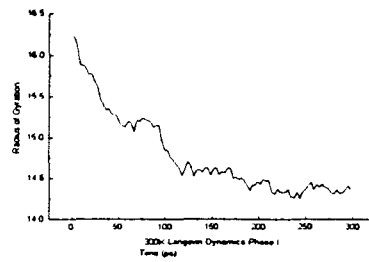
Figure 10. Radius of Gyration versus Time

- (a) 300K Newtonian Dynamics Phase I
- (b) 300K Langevin Dynamics Phase I
- (c) 500K Newtonian Dynamics Phase I
- (d) 500K Langevin Dynamics Phase I
- (e) 300K Newtonian Dynamics Phase II
- (f) 300K Langevin Dynamics Phase II
- (g) 500K Newtonian Dynamics Phase II
- (h) 500K Langevin Dynamics Phase II

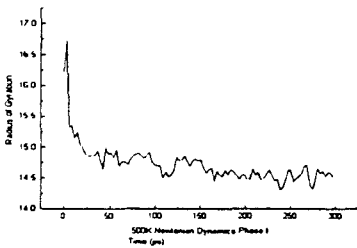
(a)



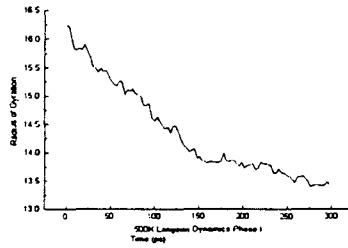
(b)



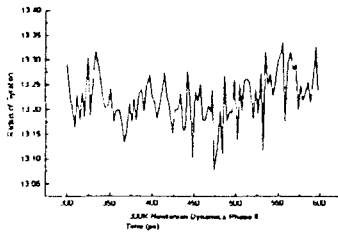
(c)



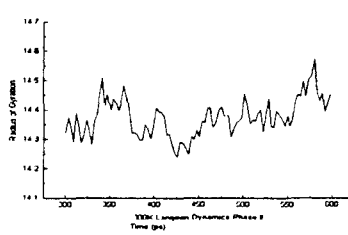
(d)



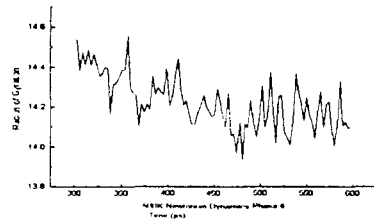
(e)



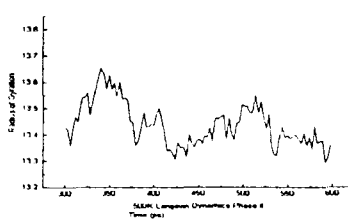
(f)



(g)

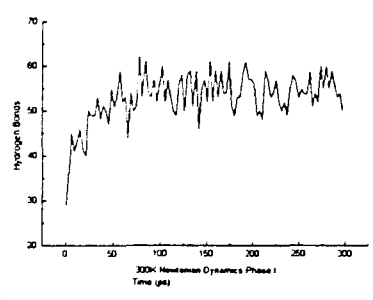


(h)

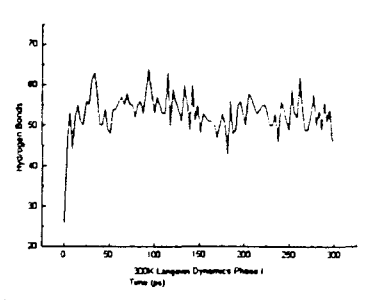


- Figure 11. Hydrogen Bonds versus Time
- (a) 300K Newtonian Dynamics Phase I
  - (b) 300K Langevin Dynamics Phase I
  - (c) 500K Newtonian Dynamics Phase I
  - (d) 500K Langevin Dynamics Phase I
  - (e) 300K Newtonian Dynamics Phase II
  - (f) 300K Langevin Dynamics Phase II
  - (g) 500K Newtonian Dynamics Phase II
  - (h) 500K Langevin Dynamics Phase II

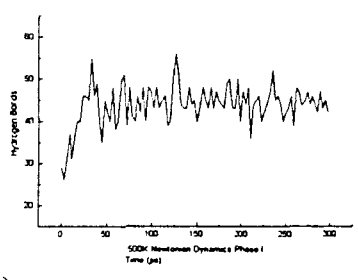
(a)



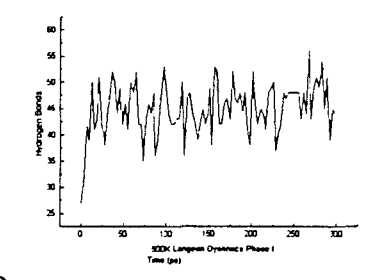
(b)



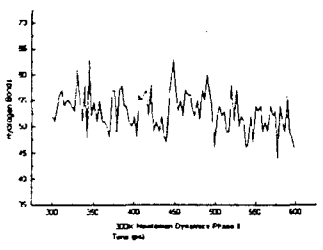
(c)



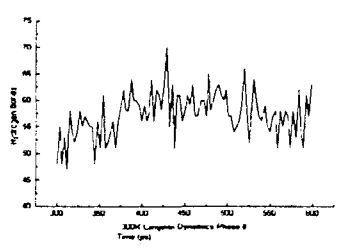
(d)



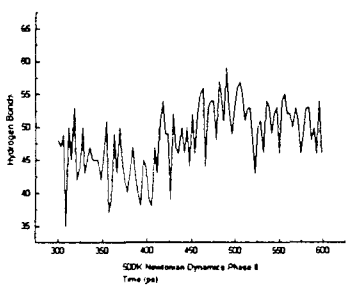
(e)



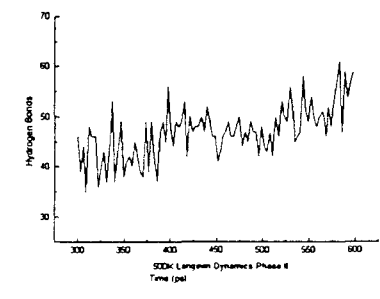
(f)



(g)

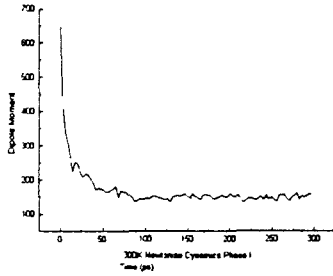


(h)

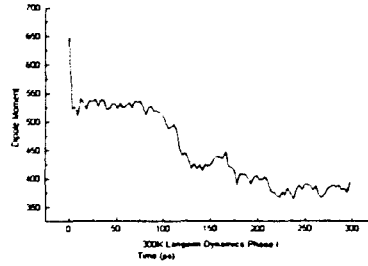


- Figure 12. Dipole Moment versus Time
- (a) 300K Newtonian Dynamics Phase I
  - (b) 300K Langevin Dynamics Phase I
  - (c) 500K Newtonian Dynamics Phase I
  - (d) 500K Langevin Dynamics Phase I
  - (e) 300K Newtonian Dynamics Phase II
  - (f) 300K Langevin Dynamics Phase II
  - (g) 500K Newtonian Dynamics Phase II
  - (h) 500K Langevin Dynamics Phase II

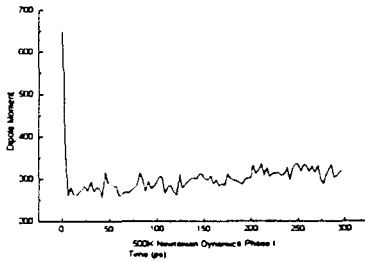
(a)



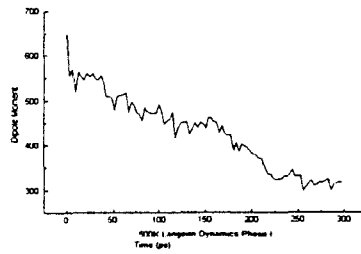
(b)



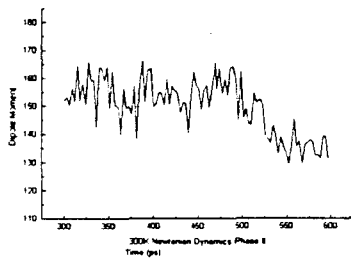
(c)



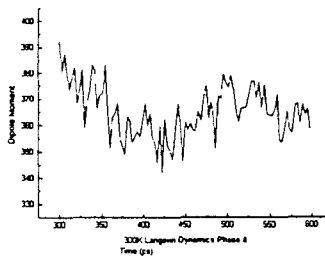
(d)



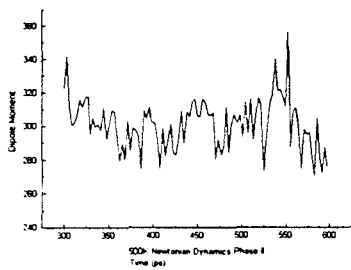
(e)



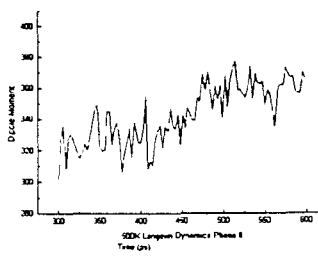
(f)



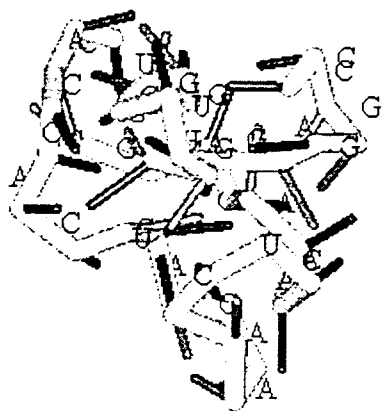
(g)



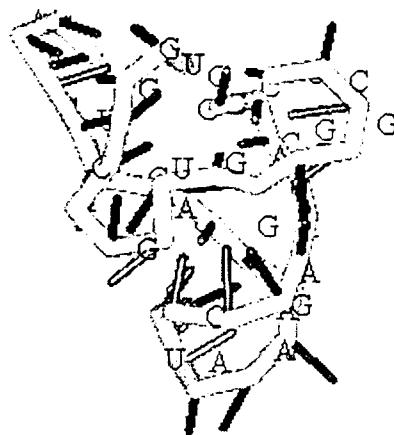
(h)



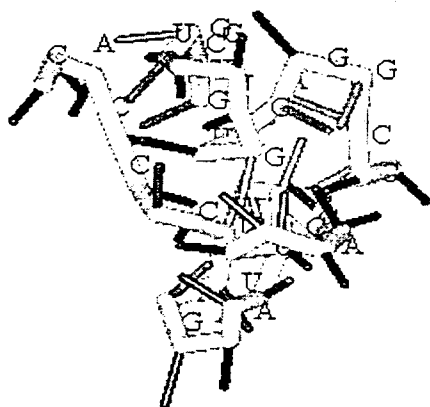
(a)



(b)



(c)



(d)

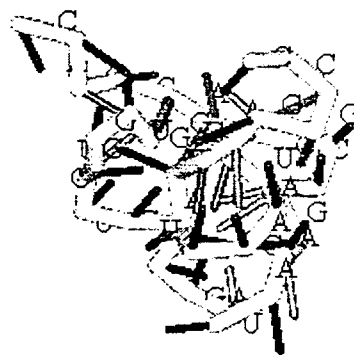


Figure 13. Final Structures at 600 ps

(a) 300K Newtonian

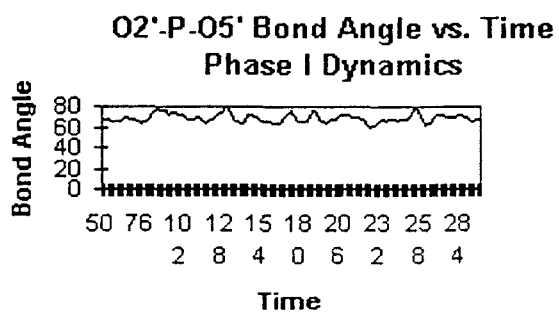
(b) 300K Langevin

(c) 500K Newtonian

(d) 500K Langevin



(a)



(b)

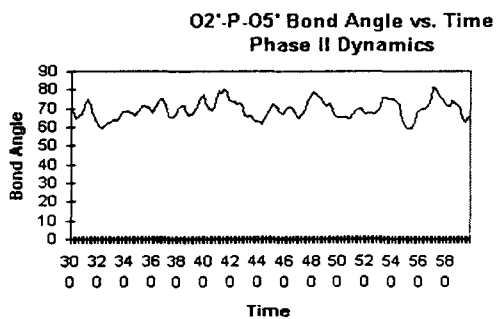


Figure 14. O2'-P-O5' Bond Angle versus Time:  
(a) Phase I Dynamics (b) Phase II Dynamics.

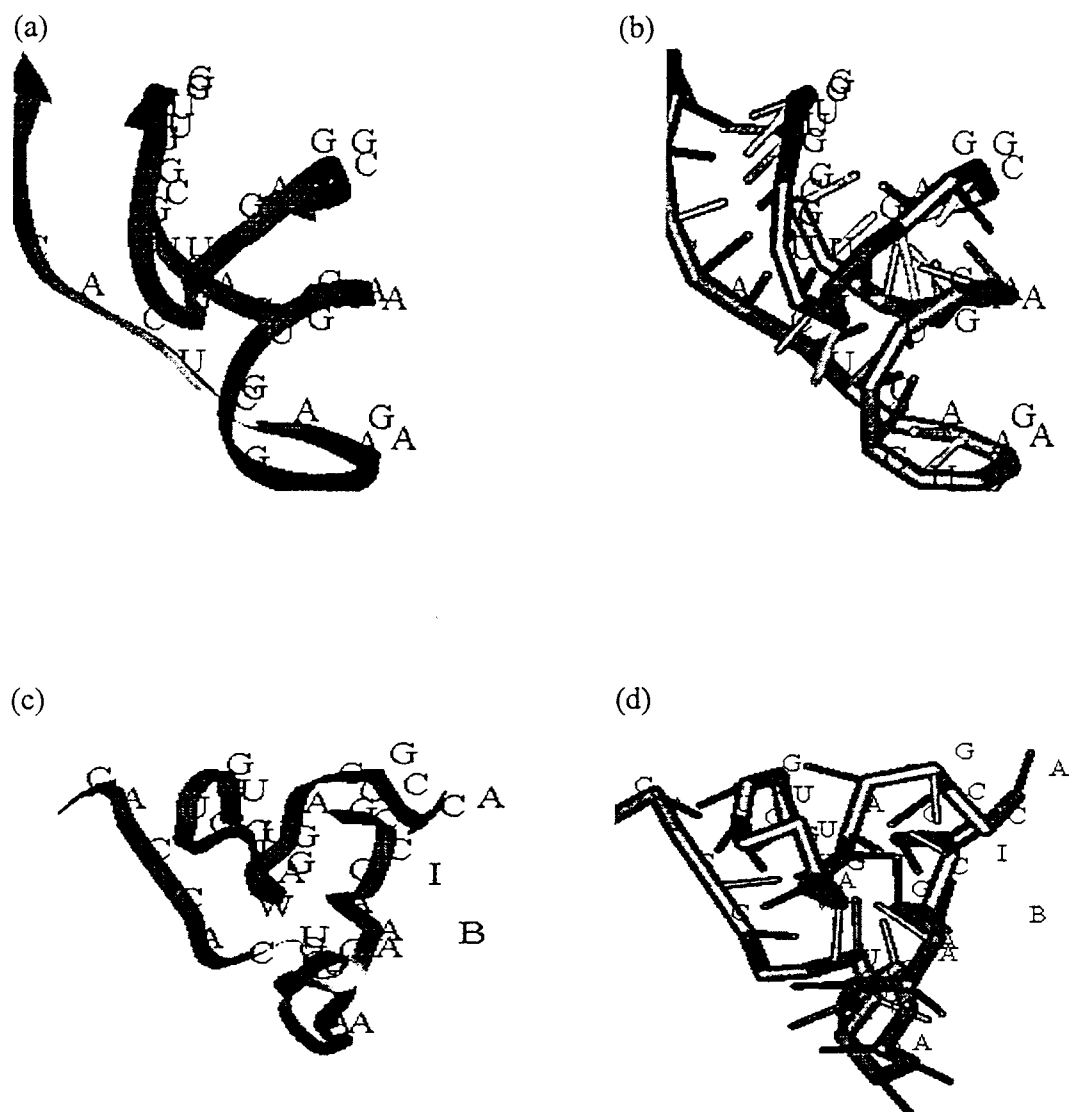


Figure 15. Minimized ribozyme structure: (a) ribbon and (b) ladder views. Average Ribozyme structure over 550 ps dynamics: (c) ribbon and (d) ladder views.

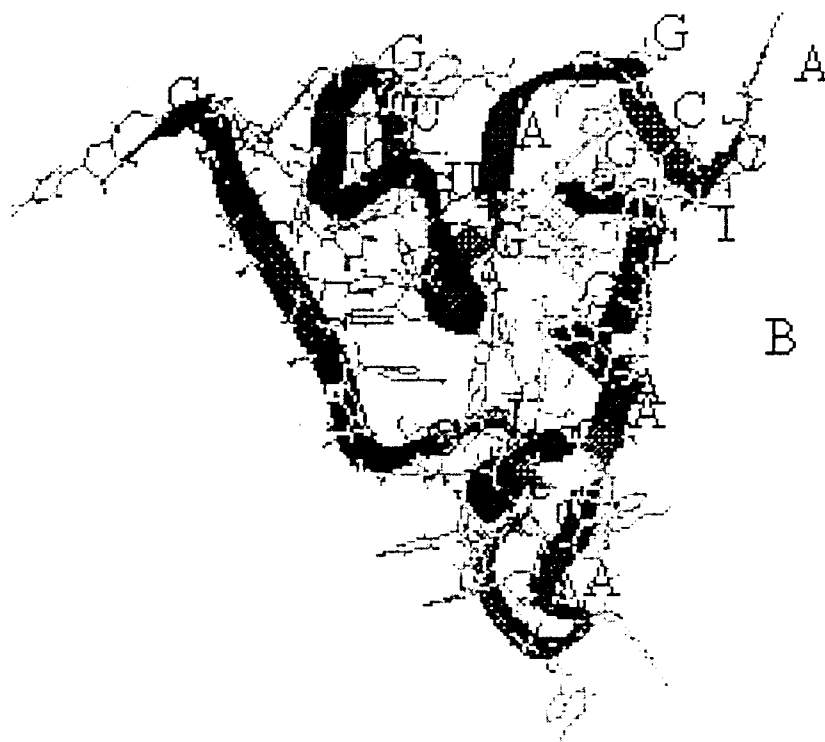
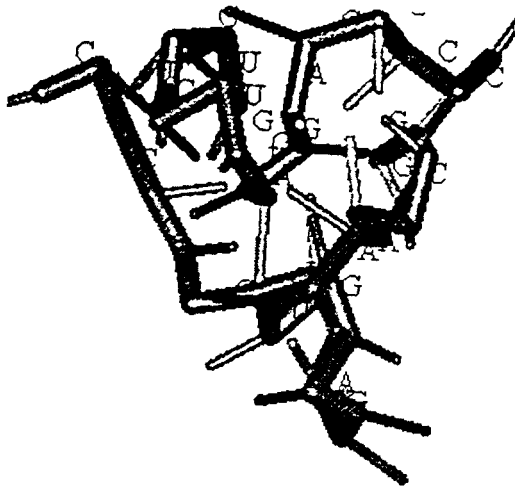
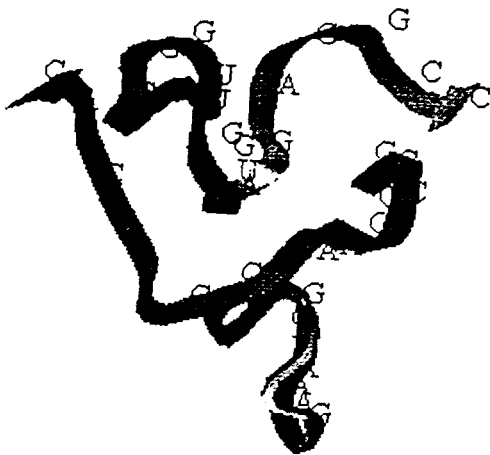


Figure 16. Average Structure of Ribozyme showing  $Mg^{2+}$  ions

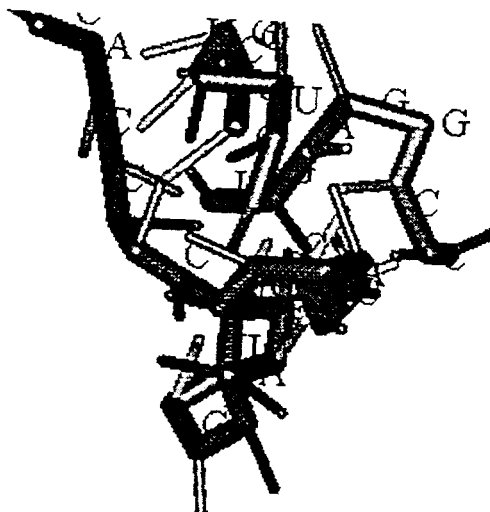
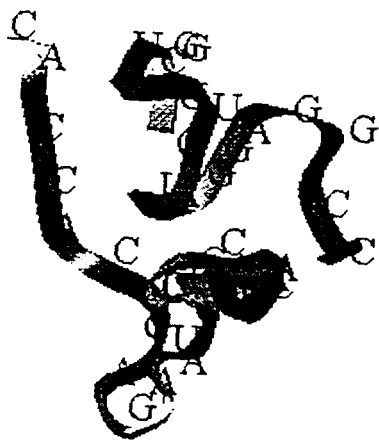
Figure 17. Ribbon and ladder views of ribozyme

- (a) Ribozyme at 50 ps
- (b) Ribozyme at 100 ps
- (c) Ribozyme at 150 ps
- (d) Ribozyme at 250 ps
- (e) Ribozyme at 350 ps
- (f) Ribozyme at 450 ps
- (g) Ribozyme at 550 ps
- (h) Ribozyme at 600 ps

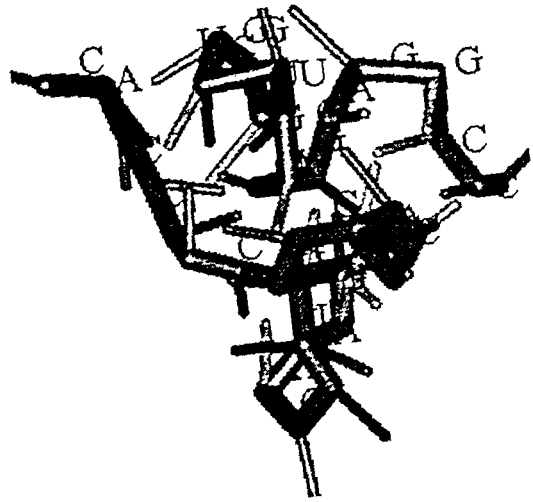
(a)



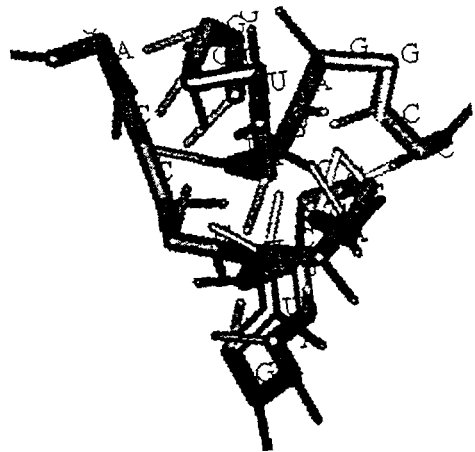
(b)



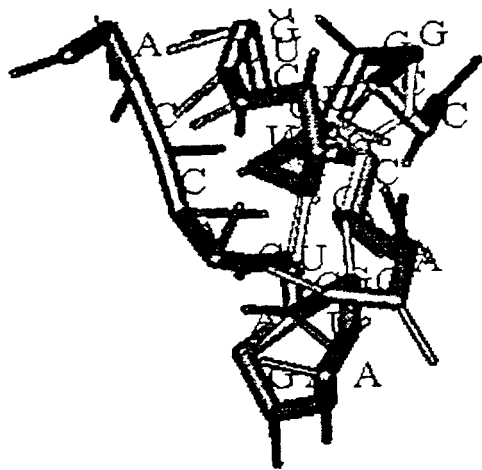
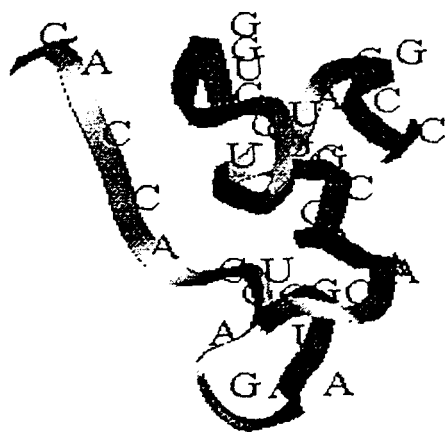
(c)



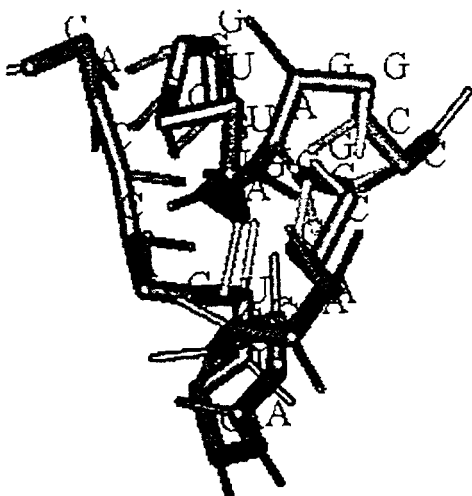
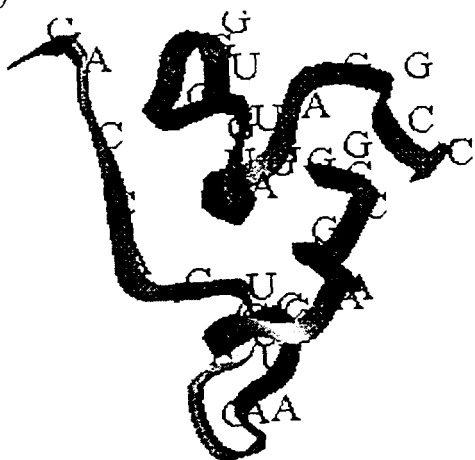
(d)



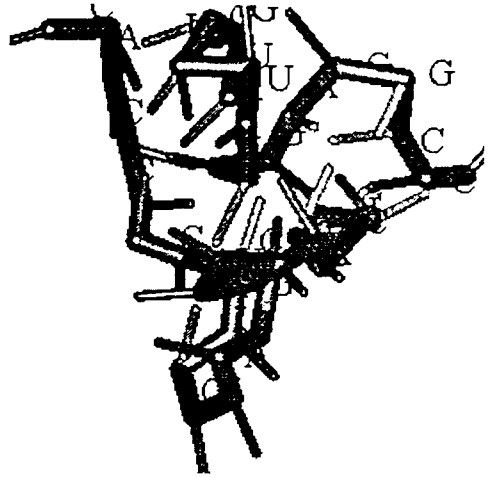
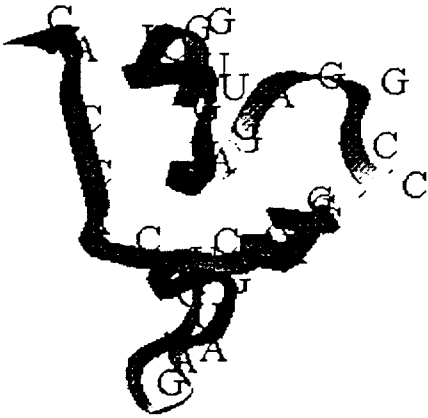
(e)



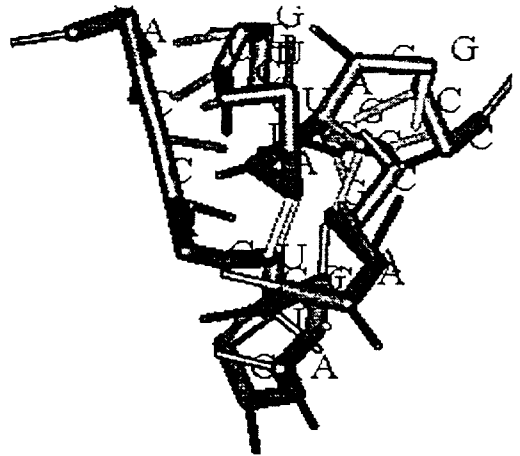
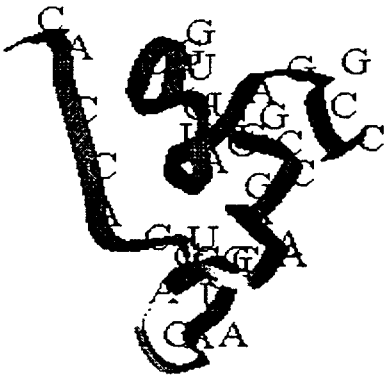
(f)



(g)



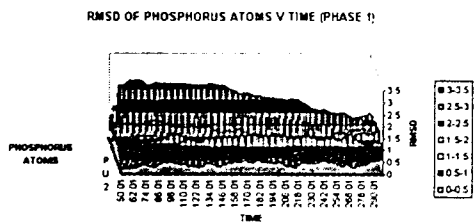
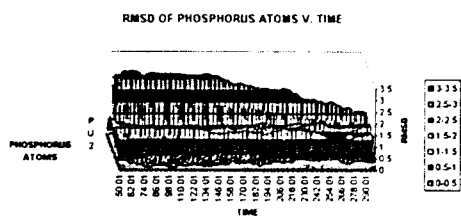
(h)



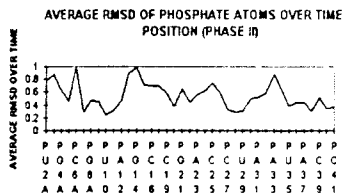
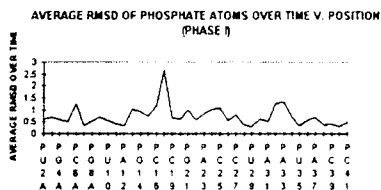


- Figure 18. RMSD with respect to Final Structure
- (a) Phosphate Atoms (Rotated View)
  - (b) Average RMSD of Phosphate Atoms Over Time versus Position
  - (c) RMSD of Phosphate Atoms versus Time (Rotated View).
  - (d) RMSD of Residues versus Time (Rotated View)
  - (e) Average RMSD of Residues over Time
  - (f) Average RMSD of Cytosine Residues Over Time versus Position
  - (g) Average RMSD of Guanine Residues Over Time versus Position
  - (h) Average RMSD of Adenine Residues Over Time versus Position
  - (i) Average RMSD of Uracil Residues Over Time versus Position

(a)

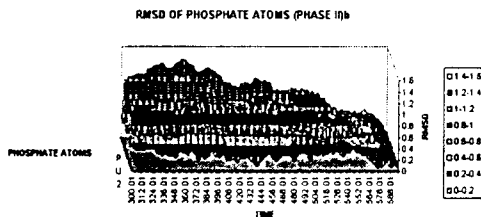
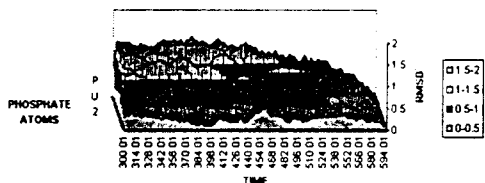


(b)



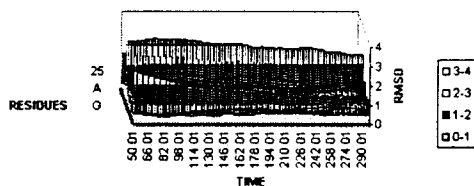
(c)

**RMSD OF PHOSPHATE ATOMS V. TIME (PHASE II)a**

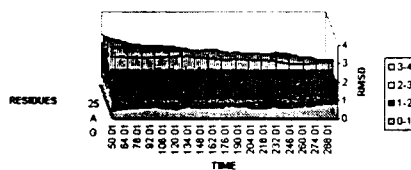


(d)

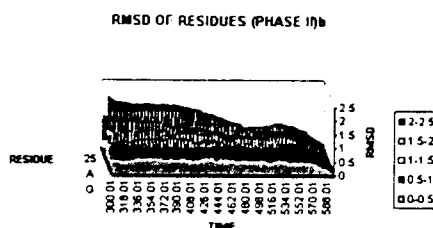
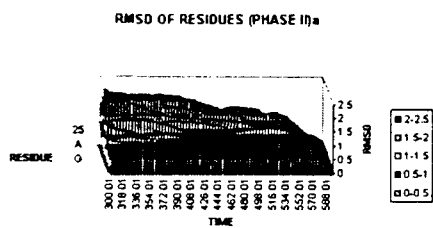
**RMSD of RESIDUES (Phase I)**



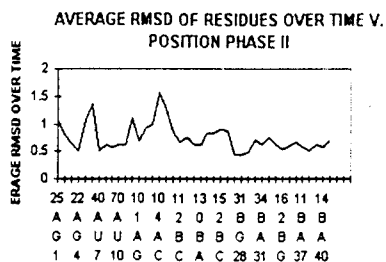
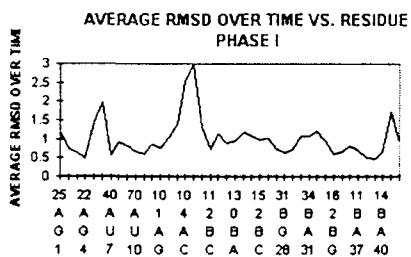
**RMSD OF RESIDUES (PHASE I)**



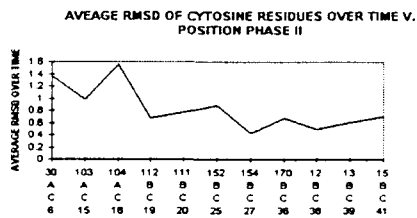
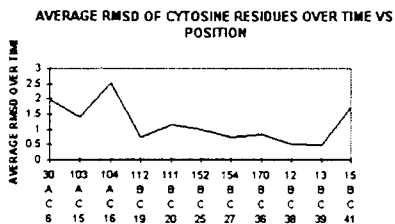
(e)



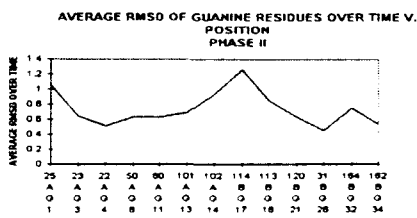
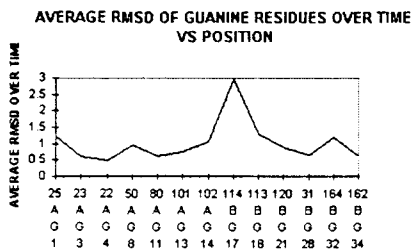
(f)



(g)

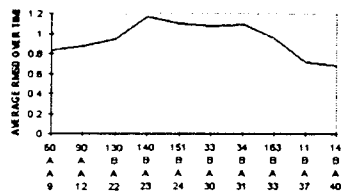


(h)

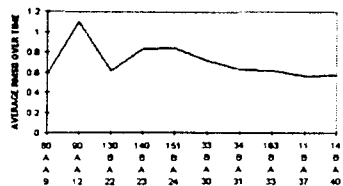


(i)

AVERAGE RMSD OF ADENINE RESIDUES OVER TIME VS. POSITION

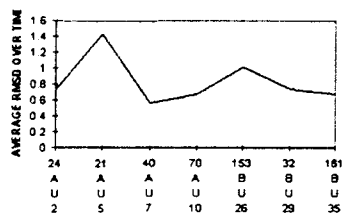


AVERAGE RMSD OF ADENINE RESIDUES OVER TIME V POSITION PHASE II



(j)

AVERAGE RMSD OF URACIL RESIDUES OVER TIME VS POSITION



AVERAGE RMSD OF URACIL RESIDUES OVER TIME V. POSITION PHASE II

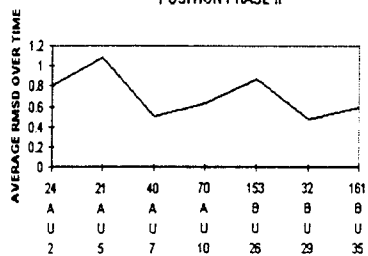


Figure 19. RMSD with respect to Average Structure  
(a) RMSD of Phosphate Atoms versus Position versus Time  
(b) Average RMSD of Phosphate Atoms versus Time  
(c) RMSD of Residues versus Position versus Time  
(d) Average RMSD of Residues versus Time.



- 
- <sup>1</sup>See generally van Gunsteren, W.F. and Mark, A.E. (1991). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 Eur. J. Biochem. 947-961.
- <sup>2</sup>Gabler, R. (1978). Electrical Interactions in Molecular Biophysics. 76-77.
- <sup>3</sup>Scott, W. and Klug, A. (1996). Ribozymes: structure and mechanism in RNA catalysis. 222.
- <sup>4</sup>Fox, T. and Kollman, P. (1996). The Application of Different Solvation and Electrostatic Models in Molecular Dynamics Simulations of Ubiquitin: How Well Is the X-Ray Structure “Maintained”? 25 PROTEINS: Structure, Function, and Genetics 318.

## CHAPTER 10

### CONCLUSION

The preliminary examination of the molecular dynamics of the hammerhead ribozyme under differing dynamic procedures and at different temperatures showed no large-scale or global conformational changes from the minimized structure throughout the 600 ps simulations or the presence of significant intermediates. The thermodynamic and conformational properties behaved as expected. Rapid equilibration was followed by a steady dynamics trajectory in each case, but the Newtonian simulations achieved equilibrium before the Langevin ones. The 500K Newtonian dynamics simulation, chosen for further RMSD analysis with respect to the final and average structures, showed localized movement in the form of slight variations in flexibility of the backbone (phosphate atoms) and the side-chains (bases).

As expected, the simulations were generally characterized by increased movement among the terminal phosphate atoms of the enzyme and substrate strands during the first phase; this behavior settles down and is replaced by increasing side-chain, conserved core, and Stem 3 movement during Phase II. The enzymatic strand showed more movement during the Phase I; this may be due to the shorter length, and decreased mass, of the enzyme strand as compared with the substrate strand. The phosphate atom associated with one of the nucleotides of the scissile bond showed significant movement during Phase I and Phase II dynamics as well.



The bases showed significant flexibility as compared to the average structure and final structure during Phase I and Phase II. The greater motions are localized to residues located in the near-terminus, terminus, core, and Stem 3 areas. The values for these motions increased through Phase I and Phase II. This observation is in contrast to the flexibility of the phosphate atoms where there is generally a decrease in flexibility as time progresses.

Additional analysis may reveal more meaningful patterns in the data including the possibility that a molten globule-like state was encountered. However, more relevant events may only be observable through a longer time frame dynamics simulation.

## APPENDIX A

CHARMm (Fortran) program for RMSD calculations

\* RMSD - rmsd calculations for hammerhead ribozyme simulations

\*

UPPER ! case for files to write

bomblevel -5

wrnlev 0

prnlev 5

! Script to read parameter, psf, and ic files

! Generate the system using the information from Quanta

open read unit 21 card name \$CHM\_DATA/MASSES.RTF

read rtf unit 21 card close unit 21

open read unit 22 card name ".charmmprm"

read param unit 22 card! close unit 22

open read unit 23 card name ".charmmpsf"

read psf unit 23 card

open read unit 24 card name ".charmmic"

ic read unit 24 card

goto (color)

label yellow

!-----

! Concatenation of Trajectory Files

! CHARMm versions differ in treatment of file headers

! Merge multiple trajectory files using traj and iread

open read unit 61 file name A.DCD

open read unit 62 file name B.DCD

open read unit 63 file name C.DCD

open read unit 64 file name D.DCD

open read unit 65 file name E.DCD

open read unit 66 file name F.DCD

open read unit 67 file name G.DCD

open write unit 68 file name COMPLETE.DCD

trajectory iread 61 nread 7 iwrite 68

set a 1

```
label start
traj read
traj write
incr a by 100
if a lt 27500 goto start
close unit 23
close unit 24
close unit 68
close unit 61
close unit 62
close unit 63
close unit 64
close unit 65
close unit 66
close unit 67
```

```
label red
```

```
! -----
! Calculate the Average Structure over the trajectory COMPLETE.DCD
```

```
! 1. Cartesian Coordinates
```

```
open read unit 32 file name COMPLETE.DCD
coor dyna pax firs 32 nuni 1 skip 10 begin 50010 stop 52750
close unit 32
```

```
! 2. View PAX analysis results
```

```
open read unit 32 file name COMPLETE.DCD
coor paxa firs 32 nuni 1 skip 10 begin 50010 stop 52750
close unit 32
```

```
! 3. List the rms fluctuations and write average coordinates
```

```
scal wmai show
print coor
open write unit 33 card name AVERAGE.CRD
write coor card unit 33
* 550 ps dynamics
* average coordinates for COMPLETE.DCD
*
```

```
close unit 32
close unit 33
```

```
! 4. Compute IC values to get dihedral angle references
```

```
ic fill
```

```
open read unit 24 card name ".charmmic"
read ic card unit 24
```

```
! 5. Compute and print IC averages
open read unit 32 file name COMPLETE.DCD
ic dyna aver firs 32 nuni 1 skip 10 begin 50010 stop 52750
print ic
close unit 32
```

```
! 6. Compute and print IC fluctuations about the average
open read unit 32 file name COMPLETE.DCD
ic dyna fluc firs 32 nuni 1 skip 10 begin 50010 stop 52750
print ic
close unit 32
close unit 24
```

```
label green
```

```
! -----
! RMSD calculations per atom - one coordinate structure set at a time
! open file for writing rmsd data sets with only one atom per residue
open write unit 57 card name RMSDP.CRD
set 2 275.5
! Fill comparison coordinate set with average coordinates
open read unit 44 card name AVERAGE.CRD
read coor card unit 44
upda cutnb 15.0 ctonnb 11.0 ctofnb 14.0 wmin 0.8
ener rdie
coor copy comp
open unit 43 read file name COMPLETE.DCD
trajectory iread 43
set 9 1
label loop
traj read
coor orient rms
coor rms sele all end
coor diff
coor dist weigh
scalar wmain average by res sele all end
print coor rms sele type P card unit 57 end
print coor rms card unit 57 sele type P end
incr 9 by 1
if 9 lt @2 goto loop
```

```
! The orient command is for coordinate orientation such that the principle
! geometric axis coincides with the x-axis, and the next largest
! coincides with the y axis. The structure is oriented about its
! center of geometry.
! RMS keyword overlaps the main coordinates
! with those in the comparison set
close unit 43
close unit 44
```

```
label purple
```

```
! -----
! RMSD calculations per residue- one coordinate structure set at a time
set 2 275.5
! Fill comparison coordinate set with average coordinates
open write unit 63 card name RESIDUE.CRD
set 2 275.5
open read unit 44 card name AVERAGE.CRD
read coor card unit 44
upda cutnb 15.0 ctonnb 11.0 ctofnb 14.0 wmin 0.8
ener rdie
coor copy comp
open unit 43 read file name COMPLETE.DCD
trajectory iread 43
set 9 1
label loop
traj read
coor orient rms
coor rms sele all end
coor diff
coor dist weigh
scalar wmain average by res sele all end
print coor rms sele type O5' end
incr 9 by 1
if 9 lt @2 goto loop
close unit 43
close unit 44
stop
```

CHARMm (Fortran) programs for data manipulation

```

C   Program table
    real R(300,1500)
    integer I,J
    open (unit=1, file ='x1.out')
    open (unit=2, file='xa.out')
    Do 100 I=1,275
    Do 100, J=1,39
100  read (1,10,end=300) R(I,J)
    10  Format (63X,F7.5)
    10  Format (63X,F7.5)
        Do 200 J=1,39
200  Write (2,20) (R(I,J),I=1,125)
    20  Format (125(1X,F7.5))
        close (unit=1)
        close (unit=2)
    end

```

```

C   Program table
    real R(300,1500)
    integer I,J
    open (unit=1, file ='x1.out')
    open (unit=2, file='xb.out')
    Do 100 I=1,275
    Do 100, J=1,39
100  read (1,10,end=300) R(I,J)
    10  Format (63X,F7.5)
300  Continue
        Do 200 J=1,39
200  Write (2,20) (R(I,J),I=126,275)
    20  Format (150(1X,F7.5))
        close (unit=1)
        close (unit=2)
    end

```

Sample Data Lines from Coordinate File

```

862  28 G   O2P  20.38689 -9.18638 35.77438 B  31  0.27038
863  28 G   O5'  20.49061 -10.86409 37.65214 B  31  0.29958
864  28 G   C5'  21.65806 -11.12931 38.39593 B  31  0.34141
865  28 G   C4'  21.39165 -11.89022 39.69059 B  31  0.31036

```

## BIBLIOGRAPHY

- Abrahams, J., van den Berg, M., van Batenburg, E., and Pleij, C. (1990). Prediction of RNA secondary structure, including pseudoknotting, by computer simulation. *18 Nucleic Acids Research* 10: 3035-3044.
- American Chemical Society. (1994). *Molecular and Biomolecular Electronics. Advances in Chemistry/Series 240*. ISBN: 0-8412-2698-9.
- Andres, P.R., D.J. Craik, and J.L. Martin. (1984). Functional Group Contributions to Drug-Receptor Interactions. *Journal of Medicinal Chemistry*, 27, 1648-1657.
- Atkins, P.W. (1994). *The Second Law: Energy, Chaos, and Form*. (New York: W.H. Freeman).
- Axel and Brunger. (1992). *X-Plor Version 3.1. A System for X-ray crystallography and NMR*. New Haven: Yale University Press.
- Baker, D., Sohl, J.L. and Agard, D. (1992). A protein folding reaction under kinetic control. 356 *Nature* 263-265.
- Balaji, P.V., Qasba, P.K. and Rao, V.R. (1993). Molecular Dynamics Simulations of Asialoglycoprotein Receptor Ligands. *Biochemistry*, 32, 12599.
- Banks, J., Brower, R., Ma, J. (1995). Effective Water Model for Monte Carlo Simulations of Proteins. 35 *Biopolymers* 331-341.
- Bartel, D. and Szostak, J. (1993). Isolation of New Ribozymes from a Large Pool of Random Sequences. 261 *Science* 1411-1418.
- Bassi, G.S., Mollegaard, N.-E., Murchie, A.I.H., von Kitzing, E. and Lilley, D.M.J. (1995). Ionic interactions and the global conformations of the hammerhead ribozyme. 2 *Nat. Struct. Biol.* 45-55.
- Baum, R. (1996). Chemists Create New Oligonucleotide Analogs for 'Antisense' Applications. *Chemical & Engineering News*. April 18, 1994: 21-22.
- Beaudry, A. and Joyce, G.F. (1992). Directed Evolution of an RNA Enzyme. 257 *Science* 635-641.
- Beaudry, D., Bussiere, F., Laureau, F., Lessard, C. and Perreault, J. (1995). The RNA of both polarities of the peach latent mosaic viroid self-cleaves in vitro solely by single hammerhead structures. 23 *Nucleic Acids Research* 5: 745-752.
- Beck, J. and Nassal, M. (1995) efficient hammerhead ribozyme-mediated cleavage of the structured hepatitis B virus encapsidation signal in vitro and in cell extracts, but not in intact cells. 23 *Nucleic Acids Research* 4954-4962.

- Beeman, D. (1976). Some Multistep Methods for Use in Molecular Dynamics Calculations. 20 *J. Comput. Phys.* 130.
- Ben-Naim, A., Ting, K.-L., Jernigan, R.O.L. (1989). I. Separation of the Volume and Surface Interactions with Estimates for Proteins. 28 *Biopolymers* 1309-1325.
- Beveridge, D.L. and DiCapua, F.M. (1989). Free Energy via Molecular Simulation: Applications to Chemical and Biomolecular Systems. 18 *Annu. Rev. Biophys. Biophys. Chem.* 431, 461.
- Beveridge, D. and Lavery, R., eds. (1990). *Theoretical Biochemistry & Molecular Biophysics, Volume 1: DNA.* (New York: Adenine Press).
- Biou, V., Gibrat, J.F., Levin, J.M., Robson, B., and Garnier, J. (1988). Secondary Structure Prediction: Combination of Three Different Methods. *Protein Engineering*, 2, 185-191.
- Blin, N., Camoin, L., Maignet, B. and Strossberg, A.D.. (1993). Structural and Conformational Features Determining Selective Signal Transduction in the Beta 3-Adrenergic Receptor. 44 *Molecular Pharmacology* 1094.
- Bohm, H., G. Klebe, T. Lorenz, T. Mietzner, and L. Siggel. (1990). Different Approaches to Conformational Analysis: A Comparison of Completeness, Efficiency, and Reliability Based on the Study of a Nine-Membered Lactam. 11 *Journal of Computational Chemistry* 9: 1021-1028.
- Bordo, D. and Argos, P. (1990). Evolution of Protein Cores - Constraints in Point Mutations as Observed in Globin Tertiary Structures. 211 *Journal of Molecular Biology* 975-988.
- Bowie, J.U., Luthy, R., and Eisenberg, D. (1991). A Method to Identify Protein Sequences That Fold into a Known Three-Dimensional Structure. 253 *Science* 164.
- Branden, C. and Tooze, J. (1991). *Introduction to Protein Structure.* (New York: Garland).
- Bregeon, A.C., Chevrier, B., Podjarny, A., Moras, D., de Bear, J.S., Gough, G.R., Gilham, P.T. and Johnson, J.E. (1988). High Resolution Structure of the RNA Duplex. 335 *Nature* 375.
- Bregeon, A.C., Chevrier, B., Podjarny, A., Johnson, J., de Bear, J.S., Gough, R.R., Gilham, P.T. and Moras, D. (1989). Crystallographic Structure of an RNA Helix. 209 *Journal of Molecular Biology* 459.
- Brennan, R.G. and Matthews, B.W. (1989). Structural Basis of DNA-Protein Recognition. 286 *Trends in Biochemical Science* 14.
- Brooks, B.R. and Hodoscek, M. (1992). Parallelization of CHARMM for MIMD Machines. 7 *Chemical Design Automation News* 12:16-22.



- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983). CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. 4 *Journal of Computational Chemistry* 187.
- Brooks, C.L., Karplus, M., and Pettitt, B.M. (1988). *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. (Advances in Chemical Physics Volume LXXI) (John Wiley and Sons: New York) 20-21.
- Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. (1995). Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. 21 *PROTEINS: Structure, Function, and Genetics* 167-195.
- Buzayan, J., van Tol, H., Feldstein, P. and Bruening, G. (1990). Identification of a non-junction phosphodiester that influences an autolytic processing reaction of RNA. 18 *Nucleic Acids Research* 15:4447-4451.
- Cantor, C.R. and Schimmel, P.R. (1980). *Biophysical Chemistry. Part I: The conformation of biological macromolecules*. (San Francisco: Freeman).
- Cantor, C.R. and Schimmel, P.R. (1980). *Biophysical Chemistry. Part II: Techniques for the study of biological structure and function*. (San Francisco: Freeman).
- Cantor, C.R. and Schimmel, P.R. (1980). *Biophysical Chemistry, Part III: The behavior of biological macromolecules*. (San Francisco: Freeman).
- Caspar, D.L.D. (1995). Problems in simulating macromolecular movements. 3 *Structure* 327-329.
- Cech, T.R. (1986). RNA as an Enzyme. 255 *Scientific American* 5:64-75.
- Cech, T.R. (1987). The Chemistry of Self-Splicing RNA and RNA Enzymes. 236 *Science* 1532-1539.
- Cech, T.R. (1988). Ribozymes and Their Medical Implications. 260 *JAMA* 20:3030-3034.
- Cech, T.R. (1989). Ribozyme self-replication? 339 *Nature* 507-508.
- Cech, T.R. (1992). Ribozyme Engineering. 2 *Current Opinion in Structural Biology* 605-609.
- Cech, T.R. (1993). Fishing for fresh catalysts. 365 *Nature* 204-205.
- Cech, T.R. and Bass, B.L. (1986). Biological Catalysis by RNA. 55 *Ann. Rev. Biochem.* 599-629.
- Cech, T.R. (1993). Structure and Mechanism of the Large Catalytic RNAs: Group I and Group II Introns and Ribonuclease P in Gesteland, R.F. and Atkins, J.F., eds. (1993). *The RNA World. The Nature of Modern RNA Suggests a Prebiotic RNA World*. (New York: Cold Spring Harbor Laboratory Press) 239-269.

- Celander, D. and Cech, T. R. (1991). Visualizing the Higher Order Folding of a Catalytic RNA Molecule. 251 *Science* 401-407.
- Chemistry with computation: An Introduction to SPARTAN. 1:1. (1995).
- Clark, M., Cramer III, R.D., and van Opdenbosch, N. (1989). Validation of the General Purpose Tripos 5.2 Force Field. 10 *Journal of Computational Chemistry* 8: 982.
- Comba, P. (1996). Inorganic Molecular Mechanics. 73 *Journal of Chemical Education* 2: 108.
- 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., and Kollman, P.A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. 117 *J. Am. Chem. Soc.* 5179-5197.
- Cotton, F. A. (1990). *Chemical Applications of Group Theory*. Third Edition. (New York: John Wiley and Sons).
- Crippen, G.M. (1982). Conformational Analysis by Energy Embedding. 3 *Journal of Computational Chemistry* 4:471-476.
- Daggett, V. and Levitt, M. (1993). Realistic Simulations of Native-Protein Dynamics in Solution and Beyond. 22 *Annu. Rev. Biophys. Biomol. Struct.* 358.
- Daggett, V. and Levitt, M. (1992). Molecular Dynamics Simulations of Helix Denaturation. 2223 *J. Mol. Biol.* 1121.
- Daggett, V. and Levitt, M. (1993). Protein Unfolding Pathways Explored Through Molecular Dynamics Simulations. 232 *J. Mol. Biol.* 600.
- Dahl, S.G., Edvardsen, O. and Sylte, I. (1991). Molecular Dynamics of Dopamine at the D2 Receptor. 88 *Proceedings of the National Academy of Sciences USA* 8111.
- Dill, K.A., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D., and Chan, H.S. (1995). Principles of protein folding-A perspective from simple exact models. 4 *Protein Science* 561-602.
- Draper, D.E. (1996). Parallel worlds. 3 *Nature Structural Biology* 5: 397-400.
- Drexler, K. E. (1994). Molecular Nanomachines: Physical Principles and Implementation Strategies. 23 *Annu. Rev. Biophys. Biomol. Struct.* 377-405.
- Edvardsen, O., Sylte, I. and Dahl, S.G. (1992). Molecular Dynamics of Serotonin and Ritanserin Interacting with the 5-HT<sub>2</sub> Receptor. *Molecular Brain Research*, 14, 166 (1992).
- Ewbank, J., Creighton, T., Hayer-Hartl, M., and Hartl, F. (1995). What is the molten globule? 2 *Structural Biology* 1:10-11.
- Fedor, M. and Uhlenbeck, O. (1990). Substrate sequence effects on "hammerhead" RNA catalytic efficiency. 87 *Proc. Natl. Acad. Sci. USA* 1668-1672.

- Fernandez, A. (1993). Simulating an exploration of RNA conformational space with an appropriate parallel-updating strategy. 48 *Physical Review E* 4: 3107-3111.
- Fersht, A. *Enzyme Structure and Mechanism* (New York: W.H. Freeman).
- Fesik, S.W. (1991). NMR Studies of Molecular Complexes as a Tool in Drug Design. 34 *Journal of Medicinal Chemistry* 10, 2937-2945.
- Findlay, J. and Eliopoulos, E. (1990). Three-Dimensional Modeling of G Protein-Linked Receptors *Trends in Pharmacological Science*, 11, 492.
- Fink, A.L. (1995). Compact Intermediate States in Protein Folding. 24 *Annual Rev. Biophys. Biomol. Struct.* 495-522.
- Fischer, D., R. Norel, H. Wolfson, and R. Nussinov. (1993). Surface Motifs by a Computer Vision Technique: Searches, Detection, and Implications for Protein-Ligand Recognition. 16 *Proteins* 278.
- Fletcher, R. (1981). *Practical Methods of Optimization*, Vol. 1: Unconstrained Optimization, Vol. 2: Constrained Optimization. (New York: John Wiley and Sons).
- Fontana, W., Stadler, P. Tarazona, P. Weinberger, E. and Schuster. (1993). RNA folding and combinatorial landscapes. 47 *Physical Review (E)* 3.
- Foresman, J.B. and Frisch, A. *Exploring Chemistry with Electronic Structure Methods: A Guide to Using Gaussian*. (Pittsburgh: Gaussian, Inc.).
- Fosdick, L.D., Jessup, E.R., Schauble, C., and Domik, G. (1996). *An Introduction to High-Performance Scientific Computing*. (Cambridge: MIT Press).
- Frauenfelder, H., Sligar, S.G., and Wolynes, P.G.. (1991). The Energy Landscapes and Motions of Proteins. 254 *Science* 1599.
- Gabler. (1978). *Electrical Interactions in Molecular Biophysics*.
- Gao, J. and Weiner, J.H. (1994). Nature of Stress on the Atomic Level in Dense Polymer Systems. 266 *Science* 748.
- Gesteland, R.F. and Atkins, J.F., eds. (1993). *The RNA World. The Nature of Modern RNA Suggests a Prebiotic RNA World*. (New York: Cold Spring Harbor Laboratory Press).
- Goodfellow, J.M. and Moss, D.S. (1992). *Computer Modelling of Biomolecular Processes*. (England: Ellis Horwood).
- Guida, W.C. (1994). Software for Structure-based Drug Design. 4 *Current Opinion in Structural Biology* 777-781.
- Hadju, J. and Andersson, I. (1993). Fast Crystallography and Time-Resolved Structures. 22 *Annu. Rev. Biophys. Biomol. Struct.* 467-98.

- Haile, J. (1992). *Molecular Dynamics Simulation: Elementary Methods*. (New York: John Wiley).
- Harris, L.F., Sullivan, M.R., Popken-Harris, P., and Hickok, D.F. (1994). Molecular Dynamics Simulations in Solvent of the Glucocorticoid Receptor Protein in Complex With a Glucocorticoid Response Element DNA Sequence. 12 *Journal of Biomolecular Structure and Dynamics* 249-270.
- Harvey, S.C. (1986). Conformational dynamics of transfer RNA. 3 *Comments on Molecular and Cellular Biophysics* 219.
- Harvey, S.C. and McCammon, J.A. (1981). Intramolecular flexibility in phenylalanine transfer RNA. 294 *Nature* 286.
- Harvey, S.C. and McCammon, J.A. (1982). Macromolecular conformational energy minimization: An algorithm varying pseudodihedral angles. 6 *Computers and Chemistry* 173.
- Harvey, S.C., Prabhakaran, M., and McCammon, J.A. (1985). Molecular-dynamics simulation of phenylalanine transfer RNA. 1. Methods and general results. 24 *Biopolymers* 1169.
- Harvey, S.C., Prabhakaran, M., Mao, B. and McCammon, J.A. (1984). Phenylalanine transfer RNA: Molecular dynamics simulation. 223 *Science* 1189.
- Harvey, S.C., Prabhakaran, M., Suddath, F.L. and McCammon, J.A. (1985). Computer graphics and moving pictures in the analysis of intramolecular motions in phenylalanine transfer RNA in *Molecular Dynamics and Protein Structure*, J. Hermans, ed. (Chapel Hill: University of North Carolina).
- Haseloff, J. and Gerlach, W.L. (1988). Simple RNA enzymes with new and highly specific endoribonuclease activities. 334 *Nature* 18:585-591.
- Kuroda, Y., Kidokoro, S., Wada, A. (1992). Thermodynamic Characterization of Cytochrome *c* at Low pH: Observation of the Molten Globule State and of the Cold Denaturation Process. 223 *J. Mol. Biol.* 1139-1153.
- Haynie, D. and Freire, E. (1993). Structural Energetics of the Molten Globule State. 16 *PROTEINS: Structure, Function, and Genetics* 115-140.
- Haynie, D. and Freire, E. (1994). Thermodynamic Strategies for Stabilizing Intermediate States of Proteins. 34 *Biopolymers* 261-271.
- Head, J.D. (1990). Partial Optimization of Large Molecules and Clusters. 11 *J. Comput. Chem.* 67-75.
- Head, J.D. and Zerner, M. C. (1985). A Broyden-Fletcher-Goldfarb-Shanno Optimization Procedure for Molecular Geometries. 122 *Chem. Phys. Lett.* 264-270.
- Hendrickson, W.A. (1995). X Rays in Molecular Biophysics. 48 *Physics Today* 11:42-43.

- Hendry, P. and McCall, M.J. (1995). A comparison of the in vitro activity of DNA-armed and all-RNA hammerhead ribozymes. 23 *Nucleic Acids Research* 3928-3936.
- Heus, H. and Pardi, A. (1991). Nuclear magnetic resonance studies of the hammerhead ribozyme domain. Secondary structure formation and magnesium ion dependence. 217 *J. Mol. Biol.* 113-124.
- Hilderbrandt, R.L. (1977). Application of Newton-Raphson Optimization Techniques in Molecular Mechanics Calculations. 1 *Computers and Chemistry* 179-186.
- Hinchliffe, A. (1988). *Computational Quantum Chemistry*. (New York: John Wiley).
- Iachello, F. and Levine, R.D. (1995). *Algebraic theory of molecules*. (New York: Oxford University Press).
- Kang, Y.K., Gibson, K.D., Nemethy, G., and Scheraga, H.A. (1987). Free Energies of Hydration of Solute Molecules. 4. Revised Treatment of the Hydration Shell Model. 92 *Journal of Physical Chemistry* 16: 4739-4742.
- Kasinos, N., Lilley, G.A., Subbarao, N., and Haneef, I. (1992). A Robust and Efficient Automated Docking Algorithm for Molecular Recognition. 5 *Protein Engineering* 69.
- Keepers, J.W., Kollman, P.A., Weiner, P.K., and James, T.L.. (1982). Molecular Mechanical Studies of DNA Flexibility: Coupled Backbone Torsion Angles and Base-Pair Openings. *Proceedings of the National Academy of Science USA*, 79, 5537.
- Kennard, O. and Salisbury, S. (1993). Oligonucleotide X-ray Structures in the Study of Conformation and Interactions of Nucleic Acids. 268 *The Journal of Biological Chemistry* 10701.
- Klimkowski, V. J., Schafer, L. and Scarsdale, J.N. (1984). Ab initio Studies of Structural Features Not Easily Amenable to Experiment. 109 *Journal of Molecular Structure (Theochem)* 311-320.
- Kochoyan, M. and Leroy, J. (1995). Hydration and solution structure of nucleic acids. 5 *Current Opinion in Structural Biology* 329-333.
- Kollman, P.A. (1996). Advances and Continuing Challenges in Achieving Realistic and Predictive Simulations of the Properties of Organic and Biological Molecules. 29 *Accounts of Chemical Research* 10: 461-469.
- Kroschwitz, J. and Winokur, M. (1990). *Chemistry: general, organic, biological*. New York: McGraw-Hill.
- Kuhl, F.S., Crippen, G.M. and Friesen, D.K.. (1984). A Combinatorial Algorithm for Calculating Ligand Binding. 5 *Journal of Computational Chemistry* 1:24-34.
- Kuntz, I.D. (1992). Structure-Based Strategies for Drug Design and Discovery 257 *Science* 1078.

- Kuntz, I.D., Meng, E.C. and Shoichet, B.K. (1994). Structure-Based Molecular Design. 27 *Accounts of Chemical Research* 117.
- Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E. (1982). A Geometric Approach to Macromolecule-Ligand Interactions. 161 *Journal of Molecular Biology* 269-288.
- Kuwajima, K. (1989). The molten globules state as a clue for understanding the folding and cooperativity of globular protein structure. 6 *Proteins* 87-103.
- Laing, L.G., Gluick, T.C., and Draper, D.E. (1994). Stabilization of RNA structure by Mg ions. 237 *J Mol Biol* 577-587.
- Landry, S.J. and Gierasch, L.M. (1994). Polypeptide Interactions with Molecular Chaperones and their Relationship to In Vivo Protein Folding. 23 *Annu. Rev. Biophys. Biomol. Struct.* 647-651.
- Langone, J.J. *Molecular Design and Modeling: Concepts and Applications. Part B: Antibodies and Antigens, Nucleic Acids, Polysaccharides, and Drugs in Methods in Enzymology, Vol. 203.* (San Diego: Academic Press).
- Latham, J. and Cech, T. (1989). Defining the Inside and Outside of a Catalytic RNA Molecule. 245 *Science*. 276-282.
- Laughton, C.A. (1994). A Study of Simulated Annealing Protocols for Use with Molecular Dynamics in Protein Structure Prediction. 7 *Protein Engineering* 235.
- Lawley, K.P. (1980). *Potential Energy Surfaces. Advances in Chemical Physics—Volume XLII.* (New York: John Wiley & Sons).
- Lee, C. and M. Levitt. (1991). Accurate Prediction of the Stability and Activity Effects of Site-Directed Mutagenesis on a Protein Core. 352 *Nature* 448.
- Lee, Y.S., Pearlstein, R. and Kador, P.F. (1994). Molecular Modeling Studies of Aldose Reductase Inhibitors. 37 *Journal of Medicinal Chemistry* 787.
- Lewis, R.A., Roe, D.C., Huang, C., Ferrin, T.E., Langridge, R. and Kuntz, I.D. (1992). Automated Site-Directed Drug Design Using Molecular Lattices. 10 *Journal of Molecular Graphics* 66.
- Lieth, C.W., R.E. Carter, D. Dolata, and T. Liljefors. (1984). RINGS - A General Program to Build Ring Systems. 2 *Journal of Molecular Graphics* 4: 117-123.
- Lilley, D.M.J., Clegg, R.M., Diekmann, S., Seeman, N.C., von Kitzing, E., and Hagerman, P.J. (1995). A nomenclature of junctions and branchpoints in nucleic acids. 23 *Nucleic Acids Research* 17: 3363-3364.
- Lin, S.L., R. Nussinov, D. Fischer, and H.J. Wolfson. (1994). Molecular Surface Representations by Sparse Critical Points. 18 *Proteins* 94.

- Lipkowitz, K. (1995). Abuses of Molecular Mechanics. 72 *Journal of Chemical Education* 12: 1070-1075.
- Long, D. and Uhlenbeck, O. (1993). Self-cleaving catalytic RNA. 7 *The FASEB Journal* 25-30.
- Lowe, J.P. *Quantum Chemistry*. (New York: Academic Press).
- Lustig, B., Lin, N.H., Smith, S., Jernigan, R.L., and Jeang, K. (1995). A small modified hammerhead ribozyme and its conformational characteristics determined by mutagenesis and lattice calculation. 23 *Nucleic Acids Research* 3531-3538.
- Luthey-Schulten, Z., Ramirez, B.E., and Wolynes, P.G. (1995). Helix-Coil, Liquid Crystal, and Spin Glass Transitions of a Collapsed Heteropolymer. 99 *J. Phys. Chem.* 2177-2185.
- MacKerell, A., Wiorkiewicz-Kuczera, J. and Karplus, M. (1995). An All-Atom Empirical Energy Function for the Simulation of Nucleic Acids. 117 *J. Am. Chem. Soc.* 11946-11975.
- Major, F., Gautheret, D., and Cedergren, R. (1993). Reproducing the three-dimensional structure of a tRNA molecule from structural constraints. 90 *Biochemistry* 9408.
- Malhotra, A., Tan, R.K-Z., and Harvey, S.C. (1994). Modeling Large RNAs and Ribonucleoprotein Particles Using Molecular Mechanics Techniques. 66 *Biophysical Journal*.
- Max, N.L., D. Malhotra, and A. Hopfinger. (1981). Computer Graphics and the Generation of DNA Conformations for Intercalation Studies. 5 *Computers and Chemistry* 19.
- McCammon, J. and Harvey, S. (1989). *Dynamics of proteins and nucleic acids*. (Cambridge: Cambridge University Press)
- McCammon, J.A., Gelin, B.R., and Karplus, M. (1977). Dynamics of folded proteins. 267 *Nature* 585.
- McRee, D. (1993). *Practical Protein Crystallography*. (San Diego: Academic Press).
- McWeeny, R. *Methods of Molecular Quantum Mechanics*, 2<sup>nd</sup> Edition. (New York: Academic Press).
- Meng, E.C., B.K. Shoichet, and I.D. Kuntz. (1992). Automated DOCKING with Grid-Based Energy Evaluation. 13 *Journal of Computational Chemistry* 505.
- Merritt, E.A. and Murphy, M.E.P. (1994). Raster3D Version 2.0: A Program for Photorealistic Molecular Graphics. *Acta Crystallographica D*50, 869.
- Mezey, P.G. (1987). *Potential Energy Hypersurfaces*. Studies in physical and theoretical chemistry, v. 53. (New York: Elsevier).
- Miller, K.J., Kowalczyk, P., Segmuller, W. and Walker, G. (1983). Interactions of Molecules with Nucleic Acids. VII. Evaluation and Presentation of Steric Contours and Molecules in Bonding Sites. 4 *Journal of Computational Chemistry* 3: 366-378.

- Mohamadi, F., N.G.J. Richards, W.C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, and W.C. Still. (1990). MacroModel - An Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. 11 *Journal of Computational Chemistry* 440.
- Mulholland, A.J., Grant, G.H., and Richards, W.J. (1993). Computer Modeling of Enzyme Catalyzed Reaction Mechanisms. 6 *Protein Engineering* 133.
- Nagasawa, M., ed. (1987). *Molecular Conformation and Dynamics of Macromolecules in Condensed Systems. Studies in Polymer Science 2.* (New York: Elsevier).
- Nakumara, S. and Doi, J. (1994). Dynamics of transfer RNAs analyzed by normal mode calculation. *Nucleic Acids Research*, 22, 514.
- Norel, R., Fischer, D., Wolfson, H.J., and Nussinov, R.. (1994). Molecular Surface Recognition by a Computer Vision-Based Technique. 7 *Protein Engineering* 39.
- Olson, W.K. and Flory, P.J. (1972). Steric configurations of polynucleotide chains, I: Steric interactions in polynucleotides: A virtual bond model. 11 *Biopolymers* 1.
- Ostlund, N.S. and Whiteside, R.A. (1985). A machine architecture for molecular dynamics-The systolic loop. 439 *Annals of the New York Academy of Sciences* 195.
- Paine, G.H. and Scheraga, H.A. (1985). Prediction of the Native Conformation of a Polypeptide by a Statistical-Mechanical Procedure. I. Backbone Structure of Enkephalin. 24 *Biopolymers* 1391.
- Pan, T. and Uhlenbeck, O.C. (1992). A small metalloribozyme with a two-step mechanism. 358 *Nature* 560-563.
- Paoletta, G., Sproat, B., and Lamond, A. (1992). Nuclease resistant ribozymes with high catalytic activity. 11 *The EMBO Journal* 5:1913-1919.
- Pardo, L., Ballesteros, J.A., Osman, R. and Weinstein, H. (1992). On the Use of the Transmembrane Domain of Bacteriorhodopsin as a Template for Modeling the Three-dimensional Structure of Guanine Nucleotide-Binding Regulatory Protein-Coupled Receptors. 89 *Proceedings of the National Academy of Science USA* 4009.
- Pearlman, D.A. and Kollman, P.A. (1991). Evaluating the Assumptions Underlying Force Field Development and Application Using Free Energy Conformational Maps for Nucleosides. 113 *J. Am. Chem. Soc.* 7177.
- Pearlman, D. and Kollman, P. (1990). Are Free Energy Calculations Necessary? A Comparison of DNA Modeling Studies. *in* Beveridge, D. and Lavery, R., eds. (1990). *Theoretical Biochemistry & Molecular Biophysics, Volume 1: DNA.* (New York: Adenine Press). 139-152.
- Peishoff, C.E., J.S. Dixon, and K.D. Kopple. (1990). Application of the Distance Geometry Algorithm to Cyclic Oligopeptide Conformation Searches. 30 *Biopolymers* 45-56.



- Perico, A., Guenza, M., Mormino, M., and Fioravanti, R. (1995). Protein Dynamics: Rotational Diffusion of Rigid and Fluctuating Three Dimensional Structures. 35 *Biopolymers* 47.
- Perreault, J., Labuda, D., Usman, N., Yang, J., and Cedergren, R. (1991). relationship between 2'-Hydroxyls and Magnesium Binding in the Hammerhead RNA Domain: A Model for Ribozyme Catalysis. 30 *Biochemistry* 4020-4025.
- Perutz, M. Protein Structure. (San Francisco: W.H. Freeman and Company).
- Pieken, W., Olsen, D., Benseler, F., Aurup, H. and Eckstein, F. (1991). Kinetic Characterization of Ribonuclease-Resistant 2'-Modified Hammerhead Ribozymes. 253 *Science* 314-317.
- Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. 372 *Nature* 68-74.
- Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. 372 *Nature* 68-74.
- Porschke, D. (1995). Modes and Dynamics of  $Mg^{2+}$ -Polynucleotide Interactions in Cowan, J.A., ed. *The Biological Chemistry of Magnesium*. (New York: VCH) 85-108.
- Prabhakaran, M., Harvey, S.C. and McCammon, J.A. (1985). Molecular-dynamics simulation of phenylalanine transfer RNA. II. Amplitudes, anisotropies, and anharmonicities of atomic motions. 24 *Biopolymers* 1189.
- Prabhakaran, M., Harvey, S.C., Mao, B. and McCammon, J.A. (1983). Molecular dynamics of phenylalanine transfer RNA. 1 *Journal of Biomolecular Structure and Dynamics* 357.
- Prabhakaran, M., McCammon, J.A. and Harvey, S.C. (1985). Atomic motions in phenylalanine transfer RNA probed by molecular dynamics simulations in *Molecular Basis of Cancer, Part A: Macromolecular Structure, Carcinogens, and Oncogenes*, R. Rein, ed. pp. 123-9. (New York: Alan R. Liss).
- Press, W.H., Flannery, B.P., Teukolsky, S. A., and Vetterling, W.T. (1986). *Numerical Recipes. The Art of Scientific Computing*. (Cambridge University Press).
- Privalov, P.L. (1996). Intermediate States in Protein Folding. 259 *J. Mol. Biol.* 707-725.
- Ptitsyn, O.B. (1996). How molten is the molten globule? 3 *Nature Structural Biology* 6: 488-490.
- Ptitsyn, O.B. (1966). Conformations of Macromolecules.
- Ptitsyn, O.B., and Uversky, V.N. (1994). The molten globule is a third thermodynamical state of protein molecules. 341 *FEBS Letters* 15-18.
- Ptitsyn, O.B. (1987). Protein folding: hypotheses and experiments. 6 *J. Protein Chem.* 273-293.

- Pulay, P. (1969). Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. I Theory Mol. Phys. 17:197-204.
- Pyle, A. and Cech, T. (1991). Ribozyme recognition of RNA by tertiary interactions with specific ribose 2'-OH groups. 350 Nature 628-631.
- Pyle, A.M. (1993). Ribozymes: a distinct class of metalloenzymes. 261 Science 709-714.
- Pyle, A.M. and Green, Justin B. (1995) RNA folding. 5 Current Opinion in Structural Biology at 303.
- Ramos, M.J. (1992). Model Structure for the Human Blood Coagulation Agent Beta-Factor XIIa. 9 Journal of Molecular Graphics 92.
- Rapaport, D.C. (1995). The Art of Molecular Dynamics Simulation. (Cambridge: Cambridge University Press).
- Rawls, R. (1996). Splicing ribozyme can 'edit' mammalian RNA.. Chemical and Engineering News. June 3, 1996 at 7.
- Reid, R.H., C.A. Hooper, and B.R. Brooks. (1989). Computer Simulations of a Tumor Surface Octapeptide Epitope. 28 Biopolymers 525-530.
- Rhodes, G. (1996). Crystallography Made Crystal Clear. (San Francisco: Academic Press).
- Rich, A. (1978). Transfer RNA: Three-dimensional structure and biological function. 3 Trends in Biochemical Sciences 34.
- Richards, F.M. (1977). Areas, volumes, packing and protein structure. 6 Annual Review of Biophysics and Bioengineering 151.
- Richards, W.G. (1984). Quantum Pharmacology. 8 Endeavour 172.
- Richardson, J.S. and Richardson, D.C. (1985). The De Novo Design of Protein Structures. 14 Trends in Biochemical Science 304.
- Roitberg, A., Gerber, R., Elber, R., Ratner, M. (1995). Anharmonic Wave Functions of Proteins: Quantum Self-Consistent Field Calculations of BPTI. 268 Science 1319.
- Rotstein, S.H. and M.A. Murcko. GroupBuild: A Fragment-Based Method for De Novo Drug. 36 Design Journal of Medicinal Chemistry 1700.
- Ruffner, D. and Uhlenbeck, O. (1990). Thiophosphate interference experiments locate phosphates important for the hammerhead RNA self-cleavage reaction. 18 Nucleic Acids Research 20:6025-6029.
- Rullmann, J.A.C. and van Duijnen, P.Th. (1990). Potential Energy Models of Biological Macromolecules: A Case for Ab Initio Quantum Chemistry. 1 Reports in Molecular Theory 1-21.
- Saenger, W. (1984). Principles of Nucleic Acid Structure. (New York: Springer-Verlag).

- Saenger, W. (1987). Structure and Dynamics of Water Surrounding Biomolecules. 16 *Ann. Rev. Biophys. Biophys. Chem.* 93-114.
- Sarma, R.H., Ed. *Nucleic Acid Geometry and Dynamics*. (New York: Pergamon Press).
- Saunders, M., K.N. Houk, Y. Wu, W.C. Still, M. Lipton, G. Chang, and W.C. Guida. (1990). Conformations of Cycloheptadecane. A Comparison of Methods for Conformational Searching. 112 *Journal of the American Chemical Society* 1419-1427.
- Sayle, R. and Bissell, A. (1992). RasMol: A Program for Fast Realistic Rendering of Molecular Structures With Shadows in *Proceedings of the 10<sup>th</sup> Eurographics UK '92 Conference*, University of Edinburgh, Scotland, April 1992.
- Scheraga, H. (1981). Influence of Interatomic Interactions on the Structure and Stability of Polypeptides and Proteins. 20 *Biopolymers* 1877.
- Schertler, G.F.X., C. Villa, and R.Henderson. (1993). Projection Structure of Rhodopsin. 362 *Nature* 770.
- Schlegel, H.B. (1987). Optimization of Equilibrium Geometries and Transition Structures. 67 *Adv. Chem. Phys.* 249-286.
- Schlick, T., Hingery, B., Peskin, C., Overton, M. and Broyde, S. (1990). Search Strategies, Minimization Algorithms, and Molecular Dynamics Simulations for Exploring Conformational Spaces of Nucleic Acids. in Beveridge, D.L. and Lavery, R., eds. *Theoretical Biochemistry & Molecular Biophysics, Volume 1: DNA*. (New York: Adenine Press) 39-58.
- Scott, W. and Klug, A. (1996). Ribozymes: structure and mechanism in RNA catalysis. 21 *TIBS* 220-224.
- Sharp, K. and Honig, B. (1995). Salt effects on nucleic acids. 5 *Current Opinion in Structural Biology* 323-328.
- Sharp, P. (1988). RNA Splicing and Genes. 260 *JAMA* 20: 3035-3041.
- Sinden, R.R. (1994). *DNA structure and function*. (San Diego: Academic Press).
- Skolnick, J., Koliniski, A., and Godzik, A. (1993). From independent modules to molten globules: Observations on the nature of protein folding intermediates. 90 *Proc. Natl. Acad. Sci. USA* 2099-2100.
- Smith, D. (1995). Magnesium as the Catalytic Center of RNA Enzymes. in Cowan, J.A. , ed. *The Biological Chemistry of Magnesium*. (New York: VCH) 113.
- Spiegel, Murray R. (1993). *Mathematical Handbook of Formulas and Tables*. (McGraw-Hill, Inc.).
- Steinbach, P.J. and Brooks, B.R. (1994). New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. 15 *Journal of Computational Chemistry* 667.

- Stewart, J.P. (1990). Semiempirical Molecular Orbital Methods in Reviews in Computational Chemistry, Vol. 1, 45, D.B. Boyd and K.B. Lipkowitz, Eds., VCH Publishers, New York.
- Still, C., Tempczyk, A., Hawley, R.C., and Hendrickson, T., (1990). Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. 112 J. Am. Chem. Soc. 6127-6129. (solvent treated as a statistical continuum).
- Strauss, H.L. (1983). Pseudorotation: A Large Amplitude Molecular Motion. 34 Annual Review of Physical Chemistry 301.
- Strosberg, A.D., Camoin, L., Blin, N. and Maigret, B. (1993). Receptors Coupled to GTP-Binding Proteins, Ligand Binding and G-Protein Activation is a Multistep Dynamic Process. 9 Drug Design and Discovery 199.
- Stryer, L. (1988). Biochemistry. New York: Freeman.
- Sybesma, C. (1977). Biophysics: An Introduction. (The Netherlands: Kluwer Academic Press).
- Szabo, A. and Ostlund, N. Modern Quantum Chemistry, First Edition (Revised). (New York: McGraw-Hill).
- Thirumalai, D. and Woodson, S.A. (1996). Kinetics of Folding of Proteins and RNA. 29 Acc. Chem. Res. 433-439.
- Thornton, J.M. and Gardner, S.P. (1989). Protein Motifs and Data-Base Searching. 14 Trends in Biochemical Science 300.
- Timms, D., Wilkinson, A.J., Kelly, D.R., Broadley, K.J., and Davies, R.H. (1992). Interactions of Tyr377 in a Ligand-Activation Model of Signal Transmission through Beta 1-Adrenoceptor Alpha-Helices. 19 International Journal of Quantum Chemistry: Quantum Biology Symposium 197.
- Ts'ao, P.O.P., ed. (1974). Basic Principles in Nucleic Acid Chemistry. (New York: Academic Press).
- Tung, Chang-Shung and Carter II, E.S. (1994). Nucleic Acid Modeling Tool (NAMOT): An Interactive Graphic Tool for Modeling Nucleic Acid Structures. 10 CABIOS 4:427-433.
- Tung, Chang-Shung and Harvey, S.C. (1986). Base Sequence, Local Helix Structure, and Macroscopic Curvature of A-DNA and B-DNA. 261 Journal of Biological Chemistry 3700.
- Tung, Chang-Shung and Harvey, S.C. (1986). Computer Graphics Program to Reveal the Dependence of the Gross Three-dimensional Structure of the B-DNA Double Helix on Primary Structure. 14 Nucleic Acids Research 381.
- Tuschl, T. Gohkle, C., Jovin, T., Westhof, E., Eckstein, F. (1994). A Three-Dimensional Model for the Hammerhead Ribozyme Based on Fluorescence Measurements. 266 Science 785-788.

- Uchimaru, T., Uebayasi, M., Tanabe, K., and Taira, K. (1993). Theoretical analyses on the role of  $Mg^{2+}$  ions in ribozyme reactions. 7 *The FASEB Journal* 137-142.
- Udgaonkar, J. and Baldwin, R. (1995). Nature of the Early Folding Intermediate of Ribonuclease A. 34 *Biochemistry* 12: 4088-4096.
- Uhlenbeck, O.C. (1987). A small catalytic oligoribonucleotide. 328 *Nature* 596-600.
- Uversky, V., Semisotnov, G., Pain, R. and Ptitsyn, O.B. (1992). 314 *FEBS Letters* 1:89-92.
- van Gunsteren, W. F. and Mark, A. E. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 948.
- van Gunsteren, W. F. and Berendsen, H.J.C. (1985). Molecular dynamics simulations: Techniques and applications to proteins. In *Molecular Dynamics and Protein Structure*, ed. J. Hermans, pp. 5-14. Chapel Hill: University of North Carolina.
- van Gunsteren, W. F. and Mark, A.E. (1991). On the interpretation of biochemical data by molecular dynamics computer simulation. 204 *Eur. J. Biochem.* 947-961.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1977). Algorithms for macromolecular dynamics and constraint dynamics. 34 *Molecular Physics* 1311.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1982). Molecular dynamics: Perspective for complex systems. 10 *Biochemical Society Transactions* 301.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1990). Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. 29 *Angew. Chem. Int. Ed. Engl.* 992-1021.
- van Gunsteren, W.F., Luque, F.J., Timms, D. and Torda, A.E. (1994). Molecular Mechanics in Biology: From Structure to Function, Taking Account of Solvation. 23 *Annu. Rev. Biophys. Biomol. Struct.* 847-863.
- Vasquez, M. and H.A. Scheraga. (1985). Use of Buildup and Energy-Minimization Procedures to Compute Low-Energy Structures of the Backbone of Enkephalin. 24 *Biopolymers* 1437.
- Venable, R.M., Brooks, B.R., and Carson, F.W. (1993). Theoretical Studies of Relaxation of a Monomeric Subunit of HIV-1 Protease in Water Using Molecular Dynamics. 15 *PROTEINS: Structure, Function, and Genetics* 374.
- Venable, R.M., Zhang, Y., Hardy, B.J., and Pastor, R. (1993). Molecular Dynamics Simulations of a Lipid Bilayer and of Hexadecane: An Investigation of Membrane Fluidity. 262 *Science* 223.
- Verlet, L. (1967). Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. 159 *Phys. Rev.* 98.

- Vidugiris, G., Markley, J., Royer, C. (1995). Evidence for a Molten Globule-like Transition State in Protein Folding from Determination of Activation Volumes. 34 *Biochemistry* 15: 4909-4912.
- Waltho, J.P., Vinter, J.G., Davis, A., and Williams, D.H. (1988). Forces in Molecular Recognition: Comparison of Experimental Data and Molecular Mechanics Calculations. 2 *Journal of Computer-Aided Molecular Design* 31.
- Wang, Y., Zhang, H., Li, W., and Scott, R. (1995). Discriminating compact nonnative structures from the native structure of globular proteins. 92 *Proc. Natl. Acad. Sci, USA* 709-713.
- Warshel, A. (1991). *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. (New York: John Wiley & Sons, Inc.).
- Watson, J.D. and Crick, F.H.C. (1953). Molecular structure of nucleic acids. 171 *Nature* 737.
- Weiner, P.K. and Kollman, P.A. (1981). AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. 2 *Journal of Computational Chemistry* 287.
- Weiner, S.J., Kollman, P.A., Case, D.A., Chandra Singh, U., Ghio, C., Alagona, G., Profeta, Jr., S., and Weiner, P. (1984). A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. 106 *Journal of The American Chemical Society* 765.
- Welch, G.R. (1986). *The Fluctuating Enzyme*. (New York: John Wiley and Sons).
- Werner, M. and Uhlenbeck, O.C. (1995). The effect of base mismatches in the substrate recognition helices of hammerhead ribozymes on binding and catalysis. 23 *Nucleic Acids Research* 12:2092-2096.
- Westhof, E. and Patel, D. (1995). Nucleic Acids: Diversity, folding, and stability of nucleic acid structures. 5 *Current Opinion in Structural Biology* 279-281.
- Westkaemper, R.B. and Glennon, R.A. (1991). Approaches to Molecular Modeling Studies and Specific Application to Serotonin Ligands and Receptors. 40 *Pharmacology Biochemistry and Behavior* 1019.
- Williams, K.A. and Deber, C.M. (1991). Proline Residues in Transmembrane Helices: Structural or Dynamic Role? *Biochemistry* 8919.
- Wilson, S. and Diercksen, G.H.F. , eds. (1992). *Methods in Computational Molecular Physics*. NATO ASI Series. Series B: Physics Vol. 293. (New York: Plenum Press).
- Winn, J.S. (1994). *Physical Chemistry*. (New York: HarperCollins).
- Withka, J.M., Swaminathan, S., Srinivasan, J., Beveridge, D.L., and Bolton, P.H.. (1992). Toward a Dynamical Structure of DNA: Comparison of Theoretical and Experimental NOE Intensities. *Science* 597.

Wong, S.S.M. (1992). *Computational Methods in Physics and Engineering*. (New Jersey: Prentice Hall).

Yarus, M. (1993). How many catalytic RNAs? Ions and the Cheshire cat conjecture. 7 *The FASEB Journal* 31-39.

Zerner, Michael C. (1991). Semiempirical Molecular Orbital Methods in *Reviews in Computational Chemistry*, Vol. 2, 313, D.B. Boyd and K.B. Lipkowitz, Eds., VCH Publishers, New York .

Zhang, D. and Weinstein, H. (1993). Ligand Selectivity and the Molecular Properties of the 5-HT<sub>2</sub> Receptor: Computational Simulations Reveal a Major Role for Transmembrane Helix. 7 *Medicinal Chemistry Research* 357.

Zhang, D. and Weinstein, H. (1993). Signal Transduction by a 5-HT<sub>2</sub> Receptor: A Mechanistic Hypothesis from Molecular Dynamics Simulations of the Three-Dimensional Model of the Receptor Complexed to Ligands. 7 *Journal of Medicinal Chemistry* 934.

Zuker, M. (1989). On Finding All Suboptimal Foldings of an RNA Molecule. 244 *Science* 48-52.

## VITA

Vasant Gandhi was born in Chicago, IL on July 27, 1962. He received undergraduate, law, and graduate taxation degrees from the University of Illinois at Urbana-Champaign. He entered Loyola University Chicago as a graduate student in the Department of Chemistry in the Fall semester of 1994.



## APPROVAL SHEET

The thesis submitted by Vasant T. Gandhi has been read and approved by the following committee:

Dr. Kenneth W. Olsen, Ph.D., Director  
Professor and Chairman, Chemistry  
Loyola University of Chicago

Dr. David S. Crumrine, Ph.D.  
Associate Professor, Chemistry  
Loyola University of Chicago

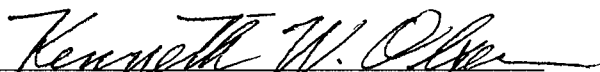
Dr. Willetta Greene-Johnson, Ph.D.  
Assistant Professor, Chemistry  
Loyola University of Chicago

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Science.

12/9/96

Date



Research Director