STRUCTURAL SIGNIFICANCE OF THE ANTICODON 3'-ADJACENT N6-(Δ^2-ISOPENTENYL) ADENOSINE AND RELATED MODIFICATIONS IN tRNA

A THESIS

submitted to the

UNIVERSITY OF PUNE

for the degree of

DOCTOR OF PHILOSOPHY

in

BIOCHEMISTRY

By

KAILAS D. SONAWANE

PHYSICAL CHEMISTRY DIVISION

NATIONAL CHEMICAL LABORATORY

PUNE - 411 008, INDIA

JULY 2002

…….Dedicated to My Family and Teachers

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled, **"STRUCTURAL SIGNIFICANCE OF THE ANTICODON 3'-ADJACENT N6-(Δ^2-ISOPENTENYL) ADENOSINE AND RELATED MODIFICATIONS IN tRNA"** submitted by **Mr. Kailas D. Sonawane,** for the Degree of Doctor of Philosophy, was carried out by the candidate under my supervision in the Physical Chemistry Division, National Chemical Laboratory, Pune, INDIA. Such material as has been obtained from other sources has been duly acknowledged in the thesis.

Dr. R. Tewari

Research Guide

CONTENTS

ACKNOWLEDGEMENT

It is a pleasure to acknowledge with respect the valuable guidance, numerous discussions, suggestions and support given by my research guide, Dr. R. Tewari, Scientist-F, NCL, Pune during the course of this work. His constant encouragement, insight and close attention to details have been lead to the successful completion of this work.

I wish my sincere thanks to Dr. S. K. Date, Head, Physical Chemistry Division, for encouragement and constant support.

I am also thankful to Dr. C. G. Suresh, Sicentist-E1, Biochemical Sciences Division, NCL, Pune for cooperation, encouragement and motivation to complete this work successfully. I also thank his students K. N. Rao, Manish, Sathya, Priya and Suresh for nice cooperation.

I wish to express my thanks to Dr. Sourav Pal, Dr. Anil Kumar, Mr. Gangopadhyaya, Mr. Koshi, Mr. Dolas, Mr. Punekar and all the other scientific and non scientific staff in the Physical Chemistry Division, NCL, for their help and cooperation given to me in completing my research work successfully..

I would like to thank my friends Uddhav, Suresh, Milind, Shivanand, Balkrishna, Bennur, and Subrhamanium, for co-operation, discussions and help through out this investigations. I would also like to thank my labmates Dr. K. Sailaja, Dr. V. Nayana, Chandrakumar, Sham, Sharan, Sajeev, Akhilesh, Prashant, Subhashini, Sophi for keeping cheerful environment while working.

No words can express my feelings towards my parents, brothers, Vikas, Babasaheb, for their love and unfailing support.

I thank Dr. P. Ratnasamy, Director, NCL for permitting me to carry out my research work at NCL.

Finally, I thank CSIR, New Delhi, for financial support.

Kailas D. Sonawane

ABBREVIATIONS

one letter code	symbol	Name
U	U	uridine
C	C	cytosine
A	A	adenine
G	G	guanosine
+	i^6A	N6-isopentenyladenosine
*	ms^2i^6A	2-methylthio-N6-isopentenyladenosine
`	io^6A	N6-(cis-hydroxyisopentenyl) adenosine
&	ncm^5U	5-carbamoylmethyluridine
}	k^2C	lysidine
<	?C	unknown modified cytidine
'	m^3C	3-methylcytidine
D	D	Dihydrouridine
J	Um	2'-O-methyluridine
!	$cmnm^5U$	5-carboxymethylaminomethyluridine
1	mcm^5U	5-methoxycarbonylmethyluridine
5	mo^5U	5-methoxyuridine
Z	psim	2'-O-methylpseudouridine
Q	Q	queuosine
I	I	inosine
M	ac^4C	N^4-Acetalcytidine
N	?U	unknown modified uridine
L	m^2G	N2-methylguanosine
R	m22G	N2,N2-dimethylguanosine
P	psi	pseudouridine
B	Cm	2'-O-methylcytidine
#	Gm	2'-O-methylguanosine
H	?A	unknown modified adenosine
V	cmo^5U	uridine 5-oxyacetic acid
=	m^6A	N6-methyladenosine
6	t^6A	N6-threonylcarbamoyladenosine
K	m^1G	1-methylguanosine

ABSTRACT

Transfer RNA (tRNA) is a multifunctional key cellular molecule which is widely known for its involvement in genetic decoding during protein biosynthesis (1-3). Modified components found in tRNA are derivatives of the four common ribonucleosides namely Adenosine (A), Guonosine (G), Uridine (U) and Cytidine (C). Most of the modifications involve simple alkylation, hydrogenation, thiolation or isomerization of these four common ribonucleosides in the base and the 2'-hydroxyl group of the ribose. However, some modifications involve complex chemical modifications which are characterized by the presence of diverse functional groups in substituents, such tRNA components are referred as hypermodified nucleosides. The hypermodified nucleic acids bases N^6-(Δ^2-isopentenyl) adenine, i^6Ade, and related derivatives like 2-methylthio- N^6-(Δ^2-isopentenyl) adenine, ms^2i^6Ade or N^6-(Δ^2- cis-hydroxyisopentenyl) adenine, cis-io^6Ade (also known as cis- zeatin) and 2-methylthio- N^6-(Δ^2 –cis-hydroxyisopentenyl) adenine, cis-ms^2io^6Ade (also known as cis-ms^2zeatin) occur at anticodon 3'-adjacent position in those tRNAs which recognize the codons starting with uridine (4). Likewise the presence of trans-zeatin and trans-ms^2zeatin in tRNA has also been reported. The change in tRNA modification from isopentenyl to hydroxyisopentenyl substituent with aerobiosis may signify turning off the anaerobic metabolism pathways and switching on to aerobic ones (5). Under aerobic conditions characterized by the presence of free molecular oxygen, in salmonella tRNA hydroxylation of ms^2i^6Ade to ms^2io^6Ade occurs (6). As free bases i^6Ade and io^6Ade act as cytokinin active compounds which promote cell division and differentiation in plant tissue culture. Similarly the hypermodified nucleosides 5-carbamoylmethyl uridine (ncm^5U) and Lysidine (k^2C) occur at the first position (34^{th}) in the anticodon loop of tRNA. The lysidine modification prevents mistranslation of AUG as isoleucine and that of AUA as methionine (7). The diverse modifications occuring at the first position of the anticodon do not hinder canonical Watson-Crick base pairing, but can restrict or enlarge the scope of Wobble base pairing for reading the third base of the codon. The anticodon 3'-adjacent modifications prevent extended Watson-Crick base pairing on the 3'-side and may thus help define the proper reading frame for anticodon-codon interactions. In order to understand the conformational preferences of these hypermodified nucleosides,

conformational energy calculations using suitable theoretical methods with computer simulations and molecular modeling visualization have been made.

OUTLINE OF THE THESIS: The thesis has been divided into the following chapters.

CHAPTER I : INTRODUCTION

The general introduction to nucleic acids and its common components, more specifically to tRNA with its usual and modified nucleic acid bases is provided. For instance N^6-(Δ^2-isopentenyl)adenine, i^6Ade with hydrophobic substituent and related derivatives like 2-methylthio- N^6-(Δ^2-isopentenyl) adenine , ms^2i^6Ade, the hydroxylated derivatives, cis or trans zeatin (io^6Ade), cis or trans ms^2zeatin (ms^2io^6Ade) besides modified nucleic acid base having hydrophilic substituent like N6-(N-glycylcarbonyl)adenine (gc^6Ade), occur at 37^{th} position in anticodon loop of specific tRNAs. Examples of modified nucleosides at 34^{th} position include 5-cabamoylmethyl uridine (ncm^5U) and 4-amino-2-(N^6-lysino)-1-(β-D-ribofuranosyl)pyrimidinium ("Lysidine") designated as k^2C. References to the published literature on occurrence of modified bases in various tRNAs of different organisms are provided. Aim of the chapter is to provide the scope of the thesis and an overall outline of the reported investigation.

CHAPTER II: METHODS

Chapter II describes the methods utilized to study the conformational preferences and the significance of modified base in tRNA anticodon loop, the methods are-

1). PCILO: The semi-empirical quantum chemical perturbative configuration interaction with localized orbitals (PCILO) method has been used for the energy calculations of various molecular conformations (8-9).

2). Automated geometry optimization calculations using quantum chemical molecular orbital methods based on Modified Neglect of Differential Overlap (MNDO) framework but various subsequent parametric refinements like Austin Model 1 (AM1) and (PM3) for comparative study.

3). Full geometry optimization with the help of molecular mechanics force field (MMFF) method specially useful for tackling large molecules.

4). Hartree-Fock (Self-Consistent Field) method has also has been used to check the salient features.

5). Molecular dynamics simulation has been used to study the structural features of tRNA anticodon stem loop. Steepest-Descent minimization is also initially applied in order to relax the starting geometry and remove steric clashes.

CHAPTER III: PART-A: CONFORMATIONAL PREFERENCES OF HYDROXYISOPENTENYL ADENOSINE AND ANALOGS.

Chapter III, part-A presents the results and discussion of hypermodified bases, N^6-(Δ^2-cis-hydroxyisopentenyl)adenine, cis-io^{6}Ade, or cis-zeatin and N^6-(Δ^2-trans-hydroxyisopentenyl) adenine, trans-io^{6}Ade or trans-zeatin and 2-methylthio derivatives of these cis-ms^{2}io^{6}Ade or cis-ms^{2}zeatin and trans-ms^{2}io^{6}Ade or trans-ms^{2}zeatin. The most stable and alternative stable conformations for these molecules are presented. Model hypermodified nucleotide diphosphate segment, trinucleotide segment, and seven nucleotide anticodon loop segments contain N^6-(Δ^2-isopentenyl)adenosine, i^{6}A and its 2-methylthio derivative ms^{2}i^{6}A Interactions of the modified base with the other bases and ribose-phosphate backbone in anticodon loop are discussed in this chapter. The important role of polar hydroxyl group in various zeatin isomers and ms^2-derivatives is also discussed.

PART B: COMPILATION AND ANALYSIS OF tRNA DATABASE:

PART-B: The tRNA sequence database of M. Sprinzl et al. (10) has been utilized for analysis. Sequence patterns and statistical analysis of various modified bases occuring at 37th and at 34th positions in anticodon loop of tRNA is presented.

CHAPTER IV: HYPERMODIFIED NUCLEOSIDE 5-CARBAMOYLMETHYL URIDINE (ncm^{5}U$_{34}$) AND ITS INTERACTION WITH IO^{6}A$_{37}$

Chapter IV presents the most stable conformation and alternative structures for the hypermodified nucleoside 5-carbamoylmethyl uridine (ncm^{5}U). In tRNA^{ser}(UGA) from Nicotiana rustica, the wobble nucleoside ncm^{5}U occurs at 34th position along with io^{6}A present at 37th position. A seven nucleotide anticodon loop structure is described here for cis or trans io^{6}A. The intermolecular interactions between adjacent nucleosides and interactions of 34th and 37th bases with ribose phosphate backbone in anticodon loop is discussed.

CHAPTER V: STUDIES ON OTHER INTERESTING MODIFIED NUCLEOSIDES

PART-A Includes the protonation induced conformational transitions of the N(6) substituents, in hypermodified nucleic acid base N6-(N-glycylcarbonyl)adenine

(gc^6Ade), on diprotonation of the adenine ring at any two of the three possible N(1), N(3) and N(7) sites.

PART-B Includes the conformational preferences of hypermodified nucleoside 4-amino-2-(N^6-lysino)-1-(β-D-ribofuranosyl)pyrimidinium ("Lysidine") designated as k^2C.

CHAPTER VI: ANTICODON LOOP STRUCTURE WITH MODIFIED COMPONENTS

Molecular dynamics study of anticodon stem loop part is described. In $tRNA^{ser}$ (UGA) from Nicotiana rustica, modified nucleosides, ncm^5U at 34^{th} position and io^6A at 37^{th} position are found. E. coli $tRNA^{ser}$(GGA) also contains io^6A at the 37th position. Molecular models of the anticodon arm of these have been utilized for molecular dynamics simulations to probe the effect of modified components on the anticodon loop structure.

REFERENCE:

1. McCloskey J. A. and Nishimura, S. Accounts Chem. Res. 10, 403-410, 1977
2. Feldmann, M. Y. Prog. Biophys. Mol. Biol. 32, 83, 1977
3. Bjork, G. R.; Ericson, J. U.; Gustafsson, C. E. D.; Hagervall, T.G.; Jonsson Y.H.; Wikstrom, P.M. Ann. Rev. Biochem. 56, 263-287, 1987.
4. Buck, M.; Griffiths, E. Nucleic Acid Res. 10, 2609-2624, 1982
5. Starzyk, R.; TIBS. 333-334, 1984.
6. Buck, M.; Ames, B.N. Cell.36, 523-531, 1984.
7. Muramatsu, T.; Nishikawa K.; Nemoto F.; Kuchino Y.; Nishimura S.; Miyazawa T.; Yokoyama S. Nature. 336, 179-181, 1988.
8. Diner, S.; Malrieu, J. P.; Jordan, F.; Gillbert, M. Theor. Chim. Acta. 15, 100-110, 1969.
9. Malrieu, J. P. In semiempirical Methods of Electronic Structure Calculations, Part A, Techniques, Segal, G. A., Ed.; Plenum: New York,; p. 69-103, 1977.
10. Sprinzl, M.; Horn, C.; Brown M.; Ioudovitch, A.; Steinberg, S. Nucleic Acid Res. 26, 148-153, 1998.
11. Tewari R. Int. J. Quantum Chem. 34, 133-142, 1988.

CHAPTER-I
INTRODUCTION

1.1 General Introduction of tRNA Molecule

Transfer RNA (tRNA) is term used by Allen and Schweet to describe a group of low molecular weight RNA molecules that play an important role in protein synthesis [1]. It is the smallest RNA molecule found in the cell, it accounts for only 5 percent of total RNA. It has been found in all animal, plant, viral, bacterial cells and in eukaryotic organisms. It has also been reported in cell organelles like mitochondria [2, 3]. It is the original adapter molecule, suggested by Crick in his adapter hypothesis [4]. It was first isolated by Hoagland and Zamecnik in 1958 [5]. When first isolated it was called soluble RNA or sRNA [6, 7], since it was found in the soluble cytoplasmic portion (centrifugal fraction) of the cell [5]. The molecule is small in size, nearly 70 to 80 nucleotides in length, with molecular weight of 25 kd. There is at least one specific tRNA for each of the 20 natural amino acids. Also, there is at least one specific aminoacyl-tRNA synthetase (activating enzyme) for each amino acid. With the help of these activating enzymes protein synthesis starts with the ATP-driven activation of amino acid.

1.1.1 Primary Structure of tRNA:

The first total sequence of a transfer RNA was determined in 1965, by Holley [8]. Nearly 100 different tRNAs have been sequenced. Nucleotide sequence information on tRNA is obtained with the help of two-dimensional electrophoresis technique developed by Sanger and co-workers using RNA labeled in vivo with ^{32}P [9]. Major tRNA sequences determined so far have depended on this technique. The sequencing of other much larger RNA molecules is also made possible with this method. Primary structure i.e. sequence analysis gives basic information about tRNA molecule. Results obtained by sequence analysis also provide valuable information about secondary structure within a “cloverleaf “ diagram.

1.1.2 Secondary Structure of tRNA Molecule:

Holley, described the folding of the tRNA sequences to a secondary structure in a “cloverleaf” diagram (fig.1.1). All the tRNA nucleoside sequences determined to date are arranged in such a secondary structure. His study of yeast phenylalanine tRNA gives the first complete sequence of any nucleic acid, as well as provides information about functions of tRNA. The transfer RNA ($tRNA^{Phe}$) molecule is a single chain of seventy-six ribonucleosides. It folds in such a way that, there is base-paired double-strand regions called stems and unpaired single strand regions called

loops. The molecule looks like a four-leaf clover and so gets its name cloverleaf diagram (fig.1.1). The 5'- terminus is always phosphorylated, the residue at 5'-terminus is usually (pG). Where as the 3'- terminus has a free hydroxyl group. The base sequence at the 3'- end of tRNA is – CCA^{OH} (fig.1.1).

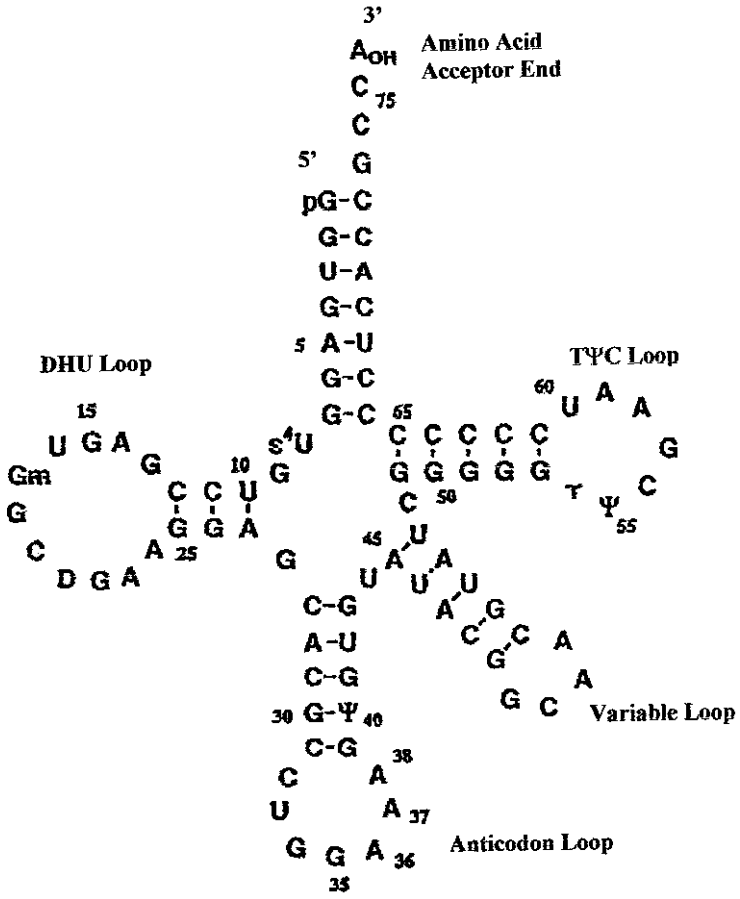


Fig. 1.1: Cloverleaf diagram for tRNA molecule.

The nucleotide sequence of the tRNA molecule contains many modified nucleosides. These modified nucleosides are produced by post-transcriptional enzymatic modifications inside the cell. Some of these involve common RNA bases modified by single methyl (-CH3) groups, while others are quite complex. According to these modified bases present some parts of the cloverleaf diagram get their name. For example, the TΨC loop which contains modified base ribothymidine (T), the 2'-dehydroxyribo form of which is commonly found in DNA but is not found in RNA species other than tRNA, pseudouridine (Ψ) and cytosine (C). The DHU loop, contains a modified base dihydrouracil (DHU). Other regions in cloverleaf diagram are the variable loop, which contains different number of residues in different tRNAs.

The anticodon loop, which contains three bases of anticodons and the amino acid acceptor stem, which accepts the amino acid specific to that particular tRNA.

In cloverleaf diagram (fig.1.1) the number of base pairs in the stem regions are found to be constant, for instance, seven base pairs in the amino acid acceptor stem, five in the TΨC and anticodon stems and three or four in the DHU stem. Base pairs in stem region are bonded by Watson-Crick hydrogen bonding except in prokaryotic initiator $tRNA^{fMet}$ where residue 1 and 70 are not base paired. In *E. coli* selenocystein tRNA has eight base pairs in amino acid acceptor stem. These characteristics are maintained in tRNA molecules from plants, animals, bacteria and viruses.

1.1.3 Three-Dimensional Structure of tRNA:

Crystallographers made efforts to crystallize tRNAs from various species [10-14]. The following tRNA molecules were investigated: $tRNA^{Phe}$, $tRNA^{Gly}$, $tRNA^{Met1}$, and $tRNA^{Asp}$ from yeast, $tRNA^{fMet}$, and $tRNA^{Arg}$ from *E.coli*. But these crystals did not give the desired high resolution. Only yeast $tRNA^{Phe}$ [15, 16], crystals provide suitable details on tRNA structure. Recently still higher resolution crystal structures of Yeast $tRNA^{Phe}$ have been obtained [17, 18].

The shape of completely folded tRNA is found to have "L-shape" (fig.1.2). It is always shown in the orientation of a " Γ ". At one end of 'L' there is amino acid attachment site and at the other end is the anticodon loop. These are quite far from each other (80 A°). The DHU and TΨC loops form one corner of 'L'. The TΨC and acceptor stem form one continuous helix. The –CCA terminus and the adjacent helical region do not interact strongly with the rest of the molecule. This part of the molecule may change its conformation during amino acid activation and also during protein synthesis on the ribosome. The two helical regions are arranged in such a way to provide the structural support for the "L-shape" folding pattern. Each region consists of about 10 base pairs, which makes one turn of the double helix.

Half of the nucleotides in tRNAs are base paired (helical regions) and form double helices. Five groups of bases are not base paired (non-helical region) these are, the 3'-CCA terminal region, the TΨC loop, the 'variable loop' the DHU loop, and the anticodon loop.

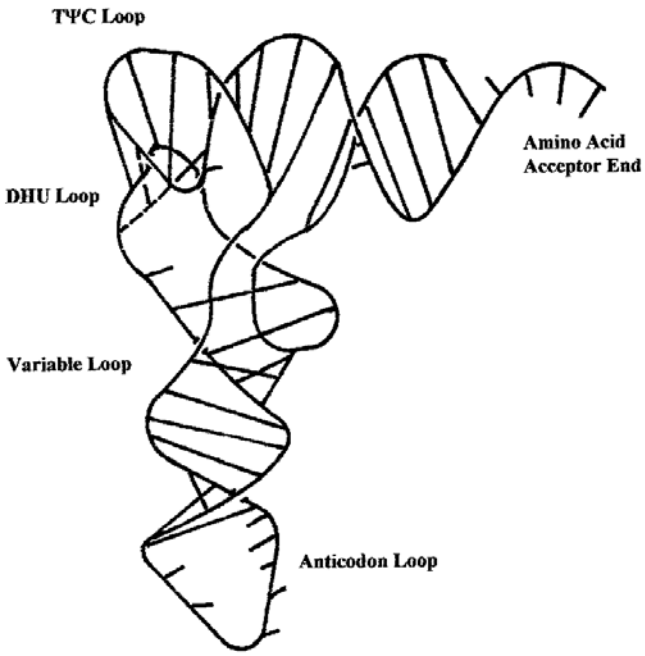


Fig 1.2: Showing L- shape model for tRNA molecule.

In non-helical region some of the bases participate in unusual tertiary hydrogen bonding interactions. These tertiary interactions between bases need not be complementary. Such non-canonical pairings include (e.g. G.G, A.A, C.C, A.C, and C.U). Mismatch pairings may occasionally be also found in the stems of some tRNAs. The distribution of nearly 250 of these non-canonical pairs was compiled by Grosjean et al. [19]. The ribose phosphate backbone interacts with some bases and it also interacts with another region of backbone itself. In these interactions the 2'-OH group of ribose ring may act as a hydrogen bond donor as well as an acceptor. In addition most of the bases are found to be stacked. These hydrophobic interactions between adjacent aromatic rings play major role in stabilizing the three-dimensional structure of tRNA.

1.1.4 Functional Importance of tRNA Molecule:

As its name indicates, its function in protein synthesis is to transfer the amino acids from free solution into its proper position in the polypeptide chain. Transfer RNA molecule is multifunctional key cellular molecule, which is widely known for its involvement in genetic decoding during protein biosynthesis [20-22]. It translates the nucleotide sequence in mRNA to amino acids sequence of protein. Thus tRNA molecule plays a central role in gene expression. Some of the tRNAs also participate in a variety of other functions in cellular metabolism such as cell wall biosynthesis,

chlorophyll and heme biosynthesis, as primers for retroviral reverse transcriptase, including the HIV (Human Immunodeficiency Virus) enzyme, amino acid transformations such as conversion of glutamic acid to glutamate semialdehyde. The conversion of amino acid serine to selenocystein, (also known as 21st amino acid), is made on the tRNA and is then inserted into proteins at specific sites [23]. Inspite of these various functions, predominant role of tRNA as a connecting link between the nucleic acid language of the genetic code and the amino acid language of the working cell is widely known. Some modified and hypermodified nucleosides usually occur at various positions in tRNA, these can modulate various cellular and metabolic processes. For its central role in decoding and multiple functions transfer RNA remains a major subject of research in molecular structural biology and biophysics.

1.2 Modified Nucleosides in tRNA

1.2.1 Historical Background and Introduction:

Transfer RNA is the most extensively modified nucleic acid in the cell, and immediately after its discovery it was shown that it contains modified nucleosides [24, 25]. First time in 1948 modified nucleosides were found in nucleic acids [26]. Modified nucleosides must have evolved early, like tRNAs, which were present in the progenitor, contained modified nucleosides. Several attempts were made during the 1960s and 1970s, to know the function of one class of modified nucleosides, the methylated nucleosides, but very little progress were made. The methods used were not sensitive or specific for revealing the specific role of modified nucleosides at that time. In the period of 1970s there was very little understanding about the role of modified nucleosides in tRNA, mainly because of difficulties in studying their functions. Now, significance of modified nucleosides in tRNA is being widely recognized.

RNA molecules are made up from four common nucleosides called adenosine (A), guanosine (G), cytidine (C) and uridine (U). Instead of these common nucleosides some modified (unusual) nucleosides also occur at various positions in tRNA. So far nearly 100 modified nucleosides have been found in nucleic acids, most of these occur in tRNA. All of these are the derivatives of common bases or are ribose modified. Most of the modifications involve simple alkylation, hydrogenation, thiolation or isomerization of these four common ribonucleosides, in the base and the 2'-hydroxyl group of the ribose.

However, some modifications involve complex chemical modifications, which are characterized by the presence of the diverse functional groups in the base substituents, such tRNA components are referred as hypermodified nucleosides. Hypermodified nucleosides occur in tRNA at various positions. Its frequency is more in tRNA anticodon loop. Anticodon loop is very dynamic part of tRNA molecule. Hypermodified nucleosides occur frequently at the 'wobble' (34^{th}) position and anticodon 3'-adjacent (37^{th}) position in tRNA anticodon loop. Except for queuosine (Q) and inosine (I) [27-28] formation, the synthesis of most modified nucleosides occurs at the polynucleotide level after the transcription of tRNA genes.

1.2.2 Modified Nucleosides in tRNA From Different Organisms:

Modified nucleosides are present in various organisms belonging to all three kingdoms [29], Archaea, Bacteria, and Eucarya (which were formerly called the kingdoms of archaebacteria, eubacteria, and eukaryotes, respectively) [30]. A few of these (D, Ψ, Cm, m^1G, Um, ac^4C, Gm, t^6A, m^1A, and I) are common in tRNAs from all three kingdoms [31], and some modified nucleosides (Ψ13, Cm32, m^1G37, t^6A37, Ψ38, Ψ39, Ψ55, and m^1A58) are also present at equivalent (comparable) tRNA locations in all three kingdoms. This indicates the common evolutionary origin [32]. Some of the modified nucleosides are specific for Archaea, Bacteria, or Eucarya, and indicate that these modified nucleosides might have evolved after the three kingdoms divided [31-33]. Largest variety and number of modified nucleosides are found in eukaryotic tRNAs.

There are also some specific modifications presents in the other two kingdoms, i.e. mnm^5s^2U for Bacteria and $m^1Ψ$ for tRNA from Archaea [32].

1.2.3 Nomenclature of Modified Nucleosides:

The abbreviations as well as name of modified nucleosides are listed on starting page. An index and an exponent indicates the number and position of the substitution, respectively, e.g., 6-dimethyladenosine is abbreviated m_2^6A. An abbreviation to the left of the nucleoside symbol denotes modification of the base, whereas a symbol to the right of the nucleoside symbol denotes modification of the ribose. Abbreviations for the groups are as follows: c, carbon; i, isopentenyl; k, lysine; m, methyl; o, oxy; s, thio; t, threonine; g, glycine.

1.2.4 Enzymatic Synthesis of Modified Nucleosides and Recognition:

Modified nucleosides, naturally occur in tRNA and are synthesized by various cellular enzymes. The tRNA-modifying enzymes, which catalyze conversion to these modified nucleosides, are sensitive to different sequence – structural identity features in tRNA molecules. Each highly specific tRNA modifying enzyme has different requirements to recognize its target nucleotide. For some tRNA modifying enzymes, sequence surrounding the target nucleotide is important and for others recognition of overall three-dimensional structure is essential.

1.2.4.1 Enzymatic Synthesis of i^6A and Related Derivatives in the Cell :

The conversion of A_{37} to isopentenylated, i^6A_{37} is done by E.coli. isopentenyl-tRNA: A_{37} transferase (35kDa protein), also known as tRNA (i^6A_{37}) synthase. The Δ^2-isopentenyl pyrophosphate: tRNA- Δ^2-isopentenyl transferase, is abbreviated as IPTT, and is also referred as tRNA prenyl transferase or dimethylallyl diphosphate: tRNA dimethylallyltransferase, DMAPP-tRNA transferase. Anna Golovko et al., (2000), discussed the cloning of human tRNA isopentenyl transferase [34]. The enzyme responsible for the formation of i^6A, was shown to catalyze the transfer of Δ^2-isopentenyl group of Δ^2-isopentenyl pyrophosphate, derived from mevalonic acid, to the amino group on C(6) of adenosine-37 [35]. The modified nucleoside N^6- (Δ^2-isopentenyl)adenosine, i^6A is found in position 37 at 3' adjacent position of anticodon loop in almost all eukaryotic and bacterial tRNAs that read codons starting from U [36]. This enzyme is very specific, which can recognize the sequence $A_{36}A_{37}A_{38}$ as well as anticodon stem with five base pairs [37]. It has also been shown that, instead of a specific anticodon stem and loop sequence, the key feature required for the recognition of E. Coli tRNAs by IPTT is the $A_{36}A_{37}A_{38}$ sequence occurring within the seven-membered anticodon loop [38]. There must be a possibility that IPTT binding induces some conformational changes in the anticodon loop and stem. Such changes have been found in some crystal structures of tRNAs with their cognate aminoacyl-tRNA synthatases [39, 40]. Likewise, the E. Coli isopentenyl-tRNA: A_{37} transferase may also change anticodon loop and anticodon – proximal stem to increase contacts with the active sites.

The base pair at positions 29-41 and 30-40 in the anticodon stem are most important ones for the efficient isopentenylation of A_{37}. There is a preference for purines at positions 29 and 30 (5'-side of the anticodon stem) and pyrimidines at

positions 40 and 41 (3'-side of the anticodon stem). If the inversion of these two base pairs occurs completely then no conversion takes place of A_{37} to i^6A_{37}. Even inversion of only one of the two base pairs considerably reduces the formation of i^6A_{37}. At positions 31-39, base pairs U-A or A-U or G-C serve equally well. In the five-base paired anticodon stem, top two positions namely 27-43 and 28-42 may not participate in the recognition of tRNA by IPTT because even the mismatch G*U base pair is tolerated in E.Coli $tRNA^{Cys}$ [35].

There are some other modified nucleoside derived from N^6-(Δ^2-isopentenyl)adenosine, i^6A_{37}. In eubacteria, like (E. Coli, Proreus, and Bacillus) and chloroplasts, N^6-(Δ^2-isopentenyl)adenosine, i^6A in tRNAs is thiomethylated to 2-methylthio derivative (ms^2i^6A), where as it is known that under some physiological conditions, ms^2i^6A in tRNAs from plants and other eubacteria (Salmonella, Pseudomonas, or Klebsiella) is furthur hydroxylated to ms^2io^6A [41, 42]. The production of these i^6A derivatives requires, miaB, miaC and miaE genes, respectively [43, 44].

1.3 Modifications at 37th Position

Anticodon loop of tRNA molecule contains most of the modified nucleosides. Nucleotide triplet sequence forming a codon in mRNA is recognized by corresponding tRNA anticodon sequence present in anticodon loop of tRNA during protein biosynthesis. One of the likely function of modified nucleosides is to enable smooth protein biosynthesis. This functional role may be played by modifications present at, anticodon 3'-adjacent 37th (position) of tRNA. These modifications may help in optimizing the codon - anticodon interaction for allowing nearly equal residence times, irrespective of codon-anticodon sequence varations [45-46].

Anticodon 3'-adjacent (37th position) nucleosides have hydrophobic as well as hydrophilic substituents [43, 47]. Examples of hydrophilic substituted bases are N^6-(N-threonylcarbonyl)adenine, t^6Ade, 2-methylthio-N^6-(N-threonylcarbonyl)adenine, ms^2t^6Ade, N^6-(N-glycylcarbonyl)adenine, gc^6Ade, and related derivatives. Examples of hydrophobic substituted bases are N^6-(Δ^2 -isopentenyl)adenine, i^6Ade, 2-methylthio-N^6-(Δ^2-isopentenyl)adenine, ms^2i^6Ade, and related derivatives like, N^6-(Δ^2-*cis*-hydroxyisopentenyl)adenine, *cis*-io^6Ade (popularly known as *cis*-zeatin) and 2-methylthio N^6-(Δ^2-*cis*-hydroxyisopentenyl)adenine, *cis*-ms^2io^6Ade (also known as *cis*- ms^2zeatin). These i^6Ade analogs occur at anticodon 3'-adjacent (37th) position in

tRNAs responding to codons starting with 5'-uridine [48]. Likewise, the presence of *trans*-zeatin and *trans*-ms^2zeatin in tRNA from plants and plant bacteria has also been reported [49, 50]. Under aerobic conditions characterized by the presence of free molecular oxygen, in Salmonella typhimurium tRNA hydroxylation of ms^2i^6Ade to ms^2io^6Ade occurs [42]. In Salmonella typhimurium the ms^2io^6Ade modification of tRNA regulates growth on citric acid cycle intermediates [51]. The change in tRNA modification with aerobiosis may serve to turn off the anaerobic metabolism pathways and switch on to aerobic ones [52].

1.3.1 N^6-(Δ^2-isopentenyl)adenine, i^6Ade, and 2-methylthio-N^6-(Δ^2-isopentenyl)adenine, ms^2i^6Ade:

Hypermodified nucleosides with hydrophobic 'isopentenyl' substituent, N^6-(Δ^2 -isopentenyl)adenine, i^6Ade and its 2-methylthio derivative, ms^2i^6Ade are found in position 37 of tRNAs that read codons starting with U. The hypermodified nucleoside $i^6A_{(37)}$ is interesting because as free base isopentenyl adenine is cytokinin active [53] and promotes cell elongation and stimulates growth-differentiation in plant tissue culture. It is component of tRNAs of all organisms, it is useful for protein synthesis, it exhibits a cytokinin-like function in animal cells in culture, i^6A has anti-tumor activity in transformed cells [54], animals [55], and man [56]. Persson et al., in 1994 showed that cytokinins are found in tRNA of plants, animals, and eubacteria but not in archaea [43]. The modification level of isopentenyl adenine and related derivatives in tRNA depends upon the growth and nutritional status. Under Fe^{+++} or cysteine deficiency 2-methylthiolation is restricted and only partial modification, i^6Ade, occurs instead of ms^2i^6Ade, ms^2io^6Ade or io^6Ade. This partial modification is implicated in de-repression of aromatic amino acid biosynthetic pathways and synthesis of high affinity iron chelating enterobactin [57].

The conformational preferences of N^6-(Δ^2 -isopentenyl)adenine, i^6Ade, and 2-methylthio-N^6-(Δ^2 -isopentenyl)adenine, ms^2i^6Ade were studied theoretically by quantum chemical PCILO method [58]. The orientation of N(6) substituents in i^6Ade and ms^2i^6Ade, is found to be 'distal' i.e. (the N(6) substituents spreads away from the five membered imidazole moiety of the adenine ring). As consequence of this orientation the N(6)H becomes inaccessible for participation in the usual Watson-Crick hydrogen bonding for base pairing. Thus the base modification 3'-adjacent to the anticodon restricts the extended base pairing between the anticodon and codon and

thus can help define the proper reading frame for codon-anticodon interactions. Theoretically obtained results are similar to crystal structure observations [59, 60].

The bulky hydrophobic modifications are necessary to strengthen the week A36-U interaction. The 2-methylthio group is a major factor in the stabilization of anticodon - anticodon interactions of $ms^2i^6A_{37}$ containing tRNAs. The $ms^2i^6A_{37}$ enhances the efficiency of the decoding tRNA and makes it less sensitive to different codon context [61-66]. It has also been shown that $ms^2i^6A_{37}$ affects codon-anticodon interactions, when A_{36} is properly matched with uridine at the first position, then $ms^2i^6A_{37}$ stabilizes the codon-anticodon interactions. In contrast, if there is mismatch in the first codon position, then $ms^2i^6A_{37}$ may lead to destabilization of codon-anticodon interactions [61]. It was suggested that the $ms^2i^6A_{37}$ modification at position 37 of a tRNA, making a A36-U base pair with the first base of the codon, would stabilize the base pairing and inhibit the wobble capacity, thus reducing the first position misreading [67, 68]. An *in vitro* experiment supports this hypothesis, in which it was shown that the presence of the $ms^2i^6A_{37}$ modification reduces first position error [69].

The $ms^2i^6A_{37}$ enhances the activity of the tRNA, makes it less sensitive to the codon context, and increases fidelity. Thus lack of $ms^2i^6A_{37}$ makes influence on the physiology of the bacterial cell and plays a fundamental role in the efficiency and fidelity of translation [43].

1.3.2 Zeatin (io^6Ade) and ms^2-Zeatin (ms^2io^6Ade):

Zeatin, an adenine derivative first found in maize, is the most active form of the naturally occuring cytokinins [70-71]. Hydroxylated derivative of N^6-(Δ^2-isopentenyl)adenosine, i^6A has been observed [53] in immature sweet corn kernels, and identified as 6-(*cis*-4-hydroxy-3-methylbut-2-enylamino)-9-β-D-ribofuranosylpurine, (*cis*-ribosylzeatin). Likewise *trans*-ribosylzeatin has been (Letham, et. al., 1964) isolated from an extract of immature corn kernels a purine base and identified as 6-(*trans*-4-hydroxy-3-methylbut-2-enylamino) purine [72]. This is further confirmed by Shaw and Wilson, 1964) [73].

N^6-(Δ^2-isopentenyl)adenine, i^6Ade and its 2-methylthio derivative, ms^2i^6Ade get converted into hydroxylated forms under aerobic condition. These are N^6-(Δ^2-*cis*-hydroxyisopentenyl)adenine, *cis*-io^6Ade, also known as *cis*-zeatin, N^6-(Δ^2-*trans*-hydroxyisopentenyl)adenine, *trans*-io^6Ade, also known as *trans*-zeatin, 2-methylthio-

N^6-(*cis*-hydroxyisopentenyl)adenine, *cis*-ms^2io^6Ade or *cis*-ms^2zeatin and 2-methylthio-N(6)-(*trans*-hydroxyisopentenyl)adenine, *trans*-ms^2io^6Ade or *trans*-ms^2zeatin. The hydroxylation of the isopentenyl side chain of ms^2i^6Ade to form ms^2io^6Ade, involve a heme containing oxygenase activity that requires molecular oxygen. This hydroxylation could influence the translational capacity of the tRNA, which may lead to an altered regulation of gene expression [42].

The ribosylzeatin (io^6A) or its 2-methylthio derivative (ms^2io^6A) are found in the tRNA of plants such as garden pea, sweet corn, spinach, tobacco callus, and wheat germ [53, 74-76]. Therefore it has been suggested that only plants have the gene for the hydroxylating enzyme, which converts i^6A and ms^2i^6A to the respective hydroxylated derivative io^6A and ms^2io^6A [77]. Bacteria such as *Rhizobium leguminosarum, Agrobacterium tumefaciens,* and *Corynebacterium facians* have been shown to contain io^6A or ms^2io^6A in their tRNA [49, 78-79]. It has also been reported that bacteria such as *Salmonella typhimurium, Klebsiella pneumoniae,* and *Serratia marcescens*, which are not associated with plants, also contain ms^2io^6A in their tRNAs [80, 81]. It is not present in *Clostridium perfringens*, which is a strict anaerobe.

In *Salmonella typhimurium*, Buck and Ames (1984) showed that during aerobiosis, ms^2i^6A gets converted into its hydroxylated form, *cis*-ms^2io^6A [42]. In aerobically grown bacteria, these tRNAs contain ms^2io^6A and in anaerobically grown bacteria have its precursor, ms^2i^6A. The hydroxylation of the isopentenyl (i^6A) side chain is not possible in absence of oxygen. Ajitkumar and Cherayil, (1985) observed the *trans*-ms^2io^6A, 2-methylthio-N^6- (4-hydroxy-3-methylbut-2-enyl) or *trans*-ms^2zeatin, in the tRNA of *Azotobacter vinelandii*, a free-living, nitrogen-fixing bacterium found in the inner rhizosphere of plant roots [50].

In (1973), the Scarbrough et al. showed that the plant pathogenic prokaryote *Corynebacterium fascians*, relase the *cis*-zeatin (nonthiolated) into the culture medium [82]. However, Chapman et al. (1976) reported the presence of, *trans*-io^6A or *trans*-ribosylzeatin (nonthiolated) in the tRNA of *Agrobacterium tumefaciens* [49]. The role of $ms^2io^6A_{37}$ modification is primarily in the codon-anticodon interaction. Ericson and Bjork [66] proposed a model where the modified nucleoside is thought to increase the stacking ability of the anticodon-codon complex. While, Bouadloun et. al. [64], proved that $ms^2io^6A_{37}$ reduces the codon-context sensitivity.

Experimental work has been done to study the stereochemistry of the zeatin isomers. Letham et al. (1964) determined by mass and p.m.r. spectra the structure of zeatin, a factor inducing cell division [72]. Shaw and Wilson, 1964 carried out the synthesis of *trans*-zeatin [73], this sample was compared with zeatin [72] and found identical by usual physical measurements, spectrometry and chromatography. The stereoselective synthesis of ribosyl-*cis*-zeatin and ribosyl-*trans*-zeatn is reported [83, 84]. They developed a simple thin layer chromatography system for the separation of the *cis* and *trans* isomers, and utilized the thin layer chromatographic method to identify the stereochemistry of the ribosylzeatins obtained from the tRNA of several plant sources. Shaw et al., 1966, described the relationship between synthetic *cis*- and synthetic *trans*-ribonucleosides as geometrical isomers by their NMR spectra [85]. The crystallographic study of *cis*-zeatin, conformer 1 and conformer 2 [86] has been carried out by (Soriano-Garcia et al. 1987).

1.3.3. N^6-(N-glycylcarbonyl)adenine, gc^6Ade:

The modified nucleic acid base N^6-(N-glycylcarbonyl)adenine, gc^6Ade, is an analogue of N^6-(N-threonylcarbonyl)adenine, (tc^6Ade). It has been isolated [87] from enzyme digests of unfractionated yeast tRNA. Extensive modifications of adenine, containing amino acid substituents like threonine (tc^6Ade) or glycine (gc^6Ade) are linked through the ureido HN-CO-NH linkage to the adenine ring and occur [22, 88] in those tRNAs which respond to codons starting with A (anticodon ending in U).

In view of the common ureido HN-CO-NH linkage in gc^6Ade as well as in tc^6Ade, a broad similarity is expected [89-91] between the structural and the functional significance of such modifications. The orientation of N(6) substituents in glycylcarbonyl adenine (gc^6Ade) is found to be 'distal' (i.e. the N(6) substituents spreads away from the five membered imidazole moiety of the adenine ring). Thus the N(6)H is not available for participation in the usual Watson-Crick base pairing. Thus the base modification 3'-adjacent to the anticodon restricts the extended base pairing between the anticodon and codon and thus can help define the proper reading frame for codon -anticodon interactions.

It is possible to simulate the influence of hydrogen bond donor - acceptor interactions as in intermolecular interactions in crystal packing [92], through protonation [93]. A change in protonation status may also spur conformational changes required for the molecular function.

In nucleoside adenosine, N(1), N(3), and N(7) are the possible sites of protonation. The alkylation studies of nucleic acids and constituents [94] has shown that the reactivity order of the N(1), N(3), and N(7) sites of adenine can be quite different depending upon the molecular environment and association of these molecules in solutions.

1.4 Modifications at 34th Position

Hypermodified nucleosides also occur at "wobble" (34th) position, in anticodon loop of tRNAs. The diverse modifications occurring at the first position of the anticodon do not hinder canonical Watson-Crick base pairing, but can restrict or enlarge the scope of wobble base pairing for reading the third base of the codon. Examples of modified nucleosides at 34th positions are 5-carbamoylmethyl uridine (ncm^{5}U) and 4-amino-2- (N^6-lysino)-1-(β-D-ribofuranosyl) pyrimidinium (Lysidine) designated as k^2C.

1.4.1 5-Carbamoylmethyl Uridine (ncm^{5}U):

The hypermodified nucleoside 5-carbamoylmethyl uridine (ncm^{5}U) is a modification of uridine, which has side chain at the C(5) position of the uridine ring. The novel modified nucleoside occurs at the 34th position of the anticodon. Yamamota et al. (1985) were first to isolate and characterize its chemical structure as 5-carbamoylmethyl uridine (ncm^{5}U) from yeast $tRNA^{val}_{2a}$ [95] by ultraviolet absorption spectroscopy, mass spectroscopy, and nuclear magnetic resonance spectroscopy (NMR).

On a two-dimensional thin-layer chromatogram of Yeast $tRNA^{ser}$(UCA), the same nucleoside was identified [96]. Likewise, an unknown modified nucleoside located in the first position of $tRNA^{pro}$ from *T. utilis* anticodon was identified as 5-carbamoylmethyl uridine [97]. Dunn and Trigg, (1975) also isolated this nucleoside from crude tRNAs of baker's yeast and wheat embryo [98]. It is noteworthy that this nucleoside was found in Yeast $tRNA^{Val}$ and $tRNA^{Ser}$ which recognize UC series of codons [96] and also in *T. utilis* $tRNA^{Pro}$ [97], all of these correspond to amino acids with four or six codon families.

Sequence analysis study of compiled tRNA database [99] revealed that, in $tRNA^{Ser}$(UGA) from Nicotiana rustica [100], the wobble nucleoside 5-carbamoylmethyl uridine (ncm^{5}U) occurs at 34th position along with io^{6}A being

present at 37th position. The primary structure of anticodon loop of $tRNA^{Pro}$ with (ncm^5U) at 34th position was reported using TLC, HPLC, UV, and GCMS [101].

The effect of base modification especially uridine at 34th position could be to regulate the flexibility or rigidity of the anticodon so as to contribute to the correct and efficient translation. Ogawa et al., (1982) reported that *T. utilis* $tRNA^{Pro}$ containing 5-carbamoylmethyl uridine recognizes the codons CCU, CCC, and CCG in the E. Coli ribosomal binding system [102].

1.4.2 Lysidine (k^2C):

The hypermodified nucleoside lysidine, occurs at 34th position in anticodon loop of tRNA. This nucleoside is derived from cytidine by substituting the oxygen atom in position 2 with ε-nitrogen atom of L-lysine. In *E. Coli*, two isoleucine tRNA species are observed, $tRNA^{Ile}_{major}$ (or $tRNA^{Ile}_1$) can recognize the AUU and AUC codons [103], while $tRNA^{Ile}_{minor}$ (or $tRNA^{Ile}_2$) recognizes the codon AUA only [104]. Harada and Nishimura, in 1974 observed that, the minor species of E. Coli $tRNA^{Ile}_2$ has a novel hypermodified nucleoside N^+ with unknown chemical structure [104]. Gupta, R. (1984), has also shown the presence of an unidentified nucleoside in the first position of the anticodon of $tRNA^{Ile}_2$ from an archaebacterium, and *Halobacterium volcanii* [105]. Similarly $tRNA^{Ile}_1$ from Spinach chloroplast also has an unidentified modified nucleoside in the first position of the anticodon [106].

The primary structure of $tRNA^{Ile}_{minor}$ was determined by Kuchino et al. in (1980), but the chemical structure of nucleoside N^+ in the first position of anticodon [107] still remained unknown. The chemical structure of this novel hypermodified nucleoside was determined by NMR spectroscopy, mass spectroscopy and chemical synthesis, and the nucleoside N^+ was found to be 4-amino-2- (N^6-lysino)-1-(β-D-ribofuranosyl) pyrimidinium (with one positive charge) [108]. This novel type of modified cytidine is present at the 34th position of anticodon loop with a lysine side chain and is called "Lysidine", denoted as (k^2C).

The nucleoside lysidine has one positive charge in the pH range 3.5-8.5. The molecule can also exist in the protonated forms like zwitterion, neutral, and tautomer form of lysidine in the physiological condition. The tautomeric form of lysidine is a reasonable structure in the first position of the anticodon of $tRNA^{Ile}_2$, which recognizes the isoleucine codon AUA rather than AUU or AUC [104]. Modification of cytidine to lysidine involves condensation of lysine to the pyrimidine ring of

cytidine. Probably, the nucleoside lysidine has an L-lysine moiety rather than a D-lysine moiety.

Some other organisms with similar isoleucine tRNA species have been found. The genome of bacteriophage T4 codes for eight tRNA species including $tRNA^{Ile}$, which is specific only to the codon AUA [109]. Database compilation and tRNA sequence analysis [99] revealed that $tRNA^{Ile}$ from *Mycoplasma capricolum* contains lysidine at 34^{th} position and N^6-methyl adenosine (m^6A) at 37^{th} position in the anticodon loop, which recognizes the AUA [110] codon.

The lysidine modification prevents the mistranslation of AUG as isoleucine and that of AUA as methionine. This modification of C to k^2C (i.e. lysidine) also changes the aminoacylation identity of the tRNA from methionine to isoleucine [111]. Lysidine can recognize A and prevent misrecognition of G. It can switch the methionine codon AUG to the isoleucine codon AUA. The codon and amino acid specificity of tRNA can be converted by a single post-transcriptional modification of the first position, from cytidine to lysidine, this means:

Kind of Modification	**tRNA charged with amino acid**
tRNA with C(34)	Methionine (AUG)
tRNA with k^2C(34)	Isoleucine (AUA)

The reported results in this thesis show that, long lysine side chain may fold back towards the ribose ring, to allow the carboxyl group (or the amino group) of the amino acid substituent to form hydrogen bond with the ribose 2'-hydroxyl group.

1.5 Functional Importance of Modified Nucleosides

Modified nucleosides are present in tRNA from all organisms [32]. All of these may play some important role in tRNA and any of its multiple functions. The presence of modified nucleosides may improve the efficiency of the tRNA during protein biosynthesis as well as sensitivity to codon context. Some of the modified nucleosides maintain reading frame of mRNA. Therefore, the presence of different modified nucleosides is likely to have different impacts on the activity of tRNA. Due to its importance in the decoding steps, the level of tRNA modification also influences the expression of several operons through an inefficient reading of attenuator region of some operon transcripts. The attenuation process by which tRNA influences transcription is one mechanism by which undermodified tRNAs can affect regulation through a translational-transcriptional coupling [112]. Thus, the lack of a modified

nucleoside may induce some effects on cell physiology. Metabolic as well as genetic links/correlations exist between the synthesis of modified nucleosides in tRNA and intermediary metabolism, development, and cell cycle. Therefore, the level of tRNA modification has been suggested to be a regulatory device. Modified nucleosides also play important role in tumor formation and their use as diagnostic tools [113]. Studies with the modification or removal of the base adjacent to the anticodon suggest that this hypermodified nucleoside is needed for the binding of tRNA to ribosome-mRNA complex, but it is not essential to the charging of the tRNA.

Gallo et al. [114] discussed that i^6A inhibits mitosis of human lymphocytes at a higher concentration (10^{-6} M) and stimulates it at lower concentrations (10^{-7} M), a phenomena termed 'cytokinesis' in plants. The major form in E. Coli, 2-methylthio isopentenyl adenine, ms^2i^6Ade, plays an important role in translational efficiency and fidelity [43]. The cytokinins are the important class of plant growth hormones. One of the interesting thing is that i^6A may participate in triggering DNA replication during the S phase of the cell cycle in hamster cells [115].

1.6 Scope and Objectives of the Thesis:

The reported literature survey gives an overall idea about tRNA with its usual and modified nucleosides. The hypermodified components of our interest (in this thesis) at 37^{th} positions are i^6A, ms^2i^6A, io^6Ade, ms^2io^6Ade, gc^6Ade. The hypermodified nucleosides at 'wobble' (34^{th}) positions are ncm^5U and lysidine. Modified nucleosides may modulate structure, multiple functions of tRNA and accuracy-efficiency of protein biosynthesis. Many studies have been carried out on the isolation and function of modified nucleoside, but very little is known about the structural and conformational preferences of these modified nucleosides or structural consequences of their presence in tRNA. The aim of the present thesis is to report the investigations made on conformational and structural preferences of these modified nucleosides by using molecular modeling, conformational energy calculations employing various theoretical methods and computer simulation approaches. The goals of present study are -

1. To predict the conformational preferences of hdroxyisopentenyl adenine and its derivatives which occur at anticodon 3'- adjacent (37^{th}) position in anticodon loop of some specific tRNA.
2. To study the interaction of the modified nucleoside with other nucleosides present

in anticodon loop and with the 34th modification.

3. To analyze the tRNA database for identifying prevailing sequence patterns.
4. To study the structural consequence of introducing modified nucleosides in anticodon loop.
5. To simulate the aqueous solvation of anticodon arm with 37th and 34th modifications present in tRNA anticodon loop.

1.7 References:

1. Allen, E. H. and Schweet, R. S. Biochim. Biophys. Acta. 1960, 39, 185-187.
2. Barnett, W. E. and Brown, D. H. Proc. Natl. Acad. Sci. (Wash) 1967, 57, 452- 458.
3. Barnett, W. E. and Brown, D. H. Proc. Natl. Acad. Sci. (Wash) 1967, 57, 1775-1781.
4. Crick, F.H.C. Symp. Soc. Exptl. Biol. 1958, 12, 138.
5. Hoagland, M. B.; Stephenson, M. L.; Scott, J. F.; Hecht, L. I. and Zamecnik, P.C. J. Biol. Chem. 1958, 231, 241- 257.
6. Hoagland, M. B.; Zamecnik, P. C. and Stephenson, M. L. Biochim. Biophys. Acta. 1957, 24, 215-216.
7. Ogata, K. and Nohara, H. Biochim. Biophys. Acta. 1957, 25, 659-660.
8. Holley, R. W.; Apgar, J.; Everett, G. A.; Madison, J. T.; Marquisee, M.; Merrill, S. H.; Penswick, J. R. and Zamir, A. Science 1965, 147, 1462-1465.
9. Sanger, F.; Brownlee, G. G. and Barrell, B. G. J. Mol. Biol. 1965, 13, 373-398.
10. Clark, B. F. C.; Doctor, B. P.; Holmes, K. C.; Klug, A.; Marcker, K. A.; Morris, S. J. and Paradies, H. H. Nature (London), 1968, 219, 1222-1224.
11. Cramer, F.; Von der Haar, F.; Saenger, W. and Schlimme. E. Angew. Chem. 1968, 80, 969-970.
12. Fresco, J. R.; Blake, R. D. and Langridge, R. Nature (London) 1968, 220, 1285-1287.
13. Hampel, A.; Labanauskas, M.; Connors, P. G.; Kirkegard, L.; RajBhandary, U. L.; Sigler, P. B. and Bock, R. M. Science 1968, 162, 1384-1387.
14. Kim, S. H. and Rich, A. Science 1968, 162, 1381-1384.
15. Robertus, J. D.; Ladner, J. E.; Rhodes, J. T. F. D.; Brown, R. S.; Clark, B. F. C. and Klug, A. Nature (London) 1974, 250, 546-551.
16. Suddath, F. L.; Quigley, G. J.; McPherson, A.; Sneden, D.; Kim, J. J.; Kim, S.

H.and Rich, A. Nature (London) 1974, 248, 20-24.
17. Shi, H.; Moore, P. B. RNA 2000, 6, 1091-1105.
18. Jovine, L.; Djordjevic, S.; Rhodes, D. J. Mol. Biol. 2000, 301, 401- 414.
19. Grosjean, H.; Cedergreen, R. J. and Mckay, M. Biochimie, 1982, 64, 387-397.
20. McCloskey, J. A. and Nishimura, S. Accounts Chem. Res. 1977, 10, 403-410.
21. Feldmann, M.Y. Prog. Biophys. Mol. Biol. 1977, 32, 83-102.
22. Bjork, G.R.; Ericson, J. U.; Gustafsson, C. E. D.; Hagervall, T. G.; Jonsson Y.H. and Wikstrom, P.M. Ann. Rev. Biochem. 1987, 56, 263-287.
23. RajBhandary, U. L. and Soll, D. tRNA: Structure, Biosynthesis, and Function., Chapter -1, 1, ASM press, (1995).
24. Dunn, D. B. Biochim. Biophys. Acta. 1959, 34, 286–288.
25. Smith, J. D. and Dunn, D. B. Biochem. J. 1959, 72, 294-301.
26. Hotchkiss, R. D. J. Biol. Chem. 1948, 175, 315-332.
27. Hankins, W. D.; Farkas, W. R. Biochim. Biophys. Acta. 1970, 213, 77–89.
28. Elliott, M. S.; Trewyn, R. W. J. Biol. Chem. 1984, 259, 2407–2410.
29. Steinberg, S.; Misch, A. and Sprinzl, M. Nucleic Acids Res. 1993, 21, 3011-3015.
30. Woese, C. R.; Kandler, O. and Wheelis, M. L. Proc. Natl. Acad. Sci. (USA) 1990, 87, 4576-4579.
31. Edmonds, C. G.; Crain, P. F.; Gupta, R.; Hashizume, T.; Hocart, C. H.; Kowalak, J. A.; Pomerantz, S. C.; Stetter, K. O. and McCloskey, J. A. J. Bacteriol. 1991, 173, 3138-3148.
32. Bjork, G. R. Chemica Scripta. 1986, 26B, 91-95.
33. McCloskey. J. A. Appl. Microbiol. 1986, 7, 246-252.
34. Golovko, A.; Hjalm, G.; Sitbon, F.; Nicander, B. Gene. 2000, 258, 85-93.
35. Motorin, Y.; Beck, G.; Tewari, R. and Grosjean, H.; RNA. 1997, 3, 721-733.
36. Sprinzl, M.; Steegborn, C.; Hubel, F.; Steinberg, S. Nucleic Acid Res. 1996, 24, 68- 72.
37. Tsang, T. H.; Buck, M.; Ames, B. N. Biochim. Biophys. Acta. 1983, 741, 180-196.
38. Hall, R. H. Prog. Nucleic Acid Res. Mol. Biol. 1970, 10, 57-86.
39. Roulds, M. A.; Perona, J. J.; Soll, D.; Steitz, T. A. Science 1989, 246, 1135-1142.
40. Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A..; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. Science 1991, 252, 1682-1689.

41. Agris, P. F.; Armstrong, D. J.; Schafer, K. P.; Soll, D. Nucleic Acid Res. 1975, 2, 691- 698.
42. Buck, M.; Ames, B. N. Cell 1984, 36, 523-531.
43. Persson, B. C.; Esberg, B.; Olafsson, O.; Bjork, G. R. Biochimie 1994, 76, 1152-1160.
44. Bjork, G. R. Prog. Nucleic Acid Res. Mol. Biol. 1995b, 50, 263-338.
45. Watts, M. T.; Tinoco, I. Biochemistry 1978, 17, 2455-2463.
46. Grosjean, H.; deHenau, S.; Crothers, D. M. Proc. Natl. Acad. Sci. (USA) 1978, 75, 610-614.
47. Davis, D. R. in Modifications and Editing of tRNA, Grosjean, H.; Benne, R. Eds.; ASM Press: Washington, 1998, p. 85-102.
48. Buck, M.; Griffiths, E. Nucleic Acid Res. 1982, 10, 2609-2624.
49. Chapman, R. W.; Morris, R. O.; Zaerr, J. B. Nature 1976, 262, 153-154.
50. Ajitkumar, P.; Cherayil, J. D. J. Bacteriol. 1985, 162, 752-755.
51. Persson, B. C.; Olafsson, O.; Lundgren, H. K.; Hederstedt, L.; Bjork, G. R. J. Bacteriol. 1998, 180(12), 3144-3151.
52. Starzyk, R. TIBS 1984, 333-334.
53. Hall, R. H.; Csonka, L.; David, H; McLennan, B. Science, 1967, 156, 69-71.
54. Suk, D.; Simpson, C. L.; Mihich, E. Proc. Am. Assoc. Cancer Res. 1968, 9, 69.
55. Suk, D.; Simpson, C. L.; Mihich, E. Proc. Am. Assoc. Cancer Res. 1968, 8, 23.
56. Jones, R.; Jr., Grace, J. T., Jr., Mittelman, A. and Gerner, R. E. Proc. Am. Assoc. Cancer Res. 1968, 9, 35.
57. McCray, Jr. J. W.; Herrmann, K. M. J. Bacteriol. 1976, 125, 608-615.
58. Tewari, R. Int. J. Quantum Chem. 1988, 34, 133-142.
59. McCMullan, R. K. and Sundaralingam, M. J. Am. Chem. Soc. 1971, 93, 7050-7054.
60. Bugg, C. E. and Thewalt, U. Biochem. Biophys. Res. Commun. 1972, 46, 779-784.
61. Diaz, I. and Ehrenberg M. J. Mol. Biol. 1991, 222, 1161-1171.
62. Petrullo, L. A.; Gallagher, P. J. and Elseviers, D. Mol. Gen. Genet. 1983, 190, 289-294.
63. Vacher, J.; Grosjean, H.; Houssier, C. and Buckingham, R. H. J. Mol. Biol. 1984, 177, 329-342.

64. Bouadloun, F.; Srichaiyo, T.; Isaksson, L. A. and Bjork, G. R. J. Bacteriol. 1986, 166, 1022-1027.
65. Hagervall, T. G.; Ericson, J. U.; Esberg, K. B.; Ji-nong,L.; and Bjork, G. R. Biochim. Biophys. Acta. 1990, 1050, 263-266.
66. Ericson, J. U. and Bjork, G. R. J. Mol. Biol. 1991, 218, 509-516.
67. Nishimura, S. Prog. Nucleic Acid Res. Mol. Biol. 1972, 12, 49-85.
68. Jukes, T. H. Nature 1973, 246, 22-26.
69. Wilson, R. K.; Roe, B. A. Proc. Natl. Acad. Sci. USA. 1989, 86, 409-413.
70. Letham, D. S. Life Sci. 1963, 8, 569-573.
71. Shaw, G. *In* DWS Mok, MC Mok, eds, Cytokinins: Chemistry, Activity and Function. CRC Press, Boca Raton, FL, 1994, pp 15-34.
72. Letham, D. S.; Shannon, J. S. and McDonald, I. R. Procedings Chem. Soc. 1964, 230-231.
73. Shaw, G. and Wilson D. V. Procedings Chem. Soc. 1964, 231.
74. Barrows, W. J.; Armstrong, D. J.; Kaminek. M.; Skoog, F.; Bock, R.M.; Hetch, S. M.; Dammann, L. G. Leonard, N. J. and Occolowitz. Biochemistry 1970, 9,1867-1872.
75. Barrows, W. J.; Skoog, F.; Leonard, N. J. Biochemistry 1971, 10, 2189-2194.
76. Verman, H. J.; Skoog, F.; Frihart C. R. and Leonard, N. J. Plant Physiology 1972, 49, 848-851.
77. Swaminathan S.; Bock, R. M. and Skoog, F. Plant Physiology 1977, 59, 558-563.
78. Cherayil, J. D. and Lipsett, M. N. J. Bacteriol. 1977, 131, 741-744.
79. Matsubara, S.; Armstrong, D. J. and Skoog, F. Plant Physiology 1968, 43, 451-453.
80. Buck, M.; McCloskey, J. A.; Basile, B. and Ames, B. N. Nucleic Acids. Res. 1982, 10, 5649-5662.
81. Janzer, J. J.; Raney, J. P. and McLennan, B. D. Nucleic Acids. Res. 1982, 10, 5663- 5672.
82. Scarbrough, E.; Armstrong, D. J.; Skoog, F.; Frihart C. R. and Leonard, N. J. Proc. Natl. Acad. Sci. USA 1973, 70, 3825-3829.
83. Playtis, A. J. and Leonard N. J. Biochem. Biophys. Res. Commun. 1971, 45, 1-5.
84. Leonard N. J.; Playtis, A. J.; Skoog, F. and Schmitz, R. Y. J. Am. Chem. Soc. 1971, 93, 3056-3058.

85. Shaw, G.; Smallwood, B. M. and Wilson, D. V. J. Chem. Soc. 1966, 921-924.
86. Soriano-Garcia, M.; Toscano, R. A.; Arroyo-Reyna, J. A. J Crystallographic Spectr. Res. 1987, 17, 221-230.
87. Schweizer, M. P.; McGrath, K and Baczynskyj, L. Biochem. Biophys. Res. Commun. 1970, 40, 1046-1052.
88. Adamiak, R. W and Gornicki. Prog. Nucleic Acid Res. Mol. Biol. 1985, 32, 27-74.
89. Tewari, R. Int. J. Quantum Chem. 1987, 31, 611- 624.
90. Tewari, R. Indian J. Biochem. Biophys. 1987, 24, 170-176.
91. Tewari, R. J. Biomol. Struct. Dyn. 1990, 8, 675-686.
92. Parthasarathy, R.; Ohrt, J. M and Chheda, G. B. Biochemistry 1977, 16, 4999-5008.
93. Tewari, R. Int. J. Quantum Chem. 1994, 51, 105-112.
94. Lawley, P. D.; Brooks, P. Biochem. J. 1964, 92, 19C-20C.
95. Yamamoto, N.; Yamaizumi, Z.; Yokoyama, S.; Miyazawa, T and Nishimura, S. J. Biochem. 1985, 97, 361-364.
96. Etcheverry, T.; Colby, D. and Guthrie, C. Cell 1979, 18, 11-26.
97. Ogawa, K.; Kondo, T.; Kawakami, M and Takemura, S. Nucl. Acids Res. Symp. Ser. 1983, 12, 131-132.
98. Dunn, D. B. and Trigg, M. D. M. Biochem. Soc. Trans. 1975, 3, 656-659.
99. Sprinzl, M.; Horn, C.; Brown M.; Ioudovitch, A.; Seinberg, S. Nucleic Acid Res. 1998, 26, 148-153.
100. Teichmann, T.; Urban, C.; Beier, H. Plant. Mol. Biol. 1994, 24, 889-901.
101. Kieth, G. Biochim. Biophys. Acta. 1990, 1049, 255-260.
102. Ogawa, K.; Kawakami, M.; Shimizu, Y. and Takemura, S. J. Biochem. 1982, 91, 1241-1248.
103. Yarus, M. and Barrell, B. G. Biochem. Biophys. Res. Commun. 1971, 43, 729-734.
104. Harada, F. and Nishimura, S. Biochemistry 1974, 13, 300-306.
105. Gupta, R. J. Biol. Chem. 1984, 259, 9461-9471.
106. Francis, M. A. and Dudock, B. S. J. Biol. Chem. 1982, 257, 11195-11198.
107. Kuchino, Y.; Watanabe, S.; Harada, F. and Nishimura, S. Biochemistry 1980, 19, 2085-2089.
108. Muramatsu, T.; Yokoyama, S.; Horie, N.; Matsuda, A.; Ueda, T.; Yamaizumi, Z.;

Kuchino, Y.; Nishimura, S. and Miyazawa, T. J. Biol. Chem. 1988, 263, 9261-9267.

109. Scherberg, N. H. and Weiss, S. B. Proc. Natl. Acad. Sci. USA 1972, 69, 1114-1118.
110. Andachi, T.; Yamao, F.; Muto, A. and Osawa, S. J. Mol. Biol. 1989, 209, 37-54.
111. Muramatsu, T.; Nishikawa, K; Nemoto, F.; Kuchino, Y.; Nishimura, S. and Miyazawa, T and Yokoyama, S. Nature (London) 1988, 336, 179-181.
112. Yanofsky, C. and Soll, D. J. Mol. Biol. 1977, 113, 663-377.
113. Nass, G. ed. Modified nucleosides and Cancer., Berlin, Springer-Verlag, 1983.
114. Gallo, R.C.; Wang-Peng, J.; Perry, S. Science 1969, 165, 400-404.
115. Quesney-Huneeus, V.; Glaick, H. A.; Siperstein, M. D.; Erickson, S. K.; Spencer, T. A. and Nelson, J. A. J. Biol. Chem. 1982, 258, 378-385.

CHAPTER-II

METHODS

2.1 Introduction

Theoretical methods have been utilized to study the conformational preferences and structural significance of hypermodified bases presents at 34^{th} and 37^{th} positions in anticodon loop of tRNA. Quantum mechanical methods and molecular mechanics force field based methods used in the investigations are described here.

2.2 Quantum Mechanical Methods

Quantum mechanics describes molecules in terms of interactions among nuclei and electrons, and molecular geometry in terms of minimum energy arrangements of nuclei. Quantum mechanics states that the energy and other related properties of a molecule may be obtained by solving the Schrodinger equation-

$$H\Psi=E\Psi$$

Where, H is Hamiltonian,

Ψ is a many-electron wavefunction and

E is total energy of many electron system.

Unfortunately, the many-electron Schrodinger equation can not be solved exactly even for the simplest many-electron systems. Approximations need to be introduced to provide practical methods.

2.2.1 *Ab initio* Methods:

2.2.1.1 Born-Oppenheimer Approximation:

One way to simplify the Schrodinger equation for molecular systems is to assume that the nuclei do not move. This is termed the Born-Oppenheimer approximation, and leads to an electronic Schrodinger equation.

2.2.1.2 LCAO Approximation:

It is simple and useful method of obtaining a wave function, which describe the molecular orbital. In LCAO method, atoms have atomic orbitals, which are described by the wavefunction. The two similar atoms 1 and 2 have atomic orbitals, which are described by the wavefunction. Atomic orbital of 1 has wavefunction $\phi 1$, atomic orbital of atom 2 has wavefunction $\phi 2$. When these atoms form a bond the electron originally in the atomic orbitals now occupy molecular orbitals. These molecular orbitals are formed by linear combination of atomic orbitals $\phi 1$ and $\phi 2$.

$$\Psi = C_1\phi_1 + C_2\phi_2$$

Where, ϕ_1, ϕ_2 = atomic orbitals

Ψ = Wavefunction (Molecular Orbital)

C_1, C_2 = Coefficient.

The idea can be extended to a polyatomic molecule and generalized so that each molecular orbital in a molecule is a linear combination of all the atomic orbitals occupied by the electrons in the constituent atoms.

$$\psi = \sum_{\mu}^{basis\ fn} C_{\mu}\,\phi_{\mu} \quad \text{.....(1)}$$

Where, *C* is the molecular orbital coefficient.

2.2.1.3 Hartree-Fock Approximation:

The second important approximation is the Hartree-fock Self-Consistent field (HF-SCF) method. The energy of a particular electron depends on the electric fields produced by the atomic nuclei and by all the other electrons.

The most obvious simplification to the Schrodinger equation involves separation of variables, that is replacements of the many electron wavefunction by a product of one electron wavefunction. The simplest acceptable replacement termed a Hartree-fock or single determinant wavefunction involves a single determinant of products of one-electron functions, termed spin orbitals. Each spin orbitals is written as a product of space part Ψ which is a function of the co-ordinates of a single electron and is referred to as a molecular orbital and one of two possible spin parts α or β. Only two electrons may occupy a given molecular orbital and they must be of opposite spin. The Hartree-fock approximation leads to a set of coupled differential equations (the Hartree-fock equations) each involving a single electron. The Hartree-fock equation is an eigenvalue equation of the form.

$$f(i)\ X(x_i) = \epsilon X(x_i)$$

Where, *f (i)* is an effective one electron operator (called fock operator) of the form.

X= Spin orbital,

x_i= Space and spin coordinates for i^{th} electron,

ϵ= Energy of i^{th} electron.

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{m} \frac{ZA}{r_i A} + V^{HF}(i)$$

Where $V^{HF}(i)$, which is the average potential experienced by the i^{th} electron due to the presence of the other electrons. The essence of the Hartree-fock approximation is to replace the complicated many-electron problem by a one electron problem in which electron-electron repulsion is treated in an average way.

The Hartree-fock potential or equivalently the 'field' seen by the i^{th} electron depends on the spin orbitals of the other electrons. Thus the H-F equation is non-linear and must be solved iteratively. The procedure for solving the Hartree-fock equation is called self-consistent-field (SCF) method.

The basic idea of the SCF method is simple. By making an initial guess at the spin orbitals. One can calculate the average field (i.e.V^{HF}) seen by each electron and then solve the eigenvalue equation for a new set of spin orbitals. Using these new spin orbitals, one can obtain new fields and repeat the procedure until self-consistency is reached (i.e. until the fields no longer change and the spin orbitals used to construct the fock operator are the same as its eigenfunctions).

The Hartree-fock and LCAO approximation eq^n (1), taken together and applied to the electronic Schrodinger equation lead to the Roothaan-Hall eq^n (2).

$$\sum(F_{\mu\nu} - \varepsilon_i S_{\mu\nu})\, C_{\nu i} = 0 \quad \ldots\ldots\ldots\ldots\ldots(2)$$

Here ε are orbitals energies,

S is the overlap matrix,

F is the fock matrix, which is analogous to the Hamiltonian in the Schrodinger equation.

Methods resulting from solution of the Roothaan-Hall equations are termed Hartree-fock or *Ab initio* models. The corresponding energy for an infinite (complete) basis set is termed the Hartree-fock energy.

2.2.2 Semi-empirical Methods

Substantial part of time required in performing *ab initio* Hartree-Fock SCF calculation is invariably spent on calculation and manipulation of integrals. The easiest way to reduce the computational effort is therefore to neglect or approximate some of these integrals. Semi-empirical methods achieve this in part by explicitly considering only the valence electrons of the system, the core electrons are subsumed into the nuclear core. The semi-empirical calculations invariably use basis sets comprising Slater-type s, p and some times d orbitals.

When the ideas of *ab initio* molecular orbital calculations were first developed, it was clear that the number of arithmetical operations required to investigate even the simplest of systems would be very large. Drastic approximations were required. Many semi-empirical theories use the zero-differential overlap approximation, (ZNDO). In

this approximation, the overlap between pairs of different orbitals is set to zero for all volume elements dv:

$$\phi_\mu \phi_\nu \, dv = 0$$

This directly leads to the following results for the overlap integrals.

$$S_{\mu\nu} = \delta_{\mu\nu}$$

If the two atomic orbitals ϕ_μ and ϕ_ν are located on different atoms than the differential overlap is referred to as diatomic differential overlap; if ϕ_μ and ϕ_ν are on the same atom then we have monoatomic differential overlap. If the ZDO approximation is applied to the two-electron repulsion integral then the integral will equal zero. It can immediately be seen that all three-and four-centre integrals are set to zero under the ZDO approximation.

In the semi-empirical approach CNDO (Complete Neglect of Differential Overlap) [1-2] the parameters required were developed from *ab initio* calculations. A similar approach was tried by Dewar [3] and called MINDO/3. Dewar followed the alternative philosophy of parameterising the method from experimental data rather than from *ab initio* calculations. This program was developed and called MNDO [4] and then AM1 [5]. PM3 is an alternative development [6]. The chief advantage of semi-empirical molecular orbital programs over *ab initio* MO program is their speed.

Under CNDO method, only the valence electrons are considered explicitly. Interactions between electrons on the same atom are handled simply. Since the atomic orbital basis is built from orthonormal atomic functions, then the interaction between different orbitals μ and ν.

2.2.2.1 MNDO:

Dewar and Thiel introduced the Modified Neglect of Diatomic Overlap (MNDO) method, which was based on NDDO [4]. The most significant advantage of MNDO over MINDO/3 is the use throughout of monoatomic parameters. MINDO/3 requires diatomic parameters in the resonance integral and the core-core repulsion. In more recent version of the MNDO method d orbitals have been explicitly included for the heavier elements [7].

2.2.2.2 AM1:

The Austin model 1 (AM1) method based on MNDO framework was the next semi-empirical approach developed by Dewar's group [5]. This method rectifies the tendency to overestimate repulsions between atoms when separated by distances

approximately equal to the sum of their van der walls radii. The strategy adopted is to modify the core-core term using Gaussian functions. Both attractive and repulsive Gaussian functions are used. The AM1 method improves the interatomic distances between atoms.

2.2.2.3 PM3:

PM3 is also based on MNDO [6]. The PM3 Hamiltonian contains essentially the same elements as that for AM1 but the parameters for the PM3 model were derived using an automated parameter optimization procedure. The PM3 method gives greater parameterization for many elements. Parameters are derived from experimental data to simplify the computation.

2.2.2.4 PCILO:

The semi-empirical quantum chemical perturbative configuration interaction with localized orbitals (PCILO) method [8-9], which is an all valence orbital approach has been used in this method.

2.2.2.4.1 Configuration Interaction (CI):

The SCF one-electron model method yields a molecular wavefunction which contains a 'built-in' error, the correlation error. To take care of this error, configuration interaction (CI) technique is used. Valence orbitals (outer orbital) contain valence electrons. Excitation of single electron from valence orbital gives singly excited wavefunctions ($\psi 1$). Two electrons excitation results in doubly excited wavefunctions ($\psi 2$). The mixing of these wavefunctions (configurations) in an appropriate manner is a way of correcting the correlation error, inherent in single-determinantal wavefunction. The linear combination of all the ψ_k is –

$$\psi = \sum_k C_k \psi_k$$

This is called configuration interaction (CI). The classical CI technique uses a variational procedure for obtaining the coefficients C_k, which minimizes the total energy. This gives a very large configuration interaction matrix and the number of possible excitations increases so rapidly with the enlargement of system. So that drastic arbitrary truncation of the CI matrix is needed.

2.2.2.4.2 Perturbative configuration Interaction (PCI):

Perturbation technique is used for solving the configuration interaction problem. The single configuration determinants ψ_k may be considered as the

eigenvectors of an unperturbed Hamiltonian H_o, which differs from the exact Hamiltonian H (including correlation) by a small perturbation term λV. In this case Hamiltonian splits into two parts H_o and λV.

$$H = H_o + \lambda V$$

The wavefunction and energy corresponding to the exact solution may thus be developed into Tayler expansions in the neighbourhood of λ=0.

$$\psi = \psi_k + \lambda\psi_1 + \lambda^2\psi_2 + \ldots\ldots\ldots\ldots$$

$$E = Ek + \lambda E'_1 + \lambda^2 E'_2 + \ldots\ldots\ldots..$$

In PCILO we retain energy corrections up to the third order for the ground state energy.

First order energy:

$$E'_1 = V_{oo}$$

Second order energy:

$$E'_2 = \sum_{K \neq 0} \frac{|V_{0K}|^2}{E_0 - E_K}$$

Third order energy:

$$E'_3 = \sum_{K \neq 0} \sum_{l} \frac{V_{K0} V_{l0} V_{0K}}{(E_0 - E_K)(E_0 - E_l)}$$

2.2.2.4.3 Localized orbitals:

The orbitals, which are localized on the chemical bonds are called localized orbitals. The molecular wave function is built from orbitals localized on the chemical bonds of the molecule and lone pairs.

2.2.2.4.4 General outline of PCILO:

The method uses the Zero differential overlap (ZDO) approximation as well as approximations in integral evaluation and parameterization similar to Complete Neglect of Differential Overlap (CNDO/2) approach. However the PCILO method goes beyond the CNDO framework by including the perturbative correction terms for configuration interaction. Antibonding orbitals corresponding to the localized bonds provide the excited states for configuration mixing.

Bonding orbitals are constructed from suitable combination of atomic orbitals.

$$i = \alpha X_1 + \beta X_2$$

While, the corresponding antibonding orbital is

$$i^* = - \beta X_1 + \alpha X_2$$

The PCILO method has been tested successfully in predicting the preferred conformational structures for a large number of biologically important molecules of various types. The method is particularly suited for conformational analysis of moderately large molecules when speed and reliability are needed.

For each conformation the polarity of the chemical bonds has been optimized, and in calculations of total ground state energy, correction terms up to the third order in configuration interaction are included.

The calculations have been performed by changing the various torsion angles over the entire range 0° - 360°, at increments of 30°, one variable at a time. The logical selection of grid points approach [10] has been used for searching the most stable structures, and the alternative stable structures in the multidimensional conformational space. Rigid body rotation has been assumed, thus fixed bond lengths and bond angles are maintained irrespective of the torsion angle values. The rigid body approximation is an important simplifying concept and has been immensely successful in investigations of conformational preferences of biomolecules [11-16].

2.3 Force Field Based Methods

2.3.1 Molecular Mechanics Force Field (MMFF):

Among the theoretical methods available to predict geometry, the molecular mechanics method is one of the most commonly used. In molecular mechanics the molecule is viewed as a collection of points (representing atomic nuclei) connected by springs (representing chemical bonds) with different elasticities (force constants). The forces holding the atoms together can be described by potential energy functions for each structural interaction components, like bond lengths, bond angles, non-bonded interactions, and so on. The molecular mechanics energy E_{MM}, is made up of a number of components. The energy due to every bond (E_{bonds}) stretching composition is added to the energy due to deviation of bond angles, (E_{angles}) from their equilibrium values. Likewise energy of all van der walls (E_{vdw}) interactions between non-bonded atoms is added. The earliest quantitative molecular mechanics models used only these terms [17-19]. It soon became clear that a term for torsion angles ($E_{torsion}$) was also required in order to explain many properties. For molecules with electronegative groups charge

interaction must also be included. These methods became sufficiently mature for a review to be necessary in 1956 [20] and the advantages of machine computation were advanced by Hendrickson in 1961 [21]. Many different groups have developed force fields, and they all follow this scheme. Although most have additional terms as well, which will be referred to as $E_{miscellaneous}$. The equation is –

$$E_{MM} = \Sigma_{Ebonds} + \Sigma_{Eangles} + \Sigma_{Evdw} + \Sigma_{Etorsion} + \Sigma_{Emiscellaneous}$$

Molecular mechanics methods may easily be applied to molecules containing 1000 or more atoms. The molecular mechanics force fields, such as MMFF94 [22], provide quantitative descriptions of organic molecular structures and conformations. The results obtained by molecular mechanics are comparable in quality with regard to geometry as those obtained by sophisticated quantum mechanical methods.

2.3.2 Molecular Dynamics (MD) Simulation:

Molecular dynamics (MD) as a method of studying the motion and the conformational space of molecular systems by integration of the classical Newtonian equations of motions given a potential energy function and its associated force field.

The first molecular dynamics (MD) simulations on nucleic acids started in 1983 on DNA [23]. After that McCammon and Harvey carried out MD simulation on the 76 nucleotide $tRNA^{Phe}$ [24]. Anticodon loop of $tRNA^{Asp}$ is also studied by MD simulation method [25].

MD simulations use potentials such as Lennard Jones potentials, in which the force between two atoms or molecules changes continuously with their separation. This requires equations of motions to be integrated by breaking the calculation into a series of very short time steps, usually around 1 femtosecond (10^{-15} s). At each step, the forces on the atoms are computed and combined with the current positions and velocities to generate new positions an updated set of forces is computed and so on. In this way a molecular dynamics simulation generates a trajectory that describe how the dynamics variables change with time.

The trajectory is obtained by solving the differential equations embodied in Newton's second law, which states that force equals the rate of change of momentum-

$$F=ma \text{ or } d^2x_i/dt^2 = Fx_i/m_i \quad \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

This equation describes the motion of a particle of mass m_i along one coordinate (x_i) with Fx_i being the force on the particle in that direction.

Operationally, and at the simplest possible level, two entities are required in an MD program, a force field and a way of integrating the equations of motions [26]. The verlet algorithm [27], also known as Leap-Frog method, for integration of the equations of motions utilized. As an MD method, implemented in a digital computer the calculations of motions are done at discrete intervals, the length of these intervals defines the time step. The displacement of an atom over a time step Δt is given by,

$$\overline{X}(t+\Delta t) - X(t) = \frac{\delta \overline{X}(t)}{\delta t}\Delta t = \overline{V}(t^*)\Delta t \quad \text{.....(4)}$$

Where $V(t^*)$ is the velocity at time t^*.

The verlet method uses the velocity at the midpoint of the time interval in equation (4), since the velocity is not constant through the interval i.e. t^* is taken as $t+\Delta t/2$. The velocity at the midpoint can be estimated from the previous time step and from the acceleration a, which in turn can be calculated directly from the force equation (3).

$$\overline{V}(t+\frac{\Delta t}{2}) = \overline{V}(t-\frac{\Delta t}{2}) + \overline{a}\Delta t \quad \text{.....(5)}$$

After calculating the new velocity, it is substituted into equation (4), and the cycle is repeated.

2.3.2.1 Parameters for MD:

Some of the important inputs a molecular dynamics simulation needs.

1) A starting structure
2) The temperature
3) The step size
4) The length of the run.

2.3.2.1.1 A starting Structure:

An MD simulation must begin with a sensible geometry for the structure of interest. This should be done with some care as the initial arrangements can often determine the success or failure of a simulation. Usually, an appropriate model will be built using any experimental data that are available, and this will be minimized so that all bond lengths, bond angles, etc. have sensible values and bad contacts between the non-bonded atoms are avoided.

2.3.2.1.2 The temperature:

The amount of kinetic energy given to the structure depends on the temperature of the system. The energy is divided between the atoms so that on the

average each atom gets more or less the same energy. Instantaneous distribution of velocities for various atoms in the system however follows Maxwell-Boltzman statistics.

2.3.2.1.3 Step size:

Because the potential energy field changes as a molecular structure is adjusted, accelerations must also change. Consequently, the time step in MD must be very small. The smaller the time step, the better the approximation. Large time steps results in unreasonably large movements of the atoms. The step in region of 1 femptoseconds (fs) is recommended for most situations. The use of SHAKE [28] algorithm is most often used to constrain the bonds involving hydrogen atoms. In such cases it is possible to use a slightly larger time step.

2.3.2.1.4 Length of the run:

It is a total molecular dynamics simulation time for which system is studied. The choice of a termination time will depend on the information required.

2.3.2.2 Steepest-Descent minimization:

Steepest-Descent is a minimization method, which use derivative information. It is usually possible to calculate the slope of a potential surface from the potential energy functions. In the one-dimensional case, this tells us how steep the line is, and which way is down.

The simplest way to use the derivative information is simply to go downhill in the steepest slope. A one-dimensional minimization is much easier than a multidimensional one. So once the steepest-descent direction has been chosen, we can do a one-dimensional minimization along that direction.

2.4 References:

1. Pople, J. A.; Santry, D. P.; Segal, G. A.; J. Chem. Phys. 1965, 43, S129-S135.
2. Pople, J. A.; Segal, G. A. J. Chem. Phys. 1965, 43, S136-S149.
3. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. J. Am. Chem. Soc., 1975, 97, 1285-1293.
4. Dewar, M. J. S.; Thiel, W. J. Am. Chem. Soc. 1977, 99, 4899-4907.
5. Dewar, M. J. S.; Zoebsch, E. G.; Healy, E. F.; Stewart, J. J. P. J. Am. Chem. Soc., 1985, 107, 3902-3909.
6. Stewart, J. J. P. J. computational Chemistry. 1989, 10, 209-220.

7. Thiel, W. and Voityuk, A. A. J. Molecular Structure (Theochem) 1994, 313, 141-154.
8. Diner, S.; Malrieu, J. P.; Jordan, F.; Gillbert, M. Theoret. Chim. Acta 1969, 15, 100-110.
9. Malrieu, J. P. in Semiempirical Methods of Electronic Structure Calculations, Part A, Techniques, Segal, G. A.; Ed.; Plenum, New York, 1977, p. 69-103.
10. Tewari, R. Int. J. Quantum Chem. 1987, 31, 611-624.
11. Pullman, B.; Pullman, A. Advan. Protein Chem. 1974, 16, 347-526.
12. Pullman, B.; Saran, A. Prog. Nucleic Acid Res. Mol. Biol. 1976, 18, 215-325.
13. Ramachandran, G. N. and Sasisekharan, V. Adv. Protein Chem. 1968, 23, 283-437.
14. Stimson, E. R.; Zimmerman, S. S. and Scheraga, H. A. Macromolecules 1977, 10, 1049-1060.
15. Tosi, C.; Clementi, E. and Matsuoka, O. Biopolymers 1978, 17, 67-84.
16. Tewari, R.; Nanda, R. K. and Govil, G. J. Theor. Biol. 1974, 46, 229-239.
17. Hill, T. L. J. Chem. Phys. 1946, 14, 465.
18. Westheimer, F. H.; Mayer, J. E. J. Chem. Phys. 1946, 14, 733-738.
19. Barton, D. H. R. J. Chem. Soc. 1948, 340-342.
20. Westheimer, F. H.; Edited by Newmann, M. S., Weiley, New York, 1956, 523-555.
21. Hendrickson, J. B. J. Am. Chem. Soc. 1961, 83, 4537-4547.
22. Clark, M.; R. D. Cramer III and Van Opdensch, N. J. Comp. Chem. 1989, 10, 982- 1012.
23. Levitt, M. Computer simulations of DNA double helix dynamics. Cold Spring Harb. Symp. Quant. Biol. 1983, 47, 251-262.
24. McCammon, J. A.; Harvey, S. C. Dynamics of Proteins and Nucleic Acids. New York, Cambridge University Press, 1987.
25. Auffinger, P.; Louise-May, S.; Westhof, E. J. Am. Chem. Soc. 1995, 117, 6720-6726.
26. Symon, K. R.; Mechanics, 3rd edition, Addison-Wesley, Reading, MA (1971).
27. Verlet. L.; Phys. Rev. 1967, 159, 98-103
28. Ryckaert, J. P.; Cicotti, G. and Berendsen, H. J. C. J. Comp. Phys. 1977, 23, 327-341.

CHAPTER-III

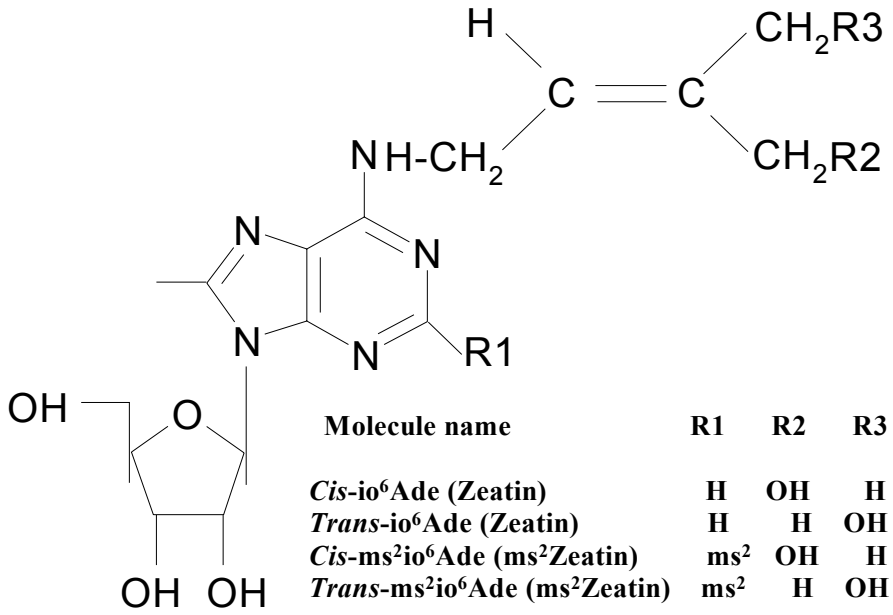
PART-A: CONFORMATIONAL PREFERENCES OF HYDROXYISOPENTENYL ADENOSINE AND ANALOGS

3.1 Introduction

Hypermodified nucleic acid base N6-(Δ^2-isopentenyl) adenine, i^6Ade, and its 2-methylthio derivative (ms^2i^6Ade) usually occurs at 3'- adjacent position of anticodon in tRNAs which recognize the codons starting with uridine [1-4]. In some bacteria grown under aerobic conditions - instead of anaerobic conditions, the hypermodified nucleic acid bases N6-(Δ^2-*cis*-hydroxyisopentenyl)adenine, *cis*-io^6Ade also known as *cis*-zeatin, and N6-(Δ^2-*trans*-hydroxyisopentenyl)adenine, *trans*-io^6Ade or *trans*-zeatin and 2-methylthio derivatives of these *cis*-ms^2io^6Ade or *cis*-ms^2-zeatin, and *trans*-ms^2io^6Ade or *trans*-ms^2zeatin, alternatively occur. As a free base isopentenyl adenine is cytokinin active [5] and promotes cell elongation and stimulates growth-differentiation in plant tissue culture.

In order to understand the functional significance of chemical modification in cytokinin active hypermodified adenine, present study on conformational preferences of *cis* and *trans* isomers of zeatin (io^6Ade) and its 2-methylthio derivative has been undertaken. Earlier, conformational preferences for modified adenine bases with hydrophilic substituents namely t^6Ade, ms^2t^6Ade, gc^6Ade [6, 7] and conformational preferences of modified adenine with hydrophobic substituents i^6Ade, ms^2i^6Ade, benzylaminopurine and furfurylaminopurine have already been reported [8, 9]. Hydroxyl group makes hydrophilic contribution in zeatin (io^6Ade) and may also enable specific intramolecular hydrogen bond donor - acceptor interactions, which may alter conformational preferences as compared to i^6Ade. The presence of the hydroxyl group may occur *cis* or *trans* to HN(6)C(10)H2, these isomers, Fig. 3.1(a), may not be readily interchangeable in view of the large barriers prohibiting rotation around C(11)=C(12) double bond. It is also of interest to compare the conformational preferences of *cis* and *trans* isomers and to further study the effects of 2-methylthiolation in isolated base as well as in model anticodon loop segment.

The hypermodified nucleoside N6-(Δ^2-isopentenyl) adenosine, i^6A, and its 2-methylthio derivative (ms^2i^6A) also have been included in the model anticodon loop segment. The ribose-phosphate backbone torsion angles are held similar to Holbrook $tRNA^{Phe}$ model [10].



Molecule name	R1	R2	R3
***Cis*-io^{6}Ade (Zeatin)**	**H**	**OH**	**H**
***Trans*-io^{6}Ade (Zeatin)**	**H**	**H**	**OH**
***Cis*-ms^{2}io^{6}Ade (ms^{2}Zeatin)**	**ms^2**	**OH**	**H**
***Trans*-ms^{2}io^{6}Ade (ms^{2}Zeatin)**	**ms^2**	**H**	**OH**

Fig. 3.1(a): Representation of *cis*-zeatin and *trans*-zeatin, as well as *cis*-ms^{2}zeatin and *trans*-ms^{2}zeatin derivatives.

3.2 Nomenclature, Convention and Procedure:

Figure 3.1(b) depicts the atom numbering and the identification of the torsion angles (which specify the internal rotation around the various acyclic chemical bonds) determining relative orientation – positioning of various atoms in zeatin and ms^2-zeatin molecules. In *cis*-zeatin the hydroxyl oxygen O(13) is attached to the carbon atom C(13) and bond C(13)-C(12) eclipses bond C(10)-C(11) (δ=0°), whereas in *trans*-zeatin the arrangement of bonds C(13)-C(12), C(12)=C(11) and C(11)-C(10) is extended (δ=180°). The *cis*- and *trans*-zeatin isomers become distinctly identifiable due to exceedingly high barriers to internal rotation around C(11)=C(12) double bond, which practically forbids rotational interconversion. The orientation of 2-methylthio group in ms^2-zeatin is specified by torsion angle ω1 [N(3)C(2)S(2)C(15)] which describes rotation of C(15) about the bond C(2)-S(2) and is measured with respect to eclipsing of the bond C(2)-N(3) by S(2)-C(15) bond, and the torsion angle ω2[C(2)S(2)C(15)H], which likewise specifies the placement of one of the three equivalent hydrogen atoms of the methyl group.

In the N(6) substituent the torsion angle α [N(1)C(6)N(6)C(10)] denotes the rotation of C(10) around bond C(6)-N(6) and is measured with respect to N(1) from the *cis* (eclipsed, 0°) position in the right hand sense of rotation. Likewise, the successive chemical bonds along the main extension of the substituent define the

subsequent torsion angles β [C(6)N(6)C(10)C(11)], γ [N(6)C(10)C(11)C(12)], δ [C(10)C(11)C(12)C(13)], ψ1 [C(11)C(12)C(13)O(13)], ψ2 [C(11)C(12)C(14)H], θ [C(12)C(13)O(13)H]. All these torsion angles are similarly measured with respect to respective *cis* (eclipsed, 0°) bonds orientation in the right hand sense of rotation.

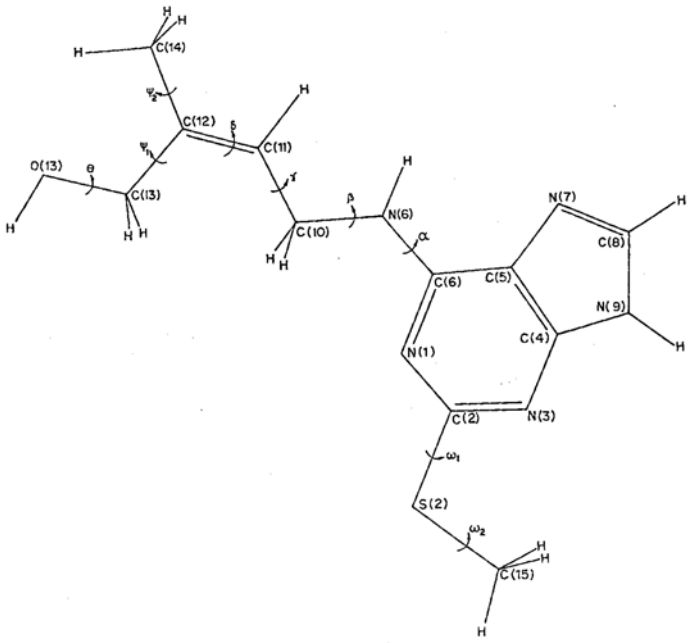


Fig. 3.1(b): Atoms numbering and torsion angles identification of *cis*-zeatin, *trans*-zeatin as well as analogous 2-methylthiolated *cis*-ms^{2}zeatin and *trans*-ms^{2}zeatin molecules.

Fig.3.1(c) model diphosphate nucleotide segment describes the atom numbering and torsion angles definition for i^{6}A, ms^{2}i^{6}A and ribose-phosphate backbone. In case of i^{6}A and ms^{2}i^{6}A hydroxyl oxygen O(13) is absent, so rest of the nomenclature, atom numbering and torsion angle definitions for α, β, γ, δ, Ψ2, ω1, and ω2 are same as in zeatin or ms^2-zeatin molecule fig.3.1(b). The torsion angle definition for Ψ1 [C(11)C(12)C(13)H] in i^{6}A and ms^{2}i^{6}A fig.3.1(c), changes due to absence of O(13). The ribose-phosphate backbone torsion angles are held similar to Holbrook tRNA^{Phe} model [10]. The torsion angles in the ribose-phosphate backbone are distinguished by the subscript b to refer to the backbone. These backbone torsion angle values retain the nomenclature as in the tRNA model [10] referring likewise the right hand sence of rotation around the central bond, measured from the eclipsed position of the outer bonds α_b [C-O3'-p-O5'], β_b [O3'-p-O5'-C5'], γ_b [p-O5'-C5'-C4'], δ_b [O5'-C5'-C4'-C3'], ε_b [C5'-C4'-C3'-O3'], and ζ_b [C4'-C3'-O3'-p], χ [O1'-C1'-N(9)-C(8)].

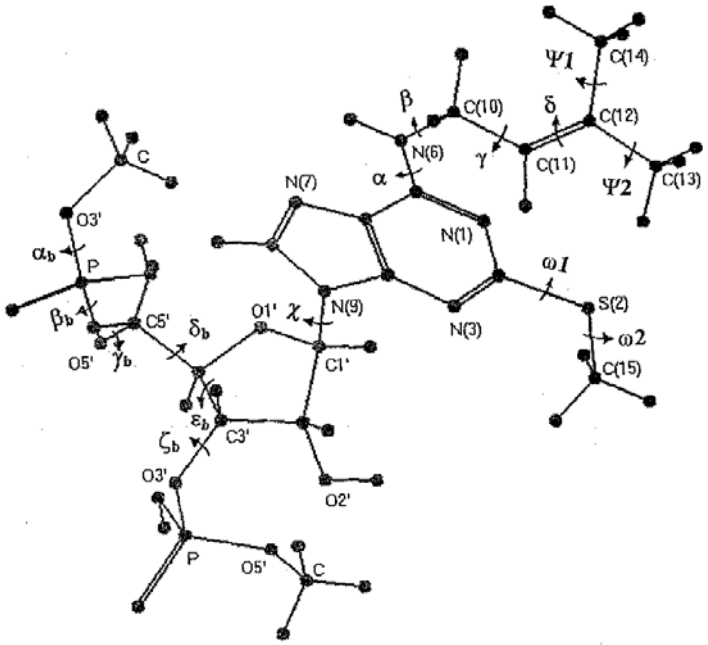


Fig.3.1(C): The atom numbering and torsion angle definition for i^6A, ms^2i^6A and ribose-phosphate backbone in model diphosphate segment are shown.

For conformation with α=180° the bond N(6)C(10) points towards the five membered imidazole moiety of adenine and is termed "proximal". For α=0° the bond N(6)C(10) points away from the cyclic five membered imidazole part of the purine ring and the conformation is termed "distal".

The quantum chemical perturbative configuration interaction with localized orbitals PCILO method [11-13] has been utilized for the energy calculations of the various molecular conformations. The observed chemical bond lengths and bond angle values in the crystal structures [14, 15] have been appropriately used in the conformational energy calculations on *cis*-zeatin, *trans*-zeatin, *cis*-ms^2zeatin, *trans*-ms^2zeatin and also for i^6A and ms^2i^6A in model anticodon loop segment. The methyl phosphate group (Me-p-) was used to truncate the model anticodon loop segments. In the case of the hydrogen atoms, the standard values from the reference source data [16], suited for the particular type of bonding environment have been used. Full geometry optimization calculations using quantum chemical MNDO [17] method and gradually refined parameterizations AM1 [18] and PM3 [19] have also been made to compare the salient features for isolated modified base. Relative preferences of the salient structures, for isolated base have also been deduced using quantum mechanical single point *ab-initio* molecular orbital Hartree - Fock SCF energy calculations (PC Spartan Pro version 6.0.6) [20] using 6-31G* basis set. For larger molecules

molecular mechanics force field (MMFF) method has been used for full geometry optimization.

3.3 Results and Discussion

3.3.1 Isolated modified base, io^{6}Ade and its analogs:

Fig. 3.2 depicts the predicted PCILO preferred most stable conformation for *cis*-zeatin. The torsion angles describing the conformation are (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°). The structure is stabilized by hydrogen bonding interactions between O(13) and HC(10)H.

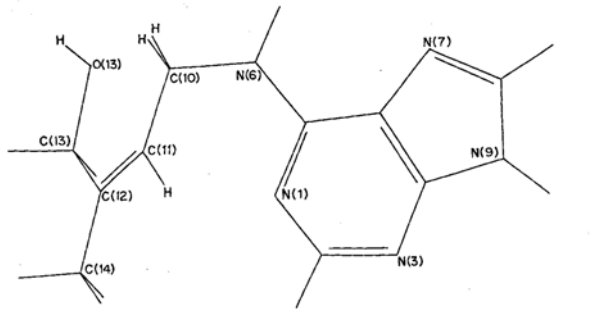


Fig. 3.2: The PCILO predicted most stable conformation for *cis*-io^{6}Ade or *cis*-zeatin.

Additional stabilization from interaction between N(1)…HC(11) is also expected. The geometrical parameters for likely hydrogen bonding interactions are given in Table 3.1. The predicted torsion angle values for io^{6}Ade may be compared to the preferred values [8] for i^{6}Ade (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°,±120°, ψ2=0°,±120°). Higher energy (0.5 kcal/mol) PCILO alternative conformation Table 3.2 arises by flipping of ψ1 to extended orientation (180°) besides change of θ to eclipsed orientation (0°). Compared to the preferred structure shown in (fig.3.2), this alternative structure lacks interaction between O(13) and HC(10) but instead has C(14) in proximity to HO(13). However, the interaction between N(1) and HC(11) is maintained alike. Next higher energy alternative conformation (1.0 kcal/mol) corresponds to θ=180°, while ψ1 takes the extended (180°) orientation. Except that the hydroxyl hydrogen is now positioned away from the C(14) and thus corresponds to higher energy, in other respects this structure is similar to the previous alternative.

Table 3.1: Geometrical parameters for hydrogen bonding in the preferred conformations of *cis*- or *trans*-zeatin and 2-methylthio analogs.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure References
N(1)…H-C(11)	2.06	1.07	90.3	3.2, 3.3, 3.4, 3.5
O(13)…H-C(10)	1.76	1.09	138.2	3.2, 3.3
N(1)…H-O(13)	2.76	0.96	133.1	3.4, 3.5
O(13)…H-C(11)	2.43	1.07	104.1	3.4, 3.5
S(2)…H-O(13)	2.15	0.96	157.9	3.5
O(13)…H-C(10)	1.79	1.09	134.8	21 (Cryst. Struct.
N(1)…H-C(11)	2.22	1.07	82.2	21 Conformer1)
O(13)…H-C(10)	2.14	1.09	110.6	21 (Cryst. Struct.
N(6)…H-O(13)	2.19	0.96	139.4	21 Conformer2)

The observed crystal structure [21] conformations of *cis*-zeatin conformer1 (α=6.2°, β=116.4°, γ=250°, δ=-2.4°, ψ1=126.4°, ψ2=130.5°, θ=220.4°) and of conformer2 (α=5.2°, β=188.0°, γ=120.8°, δ=-2.2°, ψ1=103.8°, ψ2=316.1°, θ=322.8°) differ mainly with respect to torsion angles β, ψ1, ψ2 and θ. Except for some expected minor deviations in torsion angle values the remaining torsion angles in crystal structure agree with the preferred structure in (fig.3.2). However, the predicted intramolecular interactions between O(13) and HC(10)H and between N(1) and HC(11) groups, as in (fig.3.2), are absent in crystal structure. These differences arise due to crystal packing environment which allows intermolecular hydrogen bond donor – acceptor interactions [21] involving N(9)H---N(3)', N(9)'H'---N(3), N(6)H---O(13)' and C(8)H---N(1) between neighboring primed and unprimed molecules. Compared to the predicted most stable structure by PCILO method, the observed crystal structure conformer1 turns out to be about 5.8 kcal/mol higher in energy and conformer2 is about 6 kcal/mol higher in energy.

Starting from the PCILO preferred conformation shown in (fig.3.2), automated complete geometry optimization through semi-empirical quantum chemical approaches MNDO, AM1 and PM3 methods [17-19] are included in Table 3.3.

Comparison of MNDO, AM1 and PM3 methods, results with the observed crystal structure of *cis*-zeatin (conformer1) and (conformer2) are also included in Table 3.3

Table 3.2: Predicted alternative stable conformations for *cis*- or *trans*-zeatin and 2-methylthio analogs.

Torsion Angles	Rel. Eng. (kcal/mol.)	Fig. Ref.
α=0°,β=0°,γ=±120°,δ=0°,ψ1=0°,ψ2=0°,±120°,θ=180°	0.0	3.2
α=0°,β=0°,γ=±120°,δ=0°,ψ1=180°,ψ2=0°,±120°,θ=0°	0.5	
α=0°,β=0°,γ=±120°,δ=0°,ψ1=180°,ψ2=0°,±120°,θ=180°	1.0	
α=0°,β=0°,γ=±120°,δ=0°,ψ1=0°,ψ2=0°,±120°,θ=180°,ω1=0°,ω2=180°	0.0	3.3
α=0°,β=0°,γ=±120°,δ=0°,ψ1=180°,ψ2=0°,±120°,θ=0°,ω1=0°,ω2=180°	0.4	
α=0°,β=0°,γ=±120°,δ=0°,ψ1=180°,ψ2=0°,±120°,θ=180°,ω1=0°,ω2=180°	0.9	
α=0°,β=0°,γ=±120°,δ=180°,ψ1=0°,ψ2=0°,±120°,θ=±60°	0.0	3.4
α=0°,β=0°,γ=±120°,δ=180°,ψ1=0°,ψ2=0°,±120°,θ=±300°	1.5	
α=0°,β=0°,γ=±120°,δ=180°,ψ1=180°,ψ2=0°,±120°,θ=±300°	1.7	
α=0°,β=0°,γ=±120°,δ=180°,ψ1=0°,ψ2=0°,±120°,θ=±60°,ω1=0°,ω2=180°	0.0	3.5
α=0°,β=0°,γ=±120°,δ=180°,ψ1=0°,ψ2=0°,±120°,θ=±300°,ω1=0°,ω2=180°	2.9	
α=0°,β=0°,γ=±120°,δ=180°,ψ1=180°,ψ2=0°,±120°,θ=±300°,ω1=0°,ω2=180°	3.2	

Fig.3.3 depicts the predicted PCILO preferred most stable structure for *cis*-ms^{2}zeatin. The torsion angles describing the preferred conformation are (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°). As for *cis*-zeatin, the structure is stabilized due to interactions between O(13)…HC(10) and between N(1)…HC(11). The torsion angle values may be compared with the preferred values [8] for ms^{2}i^{6}Ade (α=0°, β=180°, γ=±60°, δ=0°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=180°, ω2=180°).

The orientation of the hydroxyisopentenyl substituent is alike in *cis*-io^{6}Ade (*cis*-zeatin) and ms^{2}io^{6}Ade (*cis*-ms^{2}zeatin). Nonobstrusive orientation of the 2-methylthiogroup in (fig. 3.3) leaves the orientation of the N(6) substituent unaffected. Full geometry optimized conformation related to PCILO preferred conformation for *cis*-ms^{2}zeatin have been obtained using MNDO, AM1 and PM3 methods. The optimized values for the torsion angles are given in Table 3.3.

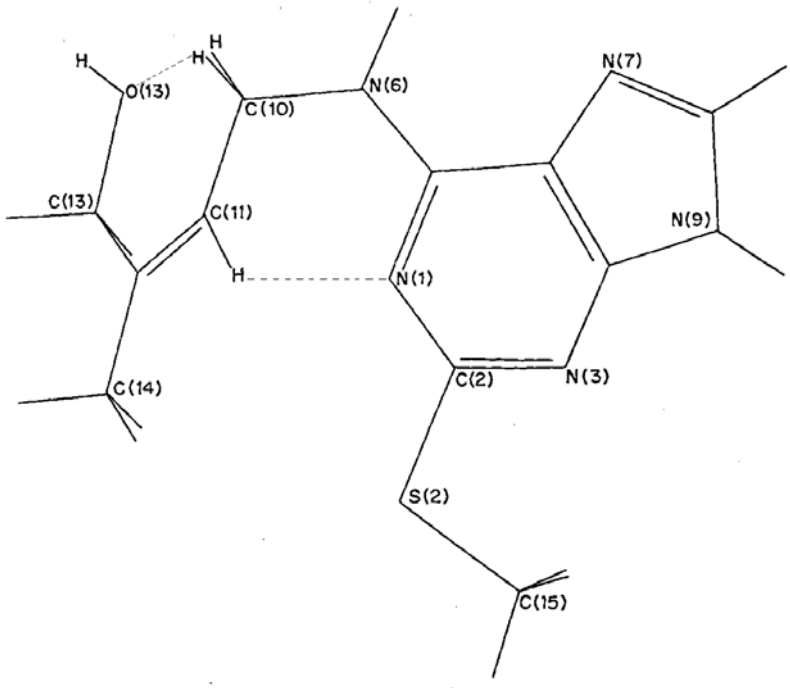


Fig. 3.3: PCILO Predicted most stable conformation for *cis*-ms^{2}io^{6}Ade *(cis*-ms^{2}zeatin).

Alternative PCILO stable conformations are shown in Table 3.2. The conformation corresponding to higher energy (0.4 kcal/mol) arises by flipping $\psi 1$ to extended orientation while keeping the eclipsed orientation (0°) for θ. Both of these torsion angles are extended for alternative stable structure at 0.9 kcal/mol higher energy.

The alternative stable structures as well as the factors stabilizing the alternative conformations of *cis*-ms^{2}io^{6}Ade are similar to that described for *cis*-io^{6}Ade. The energy differences for alternative stable structures are also alike for *cis*-zeatin and *cis*-ms^{2}zeatin.

Fig. 3.4 depicts the predicted PCILO preferred most stable structure for *trans*-zeatin. The torsion angles specifying the most stable structure are (α=0°, β=0°, γ=±120°, δ=180°, $\psi 1$=0°, $\psi 2$=0°±120°, θ=±60°). Correlation between torsion angles γ and θ is such that for the most stable structure both the torsion angles must take values with similar (positive or negative) sign.

Table 3.3: Optimized values of the torsion angles by automatic geometry optimization using PM3, MNDO and AM1 methods.

PM3	MNDO	AM1	Fig.Ref.
A] PCILO Most Stable Structures for Starting Optimization:			
α=-32°,β=3°,γ=-117°,δ=2°, ψ1=-13°,ψ2=114°,θ=-159°	α=-7°,β=9°,γ=-106°,δ= 5°, ψ1=-40°,ψ2=107°,θ=-180°	α=-15°,β=-76°,γ=-54°,δ=1° ψ1=-35°, ψ2=114°,θ=-56°	3.2
Rel.Energies: 10.5 kcal/mol.	6.3 kcal/mol.	0.1 Kcal/mol.	
α=-31°,β=2°,γ=-115°,δ=5°, ψ1=-40°,ψ2=117°,θ=-68°, ω1=0°, ω2=179°	α=-6°,β=6°,γ=-100°, δ=6°,ψ1=-86°,ψ2= 127°, θ=-175°, ω1=0°,ω2=-180°	α=-14°, β=-75°, γ=-55°, δ=1°, ψ1=-34°,ψ2= 116°, θ=-57°, ω1=0°, ω2=-180°	3.3
Rel.Energies: 8.1 kcal/mol.	3.7 kcal/mol.	0.1 Kcal/mol.	
α=25°,β=88°,γ=-138°, δ=-180°, ψ1=22°, ψ2=129°, θ=-41°	α=13°,β=78°,γ=-171°, δ= -179°, ψ1=-64°, ψ2=107°, θ=-62°	α=-18°, β=-79°, γ=-61°, δ= -179°, ψ1= -12°, ψ2= 87°, θ= -58°	3. 4
Rel.Energies: 0.0 kcal/mol.	0.0 kcal/mol.	0.0 kcal/mol	
α=23°,β=87°,γ=-135°, δ=-179°, ψ1=-60°,ψ2=123°, θ=-54°, ω1=0°, ω2=-178°	α=13°, β=77°, γ=-165°, δ=-179°,ψ1=-61°,ψ2=104°, θ=-62°, ω1=0°, ω2=179°	α=-17°, β=-78°, γ=-62°, δ=-179°,ψ1=-12°,ψ2= 87°, θ=-57°, ω1=0°, ω2=-179°	3.5
Rel.Energies: 0.0 kcal/mol.	0.0 kcal/mol.	0.0 kcal/mol	
B] Crystal Structure for Starting Optimization:			
a) Conformer1 :			
α=19°,β=95°,γ=-129°,δ=0°, ψ1=116°,ψ2=121°,θ= 296°	α=11°,β=137°,γ=-93°,δ=1° ψ1=78°,ψ2= 123°, θ=184°	α=18°,β=80°,γ=-134°,δ=2°, ψ1=176°,ψ2=133°,θ=-62°	
Rel.Energies: 0.4 kcal/mol.	0.7 kcal/mol.	0.3 kcal/mol.	
b) Conformer2 :			
α=20°,β=175°,γ=130°,δ=-2°, ψ1=113°, ψ2= -3°, θ= -58 °	α=16°,β=-174°,γ=87°,δ=-3°, ψ1=103°, ψ2= -42°,θ= -74 °	α=-8°,β=180°,γ=148°,δ=-3°, ψ1=70°, ψ2= 0°, θ= -56 °	
Rel.Energies: 1.1 kcal/mol.	0.0 kcal/mol.	1.1 kcal/mol.	

Fig. 3.4 depicts the predicted PCILO preferred most stable structure for *trans*-zeatin. The torsion angles specifying the most stable structure are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°±120°, θ=±60°). Correlation between

torsion angles γ and θ is such that for the most stable structure both the torsion angles must take values with similar (positive or negative) sign.

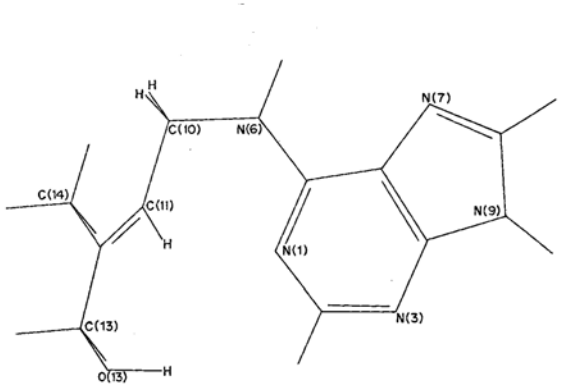


Fig.3.4 PCILO predicted most stable structure for *trans*-io^{6}Ade or *trans*-zeatin

The preferred conformation is stabilized by hydrogen bonding interactions between N(1)…HC(11). Some stabilization may also accrue due to possible hydrogen bonding between O(13) and HC(11). Still weaker interaction occurs between N(1)…HO(13). The preferred values for *trans*-zeatin may be compared with the predicted values [8] for i^{6}Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°). Presence of the hydroxyl group results in a unique preferred value of ψ1=0° for *trans*-zeatin, unlike three fold symmetry in i^{6}Ade. Full geometry optimization of PCILO preferred conformation using MNDO, AM1 and PM3 methods has been included in Table 3.3.

Alternative PCILO predicted stable conformation presented in Table 3.2, for *trans*-zeatin having 1.5 kcal/mol higher energy, arises on flipping of θ from 60° to 300° (when γ=120°, or equivalently for γ=240° the reverse flipping of θ from 300° to 60°). Although the interaction between N(1)…HC(11) is retained as in the preferred structure (fig.3.4), but weakened interactions involving HO(13) may account for the higher energy. Next alternative stable structure, having (1.7 kcal/mol) higher energy arises when flipping of θ is accompanied with the extended orientation for ψ1. Increased separation of O(13) from N(1) may account for the increased energy. In contrast to interactions stabilizing the stable structures for *cis*-zeatin, involvement of the polar O(13)H in hydrogen bonding interactions with HC(11) and N(1) in *trans*-

zeatin results in larger energy differences between the alternative stable structures. Compared to the preferred structure for *cis*-zeatin (fig.3.2), the preferred structure (fig.3.4) for *trans*-zeatin is more stable by about 1.2 kcal/mol.

Fig. 3.5 depicts the predicted PCILO preferred most stable conformation for *trans*-ms^{2}zeatin. The torsion angles specifying the most stable structure are (α=0°, β=0°, γ= ± 120°, δ=180°, ψ1=0°, ψ2=0°, ±120°, θ=± 60°, ω1=0°, ω2=180°).

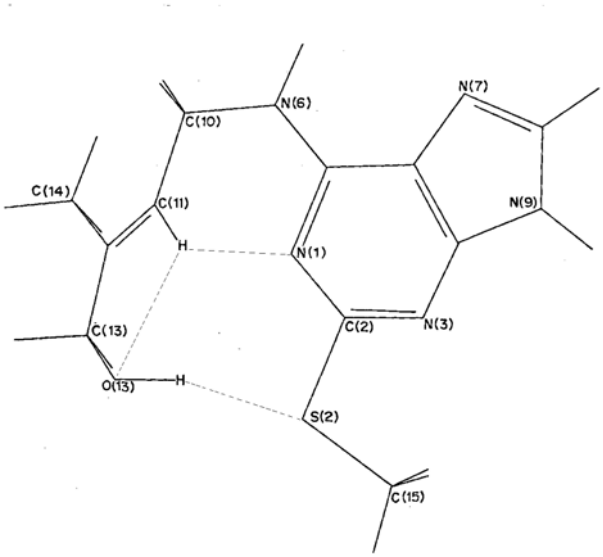


Fig. 3.5: The PCILO predicted most stable conformation for *trans*-ms^{2}io^{6}Ade or *trans*-ms^2-zeatin.

Interactions between N(1) and HC(11), besides between S(2) and HO(13) groups stabilize the preferred structure. The other likely stabilizing interactions between O(13)…HC(11) as well as between N(1)…HO(13) are also present. The predicted most stable structure may be compared with the most stable structure predicted [8] for ms^{2}i^{6}Ade (α=0°, β=180°, γ=±60°, δ=180°, ψ1=0°,±120°, ψ2=0°±120°, ω1=180°, ω2=180°). Most conspicuous difference between these two is due to the difference in the orientation of the methylthio group. Difference in the orientations of the hydroxyisopentenyl and isopentenyl groups may arise due to this factor besides due to the hydrogen bonding between O(13)H and S(2). Due to the nonobstrusive orientation of the methylthio group the orientation of the hydroxyisopentenyl substituent remains unaffected in *trans*-ms^{2}zeatin and orientation similar to *trans*-zeatin is retained. Starting from the preferred PCILO conformation full geometry optimization using MNDO, AM1 and PM3 methods is included in Table 3.3. These PCILO preferred interatomic distances are reasonable and are also

comparable with the crystal structure observations on *cis*-zeatin conformers [21] shown in Table 3.1.

Relative stability of the various PCILO preferred conformations (figs. 3.2 – 3.5) as well as the observed *cis*-zeatin conformer1 and conformer2 have also been calculated using computationally expensive quantum mechanical *ab-initio* HF SCF method using 6-31G* basis set [20] and are included in Table 3.4.

Table 3.4: single point *ab initio* HF SCF calculations with 6-31G* basis set.

Molecule Identity	PCILO preferred structures (kcal/mol)	Direct Ab initio calculations on PCILO preferred structures (kcal/mol)	*Ab initio* calculations on PM3 optimized PCILO preferred structures (kcal/mol)	Fig. Ref.
a) *Cis*-io^{6}Ade	1.3	7.7	3.6	3.2
b) *Trans*-io^{6}Ade	0.0	0.0	0.0	3.4
c) *Cis*-ms^{2}io^{6}Ade	3.0	4.0	2.6	3.3
d) *Trans*-ms^{2}io^{6}Ade	0.0	0.0	0.0	3.5
Cis-zeatin Crystal structures (Ref. 21)				
a) Conformer 1	0.0	1.3	0.0	
b) Conformer 2	0.2	0.0	3.2	

Trans-zeatin structure (fig. 3.4), is indicated stabler than *cis*-zeatin structure (fig.3.2), by 7.7 kcal/mol. Similarly *trans*-ms^{2}zeatin structure (fig. 3.5), is stabler than *cis*-ms^{2}zeatin structure (fig. 3.3), by 4.0 kcal/mol. When the *ab-initio* calculations are made on PM3 optimized geometries Table 3.3, instead of directly starting on PCILO preferred structures (figs. 3.2 – 3.5), the relative stability of *trans*-zeatin (or ms^{2}zeatin) over *cis*-zeatin (or ms^{2}zeatin) is still retained. However, the energy differences (3.6 kcal/mol and 2.6 kcal/mol respectively in case of (fig. 3.4 - Fig. 3.2) and similarly for (fig. 3.5 - Fig.3.3) get reduced. In contrast to nearly equal stability shown by PCILO for two observed crystal structures, conformer2 is indicated by *ab-initio* method to be relatively more stable than conformer1 by 1.3 kcal/mol. However, when the *ab-initio*

calculations follow PM3 optimization of the observed crystal structure conformers conformer1 is indicated to be stabler than conformer2 by 3.2 kcal/mol.

Alternative PCILO predicted stable conformation, Table 3.2, for *trans*-ms^{2}zeatin having 2.9 kcal/mol is realized by flipping θ from 60° to 300° (when γ =120°, or the reverse flipping of θ from 300° to 60° when γ =240°). The extended ψ1 orientation leads to stable structure having 3.2 kcal/mol higher energy. Due to the additional interaction of O(13)H with S(2), the energy differences between the alternative stable structures are even larger in *trans*-ms^{2}zeatin than *trans*-zeatin. The preferred *trans*-ms^{2}zeatin structure (fig.3.5) is more stable than the preferred *cis*-ms^{2}zeatin structure (fig.3.3) by nearly 3 kcal/mol, and the difference could arise from the stabilization due to O(13)H…S(2) interaction, instead of weaker O(13)H…HC(10)H interaction.

3.3.2 The model hypermodified nucleoside diphosphate segments

3.3.2.1 Me-p-i^{6}A-p-Me:

In the model hypermodified nucleotide diphosphate segment, Me-p-i^{6}A-p-Me, the ribose-phosphate backbone torsion angles are held similar to Holbrook tRNA^{Phe} model [10]. The glycosyl torsion angle value is (χ=22°). The set of torsion angle values for the preferred orientation of isopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°). These torsion angle values may be compared to the preferred values for isolated modified base i^{6}Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0,±120°, ψ2=0,±120°) [8].

The preferred most stable structure shown in (fig.3.6) is obtained when glycosyl torsion angle value is (χ=292°). The set of torsion angle values for isopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°). These preferred torsion angle values may be compared with i^{6}Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°) [8].

The intramolecular hydrogen bonding between C(11)H….N(1), Table 3.5a, remains a stabilizing factor. Interactions of ribose-phosphate backbone with isopentenyl adenine base i.e. N(3)…HC3', N(3)…HC2' are the additional stabilizing factors. The glycosyl torsion angle value (χ=22°) is held as in Holbrook tRNA^{Phe} model [10] or takes the preferred value (χ=292°). Orientation of the isopentenyl substituent remains unaffected by these changes in χ.

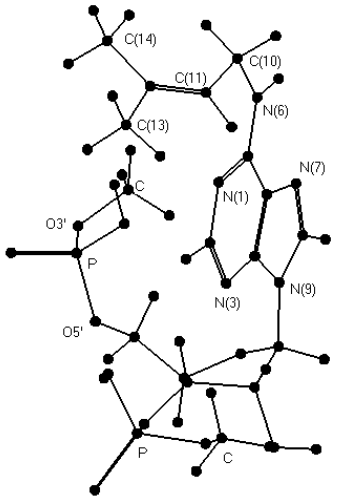


Fig. 3.6: Predicted PCILO most stable structure for Me-p-i^{6}A-p-Me in model nucleotide diphosphate segment.

Table 3.5a: Geometrical parameters for hydrogen bonding in the preferred conformations of i^{6}A, ms^{2}i^{6}A, *cis*- or *trans*- io^{6}A and *cis*- or *trans*-ms^{2}io^{6}A in model nucleotide diphosphate segment.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure Reference
N(1)…H-C(11)	2.31	1.09	90.11	3.6, 3.8, 3.10
N(3)…H-C3’	2.14	1.09	134.92	3.6, 3.8, 3.10
N(3)…H-C2’	2.18	1.09	116.55	3.6, 3.8, 3.10
N(1)…H-C(10)	2.45	1.09	96.73	3.7
N(3)…H-C5’	2.34	1.09	91.62	3.7, 3.9, 3.11
N(3)…H-C3’	2.00	1.09	136.46	3.7, 3.9, 3.11
O(13)…H-C(10)	2.99	1.09	112.59	3.9
O(13)…H-C(11)	2.17	1.09	104.59	3.10
O(13)…H-C(11)	2.34	1.09	99.75	3.11
O(13)…H-C(15)	1.17	1.09	94.22	3.11

Higher energy (1.1 kcal/mol) PCILO alternative conformation, Table 3.5b, may be reached by flipping of β to extended (β=180°) orientation. Compared to the preferred structure shown in (fig.3.6) this alternative structure lacks interaction between C(11)H and N(1). Still higher energy alternative structure at (2.9 kcal/mol) occurs when (β=90°). This alternative conformation lacks the C(11)H and N(1) interaction, but instead of this the structure is stabilized by interaction between C(10)H…N(1).

Automated geometry optimization through molecular mechanics force field (MMFF) method, starting from the preferred PCILO structure (fig. 3.6), has been carried out. The optimized torsion angles are α=-19°, β=-68°, γ=-77°, δ=-173°, ψ1=119°, ψ2=-116°, and $\chi_{(37)}$=-39.

Table 3.5b**:** Predicted alternative stable conformations for i^6A and *cis*- or *trans*- io^6A or *cis*- or *trans*-zeatin in model nucleotide diphosphate segment of anticodon loop.

Torsion Angles	Rel. Energy (kcal/mol.)	Fig. Ref.
α=0°,β=0°,γ=±120°,δ=180°,ψ1=0°,±120°,ψ2=0°,±120°,χ=292°	0.0	3.6
α=0°,β=180°,γ=±120°, δ=180°,ψ1=0°,±120°,ψ2=0°,±120°,χ=292°	1.1	
α=0°,β=90°,γ=±120°,δ=180°,ψ1=0°,±120°,ψ2=0°,±120°,χ=292°	2.9	
α=0°, β=0°, γ=±120°, δ=0°, ψ1=180°,ψ2=0°,±120°, θ=180°,χ=292°	0.0	3.8
α=0°, β=0°, γ=±120°, δ=0°, ψ1=330°,ψ2=0°,±120°, θ=180°,χ=292°	0.9	
α=0°, β=180°, γ=±120°, δ=0°,ψ1=180°,ψ2=0°,±120°,θ=180°,χ=292°	1.0	
α=0°, β=0°,γ=±120°,δ=180°,ψ1=0°,ψ2=0°,±120°,θ=180°,χ=292°	0.0	3.10
α=0°, β=0°, γ=±120°, δ=180°, ψ1=180°, ψ2=0°,±120°,θ=180°,χ=292°	0.1	
α=0°, β=180°, γ=±120°, δ=180°,ψ1=0°, ψ2=0°,±120°,θ=180°,χ=292°	0.5	

3.3.2.2 Me-p-ms^{2}i^{6}A-p-Me:

The hypermodified nucleotide 2-methylthioisopentenyl adenosine (ms^{2}i^{6}A) is incorporated in tRNA model diphosphate segment Me-p-ms^{2}i^{6}A-p-Me. The glycosyl torsion angle value, corresponding to 37th nucleoside in tRNA^{Phe} model [10] is (χ=22°). The set of torsion angles for the preferred orientation of 2-methylthioisopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0,±120°, ψ2=0,±120°, ω1=0°, ω2=180°). These torsion angle values may be compared to the preferred values [8] for isolated base ms^{2}i^{6}Ade (α=0°, β=180°, γ=±60°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=180°, ω2=180°). The orientation of the isopentenyl substituent is alike in Me-p-i^{6}A-p-Me (fig. 3.6) and Me-p-ms^{2}i^{6}A-p-Me. The 2-methylthio group orientation remains unaffected when glycosyl torsion angle (χ=22°) is held at value specified in Holbrook tRNA^{Phe} model. [10].

The PCILO most stable conformation obtained when glycosyl torsion angle takes the preferred value (χ=232°) is shown in (fig.3.7). The set of new torsion angle values are (α=0°, β=150°, γ=300°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=150°, ω2=180°). These preferred torsion angle values may be compared with [8] ms^{2}i^{6}Ade (α=0°, β=180°, γ=±60°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=180°, ω2=180°). The torsion angles β and ω1 taking different values than isolated base [8] ms^{2}i^{6}Ade.

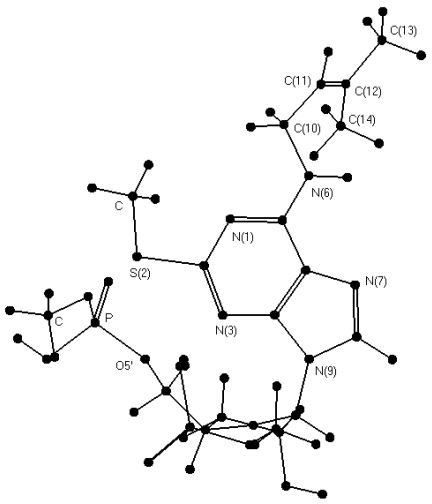


Fig.3.7: The PCILO predicted most stable conformation for Me-p-ms^{2}i^{6}A-p-Me in model nucleotide diphosphate segment.

The intramolecular hydrogen bonding between N(1)…HC(10), shown in Table 3.5a, is the stabilizing factor. However, hydrogen bonding between N(1)…HC(11) is no longer possible due to (β=150°). Interactions of ribose-phosphate backbone with the 2-methylthio isopentenyl adenine base i.e. N(3)…HC5', N(3)…HC3' are additional stabilizing factors.

When glycosyl torsion angle is held at (χ=22°), the orientation of isopentenyl substituent is specified by torsion angles (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,± 120°, ψ2=0°,±120°, ω1=0°, ω2=180°). When χ instead takes the preferred value (χ=232°), the torsion angles specifying orientation of the isopentenyl substituent are (α=0°, β=150°, γ=300°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=150°, ω2=180°). Comparing these two sets, the noticeable change is from β=0° to β=150° along with methylthio group which nearly takes *trans* orientation (ω1=150°, ω2=180°) with respect to C(2)-N(3) bond, for glycosyl torsion angle taking the preferred value (χ=232°).

When (χ=232°), the torsion angles of 2-methylthio isopentenyl substituent in model segment Me-p-ms^{2}i^{6}A-p-Me retain values similar to that predicted for the isolated modified base, ms^{2}i^{6}Ade [8], except 30° difference in β and ω1. The ribose-phosphate backbone torsion angles change to accommodate bulky 2-methylthio group. The changed ribose-phosphate backbone torsion angles are p-O5'(β_b)=180°, O5'-C5'(γ_b)=192°. Results of automated complete geometry optimization through molecular mechanics (MMFF) method starting from the preferred PCILO structure (fig. 3.7), yield α=29°, β=85°, γ=-127°, δ=174°, ψ1=5°, ψ2=11°, ω1=-153°, ω2=176°, $\chi_{(37)}$=-145° and essentially agree with the PCILO results.

3.3.2.3 *Cis*-Me-p-io^{6}A-p-Me:

In the model hypermodified nucleotide diphosphate segment *cis*-Me-p-io^{6}A-p-Me the chosen glycosyl torsion angle value from Holbrook tRNA^{Phe} model is (χ=22°). The set of torsion angles for the preferred orientation of hydroxyisopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=0°, ψ1=180°, ψ2=0°,±120°, θ=180°). These torsion angle values may be compared to the preferred values [22] for *cis*-io^{6}Ade or *cis*-zeatin (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°). The torsion angles of hydroxyisopentenyl substituent in model diphosphate hypermodified

segment retain values quite similar to isolated modified base, *cis*-io^{6}Ade [22] except for (ψ1=180°), instead of (ψ1=0°).

Allowing glycosyl torsion angle to change (fig.3.8) yields the preferred value (χ=292°). Adopting this value results in another set of preferred torsion angle values (α=0°, β=0°, γ=±120°, δ=0°, ψ1=180°, ψ2=0°,±120°, θ=180°). The torsion angles ψ1 taking different values as compared to preferred values with [22] for *cis*-io^{6}Ade or *cis*-zeatin (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0,±120°, θ=180°).

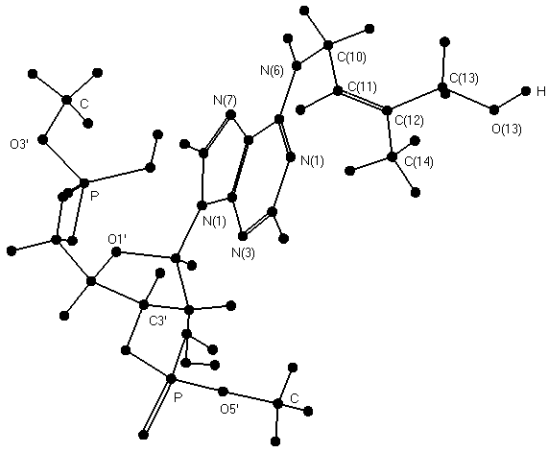


Fig. 3.8: Predicted most stable structure for *cis*-Me-p-io^{6}A-p-Me in model nucleotide diphosphate Segment.

The intramolecular hydrogen bonding between N(1)…HC(11), shown in Table 3.5a, remains a stabilizing factor. Interactions of ribose-phosphate backbone with hydroxyisopentenyl adenine base i.e. N(3)…HC3', and N(3)…HC2' are the additional stabilizing factors. The glycosyl torsion angle may be held at (χ=22°) or at the preferred values (χ=292°), but the orientation of the hydroxyisopentenyl substituent remains unaffected. Higher energy (0.9 kcal/mol) PCILO alternative conformation, presented in Table 3.5b, may be arrived by change of ψ1=330°. This structure is stabilized by interaction between N(1)…HC(11) and O(13)…HC(10). The next higher energy (1.0 kcal/mol) alternative structure may be arrived by flipping β to (β=180°). Compared to the preferred structure (fig. 3.8), this alternative conformation lacks the interaction between C(11)H and N(1). The ribose-phosphate backbone interactions C3'H…N(3), and C2'H…N(3), (Table 3.5a), may stabilize these alternative structures. Starting from PCILO most stable structure (fig. 3.8), results of geometry

optimization through molecular mechanics force field (MMFF) method are shown in Table 3.6.

Table 3.6: Optimized values of the torsion angles for model diphosphate and trinucleotide segments in anticodon loop by automatic geometry optimization using MMFF method.

Model Diphosphate Segments	Fig. Ref.	Model Trinucleotide Segments	Fig. Ref.
α=-30°,β=-80°, γ=127°, δ=-2°, ψ1=-135°, ψ2=123°,θ=177°, $\chi_{(37)}$=-44 Rel. Energy: 4.5 kcal/mol	3.8	α=31°,β=63°, γ=45°, δ=2°, ψ1=23°, ψ2=4°,θ=-179° $\chi_{(36)}$=-156, $\chi_{(37)}$=-89, $\chi_{(38)}$=-82 Rel. Energy: 3.4 kcal/mol	3.14
α=-82°,β=70°,γ=-147°, δ=-3°, ψ1=-118°, ψ2=118°,θ=-175°, ω1=-168°, ω2=177°, $\chi_{(37)}$=-74 Rel. Energy: 9.5 kcal/mol	3.9	α=25°, β=78°, γ=-125°, δ=-1°, ψ1=83° , ψ2=1°,θ=-59°, ω1=-150°, ω2=172°, $\chi_{(36)}$=-161, $\chi_{(37)}$=-88, $\chi_{(38)}$=-84 Rel. Energy: 0.0 kcal/mol	3.15
α=-36°,β=-88°,γ=-75°, δ=-173°, ψ1=4.2°, ψ2=122°,θ=177°, $\chi_{(37)}$=-1 Rel. Energy: 0.0 kcal/mol	3.10	α=-62°,β=73°,γ=-108°,δ=165°, ψ1=-94°, ψ2=-17°,θ=39°, $\chi_{(36)}$=-160, $\chi_{(37)}$=-87, $\chi_{(38)}$=-63 Rel. Energy: 0.0 kcal/mol	3.16
α=29°,β=90°γ=-129°,δ=-177°, ψ1=-17°, ψ2=125°,θ=-61°, ω1=-169°, ω2=179°, $\chi_{(37)}$=-175 Rel. Energy: 0.0 kcal/mol	3.11	α=27°, β=-64°, γ=137°, δ=-165°, ψ1=24°, ψ2=11°,θ=41°, ω1=127°, ω2=-174°, $\chi_{(36)}$=-154, $\chi_{(37)}$=-96, $\chi_{(38)}$=-69 Rel. Energy: 10.0 kcal/mol	3.17

3.3.2.4 *Cis*-Me-p-ms^{2}io^{6}A-p-Me:

The model nucleotide diphosphate anticodon loop segment *cis*-Me-p-ms^{2}io^{6}A-p-Me, has the glycosyl torsion angle value held at (χ=22°). The set of torsion angles for the preferred orientation of the base substituent in ms^{2}io^{6}A are (α=0°, β=0°, γ=±120°, δ=0°, ψ1=180°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°). These torsion angle values may be compared to the preferred values [22] for *cis*-ms^{2}io^{6}Ade (*cis*-ms^{2}zeatin) (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°). The torsion angles of hydroxyisopentenyl substituent in model diphosphate hypermodified segment retain values quite similar to isolated modified base, *cis*-ms^{2}io^{6}Ade [22] except for ψ1=180° in contrast to ψ1=0°.

Allowing glycosyl torsion angle to change yields the preferred value (χ=232°). Adopting this value results in another set of torsion angle values (α=0°, β=30°, γ=±120°, δ=0°, ψ1=270°, ψ2=0°,±120°, θ=120°, ω1=150°, ω2=180°) (fig. 3.9) for the base substituent orientation. As compared to preferred torsion angle values with [22] for *cis*-ms^{2}io^{6}Ade (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°), the torsion angle values β, ψ1, θ and ω1 preferring different values.

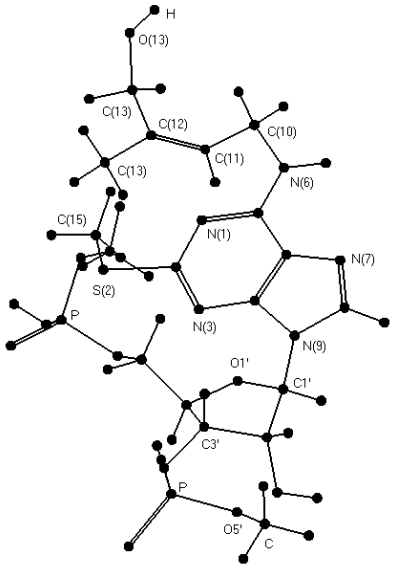


Fig. 3.9: Predicted most stable structure for *cis*-Me-p-ms^{2}io^{6}A-p-Me in model nucleotide diphosphate segment.

The expected weak hydrogen bonding, Table 3.5a, between C(10)H…O(13) may provide some stabilization, but hydrogen bonding between N(1)…HC(11) disappears when β changes to (β=30°). Interactions of ribose-phosphate backbone with 2-methylthio isopentenyl adenosine i.e. N(3)…HC5', N(3)…HC3' may be additional stabilizing factors.

The orientation of hydroxyisopentenyl substituent (α=0°, β=0°, γ=±120°, δ=0°, ψ1=180°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°) with (χ=22°) may be compared to the torsion angles (α=0°, β=30°, γ=±120°, δ=0°, ψ1=270°, ψ2=0°,±120°, θ=120°, ω1=150°, ω2=180°) suited when χ=232°. These differs in β, ψ1, θ, and ω1. The backbone torsion angles also change to accommodate the bulky 2-methylthio group when glycosyl torsion angle is held at (χ=232°). The changed backbone torsion angles are O3'-p(α_b)=47°, p-O5'(β_b)=300° and O5'-C5'(γ_b)=192°, rest of the backbone torsion angle values are same as specified in Holbrook tRNA^{Phe} model [10]. Results of

geometry optimization through molecular mechanics force field (MMFF) method starting from the PCILO most stable structure (fig.3.9) is shown in Table 3.6.

3.3.2.5 *Trans*-Me-p-io^{6}A-p-Me:

The modified nucleotide io^{6}A is incorporated in model diphosphate *trans*-Me-p-io^{6}A-p-Me segment. The glycosyl torsion angle retains the value specified in Holbrook model (χ=22°). The set of torsion angles for the preferred orientation of hydroxyisopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=180°, ψ2=0°,± 120°, θ=180°). These torsion angle values may be compared to the preferred values [22] for *trans*-io^{6}Ade (*trans*-zeatin) (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=±60°). Due to extended orientation of ψ1 and θ the –OH group of hydroxyisopentenyl substituent is placed away from the N(1) site of adenine base, thus no intramolecular interaction between O(13)H and N(1) is possible.

The predicted most stable conformation with glycosyl torsion angle (χ=292°) is shown in (fig.3.10), torsion angles for the preferred orientation of hydroxyisopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=180°). These preferred values may be compared with [22] the results for *trans*-io^{6}Ade (*trans*-zeatin) preferred orientation (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0,±120°, θ=±60°).

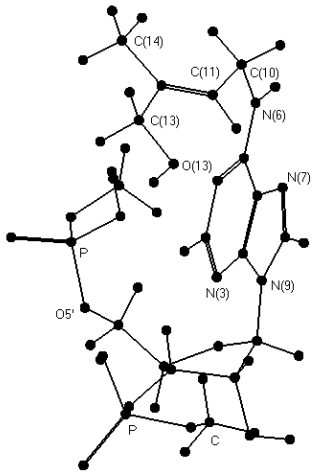


Fig.3.10: Predicted most stable conformation for *trans*-Me-p-io^{6}A-p-Me in model nucleotide diphosphate segment.

The orientation of the hydroxyisopentenyl substituent in the segment is similar to that of isolated modified base [22] *trans*-io^{6}Ade or *trans*-zeatin, except for difference in θ value. The intramolecular hydrogen bonding between N(1)…HC(11), presented in Table 3.5a, remains a stabilizing factor. Some stabilization may also accrue due to possible hydrogen bonding between O(13)…HC(11). Interactions of ribose-phosphate backbone with hydroxyisopentenyl adenine base i.e. N(3)…HC3', and N(3)…HC2' are the additional stabilizing factors.

Higher energy (0.1 kcal/mol) PCILO alternative conformation shown in Table 3.5b may be arrived at, by flipping of ψ1 to (ψ1=180°). This structure is stabilized by interaction between N(1)…HC(11), but lacks the hydrogen bonding between O(13)…HC(11). The next (0.5 kcal/mol) higher energy alternative conformation may be arrived at, by flipping of β to (β=180°). The interaction between O(13)…HC(11) is the stabilizing factor. The hydrogen bonding between C(11)H…N(1) is no longer possible. The ribose-phosphate backbone interactions between C3'H…N(3), and C2'H…N(3) remain as stabilizing factors for these alternative conformations. Compared to the preferred PCILO most stable structure for *cis*-Me-p-io^{6}A-p-Me (fig.3.8), the preferred PCILO structure (fig.3.10) for *trans*-Me-p-io^{6}A-p-Me is more stable by about 0.2 kcal/mol. By using PCILO preferred values (fig.3.10) as a starting point, results of geometry optimization through molecular mechanics force field (MMFF) method is shown in Table 3.6. The MMFF results shows that, *trans*-Me-p-io^{6}A-p-Me is stabler than *cis*-Me-p-io^{6}A-p-Me by 4.5 kcal/mol.

3.3.2.6 *Trans*-Me-P-ms^{2}io^{6}A-P-Me:

In *trans*-Me-p-ms^{2}io^{6}A-p-Me molecule, the ribose-phosphate backbone torsion angles are held at values specified in Holbrook tRNA^{Phe} model [10] for nucleotide at the 37th position. The glycosyl torsion angle value is (χ=22°), the set of torsion angles for the preferred orientation of 2-methylthiohydroxyisopentenyl substituent are (α=0°, β=0°, γ=±120°, δ=180°, ψ1=180°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°). These torsion angle values may be compared to the preferred values [22] for isolated base *trans*-ms^{2}io^{6}Ade or *trans*-ms^{2}zeatin (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=±60°, ω1=0°, ω2=180°). The torsion angles ψ1=180° and θ=180° are taking different values than preferred for isolated hypermodified base, *trans*-ms^{2}io^{6}Ade [22].

Allowing glycosyl torsion angle to change yields the preferred value (χ=232°). Taking χ=232° leads to new set of preferred torsion angle values for the base substituent orientation. These are (α=0°, β=30°, γ=±120°, δ=180°, ψ1=330°, ψ2=180,±60°, θ=30°, ω1=150°, ω2=180°) for the structure depicted in (fig. 3.11). The torsion angles β, ψ1, ψ2 θ and ω1 taking different values as compared to isolated hypermodified base *trans*-ms^{2}io^{6}Ade [22] (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=±60°, ω1=0°, ω2=180°). The geometrical parameters for hydrogen bonding in Table 3.5a show that interaction between O(13)…HC(11) and O(13)…HC(15) are the stabilizing factor. The hydrogen bonding between N(1)…HC(11) is not possible when (β=30°). Interactions of ribose-phosphate backbone with 2-methylthio isopentenyl adenine base i.e. N(3)…HC5', N(3)…HC3' may be additional stabilizing factors.

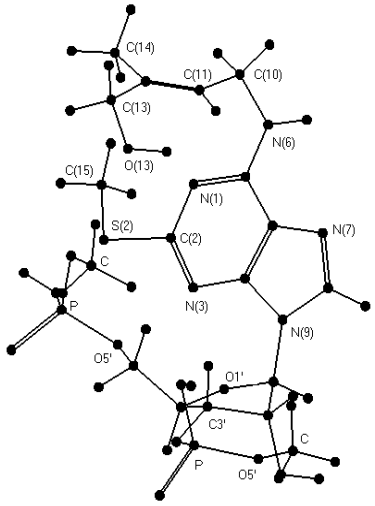


Fig.3.11: Predicted most stable conformation for *trans*-Me-p-ms^{2}io^{6}A-p-Me in model nucleotide diphosphate segment.

Comparing, the orientation of hydroxyisopentenyl substituent (α=0°, β=0°, γ=±120°, δ=180°, ψ1=180°, ψ2=0, ±120°, θ=180°, ω1=0°, ω2=180°) when glycosyl torsion angle is held at (χ=22°) with the results when χ takes the preferred value (χ=232°), the notable differences are found in β, ψ1, θ, and ω1. The ribose-phosphate backbone torsion angles also change to accommodate bulky 2-methylthio group. The changed ribose-phosphate backbone torsion angles are O3'-p(α_b)=47°, p-O5'(β_b)=300° and O5'-C5'(γ_b)=192°. However, rest of the ribose-

phosphate backbone torsion angle values remain the same as specified in Holbrook $tRNA^{Phe}$ model [10].

Although favored orientation of 2-methylthio group is (ω1=150°, ω2=180°) the hydroxyisopentenyl substituent in the model segment shows similar conformation as favored for isolated modified base *trans*-ms^2io^6Ade [22]. This is in contrast to the preferred orientation for isolated modified base ms^2i^6Ade [8].

Results of geometry optimization through molecular mechanics force field (MMFF) method presented in Table 3.6, show that, *trans*-Me-p-ms^2io^6A-p-Me is more stable (9.5 kcal/mol) than *cis*-Me-p-ms^2io^6A-p-Me. The preferred PCILO structure for *cis*-Me-p-ms^2io^6A-p-Me shown in (fig.3.9) is more stable than the preferred PCILO stable structure for *trans*-Me-p-ms^2io^6A-p-Me (fig.3.11).

3.3.3 The model trinucleotide hypermodified segments

3.3.3.1 Me-p-$A_{(36)}$-p-$i^6A_{(37)}$-p-$A_{(38)}$-p-Me:

In order to consider the interactions of i^6A_{37} with 5' and 3' adjacent nucleosides in tRNA, i^6A is incorporated into the model trinucleotide hypermodified segment. The corresponding glycosyl torsion angles are taken from the crystal structure data of Holbrook ($\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The ribose-phosphate backbone torsion angle values are also likewise adopted from the crystal structure data. The set of torsion angle values for the preferred orientation of isopentenyl substituent are (α=0°, β=90°, γ=60°, δ=180°, ψ1=0°,±120°, ψ2=30°,150°,270°). These predicted torsion angle values may be compared with the preferred values [8] for i^6Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120). The conformation of isopentenyl substituent (i^6A_{37}) in trinucleotide segment changes remarkably with respect to β and γ. The torsion angle values (β=90° and γ=60°) for trinucleotide segment in contrast to (β=0°, γ=±120°) for isolated hypermodified base, i^6Ade [8]. This structure does not show the intramolecular hydrogen bonding between $N(1)_{(37)}$…HC(11) .

When the glycosyl torsion angle for nucleotides in the segment is varied freely, the resulting preferred values are ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=243°). Adopting these preferred values results in a new set of torsion angle values for the preferred isopentenyl substituent orientation (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=180°,±60°) which is shown in (fig.3.12). These new preferred torsion angle values for isopentenyl substituent may be compared with the preferred orientation of the

hypermodified isolated base [8] i^6Ade (α=0°, β=0°, γ=±120°, δ=180°,ψ1=0°,±120°, ψ2=0°, ±120°).

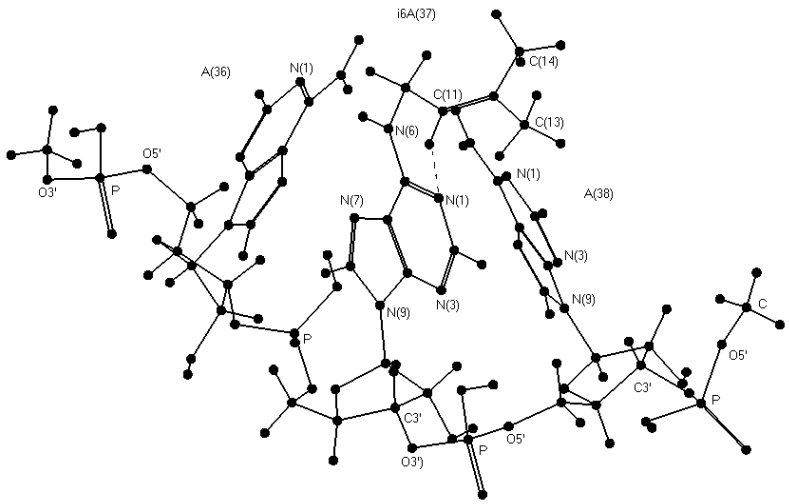


Fig. 3.12: Predicted PCILO most stable conformation for Me-p-$A_{(36)}$-p-$i^6A_{(37)}$-p-$A_{(38)}$-p-Me in trinucleotide segment.

The intramolecular hydrogen bonding Table 3.7(a) between $N(1)_{(37)}$…HC(11) is the stabilizing factor for isopentenyl substituent. Other interesting interactions between $N(6)_{(36)}$…$HN(6)_{(37)}$, $N(1)_{(36)}$…$HN(6)_{(37)}$, $N(7)_{(36)}$…$HC(8)_{(37)}$, are additional stabilizing factors. The interactions between ribose-phosphate backbone and bases, noted in Table 3.7a, namely $N(3)_{(37)}$…$HC3'_{(37)}$, $N(3)_{(37)}$…$HC2'_{(37)}$, $N(3)_{(36)}$…$HC3'_{(36)}$, $N(3)_{(36)}$…$HC5'_{(36)}$ and $N(3)_{(38)}$…$HC3'_{(38)}$ may also stabilize the structure. So the orientation of isopentenyl substituent in model trinucleotide segment having i^6A_{37} with the preferred values for glycosyl torsion angles ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=243°) results in isopentenyl substituent orientation similar to that of isolated hypermodified base i^6Ade [8], except torsion angle (ψ2=180°,±60°) taking different values.

Higher energy (1.1 kcal/mol), PCILO most stable alternative conformation, Table 3.7(b), may be arrived at by flipping of β to β=180° along with taking γ=120°. In this case the intramolecular interactions between $N(1)_{(37)}$…HC(11) is not indicated. Next higher energy alternative conformation at (1.7 kcal/mol) has β=180° and γ=240°. Interactions between ribose-phosphate backbone and adjacent bases are $N(3)_{(37)}$…$HC3'_{(37)}$, $N(3)_{(37)}$…$HC2'_{(37)}$, $N(3)_{(36)}$…$HC3'_{(36)}$, $N(3)_{(36)}$…$HC5'_{(37)}$ and

$N(3)_{(38)}$…$HC3'_{(38)}$ as included in Table 3.7(a). These may stabilize some of the alternative conformations.

Table 3.7(a): Geometrical parameters for hydrogen bonding in the preferred conformations of i^6A and ms^2i^6A in model trinucleotide segment.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure References
$N(1)_{(37)}$…H-$C(11)_{(37)}$	2.11	1.09	89.14	3.12
$N(6)_{(36)}$…H-$N(6)_{(37)}$	2.68	1.09	153.26	3.12
$N(1)_{(36)}$…H-$N(6)_{(37)}$	1.79	1.09	144.01	3.12
$N(7)_{(36)}$…H-$C(8)_{(37)}$	2.52	1.09	101.02	3.12
$N(3)_{(37)}$…H-$C3'_{(37)}$	2.32	1.09	138.57	3.12
$N(3)_{(37)}$…H-$C2'_{(37)}$	2.41	1.09	122.92	3.12
$N(3)_{(36)}$…H-$C3'_{(36)}$	2.70	1.09	115.57	3.12
$N(3)_{(36)}$…H-$C5'_{(36)}$	1.94	1.09	129.42	3.12
$N(3)_{(38)}$…H-$C3'_{(38)}$	2.87	1.09	105.31	3.12
$N(1)_{(37)}$…H-$C(10)_{(37)}$	2.49	1.09	96.58	3.13
$N(7)_{(38)}$…H-$C(15)_{(37)}$	2.28	1.09	134.59	3.13
$N(1)_{(37)}$…H-$N(6)_{(38)}$	2.56	1.09	109.59	3.13
$S(2)_{(37)}$…H-$N(6)_{(38)}$	2.34	1.09	171.43	3.13
$N(6)_{(38)}$…H-$C(14)_{(37)}$	2.75	1.09	141.17	3.13
$N(7)_{(37)}$…H-$C(8)_{(36)}$	2.36	1.09	94.79	3.13
$N(6)_{(37)}$…H-$N(6)_{(36)}$	2.71	1.09	149.27	3.13
$N(1)_{(37)}$…H-$N(6)_{(36)}$	2.19	1.09	110.73	3.13
$N(3)_{(38)}$…H-$C2'_{(38)}$	2.70	1.09	115.18	3.13
$O5'_{(37)}$…H-$C(8)_{(37)}$	2.52	1.09	159.84	3.13

Results of geometry optimization through molecular mechanics force field (MMFF) method for isopentenyl substituent orientation are α=26°, β=-71°, γ=122°, δ=-168°, ψ1=0.3°, ψ2=19°, and for glycosyl torsion angles are $\chi_{(36)}$=-162°, $\chi_{(37)}$=-85°, $\chi_{(38)}$=-75° and are in broad agreement with the PCILO results.

Table 3.7(b): Predicted alternative stable conformations for i^6A and 2-methylthio analog, ms^2i^6A, in model trinucleotide hypermodified segment.

Torsion Angles	Rel. Energy (kcal/mol.)	Fig. Ref.
α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ±120°, ψ2=180°, ±60°	0.0	3.12
α=0°, β=180°, γ=120°, δ=180°, ψ1=0°, ±120°, ψ2=180°,±60°	1.1	
α=0°, β=180°, γ=240°, δ=180°, ψ1=0°, ±120°, ψ2=180°, ±60°	1.7	
α=0°,β=210°,γ=90°,δ=180°,ψ1=0°,±120°,ψ2=0°,±120°,ω1=0°,ω2=180°	0.0	3.13
α=0°,β=210°,γ=180°,δ=180°,ψ1=0°,±120°,ψ2=0°,±120°,ω1=0°,ω2=180°	0.1	
α=0°,β=210°,γ=90°,δ=180°,ψ1=0°,±120°,ψ2=0°,±120°,ω1=180°,ω2=180°	4.2	

3.3.3.2 Me-p-$A_{(36)}$-p-$ms^2i^6A_{(37)}$-p-$A_{(38)}$-p-Me:

The hypermodified nucleoside $ms^2i^6A_{(37)}$ is included in model trinucleotide segment, Me-p-$A_{(36)}$-p-$ms^2i^6A_{(37)}$-p-$A_{(38)}$-p-Me. The adopted glycosyl torsion angle values are as in Holbrook $tRNA^{Phe}$ crystal structure data ($\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angle values for the preferred orientation of 2-methylthioisopentenyl substituent are (α=0°, β=90°, γ=60°, δ=180°, ψ1=0°,±120°, ψ2=30°,150°,270°, ω1=60°, ω2=150°). These torsion angle values may be compared with the preferred values [8] for isolated ms^2i^6Ade (α=0°, β=180°, γ=±60°, δ=180°, ψ1=0°,±120, ψ2=0°,±120°, ω1=180°, ω2=180°). The conformation of 2-methylthio isopentenyl substituent $ms^2i^6A_{37}$ in trinucleotide segment changes remarkably with respect to torsion angles β, ω1 and ω2 (β=90°, ω1=60°, ω2=150°) in contrast to (β=180°, ω1=180°, ω2=180°) for isolated hypermodified base, ms^2i^6Ade [8].

The preferred most stable structure shown in (fig.3.13) corresponds to glycosyl torsion angles taking preferred values ($\chi_{(36)}$=61°, $\chi_{(37)}$=52°, $\chi_{(38)}$=273°). Adopting these preferred glycosyl torsion angle values, leads to another preferred isopentenyl substituent orientation shown in (fig. 3.13) having torsion angle values (α=0°, β=210°, γ=90°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=0°, ω2=180°). The torsion angles β, γ and ω1 taking different values as compared with the preferred orientation of the isolated hypermodified base [8] ms^2i^6Ade (α=0°, β=180°, γ=±60°, δ=180°,ψ1=0°,±120°, ψ2=0°,±120, ω1=180°, ω2=180°).

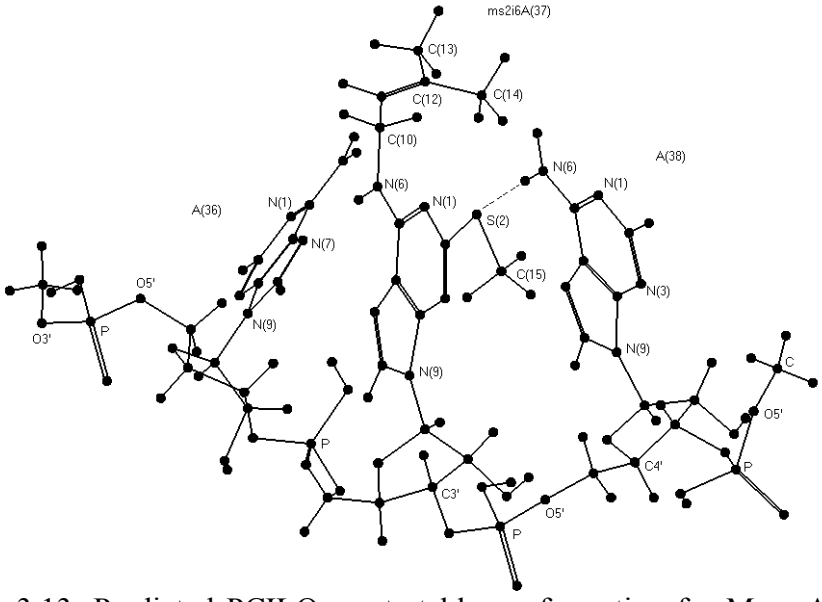


Fig. 3.13: Predicted PCILO most stable conformation for Me-p-$A_{(36)}$-p-$ms^2i^6A_{(37)}$-p-$A_{(38)}$- p-Me in trinucleotide segment.

The intramolecular hydrogen bonding included in Table 3.7(a) between C(10)H…$N(1)_{(37}$ is a stabilizing factor for the preferred 2-methyltho isopentenyl substituent orientation. Other interesting interactions between adjacent bases in trinucleotide are $N(7)_{(38)}$…$HC(15)_{(37)}$, $N(1)_{(37)}$…$HN(6)_{(38)}$, $S(2)_{(37)}$…$HN(6)_{(38)}$, $N(6)_{(38)}$…$HC(14)_{(37)}$, $N(7)_{(37)}$…$HC(8)_{(36)}$, $N(6)_{(37)}$…$HN(6)_{(36)}$ and $N(1)_{(37)}$…$HN(6)_{(36)}$. These may be additional stabilizing factors. Besides these, the interactions between ribose-phosphate backbone and bases in Table 3.7(a) namely $N(3)_{(38)}$…$HC2'_{(38)}$ and $O5'_{(37)}$…$HC(8)_{(37)}$, may also stabilize the structure shown in (fig. 3.13). Thus the preferred orientation of isopentenyl substituent in trinucleotide segment having $ms^2i^6A_{37}$, with the preferred values for glycosyl torsion angles ($\chi_{(36)}$=61°, $\chi_{(37)}$=52°, $\chi_{(38)}$=273°), is quite similar (except β, γ and ω1), to that of isolated modified base ms^2i^6Ade [8]. However, the 2-methylthio group takes different orientation (ω1=0°, ω2=180°). For trinucleotide segment having i^6A_{37}, the preferred glycosyl torsion angles are ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=243°) but the corresponding values for trinucleotide segment having $ms^2i^6A_{37}$ are ($\chi_{(36)}$=61°, $\chi_{(37)}$=52°, $\chi_{(38)}$=273°). These differences in orientation of the bases are due to the presence of bulky 2-methylthio group in trinucleotide segment having $ms^2i^6A_{37}$.

Higher energy (0.1 kcal/mol), PCILO stable alternative conformation presented in Table 3.7(b) is arrived at by flipping of γ to γ=180°. The intramolecular interactions between $N(1)_{(37)}$…HC(10) and $N(6)_{(38)}$…$HC(15)_{(37)}$, may stabilize

orientation of 2-methylthio isopentenyl substituent. The next higher energy alternative conformation at (4.2 kcal/mol) corresponds to extended ω1 orientation, ω1=180°. Interactions between $N(1)_{(37)}$…HC(10) and $N(6)_{(38)}$…$HC(15)_{(37)}$ are stabilizing factors. Starting from PCILO preferred values (fig. 3.13), geometry optimization through molecular mechanics force field (MMFF) method yields torsion angles values α=13°, β=152°, γ=80°, δ=-179°, ψ1=-0.8°, ψ2=-49°, ω1=-71°, ω2=167°, and the glycosyl torsion angles are $\chi_{(36)}$= -31°, $\chi_{(37)}$= 32°, $\chi_{(38)}$= -36°.

3.3.3.3 *Cis*-Me-p-$A_{(36)}$-p-$io^6A_{(37)}$-p-$A_{(38)}$-p-Me:

In the model trinucleotide hypermodified segment, *cis*-Me-p-$A_{(36)}$-p-$io^6A_{(37)}$-p-$A_{(38)}$-p-Me, for nucleotides at the 36th, 37th, and 38th positions in tRNA, the glycosyl torsion angles take corresponding values as specified in Holbrook model ($\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angle values for the preferred orientation of hydroxyisopentenyl substituent are (α=0°, β=90°, γ=60°, δ=0°, ψ1=210°, ψ2=0°,±120°, θ=90°). These torsion angle values may be compared with the preferred [22] values for isolated modified base, *cis*-io^6Ade (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°). The conformation of hydroxyisopentenyl substituent (io^6A_{37}) in trinucleotide segment is remarkably different with respect to β, γ and ψ1 (β=90°, γ=60°, and ψ1=210°) as compared to (β=0°, γ=±120°, ψ1=0°) for isolated hypermodified base, *cis*-io^6Ade [22]. The preferred orientation of the model trinucleotide segment does not show the intramolecular hydrogen bonding between $N(1)_{(37)}$…HC(11).

When the glycosyl torsion angles for nucleotides in the model trinucleotide segment are varied freely, the resulting preferred values are ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=243°). Adopting these preferred glycosyl torsion angle values leads to another set of torsion angles for the preferred hydroxyisopentenyl substituent orientation (in fig. 3.14) (α=0°, β=30°, γ=90°, δ=0°,ψ1=330°, ψ2=0°,±120°, θ=180°). The torsion angle values β, γ and ψ1 taking different values as compared with the preferred orientation of the isolated hypermodified base [22] *cis*-io^6Ade (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°).

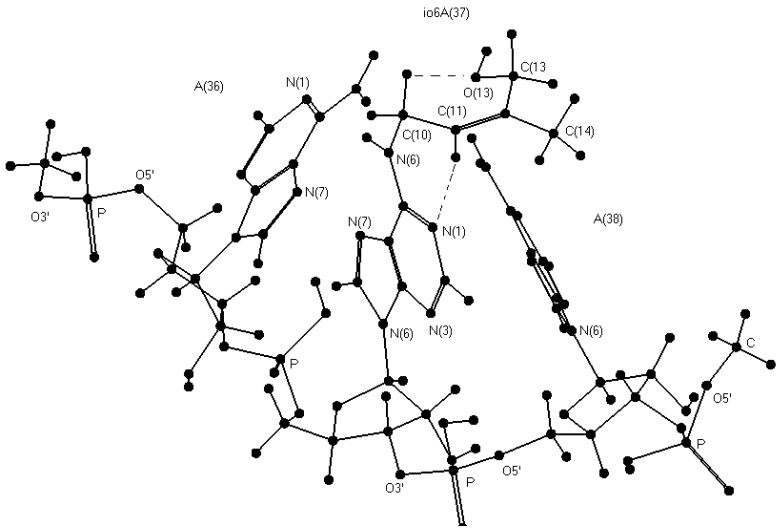


Fig. 3.14: Predicted PCILO most stable structure for *cis*-Me-p-$A_{(36)}$-p-$io^6A_{(37)}$-p-$A_{(38)}$-p-Me in trinucleotide segment.

The intramolecular hydrogen bonding shown in Table 3.8(a) between $N(1)_{(37)}$…$HC(11)_{(37)}$, $O(13)$…$HC(10)_{(37)}$, $O(13)_{(37)}$…$HN(6)_{(38)}$, and $N(1)_{(38)}$…$HC(14)_{(37)}$ may stabilize hydroxyisopentenyl substituent orientation. Other interesting interactions between adjacent bases $N(6)_{(36)}$…$HN(6)_{(37)}$, $N(6)_{(37)}$…$HN(6)_{(38)}$, $N(1)_{(36)}$…$HN(6)_{(37)}$, $N(7)_{(36)}$…$HC(8)_{(37)}$ may provide additional stabilization. The interactions between ribose-phosphate backbone and bases in Table 3.8(a), namely, $N(3)_{(37)}$…$HC3'_{(37)}$, $N(3)_{(37)}$…$HC2'_{(37)}$, $N(3)_{(36)}$…$HC3'_{(36)}$, $N(3)_{(36)}$…$HC5'_{(36)}$, $N(3)_{(38)}$…$HC5'_{(38)}$, and $N(3)_{(38)}$…$HC3'_{(38)}$ may also stabilize the structure. The orientation of hydroxyisopentenyl substituent io^6A_{37} keeping the preferred values for glycosyl torsion angles ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=243°) results in orientation somewhat similar to that of isolated hypermodified base [22] *cis*-io^6Ade, but torsion angles β, γ, and ψ1 have different values.

Higher energy (1.8 kcal/mol) PCILO most stable alternative conformation shown in Table 3.8(b) may be reached to by flipping to γ=210° along with taking θ=300°. The intramolecular interactions between $N(1)_{(37)}$…HC(11), $O(13)$…$HC(10)_{(37)}$ and $N(1)_{(38)}$…$HC(11)_{(37)}$ provide stability to the hydroxyisopentenyl substituent. The next (1.9 kcal/mol) higher energy alternative conformation is arrived with (ψ1=210°, θ=300°). The hydrogen bonding between N(1)…$HC(11)_{(37)}$ remains as a stabilizing factor, but the interaction between $O(13)$…$HC(10)_{(37)}$ is not feasible. The interactions between ribose-phosphate

backbone remain possible stabilizing factors for these alternative conformations shown in Table 3.8(a). Results of geometry optimization through molecular mechanics force field (MMFF) method are included in Table 3.6

Table 3.8(a): Geometrical parameters for hydrogen bonding in the preferred conformations of *cis*- or *trans*- io^{6}A in model trinucleotide segment.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure Reference
$N(1)_{(37)}$…H-$C(11)_{(37)}$	2.39	1.09	85.68	3.14
$O(13)_{(37)}$…H-$C(10)_{(37)}$	2.29	1.09	106.40	3.14
$O(13)_{(37)}$…H-$N(6)_{(38)}$	1.67	1.09	152.86	3.14
$N(1)_{(38)}$…H-$C(14)_{(37)}$	2.88	1.09	108.91	3.14
$N(1)_{(36)}$…H-$N(6)_{(37)}$	1.79	1.09	144.00	3.14, 3.16
$N(6)_{(37)}$…H-$N(6)_{(38)}$	2.64	1.09	123.28	3.14, 3.16
$N(7)_{(36)}$…H-$C(8)_{(37)}$	2.52	1.09	101.02	3.14, 3.16
$N(6)_{(36)}$…H-$N(6)_{(37)}$	2.68	1.09	153.26	3.14, 3.16
$N(1)_{(37)}$…H-$C(11)_{(37)}$	2.32	1.09	89.80	3.16
$N(1)_{(38)}$…H-$C(11)_{(37)}$	2.35	1.09	152.76	3.16
$N(1)_{(38)}$…H-$O(13)_{(37)}$	2.73	1.09	124.10	3.16
$O(13)_{(37)}$…H-$C(2)_{(38)}$	2.83	1.09	113.29	3.16
$O(13)_{(37)}$…H-$C(11)_{(37)}$	2.32	1.09	108.79	3.16
$N(3)_{(37)}$…H-$C3'_{(37)}$	2.32	1.09	138.23	3.14, 3.16
$N(3)_{(37)}$…H-$C2'_{(37)}$	2.40	1.09	120.39	3.14, 3.16
$N(3)_{(36)}$…H-$C3'_{(36)}$	2.70	1.09	115.57	3.14, 3.16
$N(3)_{(36)}$…H-$C5'_{(36)}$	1.94	1.09	129.42	3.14, 3.16
$N(3)_{(38)}$…H-$C5'_{(38)}$	2.66	1.09	127.93	3.14, 3.16
$N(3)_{(38)}$…H-$C3'_{(38)}$	2.53	1.09	101.68	3.14, 3.16

Table 3.8(b): Predicted alternative stable conformations for *cis*- or *trans*- io^6A and their 2-methylthio analogs in model trinucleotide segment.

Torsion Angles	Rel. Energy (kcal/mol.)	Fig. Ref.
α=0°, β=30°, γ=90°, δ=0°, ψ1=330°, ψ2=0°, ±120°, θ=180°,	0.0	3.14
α=0°, β=30°, γ=210°, δ=0°, ψ1=330°, ψ2=0°, ±120°, θ=300°,	1.8	
α=0°, β=30°, γ=90°, δ=0°, ψ1=210°, ψ2=0°, ±120°, θ=300°,	1.9	
α=0°,β=90°,γ=240°,δ=0°,ψ1=60°,ψ2=0°,±120°,θ=330°,ω1=240°,ω2=180°	0.0	3.15
α=0°,β=30°,γ=240°,δ=0°,ψ1=180°,ψ2=0°,±120°,θ=330°,ω1=240°,ω2=180°	0.2	
α=0°,β=30°,γ=240°,δ=0°,ψ1=180°,ψ2=0°,±120°,θ=180°,ω1=240°,ω2=180°	0.7	
α=0°, β=30°, γ=210°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=300°	0.0	3.16
α=0°, β=30°, γ=90°, δ=180°, ψ1=150°, ψ2=0°,±120°, θ=300°	1.1	
α=0°, β=30°, γ=90°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=300°	1.8	
α=0°,β=0°,γ=120°,δ=180°,ψ1=0°,ψ2=180°,±60°,θ=60°,ω1=90°,ω2=150°	0.0	3.17
α=0°,β=0°,γ=240°,δ=180°,ψ1=0°,ψ2=180°,±60°,θ=300°,ω1=±90°,ω2=150°	0.6	
α=0°,β=0°,γ=120°,δ=180°,ψ1=0°,ψ2=180°,±60°,θ=60°,ω1=270°,ω2=150°	1.7	

3.3.3.4 *Cis*-Me-p-$A_{(36)}$-p-$ms^2io^6A_{(37)}$-p-$A_{(38)}$-p-Me:

The hypermodified nucleoside $ms^2io^6A_{37}$, is included in the model trinucleotide segment, *cis*-Me-p-$A_{(36)}$-p-$ms^2io^6A_{(37)}$-p-$A_{(38)}$-p-Me. The glycosyl torsion angles take values as specified in Holbrook model ($\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angle values for the preferred orientation of the base substituent are (α=0°, β=90°, γ=240°, δ=0°, ψ1=60°, ψ2=0°,±120°, θ=330°, ω1=60°, ω2=90°). These torsion angle values may be compared with [22] the preferred values for isolated *cis*-ms^2io^6Ade (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°, ω1=0°, ω2=180°). The conformation of hydroxyisopentenyl substituent (io^6A_{37}) in trinucleotide segment changes remarkably with respect to β, ψ1, θ, ω1 and ω2 for trinucleotide in contrast to isolated hypermodified base, *cis*-ms^2io^6Ade [22].

When the glycosyl torsion angles for nucleotides in the segment are varied freely, the resulting preferred values are ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=273°). Adopting these preferred values leads to change in the preferred orientation of the base substituent (fig.3.15), this is specified by new set of torsion angle values (α=0°, β=90°, γ=240°, δ=0°, ψ1=60°, ψ2=0°±120°, θ=330°, ω1=240°, ω2=180°). As compared to the preferred orientation of the isolated hypermodified base [22] *cis*-ms^2io^6Ade (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°±120°, θ=180°, ω1=0°, ω2=180°), the torsion angles β, ψ1, θ and ω1 taking different values in trinucleotide hypermodified segment..

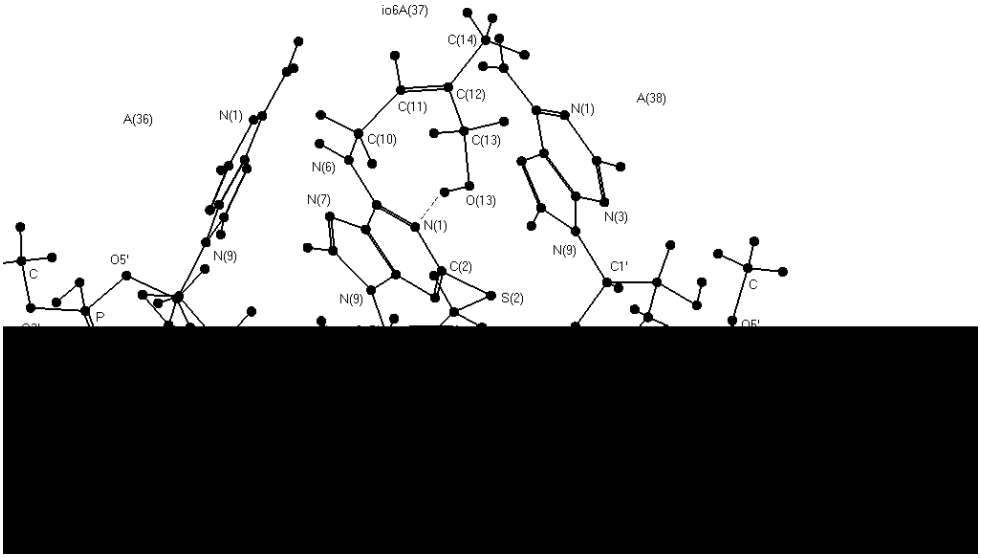


Fig.3.15: Predicted PCILO most stable conformation for *cis*-Me-p-$A_{(36)}$-p-$ms^2io^6A_{(37)}$-p-$A_{(38)}$-p-Me in trinucleotide segment.

The intramolecular hydrogen bonding interaction features noted in Table 3.9 between $N(1)_{(36)}$…$HC(10)_{(37)}$, $N(1)_{(37)}$…$HC(10)_{(37)}$, $O(13)_{(37)}$…$HC(10)_{(37)}$, $O(13)_{(37)}$…$HC(15)_{(37)}$, $S(2)_{(37)}$…$HO(13)_{(37)}$, $N(7)_{(36)}$…$HC(8)_{(37)}$ and strong interaction between $N(1)_{(37)}$…$HO(13)_{(37)}$ are the stabilizing factors for 2-methylthiohydroxyisopentenyl substituent orientation. Other interesting interactions between adjacent bases $N(6)_{(36)}$…$HN(6)_{(37)}$, and $N(1)_{(36)}$…$HN(6)_{(37)}$ are the additional stabilizing factors. The interactions between ribose-phosphate backbone and bases, Table 3.9, namely $N(3)_{(37)}$…$HC3'_{(37)}$, $N(3)_{(37)}$…$HC2'_{(37)}$, $N(3)_{(36)}$…$HC3'_{(36)}$, $N(3)_{(36)}$…$HC5'_{(36)}$, $N(3)_{(38)}$…$HC2'_{(38)}$, and $N(3)_{(38)}$…$HC3'_{(38)}$ may also stabilize the structure.

Table 3.9: Geometrical parameters for hydrogen bonding in the preferred conformations of *cis*- or *trans*- ms^2io^6A in model trinucleotide segment.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure Reference
$N(1)_{(36)}$…H-$C(10)_{(37)}$	2.66	1.09	104.54	3.15
$N(1)_{(37)}$…H-$C(10)_{(37)}$	2.49	1.09	96.58	3.15
$O(13)_{(37)}$…H-$C(10)_{(37)}$	2.46	1.09	122.49	3.15
$O(13)_{(37)}$…H-$C(15)_{(37)}$	2.43	1.09	119.88	3.15
$S(2)_{(37)}$…H-$O(13)_{(37)}$	2.84	0.96	92.18	3.15
$N(7)_{(36)}$…H-$C(8)_{(37)}$	2.52	1.09	101.00	3.15, 3.17
$N(1)_{(37)}$…H-$O(13)_{(37)}$	1.67	0.96	156.30	3.15
$N(1)_{(36)}$…H-$N(6)_{(37)}$	1.72	1.09	138.56	3.15, 3.17
$N(1)_{(37)}$…H-$C(11)_{(37)}$	2.11	1.09	89.14	3.17
$N(1)_{(37)}$…H-$O(13)_{(37)}$	2.75	0.96	133.85	3.17
$N(1)_{(36)}$…H-$C(10)_{(37)}$	2.87	1.09	92.44	3.17
$O(13)_{(37)}$…H-$C(11)_{(37)}$	2.32	1.09	108.79	3.17
$S(2)_{(37)}$…H-$O(13)_{(37)}$	2.07	1.09	157.28	3.17
$N(6)_{(36)}$…H-$N(6)_{(37)}$	2.57	1.09	151.55	3.15, 3.17
$N(3)_{(37)}$…H-$C3'_{(37)}$	2.41	1.09	140.10	3.15, 3.17
$N(3)_{(37)}$…H-$C2'_{(37)}$	2.52	1.09	123.67	3.15, 3.17
$N(3)_{(36)}$…H-$C3'_{(36)}$	2.70	1.09	115.57	3.15, 3.17
$N(3)_{(36)}$…H-$C5'_{(36)}$	1.94	1.09	129.42	3.15, 3.17
$N(3)_{(38)}$…H-$C3'_{(38)}$	2.83	1.09	90.53	3.15, 3.17
$N(3)_{(38)}$…H-$C2'_{(38)}$	2.20	1.09	109.50	3.15, 3.17

Orientation of hydroxyisopentenyl substituent remains unaffected when the preferred values for glycosyl torsion angles are used ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=273°), instead of keeping the glycosyl torsion angles ($\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°) as in tRNA crystal structure model. Only difference is in the orientation of the bulky 2-methylthio group (ω1=240°, ω2=180°), when glycosyl torsion angles take the preferred values ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=273°). In either case, regardless of

the choice of glycosyl torsion angles, interaction between $N(1)_{(37)}\ldots HC(11)_{(37)}$ is absent.

Higher energy (0.2 kcal/mol) PCILO stable alternative conformation shown in Table 3.8(b) may be arrived with $\beta=30°$, $\psi1=180°$. The hydroxyisopentenyl substituent in this case is stabilized by interactions between $O(13)_{(37)}\ldots HC(15)_{(37)}$. Next higher energy (0.7 kcal/mol) alternative conformation is reached by changing $\beta=30°$, $\psi1=180°$ along with $\theta=180°$. The interaction between $O(13)_{(37)}\ldots HC(15)_{(37)}$ remains a stabilizing factor. Other ribose-phosphate backbone interactions in Table 3.9 remain additional stabilizing factors for these alternative conformations. Results of geometry optimization through molecular mechanics force field (MMFF) method are included in Table 3.6.

3.3.3.5 *Trans*-Me-p-$A_{(36)}$-p-$io^6A_{(37)}$-p-$A_{(38)}$-p-Me:

In the model trinucleotide hypermodified segment, *trans*-Me-p-$A_{(36)}$-p-$io^6A_{(37)}$-p-$A_{(38)}$-p-Me, the glycosyl torsion angles retain the specified values in Holbrook model ($\chi_{(36)}=1°$, $\chi_{(37)}=22°$, $\chi_{(38)}=3°$). The set of torsion angle values for the preferred orientation of the base substituent are ($\alpha=0°$, $\beta=90°$, $\gamma=60°$, $\delta=180°$, $\psi1=0°$, $\psi2=30°,150°,270°$, $\theta=300°$). These torsion angle values may be compared with the preferred orientation [22] for isolated *trans*-io^6Ade ($\alpha=0°$, $\beta=0°$, $\gamma=\pm120°$, $\delta=180°$, $\psi1=0°$, $\psi2=0°,\pm120°$, $\theta=\pm60°$). The conformation of hydroxyisopentenyl substituent io^6A_{37} in trinucleotide segment changes remarkably with respect to β and γ ($\beta=90°$ and $\gamma=60°$) instead of ($\beta=0°$, $\gamma=\pm120°$) in isolated hypermodified base, *trans*-io^6Ade [22]. The intramolecular hydrogen bonding between $N(1)_{(37)}\ldots HC(11)$ is not feasible in this case.

The preferred most stable conformation occurs, when the glycosyl torsion angles for nucleotides in the trinucleotide segment are varied freely (fig.3.16). The resulting preferred glycosyl torsion angle values are ($\chi_{(36)}=211°$, $\chi_{(37)}=292°$, $\chi_{(38)}=243°$). Adopting these preferred values in the segment leads to results a new set of torsion angle values for the preferred hydroxyisopentenyl substituent orientation (fig.3.16) ($\alpha=0°$, $\beta=30°$, $\gamma=210°$, $\delta=180°$,$\psi1=0°$, $\psi2=0°,\pm120$, $\theta=300°$). These torsion angles may be compared with the values specifying preferred orientation of the isolated hypermodified base [22] *trans*-io^6Ade ($\alpha=0°$, $\beta=0°$, $\gamma=\pm120°$, $\delta=180°$, $\psi1=0°$,

ψ2=0°,±120°, θ=±60°). The torsion angles β and γ have different values in trinucleotide hypermodified segment.

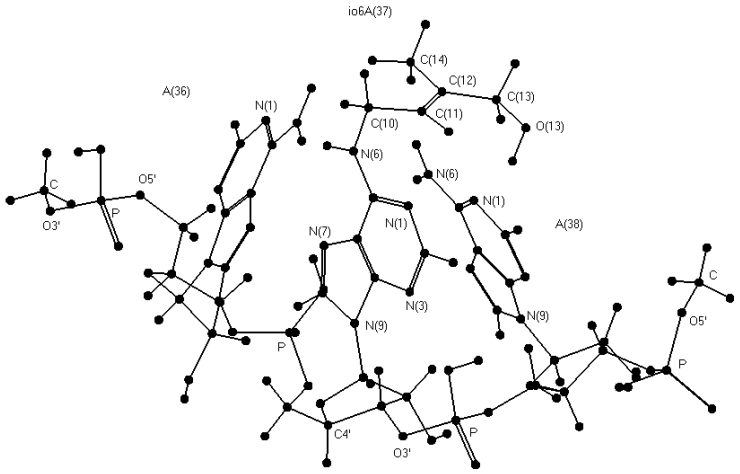


Fig.3.16: Predicted PCILO most stable conformation for *trans*-Me-p-$A_{(36)}$-p-$io^6A_{(37)}$-p- $A_{(38)}$-p-Me in trinucleotide segment

The intramolecular hydrogen bonding parameters are included in Table 3.8(a). Interactions $N(1)_{(37)}\ldots HC(11)_{(37)}$, $N(1)_{(38)}\ldots HC(11)_{(37)}$, $N(1)_{(38)}\ldots HO(13)_{(37)}$, $O(13)_{(37)}\ldots HC(2)_{(38)}$, and $O(13)_{(37)}\ldots HC(11)_{(37)}$ may provide stabilization for the preferred hydroxyisopentenyl substituent orientation. Other interesting interactions between adjacent bases $N(6)_{(36)}\ldots HN(6)_{(37)}$, $N(6)_{(37)}\ldots HN(6)_{(38)}$, $N(1)_{(36)}\ldots HN(6)_{(37)}$, and $N(7)_{(36)}\ldots HC(8)_{(37)}$ may also contribute to stability. The interactions between ribose-phosphate backbone and bases in Table 3.8(a), namely $N(3)_{(37)}\ldots HC3'_{(37)}$, $N(3)_{(37)}\ldots HC2'_{(37)}$, $N(3)_{(36)}\ldots HC3'_{(36)}$, $N(3)_{(36)}\ldots HC5'_{(36)}$, $N(3)_{(38)}\ldots HC5'_{(38)}$, and $N(3)_{(38)}\ldots HC3'_{(38)}$, also stabilize the structure. With glycosyl torsion angles taking the preferred values ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=243°) yields favoured base substituent orientation similar (except the torsion angles β and γ) to that of isolated hypermodified base *trans*-io^6Ade [22]. This preferred most stable conformation in the model trinucleotide segment shows the interaction between $N(1)_{(37)}\ldots HC(11)_{(37)}$, as also found in isolated hypermodified base *trans*-io^6Ade [22].

Higher energy stable alternative conformation indicated by PCILO at (1.1 kcal/mol), shown in Table 3.8(b), is reached by flipping to γ=90°, ψ1=150°. The intramolecular interactions $N(1)_{(37)}\ldots HC(11)_{(37)}$ and $N(1)_{(38)}\ldots HO(13)_{(37)}$ stabilize hydroxyisopentenyl substituent orientation. Next higher energy alternative conformation at (1.8 kcal/mol) is reached for (γ=90°). The interaction

$N(1)_{(37)}$…$HC(11)_{(37)}$ remains a stabilizing factor. Interaction $O(13)_{(37)}$…$HC(11)_{(37)}$ is the additional stabilizing factor. The ribose-phosphate backbone interactions Table 3.8(a) may provide additional stabilization for these alternative conformations. Results of geometry optimization through molecular mechanics force field (MMFF) method are shown in Table 3.6. The MMFF results shows that the *trans*-io^6A in trinucleotide hypermodified segment is more stable by 3.4 kcal/mol than *cis*- io^6A shown in Table 3.6. Compared to the preferred PCILO structure (fig.3.14) *cis*- io^6A, is more stable by about 0.3 kcal/mol, than the preferred structure (fig.3.16) for *trans*-io^6A in trinucleotide hypermodified segment.

3.3.3.6 *Trans*-Me-p-$A_{(36)}$-p-$ms^2io^6A_{(37)}$-p-$A_{(38)}$-p-Me:

The hypermodified nucleoside $ms^2io^6A_{37}$ is included in the model trinucleotide segment, *trans*-Me-p-$A_{(36)}$-p-$ms^2io^6A_{(37)}$-p-$A_{(38)}$-p-Me. The glycosyl torsion angles take the specified values in Holbrook model ($\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angles for the preferred base substituent orientation are (α=0°, β=0°, γ=120°, δ=180°, ψ1=0°, ψ2=180°,±60°, θ=300°, ω1=60°, ω2=90°). These torsion angle values may be compared with the preferred orientation of isolated base *trans*-ms^2io^6Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120, θ=±60°, ω1=0°, ω2=180°) [22]. Except for the torsion angles ω1=60° and ω2=90° specifying orientation of the 2-methylthioisopentenyl group, the conformation of 2-methylthio hydroxyisopentenyl substituent in trinucleotide segment is quite similar to isolated hypermodified base *trans*-ms^2io^6Ade [22].

When the glycosyl torsion angles for nucleotides in the segment are varied freely, the resulting preferred values are ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=273°). Adopting these preferred glycosyl torsion angles results in a new set of torsion angle values for the preferred 2-methylthio hydroxyisopentenyl substituent orientation (fig.3.17) (α=0°, β=0°, γ=120°, δ=180°, ψ1=0°, ψ2=180°,±60, θ=60°, ω1=90°, ω2=150°). As compared to the preferred orientation of the isolated hypermodified base *trans*-ms^2io^6Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120, θ=±60°, ω1=0°, ω2=180°) [22], the torsion angles ω1 and ω2 taking different values in trinucleotide hypermodified segment. The intramolecular hydrogen bonding interactions, shown in Table 3.9, namely $N(1)_{(37)}$…$HC(11)_{(37)}$, $N(1)_{(36)}$…$HC(10)_{(37)}$, $O(13)_{(37)}$…$HC(11)_{(37)}$, $N(7)_{(36)}$…$HC(8)_{(37)}$, $N(1)_{(37)}$…$HO(13)_{(37)}$, and $S(2)_{(37)}$…$HO(13)_{(37)}$ may stabilize the favoured 2-methylthio hydroxyisopentenyl substituent orientation. The other

interesting interactions between adjacent bases $N(6)_{(36)}\ldots HN(6)_{(37)}$, $N(1)_{(36)}\ldots HN(6)_{(37)}$, may also contribute to stability.

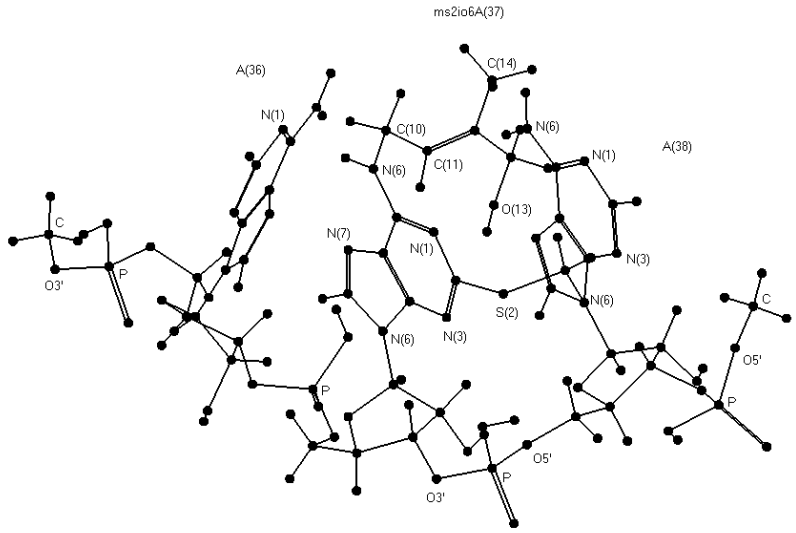


Fig.3.17: Predicted PCILO most stable conformation for *trans*-Me-p-$A_{(36)}$-p-$ms^2io^6A_{(37)}$-p-$A_{(38)}$-p-Me in trinucleotide segment.

The interactions between ribose-phosphate backbone and bases in Table 3.9 namely $N(3)_{(37)}\ldots HC3'_{(37)}$, $N(3)_{(37)}\ldots HC2'_{(37)}$, $N(3)_{(36)}\ldots HC3'_{(36)}$, $N(3)_{(36)}\ldots HC5'_{(36)}$, $N(3)_{(38)}\ldots HC3'_{(38)}$, and $N(3)_{(38)}\ldots HC2'_{(38)}$ may also contribute to the stability of the predicted structure. Taking the preferred values for glycosyl torsion angle ($\chi_{(36)}$=211°, $\chi_{(37)}$=292°, $\chi_{(38)}$=273°), results in similar orientation for the hydroxyisopentenyl substituent to that of the isolated modified base (*trans*-$ms^2io^6A_{37}$) [22], but the bulky 2-methylthio group is oriented differently with values (ω1=90°, ω2=150°).

Alternative PCILO stable conformation at higher (0.6 kcal/mol) energy, in Table 3.8(b), may be reached by flipping of γ=240°, θ=300°, and ω1=±90°. This structure is stabilized by $N(1)_{(37)}\ldots HC(11)_{(38)}$, $N(1)_{(37)}\ldots HO(13)_{(37)}$, $O(13)_{(37)}\ldots HC(11)_{(37)}$, and $S(2)_{(37)}\ldots HO(13)_{(37)}$. Another higher energy alternative conformation at (1.7 kcal/mol) is arrived by flipping of ω1=270°. This structure is stabilized by $N(1)_{(37)}\ldots HC(11)_{(38)}$, $N(1)_{(37)}\ldots HO(13)_{(37)}$, $O(13)_{(37)}\ldots HC(11)_{(37)}$, $S(2)_{(37)}\ldots HO(13)_{(37)}$, and $O(13)_{(37)}\ldots HC(15)_{(37)}$ interactions. Other interactions between ribose-phosphate backbone are also shown in Table 3.9.

The results of geometry optimization through molecular mechanics force field (MMFF) method shows that *cis*-ms^2io^6A is stabler by (10.0 kcal/mol) than *trans*-ms^2io^6A and these results are included in Table 3.6. As compared to the PCILO

preferred structure for *cis*-ms^{2}io^{6}A trinucleotide segment shown in (fig.3.15), the preferred structure for *trans*-ms^{2}io^{6}A trinucleotide segment shown in (fig.3.17) is more stable by about 1.2 kcal/mol.

3.3.4 The model anticodon loop hypermodified segment:

3.3.4.1 Me-p-C$_{(32)}$-p-U$_{(33)}$-p-G$_{(34)}$-p-G$_{(35)}$-p-A$_{(36)}$-p-i^{6}A$_{(37)}$-p-A$_{(38)}$-p-Me:

The hypermodified nucleoside i^{6}A is incorporated in the model anticodon loop segment. The glycosyl torsion angles retain the specified values in Holbrook model ($\chi_{(32)}$=31°, $\chi_{(33)}$=24°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angle values for orientation of isopentenyl adenine are (α=120°, β=270°, γ=0°, δ=180°, ψ1=330°,210°,90°, ψ2=330°,210°,90°). The structure may be compared with the preferred values [8] for isolated i^{6}Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°). The conformation of isopentenyl substituent (i^{6}A$_{37}$) in anticodon loop segment changes remarkably to (α=120°, β=270° and γ=0°) instead of (α=0°, β=0°, γ=±120°) for isolated hypermodified base, i^{6}Ade [8]. The intramolecular hydrogen bonding between N(1)$_{(37)}$…HC(11) is not feasible in this case.

Figure 3.18, depicts the preferred most stable conformation for anticodon loop segment i^{6}A. When the glycosyl torsion angles for nucleosides in the anticodon loop segment are varied freely, the resulting preferred values are ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°). Adopting these preferred values leads to new set of torsion angle values for the preferred orientation of isopentenyl substituent (fig. 3.18) (α=60°, β=300°, γ=150°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°). These torsion angle values may be compared with the preferred orientation of the isolated hypermodified base [8] i^{6}Ade (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°). The torsion angles α, β and γ of isopentenyl substituent in seven nucleotide anticodon loop structure, preferring different values than isolated base substituent, i^{6}Ade [8].

The intramolecular hydrogen bonding interactions shown in Table 3.10 are N(1)$_{(37)}$…HC(11), N(7)$_{(36)}$…HC(10)$_{(37)}$, O2'$_{(32)}$…HC(10)$_{(37)}$, O(6)$_{(35)}$…HC(14)$_{(37)}$ and O1'$_{(33)}$…HC(14)$_{(37)}$, may be the stabilizing factors for isopentenyl substituent orientation. Other interesting interactions shown in Table 3.10 may also stabilize the structure. Except for slight changes in (α=60°, β=300°, γ=150°) keeping the preferred values for glycosyl torsion angles ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°,

$\chi_{(37)}$=22°, $\chi_{(38)}$=33°) yields preferred isopentenyl substituent orientation. The intramolecular hydrogen bonding N(1)$_{(37)}$…HC(11) is indicated as for isolated hypermodified base [8] i^{6}Ade. The optimum values for glycosyl torsion angles are ($\chi_{(32)}$=271°, $\chi_{(34)}$=93°, and $\chi_{(38)}$=33°) different than Holbrook tRNA^{Phe} [10] model.

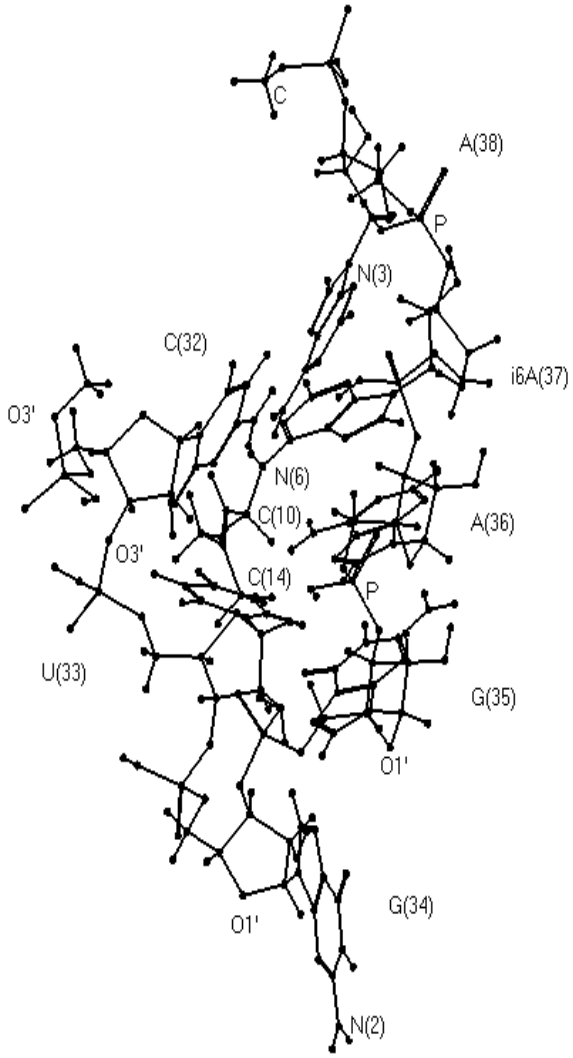


Fig 3.18: Predicted most stable conformation for i^{6}A in model anticodon loop segment, glycosyl torsion angles have take the preferred values ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°).

Table 3.10: Geometrical parameters for hydrogen bonding in the preferred conformations of i^6A in model anticodon loop hypermodified segment.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)
$N(1)_{(37)}$…H-$C(11)_{(37)}$	2.77	1.09	91.99
$N(1)_{(37)}$…H-$N(6)_{(38)}$	2.12	1.09	154.04
$N(6)_{(38)}$…H-$C(2)_{(37)}$	1.76	1.09	131.62
$N(1)_{(32)}$…H-$N(6)_{(37)}$	1.95	1.09	137.09
$N(7)_{(36)}$…H-$C(10)_{(37)}$	2.38	1.09	139.78
$N(6)_{(36)}$…H-$C(10)_{(37)}$	2.42	1.09	116.46
$N(6)_{(36)}$…H-$C(14)_{(37)}$	2.78	1.09	88.85
$O1'_{(33)}$…H-$C(10)_{(37)}$	2.32	1.09	98.95
$O1'_{(33)}$…H-$C(14)_{(37)}$	1.76	1.09	131.01
$N(6)_{(37)}$…H-$C2'_{(32)}$	2.80	1.09	99.95
$N(6)_{(37)}$…H-$C1'_{(32)}$	2.43	1.09	104.07
$O2'_{(32)}$…H-$C(10)_{(37)}$	2.63	1.09	84.74
$O2'_{(32)}$…H-$C(11)_{(37)}$	1.63	1.09	84.88
$O2'_{(32)}$…H-$C(13)_{(37)}$	2.53	1.09	92.79
$N(7)_{(38)}$…H-$C(2)_{(37)}$	1.52	1.09	84.46
$O(6)_{(35)}$…H-$C(14)_{(37)}$	2.84	1.09	129.15
$O5'_{(36)}$…H-$N(3)_{(33)}$	2.52	1.09	134.54
$O(2)_{(33)}$…H-$C3'_{(35)}$	2.68	1.09	158.57
$O(2)_{(32)}$…H-$C3'_{(32)}$	1.81	1.09	131.03
$O(2)_{(32)}$…H-$C(5)_{(33)}$	2.41	1.09	81.82
$N(7)_{(36)}$…H-$C1'_{(33)}$	2.33	1.09	150.26
$N(3)_{(36)}$…H-$N(2)_{(34)}$	2.53	1.09	111.99
$N(7)_{(34)}$…H-O2'(33)	2.56	1.09	163.48

Results of geometry optimization through molecular mechanics force field (MMFF) method lead to α=70°, β=-115°, γ=158°, δ=-176°, ψ1=9.3°, ψ2=12.4°. The glycosyl torsion angles in the anticodon loop structure are $\chi_{(32)}$=-46°, $\chi_{(33)}$=35°, $\chi_{(34)}$=-

4°, $\chi_{(35)}$=4°, $\chi_{(36)}$=-22°, $\chi_{(37)}$= 30°, $\chi_{(38)}$=65° in broad agreement with the PCILO results.

3.3.4.2 Me-p-$C_{(32)}$-p-$U_{(33)}$-p-$G_{(34)}$-p-$G_{(35)}$-p-$A_{(36)}$-p-$ms^2i^6A_{(37)}$-p-$A_{(38)}$-p-Me:

The hypermodified nucleoside ms^2i^6A is included in the anticodon loop segment. The glycosyl torsion angles are ($\chi_{(32)}$=31°, $\chi_{(33)}$=24°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°) retained as in the $tRNA^{Phe}$ model [10] of Holbrook. The set of torsion angle values for the orientation of base substituent are (α=120°, β=120°, γ=300°, δ=180°, ψ1=180°,±60°, ψ2=180°,±60°, ω1=60°, ω2=210°). These torsion angle values may be compared with the preferred values [8] for isolated ms^2i^6Ade (α=0°, β=180°, γ=±60°, δ=180°, ψ1=0°,±120, ψ2=0°,±120°, ω1=180°, ω2=180°). The conformation of ($ms^2i^6A_{37}$) in anticodon loop segment is remarkably different for (α=120°, β=120° and γ=300°, ω1=60°, ω2=210°) as compared to (α=0°, β=180°, γ=±60°, ω1=60°, ω2=180°) for isolated hypermodified base, ms^2i^6Ade [8].

When the glycosyl torsion angle for nucleotides in the anticodon loop segment are varied freely, the resulting preferred values are ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=52°, $\chi_{(38)}$=303°). Adopting these preferred values leads to a new set of torsion angle values for the preferred orientation of 2-methylthio isopentenyl substituent (fig.3.19) (α=60°, β=300°, γ=330°, δ=180°, ψ1=0°,±120°, ψ2=330°,210°,90°, ω1=0°, ω2=180°). As compared to the preferred orientation of the isolated hypermodified base [8] ms^2i^6Ade (α=0°, β=180°, γ=±60°, δ=180°, ψ1=0°,±120°, ψ2=0°,±120°, ω1=180°, ω2=180°), the torsion angles α, β, γ, ω1 and ω2 taking different values in seven nucleotide anticodon loop structure. The intramolecular hydrogen bonding interactions shown in Table 3.11 namely, $N(1)_{(37)}$…$HC(14)_{(37)}$, $N(7)_{(36)}$…$HC(10)_{(37)}$, $O1'_{(33)}$…$HC(11)_{(37)}$, $S(2)_{(37)}$…$HN(6)_{(36)}$, and $O1'_{(33)}$…H-$C(10)_{(37)}$ may provide stability to the preferred orientation of 2-methylthioisopentenyl substituent in the anticodon loop.

Other interesting interactions between adjacent bases and with ribose-phosphate backbone are also included in Table3.11 and may also contribute to stability. Adopting the preferred values for glycosyl torsion angles ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=52°, $\chi_{(38)}$=303°), results in slight changes for (α, β, γ, and ω1, ω2). but the other torsion angles specifying base substituent

orientation of ms^2i^6A in the loop, retain values similar to that of isolated hypermodified base [8] ms^2i^6Ade. The preferred orientation of 2-methylthio group in the anticodon loop has (ω1=0°, ω2=180°), these may be compared with (ω1=180°, ω2=180°) favoured for isolated hypermodified base [8] ms^2i^6Ade.

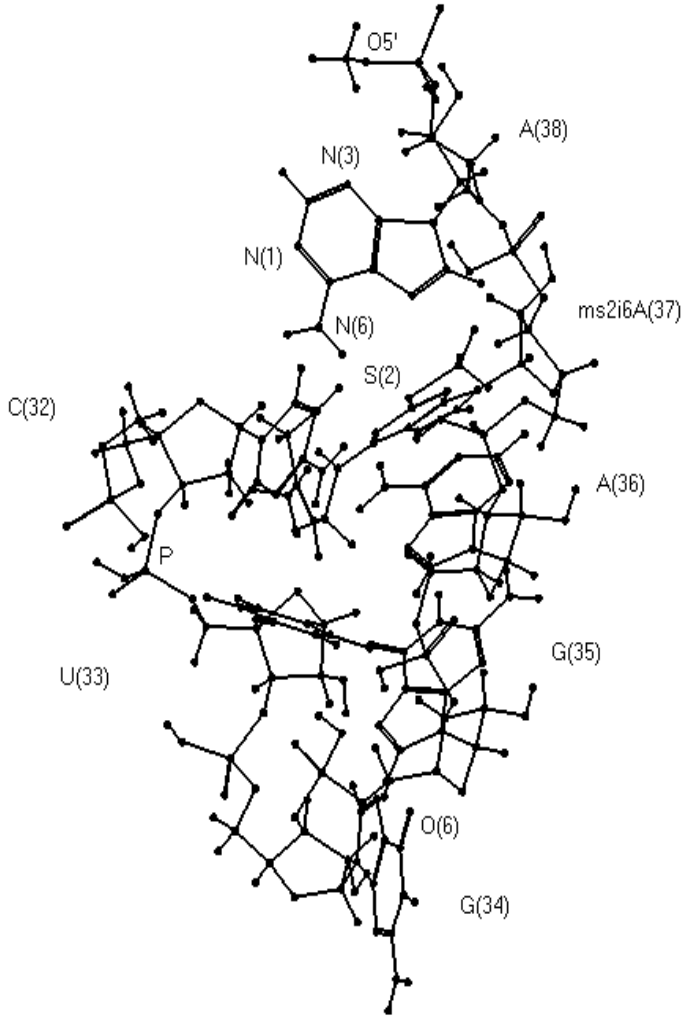


Fig 3.19: Predicted most stable conformation for ms^2i^6A in model anticodon loop segment, glycosyl torsion angles have take the preferred values ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=52°, $\chi_{(38)}$=303°).

Table 3.11: Geometrical parameters for hydrogen bonding in the preferred conformation of ms^2i^6A in model anticodon loop hypermodified segment.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)
$N(1)_{(37)}$…H-$C(14)_{(37)}$	1.89	1.09	124.61
$N(1)_{(37)}$…H-$N(6)_{(38)}$	2.66	1.09	141.99
$N(6)_{(38)}$…H-$C(14)_{(37)}$	2.41	1.09	126.33
$S(2)_{(37)}$…H-$N(6)_{(38)}$	2.49	1.09	157.44
$N(7)_{(38)}$…H-$C(15)_{(37)}$	2.49	1.09	132.12
$N(6)_{(37)}$…H-$N(6)_{(36)}$	2.71	1.09	104.59
$N(1)_{(32)}$…H-$N(6)_{(37)}$	2.19	1.09	158.19
$O(2)_{(33)}$…$HC(8)_{(36)}$	2.13	1.09	155.45
$N(7)_{(36)}$…H-$C(10)_{(37)}$	1.68	1.09	141.44
$S(2)_{(37)}$…H-$N(6)_{(36)}$	2.06	1.09	128.77
$N(1)_{(36)}$…H-$C(15)_{(37)}$	2.62	1.09	106.47
$N(7)_{(35)}$…H-$C(8)_{(34)}$	2.12	1.09	141.34
$O(6)_{(35)}$…H-$C(11)_{(37)}$	2.86	1.09	135.37
$N(1)_{(33)}$…H-$C(10)_{(37)}$	2.71	1.09	109.81
$O(2)_{(32)}$…H-$C(5)_{(33)}$	2.41	1.09	81.82
$O1'_{(33)}$…H-$C(10)_{(37)}$	2.02	1.09	83.91
$O1'_{(33)}$…H-$C(11)_{(37)}$	1.79	1.09	89.76
$N(6)_{(37)}$…$HC2'_{(33)}$	2.83	1.09	108.06
$O(2)_{(32)}$…H-$C3'_{(32)}$	1.81	1.09	131.03
$O2'_{(32)}$…H-$C(14)_{(37)}$	1.72	1.09	88.29
$N(6)_{(36)}$…H-$C(10)_{(37)}$	2.28	1.09	135.43
$O5'_{(36)}$…H-$N(3)_{(33)}$	2.52	1.09	134.54
$O(2)_{(33)}$…H-$C3'_{(35)}$	2.67	1.09	158.57
$N(1)_{(33)}$…H-$Op_{(34-35)}$	2.52	0.95	132.34

The optimum glycosyl torsion angles taking different values than Holbrook $tRNA^{Phe}$ [10] model are ($\chi_{(32)}$=271°, $\chi_{(34)}$=93°, $\chi_{(37)}$=52°, and $\chi_{(38)}$=303°). In presence of i^6A in anticodon loop structure, the optimum glycosyl torsion angle values are

($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°) and with ms^2i^6A ($\chi_{(32)}$=271°, $\chi_{(33)}$=24°, $\chi_{(34)}$=93°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=52°, $\chi_{(38)}$=303°). This difference in glycosyl torsion angles may be due to presence of 2-methylthio group. Geometry optimization through molecular mechanics force field (MMFF) method leads to α=89°, β=-125°, γ=24°, δ=177°, ψ1=1°, ψ2=68°, ω1=18°, ω2=170°, and the optimized glycosyl torsion angles are $\chi_{(32)}$=-59°, $\chi_{(33)}$=55°, $\chi_{(34)}$=-7°, $\chi_{(35)}$=13°, $\chi_{(36)}$=-20°, $\chi_{(37)}$=19°, $\chi_{(38)}$=64°.

3.4 Conclusion:

The presence of hydroxyl group in various zeatin isomers and ms^2-derivatives may even lead to stronger interaction between N(1) and HC(11). Thus the hydroxyisopentenyl substituent orientation is held relatively firmly in contrast to easily adaptable orientation [8] of isopentenyl substituent in i^6Ade and ms^2i^6Ade. Increased shielding of the N(1) site in zeatin isomers and ms^2-derivatives and relatively firmly held hydroxyisopentenyl orientation makes zeatin (io^6Ade) analogs somewhat similar to ureidopurines [6] namely, gc^6Ade, t^6Ade, and ms^2t^6Ade. These features are maintained in model nucleotide diphosphate and trinucleotide segments. Thus unlike adaptable isopentenyl substituent [8] in i^6Ade and ms^2i^6Ade, zeatin analogs need not be so involved in regulating aromatic amino acid biosynthesis [23-25]. These differences may also account for the improved efficiency and reduced codon context sensitivity [26] of suppressor tRNAs having ms^2io^6Ade modification.

The isopentenyl and hydroxyisopentenyl substituent alike retain the 'distal' conformation, which may prevent the N(1) and N(6)H from extended Watson-Crick base pairing. The predicted most stable orientation of the isopentenyl substituent, i^6A and ms^2i^6A in model diphosphate and trinucleotide segments significantly differ with each other in the preferred values for the torsion angles β and γ. Flipping of β from the preferred *cis*-orientation in i^6A to the *trans* orientation predicted in ms^2i^6A is caused by the bulky 2-methylthio substituent. But in case of seven nucleotide model segment the isopentenyl substituent orientation between these two analogs differs only in value of torsion angle γ. The glycosyl torsion angle takes different values in nucleoside diphosphate and trinucleotide model segment for these hypermodified nucleosides io^6A, i^6A, and their 2-methylthio analogs. Allowing glycosyl torsion angles to change freely, the interaction between N(1) and HC(11) is maintained for i^6A, io^6A and ms^2io^6A nucleoside in trinucleotide segment and for i^6A, ms^2i^6A in seven nucleotide

segment. The study of nucleoside diphosphates and trinucleotide segments has paved the way for probing of model anticodon loop segment. For the preferred anticodon loop structure with i^6A_{37}, the 32nd, 34th, and 38th glycosyl torsion angles favor values, different than those specified in Holbrook $tRNA^{Phe}$ model [10]. In contrast to this, 32nd, 34th, 37th, and 38th nucleotides glycosyl torsion angles favour optimum values different than specified in tRNA crystal structure data of Holbrook.

3.5 References:

1. Nishimura, S.; Yamada, Y.; Ishikura, Biochim. Biophys. Acta., 1969, 179, 517-520.
2. Hall, R. H.; The Modified Nucleosides in Nucleic Acids (Columbia University Press, New York, 1971).
3. Sprinzl, M. and Gauss, D. H. Nucleic Acid Res. 1982, 10, r1- r56.
4. Kersten, H. Prog. Nucl.Acid. Res. Mol. Biol. 1984, 31, 59-114.
5. Hall, R. H.; Csonka, L.; David, H.; McLennan, B. Science 1967, 156, 69-71.
6. Tewari, R. J. Biomol. Struct. Dyn. 1990, 8, 675-686.
7. Tewari, R. Int. J. Quantum Chem. 1987, 31, 611-623.
8. Tewari, R. Int. J. Quantum Chem. 1988, 34, 133-142.
9. Tewari, R. Int. J. Quantum Chem. 1992, 41, 709-718.
10. Holbrook, S. R.; Sussman, J. L.; Warrant, R.W. and Kim, S. H. J. Mol. Biol. 1978, 123, 631-660.
11. Diner, S.; Malrieu, J. P.; Claverie, P. Theoret. Chim. Acta 1969, 13, 1-17.
12. Diner, S.; Malrieu, J. P.; Jordan, F.; Gilbert, M. Theoret. Chim. Acta 1969, 15, 100-110.
13. Malrieu, J. P. in Semiempirical Methods of Electronic Structure Calculations, Part A, Techniques, Segal, G. A.; Ed.; Plenum, New York, 1977, p. 69-103.
14. Bugg, C. E.; Thewalt, U. Biochem. Biophys. Res. Commun. 1972, 46, 779-784.
15. McMullan, R. K.; Sundaralingam, M. J. Am. Chem. Soc. 1971, 93, 7050-7054.
16. Kennard, O. in CRC Handbook of Chemistry and Physics, 61st ed., Weast, R. C.; Astle, M. J. Eds. (CRC Press, Boca Raton, 1980-81), pp. F208-F211.
17. Dewar, M. J. S.; Thiel, W. J. Am. Chem. Soc. 1977, 99, 4899-4907.
18. Stewart, J. J. P.; J. Comp. Chem. 1989, 10, 209-220.
19. Stewart, J. J. P.; J. Comp. Chem. 1991, 1, 320-341.

20. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A.; in Ab Initio Molecular Orbital Theory, Wiley, New York, 1986.
21. Soriano-Garcia, M.;Toscano, R. A.; Arroyo-Reyna, J. A. J. Crystallographic Spectr. Res. 1987, 17, 221-230.
22. Sonawane, K. D.; Sonavane, U. B. and Tewari, R. J. Biomol. Struct. Dyn. 2002, 19 (4), 637-648.
23. Buck, M.; Griffiths, E. Nucleic Acids Res. 1982, 10, 2609-2624.
24. McCray, J. W.; Herrmann, K. M. J. Bacteriol. 1976, 125, 608-615.
25. Persson, B. C.; Bjork, G. R. J. Bacteriol. 1993, 175, 7776-7785.
26. Ericson, J. U.; Bjork, G. R. J. Mol. Biol. 1991, 218, 509-516.

PART-B: COMPILATION AND ANALYSIS OF tRNA DATABASE

3.6 Introduction

Several compilations and reviews of modified nucleosides in tRNAs have been published [1-6]. Knowledge of tRNA sequences along with the information on presence of modified nucleosides, makes tRNA one of the most interesting classes of RNA molecules, from which phylogenetic information on modified and unmodified nucleotides can be extracted. Grosjean et.al (1995) has shown the distribution of modified nucleosides by the types of organisms [6]. The location and distribution of modified nucleosides in tRNA have also been reported [7]. Bjork, G. R. [2] has predicted coding capacities of modified nucleosides at 34^{th} and 37^{th} positions of tRNA.

The present analysis of compiled tRNA database [8] gives an overview of sequence patterns and statistical analysis of various modified bases occurring at 37^{th} and 34^{th} positions in anticodon loop of tRNAs. The modified nucleosides of interest which occur at anticodon 3'-adjacent (37^{th}) position are N6-(Δ^2-isopentenyl) adenosine (i^6A), 2-methylthio-N6-(Δ^2-isopentenyl) adenosine (ms^2i^6A), N6-(Δ^2-hydroxyisopentenyl) adenosine (io^6A). Modified nucleosides of interest which occur at anticodon first position corresponding to the 34^{th} (wobble) position in anticodon loop of tRNAs are 5-carbomylmethyl uridine (ncm^5U) and 4-amino-2-(N6-lysino)-1-(β-D-ribofuranosyl)pyrimidinium ("Lysidine"), which is usually denoted as k^2c. The latest tRNA sequence database [8] updated in January 2002 has been utilized as a source for tRNA sequences. The anticodon stem-loop part of various tRNAs having these modified nucleosides at 37^{th} and 34^{th} positions have been considered. The distribution of modified nucleosides at positions 37^{th} and 34^{th}, in relation to the amino acid charged by tRNA has also been discussed in this compilation.

Figure 3.20(a) shows the conventional numbering of anticodon stem/loop part of tRNA. The results of statistical analysis are presented in the anticodon stem-loop part of cloverleaf diagram. All the sequence analysis of tRNA database has been performed through direct visual inspection of the sequence catalog.

3.7 Abbreviation and Nomenclature of modified nucleosides:

One letter (or one symbol) code as used in the tRNA database [8] compiled by Sprinzl et. al. (1998) has been adopted in the present report. Terminology and

structure of the modified nucleosides are according to [9]. Abbreviation and nomenclature used in this chapter are listed on the starting page.

3.8 Analysis of tRNA sequence database:

3.8.1 Hypermodified nucleosides at 37th positions:

3.8.1.1 N6-isopentenyl adenosine (i^{6}A):

There are 29 sequences containing i^{6}A at 37th position, these are included in Table 3.12a. The consensus sequence pattern for i^{6}A (+) hypermodified nucleoside is shown in Table 3.12b. It is known that, i^{6}A generally occurs when anticodon ends in A, however, i^{6}A still occurs when anticodon ends in C in the tRNA sequence from Saccharomyces cerevisae (ID No. RG4000), also in Candida cylindra (ID No. RS7664) and Schizosaccharomyces pomb (ID No. RX7640) i^{6}A unexpectedly occurs with anticodon ending in U instead (Table 3.12a).

In the present sequence analysis study it is found that whenever i^{6}A is present at 37th position, A is usually present at 38th position. But in tRNA^{Ser}(&GA) from Candida cylindra (ID No. RS7661), C is present at 38th position instead of A (Table 3.12a). It is noteworthy that the hypermodified nucleoside 5-carbamoylmethyl uridine (ncm^{5}U) occurs at 34th position in this sequence.

The frequency of bases and base pairs occurring at various positions in the anticodon stem/loop of 29 i^{6}A containing tRNA sequences is shown in fig. 3.20(b). The sequence motif is composed of the conserved $A_{36}A_{37}A_{38}$ anticodon loop sequence. The anticodon stem with five Watson-Crick base pairs have remarkable conservation at positions 28-42, 29-41, 30-40, and 31-39 fig. 3.20(b). At positions 29-41 (A/G-U/C) base pair, at 30-40 location G-C base pair is found instead and may impart stability to the anticodon stem. The base pairs at positions 29-41 and 30-40 in the anicodon stem are the most important ones for the efficient isopentenylation of A_{37} [10]. In both of these instances the purines are preferred at positions 29 and 30 (5'-side of the anticodon stem) where as the pyrimidines preferably occur at positions 40 and 41 (3'-side of the anticodon stem). At positions 32nd and 33rd only pyrimidines occur and purine bases are absent. At the 34th positions both purines and pyrimidines bases are equally distributed as shown in fig. 3.20(b).

Table 3.12a: tRNA sequences having $i^6A(+)$ at 37^{th} position

ID No.	Organism	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	A.A.
RC7630	S.CEREVISAE	G	C	A	G	A	P	U	G	C	A	+	A	P	C	U	G	U	Cys
RF1580	T.THERMOPHI.	U	G	C	G	A	C	U	G	A	A	+	A	P	C	G	C	A	Phe
RG4000	S.CEREVISAE	G	A	U	G	P	C	U	N	C	C	+	A	A	C	A	U	U	Gly
RS1540	B.SUBTILIS	U	C	G	G	U	C	U	5	G	A	+	A	A	C	C	G	A	Ser
RS5362	BOVINE	G	P	P	G	G	'	U	U	G	A	+	A	C	C	A	A	U	Ser
RS7630	S.CEREVISAE	A	C	A	G	A	<	U	N	G	A	+	A	P	C	U	G	U	Ser
RS7631	S.CEREVISAE	A	A	A	G	A	P	U	I	G	A	+	A	P	C	U	U	U	Ser
RS7632	S.CEREVISAE	A	A	A	G	A	P	U	I	G	A	+	A	P	C	U	U	U	Ser
RS7661	C.CYLINDRA	A	C	G	G	A	P	U	&	G	A	+	C	P	C	C	G	U	Ser
RS7662	C.CYLINDRA.	A	A	A	G	A	P	U	I	G	A	+	A	P	C	U	U	U	Ser
RS7663	C.CYLINDRA.	U	G	A	C	A	'	U	C	G	A	+	A	P	G	U	C	A	Ser
RS7664	C.CYLINDRA.	C	U	U	C	C	'	U	G	C	U	+	A	G	G	A	A	G	Ser
RS8620	S.OLERACEA	C	A	U	G	A	'	U	I	G	A	+	A	P	C	A	U	G	Ser
RS9160	RAT LIVER	A	P	G	G	A	'	U	I	G	A	+	A	Z	C	C	A	U	Ser
RS9162	RAT M.HEPA.	A	P	G	G	A	'	U	I	G	A	+	A	Z	C	C	A	U	Ser
RS9241	D.MELANO.	U	C	U	G	A	'	U	I	G	A	+	A	P	C	A	G	A	Ser
RS9242	D.MELANO.	U	C	U	G	A	'	U	C	G	A	+	A	P	C	A	G	A	Ser
RW1540	B.SUBTILIS	G	A	G	G	P	C	U	C	C	A	+	A	A	C	C	U	C	Try
RW3840	T.THERM.	A	C	G	G	U	C	U	!	C	A	+	A	A	P	C	G	U	Try
RX7640	S.POMBE.	A	C	G	G	G	C	U	C	A	U	+	A	C	C	C	G	U	Met
RY1540	B.SUBTILIS	G	C	G	G	A	C	U	Q	U	A	+	A	P	C	C	G	C	Tyr
RY4000	S.CEREVISAE.	P	U	G	A	G	C	U	G	U	A	+	A	C	U	C	A	A	Tyr
RY7630	S.CEREVISAE.	C	A	A	G	A	C	U	G	P	A	+	A	P	C	U	U	G	Tyr
RY7640	S.POMBE.	C	C	C	G	G	C	U	G	P	A	+	A	C	C	G	G	U	Tyr
RY7650	T.UTILIS	P	C	A	G	A	C	U	G	P	A	+	A	P	C	U	G	A	Tyr
RZ1665	E.COLI	C	U	G	G	A	C	U	U	C	A	+	A	U	C	C	A	G	Selcy
RZ9330	MOUSE LIVER	G	C	A	G	G	C	U	N	C	A	+	A	C	C	U	G	U	Selcy
RZ9280	BOVINE LIVER	G	C	A	G	G	C	U	1	C	A	+	A	C	C	U	G	U	Selcy
RZ9990	H.HELA-CELLS	G	C	A	G	G	C	U	N	C	A	+	A	C	C	U	G	U	Selcy

Consensus sequence for i^6A_{37} from 29 sequences shows that, various modified nucleosides occur at 34th position. Some modified nucleosides also occur at 32nd and 39th positions and are presented in Table: 3.12b.

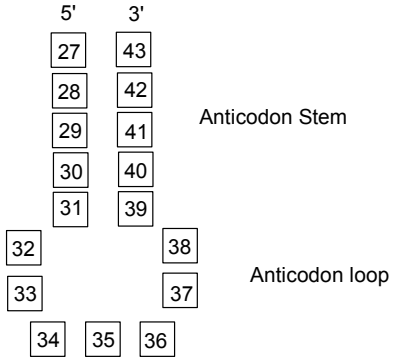
Table: 3.12b: Consensus sequences for i^6A_{37} from total 29 sequences:

27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
G U A P C	C G A P U	A C U G P	G C A	A P U G C	P C ' <	U	G N 5 U I & Q ! C 1	C A G U P	A C U	+	A C	P A C Z U G	C G P U	U G A C	G C U A	U A G C

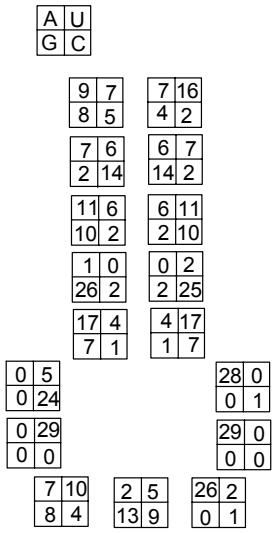
Table 3.13: Distribution of modified nucleosides at positions 37th and 34th, along with the amino acid charged by tRNA.

Modified Nucleotides	Amino Acids charged by tRNA												Total
	Cys	Phe	Gly	Ser	Try	Met Init.	Tyr	Sel. Cys.	Leu	Val	Ile	Pro	
i^6A (+)	1	1	1	14	2	1	5	4	-	-	-	-	29
Ms^2i^6A (*)	1	9	-	3	3	-	5	-	3	-	-	-	24
Io^6A (`)	-	-	-	5	-	-	-	-	-	-	-	-	5
Ncm^5U (&)	-	-	-	2	-	-	-	-	-	1	-	1	4
K^2C (})	-	-	-	-	-	-	-	-	-	-	3	-	3

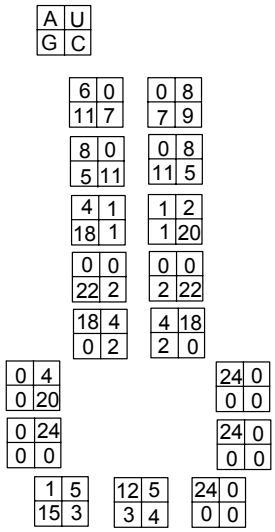
Table 3.13 identifies aminoacid specificity of tRNA having i^6A present at 37th position. The modified nucleoside i^6A is presents in $tRNA^{Ser}$, $tRNA^{Phe}$, $tRNA^{Tyr}$, $tRNA^{Trp}$, $tRNA^{Cys}$, $tRNA^{Gly}$, $tRNA^{Met}_{(Initiator)}$, $tRNA^{Sle.cys}$. Out of 29 tRNA sequences for i^6A, 14 tRNA sequences belong to amino acid serine. Followed by $tRNA^{Tyr}$ with 5 sequences, and $tRNA^{Sel.cys}$ has 4 sequences. Others with lower frequency are included in Table 3.13.



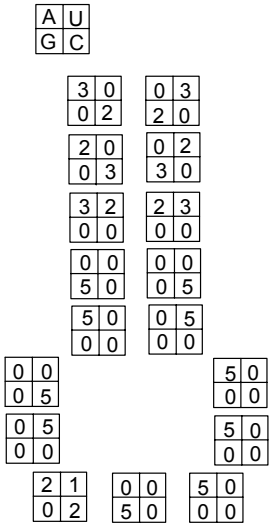
(a) convensional numbering in stem/loop.



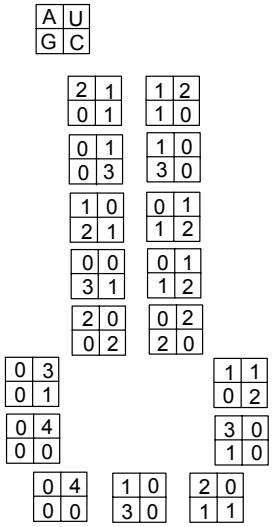
(b) i^6A at 37^{th} position.



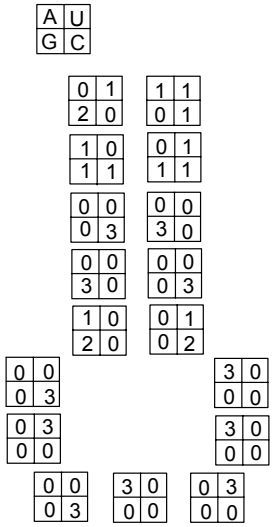
(c) ms^2i^6A at 37^{th}



(d) io^6A at 37^{th} Position



(e) ncm^5U at 34^{th} Position



(f) k^2c at 34^{th} Position

Fig. 3.20: Anticodon stem/loop part of tRNA cloverleaf diagrams showing frequency of bases and base pairs at various sites related to occurrence of the hypermodified nucleosides at 37^{th} and 34^{th} positions.

3.8.1.2 2-methylthio-N6-isopentenyl adenosine (ms^2i^6A):

The hypermodified base (ms^2i^6A) is a derivative of i^6A, which occurs at 37^{th} position. In the present compilation there are in all 24 tRNA sequences, Table 3.14a, which contain ms^2i^6A at 37^{th} position.

Table 3.14a: tRNA sequences having ms^2i^6A(*) at 37^{th} position.

		Anticodon Domain																
ID No.	Organism	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	A.A.
RC1660	E.COLI	G	C	G	G	A	P	U	G	C	A	*	A	P	C	C	G	U	Cys
RF1540	B.SUBTILIS	A	C	G	G	A	C	U	#	A	A	*	A	P	C	C	G	U	Phe
RF1660	E.COLI	G	G	G	G	A	P	U	G	A	A	*	A	P	C	C	C	C	Phe
RF2020	R. RUB.	C	G	U	G	A	C	U	G	A	A	*	A	P	C	A	C	G	Phe
RF2060	A.QUADR.	G	A	G	G	A	C	U	G	A	A	*	A	P	C	C	U	C	Phe
RF2120	B.STEARO.	A	A	G	G	A	C	U	#	A	A	*	A	P	C	C	U	U	Phe
RF2520	E.GRACILIS	G	A	G	G	A	C	U	G	A	A	*	A	P	C	C	U	U	Phe
RF3160	P.VULGARIS	G	A	G	G	A	C	U	G	A	A	*	A	P	C	C	U	C	Phe
RF3280	S.OLERACEA	G	A	G	G	A	C	U	G	A	A	*	A	U	C	C	U	C	Phe
RF5280	RAT M. HEPA.	A	A	G	C	A	C	U	G	A	A	*	A	P	G	C	U	U	Phe
RL0220	PHAGE T4	C	A	G	C	A	C	U	N	A	A	*	A	P	G	C	U	G	Leu
RL1660	E.COLI	A	G	G	G	A	P	U	H	A	A	*	A	P	C	C	C	U	Leu
RL2120	B. STEARO.	C	A	C	G	A	C	U	B	A	A	*	A	P	C	G	U	G	Leu
RS0220	PHAGE T4	C	C	G	G	U	B	U	N	G	A	*	A	A	C	C	G	G	Ser
RS1660	E.COLI	C	C	G	G	U	C	U	C	G	A	*	A	A	C	C	G	G	Ser
RS1664	E.COLI	C	C	G	G	U	B	U	V	G	A	*	A	A	C	C	G	G	Ser
RW1660	E.COLI	C	C	G	G	U	B	U	C	C	A	*	A	A	C	C	G	G	Try
RW5280	RAT M.HEPA	A	G	A	G	C	C	U	N	C	A	*	A	G	C	C	C	U	Try
RW5360	B. LIVER	A	G	A	G	C	C	U	N	C	A	*	A	G	C	C	C	U	Try
RY1541	B. SUBTILIS	G	C	G	G	A	C	U	Q	U	A	*	A	P	C	C	G	C	Tyr
RY1660	E.COLI	G	C	A	G	A	C	U	Q	U	A	*	A	P	C	U	G	C	Tyr
RY1661	E.COLI	G	C	A	G	A	C	U	Q	U	A	*	A	P	C	U	G	C	Tyr
RY2120	B. STEARO.	G	C	G	G	A	C	U	Q	U	A	*	A	P	C	C	G	C	Tyr
RY2560	S. OBLIQ.	G	C	G	G	A	P	U	G	U	A	*	A	P	C	C	G	C	Tyr

The consensus sequence pattern for ms^2i^6A is shown in Table 3.14b. The hypermodified nucleoside (ms^2i^6A) occurs when anticodon ends in A. In contrast to hypermodified nucleoside i^6A, the $A_{36}A_{37}A_{38}$ sequence is uniformly found in case of ms^2i^6A. through all the 24 tRNA sequences.

Figure 3.20(c) depicts the frequency of bases and base pairs in the anticodon stem/loop of 24 tRNA sequences. The anticodon stem is stabilized by five Watson-Crick base pairs namely 27-43 (A/G-U/C), 28-42 (A/C-U/G), 29-41 (G-C), 30-40 (G-C), and 31-39 (A-U). At positions 32nd and 33rd pyrimidine nucleosides occur exclusively and purine nucleosides are absent. At 33rd positions 24 uridine bases are present in ms^2i^6A, just as in case of i^6A_{37}, all 29 sequences have uridine at 33rd position.

Table 3.13 shows the number of tRNA charged by corresponding amino acids, when ms^2i^6A is present at 37th position. The ms^2i^6A occurs in $tRNA^{Cys}$, $tRNA^{Phe}$, $tRNA^{Ser}$, $tRNA^{Try}$, $tRNA^{Tyr}$ and $tRNA^{Leu}$. Out of the 24 tRNA sequences with $ms^2i^6A_{37}$ 9 are specific to amino acid phenylalanine, 5 tRNA sequences are specific to amino acid tyrosine, 3 tRNA sequences each are specific for amino acids serine, tryptophan and leucine , one tRNA sequence is specific to cysteine amino acid.

Table: 3.14b: Consensus sequence for ms^2i^6A at 37th position from total 24 sequences:

27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
G A C	C A G	G U C A	G C	A U C	P C B	U	G # N B H C V Q	C A G U	A	*	A	P U A G	C G	C A G U	G U C	U C G

3.8.1.3 Comparison between i^6A and ms^2i^6A tRNA sequences:

Comparing the sequence patterns for i^6A, Table 3.12b and ms^2i^6A, Table 3.14b, unknown modified uridine (N) and queuosine (Q) are found to occur at 34th position for both of these hypermodified nucleosides. Hypermodified nucleoside 5-carbamoylmethyl uridine does not occur at 34th position when ms^2i^6A is present at 37th

position. As compared to $ms^2i^6A_{37,}$ when i^6A is present at 37th position, more number of modified nucleosides occur at 32nd, 34th and 39th positions. When ms^2i^6A occurs at 37th position, no modifications are found at 29th and 35th positions. Whereas, modified nucleoside pseudouracil is present at 29th and 35th positions, when i^6A occurs at 37th position. At 34th position, in case of $ms^2i^6A_{37}$ the ratio of purine to pyrimidine bases is nearly 2:1, figure 3.20(c), but in case of i^6A_{37} the ratio of purine to pyrimidine bases at 34th position is nearly 1:1, figure 3.20(b).

3.8.1.4 N6-(*cis*-hydroxyisopentenyl) adenosine (io^6A):

In the current updated, January 2002 tRNA database [8] version, only 5 tRNA sequences are found for hypermodified nucleotide io^6A (`) which occurs at anticodon 3'-adjacent (37th) position in anticodon loop, Table 3.15(a). All the 5 tRNA sequences belong to amino acid serine and are included in Table 3.13. The consensus sequence pattern for this hypermodified nucleotide is shown in Table 3.15(b). Out of these five tRNA sequences, only in $tRNA^{Ser}$(UGA) (ID No. RS 8600) from Nicotiana rustica, the wobble nucleotide 5-carbamoylmethyl uridine (ncm^5U) occurs at 34th position along with io^6A present at 37th position. The modified base 3-methylcytidine (m^3C_{32}) is common to all the five tRNA sequences having io^6A_{37}. The $tRNA^{Ser}$(UGA) sequence is very interesting, because only in this sequence major modifications like io^6A, ncm^5U and m^3C are present at 37th, 34th and 32nd positions respectively. This has prompted the detailed conformational energy calculations on anticodon loop reported in CHAPTER-IV and molecular dynamics simulation (MD) of anticodon arm (stem/loop part) in CHAPTER-VI.

Table 3.15(a): tRNA sequences having io^6A (`) at 37th position

		Anticodon Domain																
ID No.	Organism	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	A.A.
RS8560	L. LUTEUS	A	C	A	G	A	'	U	C	G	A	`	A	P	C	U	G	U	Ser
RS8600	N. RUSTICA	A	C	A	G	A	'	U	&	G	A	`	A	P	C	U	G	U	Ser
RS8601	N. RUSTICA	A	C	A	G	A	'	U	C	G	A	`	A	P	C	U	G	U	Ser
RS8603	N. RUSTICA	C	A	U	G	A	'	U	I	G	A	`	A	P	C	A	U	G	Ser
RS8604	N. RUSTICA	C	A	U	G	A	'	U	I	G	A	`	A	P	C	A	U	G	Ser

Table: 3.15(b): Consensus sequences for io^6A at 37^{th} position from total 5 sequences:

27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
A C	C A	A U	G	A	'	U	C I &	G	A	`	A	P	C	U A	G U	U G

The fig.3.20(d) shows the frequency of occurrence of bases in anticodon loop and of the base-pairs in anticodon arm. In all the 5 tRNA sequences $A_{36}A_{37}A_{38}$ anticodon loop sequence is conserved. Which is essential for effective isopentenylation at 37^{th} position. At position 32^{nd} in each sequence, a modified nucleotide (m^3C) occurs. Unmodified U occurs invariably at 33^{rd} position and at 35^{th} position only guanosine (G) occurs. In 34^{th} position another modified nucleotide Inosine occurs twice, also unmodified C (cytosine) occurs twice. The anticodon stem with five Watson-Crick base pairs maintains A/C-U/G at 27-43 position, A/C-U/G at 28-42, A-U at 29-41, G-C at 30-40 and A-U at 31-39 positions as shown in fig.3.20(d). The base pairs at 29-41 and 30-40 in anticodon stem have purines at positions 29 and 30 (5'-side of anticodon stem) and the pyrimidines at position 40 and 41 (3'-side of the anticodon stem). Such kind of Watson-Crick base pairing in anticodon stem is useful for effective iospentenylation at 37^{th} position [10].

3.8.2 Hypermodified nucleosides at 34^{th} position:

3.8.2.1 5-carbamoylmethyl uridine (ncm^5U):

The hypermodified nucleoside 5-carbamoylmethyl uridine (ncm^5U) occurs at 34^{th} "wobble" position in anticodon loop of tRNA. There are 4 tRNA sequences obtained from the current updated version of tRNA sequence database presented in Table 3.16(a), which contain ncm^5U (&) at 34^{th} position. The hypermodified nucleosides m^1G (k), i^6A (+) and io^6A (`) are also presents at 37^{th} position in these tRNA sequences. Out of these 4 tRNA sequences, two sequences belongs to amino acid serine, one tRNA sequence is related to amino acid proline and valine each, as shown in Table 3.13.

The $tRNA^{Ser}$(&GA) from Candida cylindra contains i^6A at 37^{th} position. In this tRNA sequence the $A_{36}A_{37}C_{38}$ sequence is present. This kind of anticodon sequence is observed only when ncm^5U is present at 34^{th} and i^6A at 37^{th} position. At 32^{nd} position pseudouridine is present. This means that tRNA sequence contains ncm^5U at 34^{th}

position and i^6A at 37^{th} position in anticodon loop, the anticodon loop terminates with pyrimidine.

Table 3.16(a): tRNA sequences having ncm^5U(&) at 34^{th} position

		Anticodon Domain																
ID No.	Organism	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	A.A
RP7650	T.UTILIS	C	U	C	G	C	U	U	&	G	G	K	P	G	P	G	A	G	Pro
RS7661	C.CYLINDRA	A	C	G	G	A	P	U	&	G	A	+	C	P	C	C	G	U	Ser
RS8600	N.RUSTICA	A	C	A	G	A	'	U	&	G	A	`	A	P	C	U	G	U	Ser
RV7632	S.CEREVISAE.	P	C	G	C	C	P	U	&	A	C	A	C	G	G	C	G	A	Val

The consensus sequence for ncm^5U (&) is shown in Table 3.16(b). At 29^{th}, 30^{th} positions (5'-side of anticodon stem) no modified nucleotideis found. But at 40^{th} position modified nucleoside pseudouridine is present.

Table: 3.16b: Consensus sequences for ncm^5U at 34^{th} position from total 4 sequences:

27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
C A P	U C	C G A	G C	C A	U P '	U	&	G A	G A C	K + ` A	P C A	G P	P C G	G C U	A G	G U A

The frequency of bases and base pairs at various locations in the anticodon stem/loop of these 4 tRNA sequences is depicted in fig.3.20(e). In the anticodon loop at 32^{nd}, 33^{rd}, and 34^{th} positions pyrimidines are dominant. At 35^{th} and 37^{th} positions only purines are present. The anticodon arm has five Watson-Crick base pairs and is shown in fig.3.20(e).

3.8.2.2 Lysidine (k^2C):

During the analysis of the latest compiled tRNA database [8], another interesting hypermodified nucleoside "Lysidine" was noticed at 34^{th} 'wobble' position in anticodon loop of $tRNA^{Ile}$. There are 3 such tRNA sequences shown in Table 3.17(a), found one each, from Mycoplasma capricolum, E. coli and Bacillus subtilis $tRNA^{Ile}$. The $tRNA^{Ile}$ from Mycoplasma capricolum, and Bacillus subtilis have lysidine at 'wobble' position and N6-methyl adenosine (m^6A) at 37^{th} position. Both these tRNA sequences (ID No. RI1140 and RI1540) have similar sequence pattern

from 29th to 41st position. The E.Coli $tRNA^{Ile}$ has threonylcarbonyl adenosine (t^6A) at 37th position, the other anticodon loop components are same in $tRNA^{Ile}$ from Mycoplasma capricolum and Bacillus subtilis.

The consensus tRNA sequence having 'wobble' nucleoside lysidine at 34th position is shown in Table 3.17(b). The table shows that at 29th, 30th, 32nd, 33rd, 34th, 35th, 36th, 38th, 40th and 41st positions have only single nucleotide. Anticodon loop at four positions contains pyrimidine bases i.e. 32nd, 33rd, 34th, and 36th and at three positions 35th, 37th and 38th contains purine bases. The presence of these bulky hypermodified nucleosides in the anticodon loop may affect the structure of the anticodon loop.

Table 3.17(a): tRNA sequences having k^2C (}) at 34th position

		Anticodon Domain																
ID No.	Organism	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	A.A
RI1140	M.CAPRI.	U	C	C	G	G	C	U	}	A	U	=	A	C	C	G	G	A	Ile
RI1540	B.SUBTILIS	G	A	C	G	G	C	U	}	A	U	=	A	C	C	G	U	C	Ile
RI1662	E.COLI	G	G	C	G	A	C	U	}	A	U	6	A	P	C	G	C	U	Ile

Table: 3.17(b): Consensus sequence for lysidine (k^2C) at 34th position from total 3 sequences:

27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
U G	C G A	C	G	G A	C	U	}	A	U	= 6	A	C P	C	G	G C U	A C U

Figure 3.20(f) depicts the frequency of bases and base pairs at various locations in anticodon stem/loop of the three lysidine containing tRNA sequences. The anticodon loop is rich with pyrimidine bases. The anticodon stem part has five Watson-Crick base pairs, specifically these are at positions 29-41 (C-G), 30-40 (G-C), and 31-39 (A/G-U/C). The presence of lysidine modification is associated with tRNA specific to amino acid isoleucine, Table 3.13.

Such an important modification, 'lysidine', occurs at the first anticodon 'wobble' position, and may play functionally very important role. The lysidine

modification prevents mistranslation of AUG as isoleucine and that of AUA as a methionine. The modification of C to k^2C also switches the aminoacylation identity of the tRNA from methionine to isoleucine [11]. So lysidine plays very important role by strict recognition of A, with no misrecognition of G. This kind of modification may be useful in order to have smooth and in phase protein biosynthesis. Conformational preferences of hypermodified nucleoside lysidine and structural implications of its inclusion in anticodon loop have been studied using various energy calculation methods and are reported in CHAPTER-V, PART-B.

3.9 Conclusion:

The sequence in anticodon loop $A_{36}A_{37}A_{38}$ is conserved for i^6A, ms^2i^6A, and io^6A modifications. The i^6A, ms^2i^6A, io^6A and ncm^5U are associated with amino acid serine. Lysidine is present only in $tRNA^{Ile}$. Anticodon stem part is stabilized by five Watson-Crick base pairs. Anticodon loop positions 32^{nd} and 33^{rd} are always occupied by pyrimidines whenever any of these five hypermodified nucleosides is present. At the 35^{th} position purines are dominant over pyrimidines. Whenever, hypermodified nucleosides i^6A, ms^2i^6A, io^6A at 37^{th} position and ncm^5U, k^2C at 34^{th} position occur, only unmodified uridine (U) occurs at 33^{rd} position. Unmodified adenines are not found at positions 34^{th}. Inosine (I) and unknown modified adenine (H) are the only modified adenines present at 34^{th} positions. Whenever, i^6A and io^6A are present at 37^{th} position, then only modified nucleoside Inosine (I) occurs at 34^{th} position. When hypermodified nucleoside ms^2i^6A is present at 37^{th} position, no occurrence of Inosine at 34^{th} position is found. The modified nucleotide Queuosine occurs at 34^{th} position, in presence of the hypermodified nucleoside i^6A and ms^2i^6A at 37^{th} positions, but queuosine is not found along with io^6A. In $tRNA^{Pro}$ from Torulopsis utilis, the modified nucleoside m^1G (K) at 37^{th} position occurs along with hypermodified nucleoside ncm^5U being present at 34^{th} position.

3.10 References:

1. Agris, P.F. Nucleic Acid Res. Mol. Biol. 1996, 53, 79-129.
2. Bjork, G. R. In tRNA: Structure, Biosynthesis, and Function., Edited by Uttam RajBhandary and Dieter Soll, ASM press, 1995, 165-205.
3. Dirheimer, G. Recent Results Cancer Res. 1983, 84, 15-46.

4. Dirheimer, G.; Keith, G.; Dumas, P.; and Westhof, E. In tRNA: Structure, Biosynthesis, and Function., Edited by Uttam RajBhandary and Dieter Soll, ASM press, 1995, 93-126.
5. Grosjean, H.; Cedergren, R. J. and McKary, W. Biochimie 1982, 64, 387-391.
6. Grosjean, H.; Sprinzl, M. and Steinberg, S. Biochimie 1995, 77, 139-141.
7. Auffinger, P. A. and Westhof, E. In tRNA: Structure, Biosynthesis, and Function., Edited by Uttam RajBhandary and Dieter Soll, ASM press, 1995.
8. Sprinzl, M.; Horn, C.; Brown M.; Ioudovitch, A.; Steinberg, S. Nucleic Acid Res.1998, 26, 148-153.
9. Limbach, P. A.; Crain, P. F. and McCloskey, J. A., Nucleic Acid Res. 1994, 22, 2183-2196.
10. Motorin, Y.; Bec, G.; Tewari, R. and Grosjean, H, RNA, 1997, 3, 721-733.
11. Muramatsu, T.; Nishikawa K.; Nemoto F.; Kuchino Y.; Nishimura S.; Miyazawa T.; Yokoyama S. Nature 1988, 336, 179-181.

CHAPTER-IV

HYPERMODIFIED NUCLEOSIDE 5-CARBAMOYLMETHYL URIDINE (ncm^5U_{34}) AND ITS INTERACTION WITH $IO6A_{37}$

4.1 Introduction

In $tRNA^{Ser}$(UGA) from Nicotiana rustica, the wobble nucleoside 5-carbamoylmethyl uridine (ncm^5U) occurs at 34^{th} 'wobble' position along with hydroxyisopentenyl adenosine (io^6A) present at 37^{th} position [1,2]. The 5-carbamoylmethyl uridine is a modified nucleoside, having substituent in the C-5 position of 'wobble' uridine. The diverse modifications occurring at the first position of anticodon do not hinder canonical Watson-Crick base pairing for reading the third base of the codon.

The hypermodified nucleoside N6-(Δ^2-*cis*-hydroxyisopentenyl)adenosine, *cis*-io^6A, also known as *cis*-ribosylzeatin and N6-(Δ^2-*trans*-hydroxyisopentenyl)adenosine, *trans*-io^6A, also known as *trans*-ribosylzeatin, usually occurs at 3'-adjacent position of anticodon ending in 'A', in some bacteria grown under aerobic conditions. Where as under anaerobic conditions related N6-(Δ^2-isopentenyl)adenosine, i^6A occurs instead at anticodon 3'-adjacent position. The chemical nature of the anticodon 3'-adjacent modifications is strongly correlated with the third base in the anticodon, however, all these modifications prevent extended Watson-Crick base pairing on the 3'-side and may thus help define the proper reading frame for anticodon-codon interactions.

To study the conformational preferences of 5-carbamoylmethyl uridine (ncm^5U_{34}) along with hydroxyisopentenyl adenosine (io^6A_{37}) suitable computational techniques have been used.

4.2 Nomenclature, Convention and Procedure:

Atoms numbering and identification of the torsion angles determining relative orientation and positioning of various atoms for 5-carbamoylmethyl uridine (ncm^5U_{34}) are depicted in (fig. 4.1). Atom C(7) of the base substituent links with C(5) of uridine. The torsion angles η[C(4)C(5)C(7)C(8)] denotes the rotation of C(8) around bond C(5)C(7) and is measured with respect to C(4). The successive chemical bonds along the main extension of the substituent define the subsequent torsion angles ε[C(5)C(7)C(8)N(9)], and ϕ[C(7)C(8)N(9)H]. The glycosyl torsion angle is defined as χ[O1'C1'N(1)C(6)].

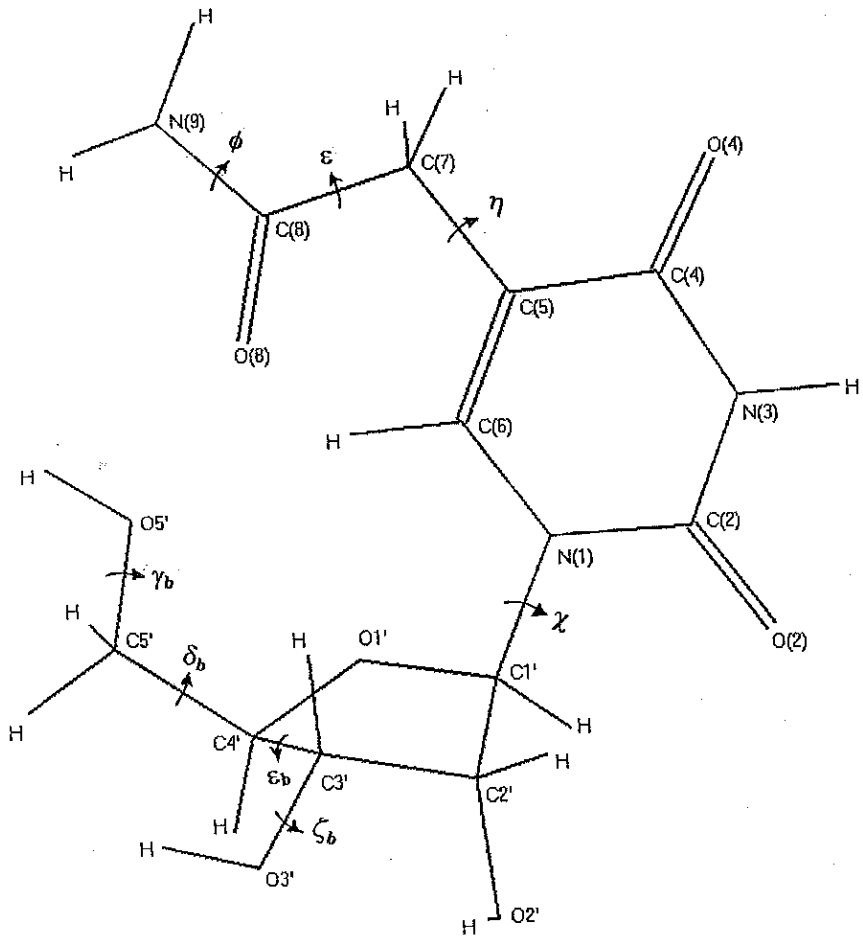


Fig.4.1: Structure of 5-carbamoylmethyl uridine (ncm^5U_{34}). Numbering of the atoms and the various torsion angles are shown.

The atom numbering and torsion angles identification for *cis*-io^6A and *trans*-io^6A is discussed in CHAPTER-III, PART-A, fig.3.1(b). The ribose-phosphate backbone torsion angles and glycosyl torsion angles are held at values specified in Holbrook $tRNA^{Phe}$ model [3] for anticodon loop structure. The 5' and 3' end of anticodon loop segments are truncated using methyl (-CH3) groups.

The quantum chemical perturbative configuration interaction with localized orbitals PCILO method [4-6] has been utilized for the energy calculations of the various molecular conformations for predicting the most stable and alternative stable structures. Full geometry optimization by *ab initio* molecular orbital SCF – HF calculations [7] is carried out to compare the salient features of single nucleoside 5-carbamoylmethyl uridine using 3-21G* basis set (PC SPARTAN Pro Version 6.0.6).

Automated geometry optimization calculations using molecular mechanics force field (MMFF) method is used to compare the salient features for larger molecules.

4.3 Results and Discussion

4.3.1 Hypermodified nucleoside (ncm^5U_{34}):

The PCILO preferred orientation of hypermodified nucleoside 5-carbamoylmethyl uridine (ncm^5U) is specified by the set of torsion angles (η=270°, ε=60°, ϕ=180°), when the base orientation is held at (χ=3°) as in tRNA crystal structure (Holbrook) model. The 5-carbamoylmethyl substituent orientation is stabilized by intramolecular interactions between O(4)…HN(9), O(4)…HC(7), O5'…HC(6), O1'…HC(6) and O(2)…HC2', which are included in Table 4.1. Thus the orientation of 5-carbamoylmethyl substituent allows formation of intramolecular hydrogen bonds involving O(4) of uridine. This hydrogen bonding takes place on the opposite side of usual Watson-Crick base pairing.

The PCILO most stable structure (fig.4.2) is obtained, when glycosyl torsion angle takes the preferred value (χ=333°). With this preferred glycosyl torsion angle value, the favored set of torsion angles for 5-carbamoylmethyl substituent orientation are (η=270°, ε=60°, ϕ=180°). These preferred values are similar to the PCILO structure when (χ=3°).

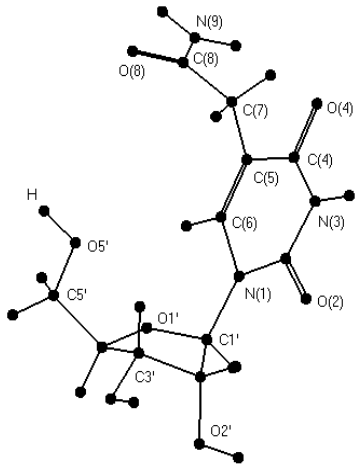


Fig.4.2: The PCILO predicted most stable conformation for 5-carbamoylmethyl uridine (ncm^5U) (η=270°, ε=60°, ϕ=180°, χ=333°).

This preferred most stable conformation (fig. 4.2), is stabilized by comparatively stronger interaction between O(2)…HC2' (Table 4.1), than for PCILO structure at (χ=3°). However, the interaction between O5'…HC(6) is no longer possible due to change in glycosyl torsion angle to (χ=333°). The intramolecular interactions between O(4)…HN(9), O(4)…HC(7) remain as stabilizing factors. The interaction between O1'…HC(6) may also stabilize the structure. The hydrogen bonding between O(4)…HN(9) takes place through opposite side of Watson-Crick base pairing edge.

Table 4.1: Geometrical parameters for hydrogen bonding in the preferred conformation of hypermodified nucleoside, 5-carbamoylmethyl uridine.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Glycosyl Torsion Angle (χ°)
O(4)…H-N(9)	2.22	1.09	132.60	3, 333
O(4)…H-C(7)	2.27	1.09	97.16	3, 333
O5'…H-C(6)	2.73	1.09	118.53	3
O1'…H-C(6)	2.36	1.09	102.92	3
O(2)…H-C2'	2.72	1.09	107.93	3
O1'…H-C(6)	2.49	1.09	99.41	333
O(2)…H-C2'	2.10	1.09	122.96	333

Table 4.2: Predicted alternative stable conformation for hypermodified nucleoside, 5-carbaomylmethyl uridine (ncm^{5}U).

Torsion angles	Rel. Eng. (kcal/mol)	Fig.Ref.
η=270°, ϵ=60°, ϕ=180°, χ=333°	0.0	4.2
η=30°, ϵ=60°, ϕ=180°, χ=333°	1.5	
η=270°, ϵ=330°, ϕ=180°, χ=333°	2.3	

Higher energy (1.5 kcal/mol) PCILO alternative stable conformation, Table 4.2, occurs by flipping of (η=30°). Interaction between O(4)…HN(9), and O(4)…HC(7) is no longer possible. Other interactions between O(2)…HC2', O1'…HC(6) remain as stabilizing factors. Next higher energy (2.3 kcal/mol) alternative conformation arises with ε altered to (ε=330°). The interaction between O(4)…HN(9) disappears. Other interactions between O(4)…HC(7), O(2)…HC2' and O1'…HC(6) may contribute to structural stability.

The automated geometry optimization through molecular mechanics force field (MMFF) method yields (η=-69.5°, ε=101.2°, ϕ=-171.2°, χ=16.3°). The values of torsion angles η, ϕ and χ match closely with PCILO results. The torsion angle ε deviates by nearly 40° as compared to PCILO most stable structure. The full geometry optimization using Hartree-Fock SCF quantum mechanical energy calculations has also been carried out to compare the salient features. The resulting optimized torsion angle values are (η=-85.7°, ε=78.5°, ϕ=-176.2°, χ=-6.6°). The torsion angles η, ϕ show minor differences than PCILO most stable structure (fig. 4.2). The torsion angle ε deviates by nearly 18° from the PCILO results and the glycosyl torsion angle χ shows 20° difference. The orientation of 5-carbamoylmethyl uridine (ncm^{5}U) retains conformation similar to the PCILO preferred conformation (fig. 4.2).

4.3.2 Model anticodon loop hypermodified segments

4.3.2.1 *Cis*-Me-p-m^{3}C-p-U$_{(33)}$-p-ncm^{5}U$_{(34)}$-p-G$_{(35)}$-p-A$_{(36)}$-p-io^{6}A$_{(37)}$-p-A$_{(38)}$-p-Me:

Hypermodified nucleosides *cis*-io^{6}A at 37th position and ncm^{5}U at 34th position are included in model anticodon loop segment. The ribose-phosphate backbone torsion angles are held at values specified in Holbrook tRNA^{Phe} [3] structure for corresponding nucleotides at 32nd, 33rd, 34th, 35th, 36th, 37th and 38th positions. The glycosyl torsion angles also retain the values specified in Holbrook model ($\chi_{(32)}$=31°, $\chi_{(33)}$=24°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angle values for the preferred hydroxyisopentenyl substituent orientation, io^{6}A are (α=0°, β=90°, γ=270°, δ=0°, ψ1=150°, ψ2=0°,±120°, θ=60°) and for 5-carbamoylmethyl substituent (ncm^{5}U) are (η=150°, ε=270°, ϕ=180°). The torsion angle values for hydroxyisopentenyl substituent (*cis*-io^{6}A$_{37}$) in anticodon loop may be compared with (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°) for isolated hypermodified base *cis*-io^{6}Ade or *cis*-zeatin [8]. The torsion angles β, γ, ψ1 and θ are

taking different values than isolated base [8]. The hydroxyisopentenyl substituent in model anticodon loop segment does not show intramolecular interactions between $N(1)_{(37)}$…HC(11).

The 5-carbamoylmethyl uridine, ncm^5U_{34}, is stabilized by intramolecular interactions between $O(4)_{(34)}$…HC(7), O(8)…$HC(6)_{(34)}$. Interactions of $ncm^5U_{(34)}$ with ribose-phosphate backbone like $pO_{(33\text{-}34)}$…HC(7) and $pO_{(33\text{-}34)}$…HC(6) also provide stability. There is no interaction between hydroxyisopentenyl substituent at 37th position (io^6A_{37}) and 5-carbamoylmethyl uridine (ncm^5U_{34}). Both the hypermodified substituents are far away from each other. The conformation of 5-carbamoylmethyl uridine (ncm^5U_{34}) takes extended orientation. The torsion angle (η=150°) takes '*trans*' orientation. The 5-carbamoylmethyl substituent does not obstruct access to Watson-Crick base pairing sites of U_{34}.

When the glycosyl torsion angle values for nucleotides in the anticodon loop segment are varied freely the resulting preferred values are ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°). Adopting these preferred values, the new set of torsion angle values for the preferred hydroxyisopentenyl substituent orientation (fig. 4.3) (α=60°, β=300°, γ=150°, δ=0°, ψ1=240°, ψ2=0°,±120°, θ=300°) and the set of torsion angles for the preferred orientation of 5-carbamoylmethyl substituent are (η=150°, ε=270°, ϕ=180°). These new preferred torsion angle values for hydroxyisopentenyl substituents orientation in model anticodon loop segment may be compared with the preferred orientation of the isolated hypermodified base [8] *cis*-io^6Ade or *cis*-zeatin (α=0°, β=0°, γ=±120°, δ=0°, ψ1=0°, ψ2=0°,±120°, θ=180°). The torsion angles α, β, γ and θ are preferring different values than isolated hypermodified base [8] *cis*-io^6Ade or *cis*-zeatin. Interactions stabilizing favored hydroxyisopentenyl substituent orientation in model anticodon loop segment are $N(1)_{(37)}$…HC(11), $O1'_{(33)}$…HC(10), $O1'_{(33)}$…HC(13), $O2'_{(32)}$…HC(14), $N(7)_{(36)}$…HC(10), $N(6)_{(36)}$…HC(10), $O(6)_{(35)}$…HC(13) and O(13)…$HN(6)_{(36)}$ which are also included in Table 4.3. These interactions show that hydroxyisopentenyl substituent comes close to the ribose ring of 32nd and 33rd nucleotides in the anticodon loop. The hydroxyl oxygen O(13) interacts with HN(6) of 36th adenosine, where as the C(14)H interacts with $O2'_{(32)}$. The interaction between $N(1)_{(37)}$…HC(11) is similar to that of isolated modified base *cis*-io^6Ade or *cis*-zeatin [8].

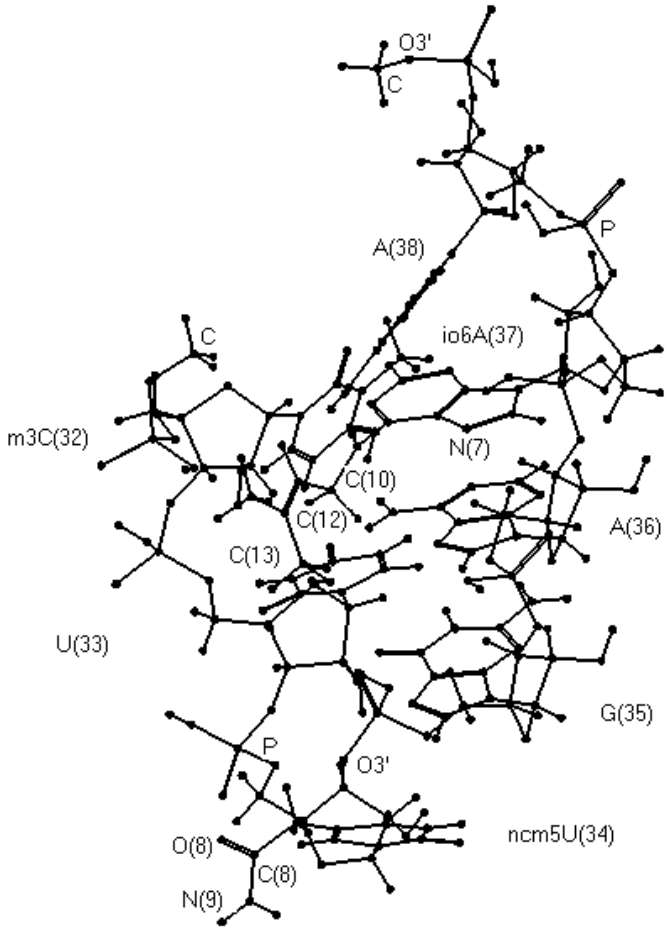


Fig. 4.3: Predicted PCILO most stable structure for *cis*-io^{6}A at 37th position and ncm^{5}U at 34th position in model anticodon loop segment, preferred glycosyl torsion angle values are ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°).

The 5-carbamoylmethyl substituent occurs at 34th position, its orientation is stabilized by intramolecular interactions between O(4)$_{(34)}$…HC(7) and O(8)…HC(6)$_{(34)}$. Interactions between 5-carbamoylmethyl substituent and ribose-phosphate backbone, like pO$_{(33-34)}$…HC(7), and pO$_{(33-34)}$…HC(6) also provide stabilization. Interactions between adjacent bases and ribose-phosphate backbone in the model anticodon loop N(7)$_{(38)}$…HC(2)$_{(37)}$, N(1)$_{(37)}$…HN(6)$_{(38)}$, N(6)$_{(38)}$…HC(2)$_{(37)}$, O(2)$_{(32)}$…HN(6)$_{(37)}$, N(3)$_{(36)}$…HN(2)$_{(35)}$, N(3)$_{(33)}$…HC(5)$_{(32)}$, O(4)$_{(33)}$…HN(4)$_{m3c(32)}$, O(4)$_{(33)}$…HC(5)$_{(32)}$, O5'$_{(36)}$…HC(5)$_{(33)}$ and pO$_{(37)}$…HC(3)$_{m3c(32)}$ may provide additional stability and are shown in Table 4.3.

Table 4.3:Geometrical parameters for hydrogen bonding in the preferred conformations of *cis*- and *trans*- $io^6A_{(37)}$, and $ncm^5U_{(34)}$, in model anticodon loop structure.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure Reference
$N(1)_{(37)}$…H-C(11)	2.73	1.09	91.65	4.3, 4.4
$O1'_{(33)}$…H-C(10)	2.26	1.09	101.34	4.3, 4.4
$O1'_{(33)}$…H-C(13)	1.88	1.09	132.10	4.3
O(13)…H-C(11)	2.32	1.09	108.78	4.4
$O1'_{(33)}$…H-C(14)	1.88	1.09	132.20	4.4
$O2'_{(32)}$…H-C(14)	2.68	1.09	90.40	4.3
$O2'_{(32)}$…H-O(13)	2.28	0.96	121.36	4.4
$O(6)_{(35)}$…H-C(14)	2.82	1.09	128.99	4.4
$N(7)_{(36)}$…H-C(10)	2.39	1.09	139.64	4.3, 4.4
$N(6)_{(36)}$…H-C(10)	2.37	1.09	116.95	4.3, 4.4
$O(6)_{(35)}$…H-C(13)	2.82	1.09	128.91	4.3
O(13)…H-$N_{(36)}$	2.30	1.09	99.91	4.3
$O(4)_{(34)}$…H-$C(7)_{(34)}$	2.27	1.09	97.19	4.3, 4.4
O(8)…H-$C(6)_{(34)}$	2.68	1.09	118.69	4.3, 4.4
$PO_{(33\text{-}34)}$…H-C(7)	2.69	1.09	95.49	4.3, 4.4
$PO_{(33\text{-}34)}$…H-C(6)	2.63	1.09	123.93	4.3, 4.4
$N(7)_{(38)}$…H-$C(2)_{(37)}$	1.52	1.09	84.46	4.3, 4.4
$N(1)_{(37)}$…H-$N(6)_{(38)}$	2.12	1.09	154.06	4.3, 4.4
$N(6)_{(38)}$…H-$C(2)_{(37)}$	1.76	1.09	131.62	4.3, 4.4
$O(2)_{(32)}$…H-$N(6)_{(37)}$	2.42	1.09	128.85	4.3, 4.4
$N(3)_{(36)}$…H-$N(2)_{(35)}$	2.58	1.09	105.08	4.3, 4.4
$N(3)_{(33)}$…H-$C(5)_{(32)}$	2.58	1.09	131.53	4.3, 4.4
$O(4)_{(33)}$…H-$N(4)_{m3c(32)}$	2.63	1.09	138.76	4.3, 4.4
$O(4)_{(33)}$…H-$C(5)_{(32)}$	2.25	1.09	149.39	4.3, 4.4
$O5'_{(36)}$…H-$C(5)_{(33)}$	2.33	1.09	124.21	4.3, 4.4
PO(37)…H-$C(3)_{m3c(32)}$	2.27	1.09	142.44	4.3, 4.4

Thus the orientation of hydroxyisopentenyl substituent (io^6A_{37}) in anticodon loop segment, with the preferred values for glycosyl torsion angles ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°) favors orientation somewhat similar to that of isolated hypermodified base [8] *cis*-io^6Ade or *cis*-zeatin. The torsion angle values (α=60°, β=300°, γ=150°) are different than for the isolated base (α=0°, β=0°, γ=±120°). But it retains the intramolecular hydrogen bonding between $N(1)_{(37)}$…HC(11) as for the isolated hypermodified base [8] *cis*-io^6Ade or *cis*-zeatin. This kind of interaction is absent in anticodon loop segment when glycosyl torsion angle values are held as in Holbrook $tRNA^{Phe}$ [3] model. Inspite of allowing glycosyl torsion angle to change the N(1) site of adenine at 37th position remains shielded.

Thus possible Watson-Crick base pairing of io^6A_{37} is prevented. This prevents misreading of mRNA. The orientation of 5-carbamoylmethyl substituent (ncm^5U) retains similar conformation irrespective of freely varying glycosyl torsion angles to take the preferred values ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°) or being held at values specified in Holbrook $tRNA^{Phe}$ [3] model. As in isolated nucleoside ncm^5U (fig. 4.2), the interaction between $O(4)_{(34)}$…HC(7) is retained in anticodon loop as well. However, likely hydrogen bonding between $O(4)_{(34)}$…HN(9) in isolated nucleoside is not possible when incorporated in anticodon loop structure. This may be due to the trans orientation (η=150°, ε=270°, ϕ=180°) favored for 5-carbamoylmethyl substituent in the loop. The preferred glycosyl torsion angle for ncm^5U_{34} retains the 'anti' conformation and is the same as in Holbrook data ($\chi_{(34)}$=3°). There is no interaction observed between io^6A_{37} and ncm^5U_{34}, even after allowing glycosyl torsion angles to change.

Higher energy (0.4 kcal/mol) PCILO stable alternative conformation included in Table 4.4, may be arrived by flipping of (θ=120°). Intramolecular interaction between $N(1)_{(37)}$…HC(11), remains a stabilizing factor. The $O(13)_{(37)}$…$HN(6)_{(36)}$ interaction may provide additional stability. Next higher energy (4.4 kcal/mol) alternative is reached when (ψ1=150° and θ=120°). This structure is stabilized by $N(1)_{(37)}$…HC(11) and $N(6)_{(36)}$…HC(13) interactions. Interaction between $O(13)_{(37)}$…$HN(6)_{(36)}$ is no longer indicated.

Table 4.4: Predicted alternative conformation for hypermodified nucleosides, io^{6}A (*cis*- or *trans*-io^{6}A) in anticodon loop structure.

Torsion Angles	Rel. Eng. (kcal/mol)	Fig. Ref.
α=60°, β=300°, γ=150°, δ=0°, ψ1=240°, ψ2=0°,±120°, θ=300°	0.0	4.3
α=60°, β=300°, γ=150°, δ=0°, ψ1=240°, ψ2=0°,±120°, θ=120°	0.4	
α=60°, β=300°, γ=150°, δ=0°, ψ1=150°, ψ2=0°,±120°, θ=120°	4.4	
α=60°, β=300°, γ=150°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=300°	0.0	4.4
α=60°, β=300°, γ=150°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=60°	0.4	
α=60°, β=300°, γ=150°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=180°	0.7	
α=60°,β=300°, γ=150°,δ=180°, ψ1=180°, ψ2=0°,±120°,θ=60°	0.8	

Automated geometry optimization through molecular mechanics force field (MMFF) method have been carried out to compare the salient feature and are shown in Table 4.5. The MMFF results also show orientation of hydroxyisopentenyl substituent and 5-carbamoylmethyl substituent similar to that found by PCILO method. The torsion angles values for α, γ, and ψ1 differ within 25° from PCILO preferred values, where as the torsion angles β, θ, η and ε show somewhat larger differences from PCILO results. The torsion angles δ and ψ2 have nearly similar values.

4.3.2.2 *Trans*-Me-p-m^{3}C-p-U$_{(33)}$-p-ncm^{5}U$_{(34)}$-p-G$_{(35)}$-p-A$_{(36)}$-p-io^{6}A$_{(37)}$-p-A$_{(38)}$-p-Me

The model anticodon loop segment contains *trans*-io^{6}A at 37th position and ncm^{5}U at 34th position. The ribose-phosphate backbone and glycosyl torsion angle values are as specified in Holbrook tRNA^{Phe} [3] model. The glycosyl torsion angle values for corresponding to nucleotides at the 32nd, 33rd, 34th, 35th, 36th, 37th, and 38th positions are ($\chi_{(32)}$=31°, $\chi_{(33)}$=24°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=3°). The set of torsion angle values for the preferred orientation of the hydroxyisopentenyl substituent in the anticodon loop segment for *trans*-io^{6}A are (α=0°, β=90°, γ=270°, δ=180°, ψ1=210°, ψ2=180°,±60°, θ=300°) and for 5-carbamoylmethyl substituent are (η=150°, ε=270°, ϕ=180°). These torsion angles of hydroxyisopentenyl substituent favored in anticodon loop segment may be compared with the values for the isolated

hypermodified base *trans*-io^{6}Ade or *trans*-zeatin [8] (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=±60°). The torsion angles β, γ and ψ1 of hydroxyisopentenyl substituent taking different values in anticodon loop. This PCILO favored structure does not show intramolecular interaction between $N(1)_{(37)}$…HC(11). The wobble nucleoside however has stabilizing intramolecular interactions $O(4)_{(34)}$…HC(7) and O(8)…$HC(6)_{(34)}$. The ncm$^5U_{(34)}$ orientation is also stabilized by interactions with ribose-phosphate backbone like $pO_{(33\text{-}34)}$…HC(7) and $pO_{(33\text{-}34)}$…HC(6). There is no interaction between hydroxyisopentenyl substituent at 37th position (io^{6}A) and 5-carbamoylmethyl substituent (ncm^{5}U) at 34th position. Both these hypermodified nucleosides keep apart. The torsion angle (η=150°) takes '*trans*' orientation. The 5-carbamoylmethyl substituent orientation does not limit access to potential Watson-Crick base pairing sites of $U_{(34)}$.

After allowing the glycosyl torsion angles in the anticodon loop segment to change freely the resulting preferred values are ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°). With these preferred values of glycosyl torsion angles adopted, search for the preferred orientation of *trans*-io^{6}A (fig. 4.4) at 37th position in anticodon loop segment yields (α=60°, β=300°, γ=150°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=300°) and for 5-carbamoylmethyl substituent, the preferred set of torsion angles are (η=150°, ε=270°, ϕ=180°). The torsion angles α, β and γ of hydroxyisopentenyl substituent in anticodon loop are preferring different values as compared to the preferred orientation of isolated hypermodified base [8] *trans*-io^{6}Ade or *trans*-zeatin (α=0°, β=0°, γ=±120°, δ=180°, ψ1=0°, ψ2=0°,±120°, θ=±60°).

The interactions contributing to stability of anticodon loop structure (Table 4.3), having *trans*-io^{6}A, are, $N(1)_{(37)}$…HC(11), $N(7)_{(36)}$…HC(10), $N(6)_{(36)}$…HC(10), $O(13)_{(37}$…HC(11), $O1'_{(33)}$…HC(10), $O1'_{(33)}$…HC(14), $O2'_{(32)}$…HO(13) and $O(6)_{(35)}$…HC(14). Other interactions between adjacent bases and ribose-phosphate backbone are $N(7)_{(38)}$…$HC(2)_{(37)}$, $N(1)_{(37)}$…$HN(6)_{(38)}$, $N(6)_{(38)}$…$HC(2)_{(37)}$, $O(2)_{(32)}$…$HN(6)_{(37)}$, $N(3)_{(36)}$…$HN(2)_{(35)}$, $N(3)_{(33)}$…$HC(5)_{(32)}$, $O(4)_{(33)}$…$HN(4)_{m3c(32)}$, $O(4)_{(33)}$…$HC(5)_{(32)}$, $O5'_{(36)}$…$HC(5)_{(33)}$ and $pO_{(37)}$…$HC(3)_{m3c(32)}$. These also stabilize the structure and are included in (Table 4.3).

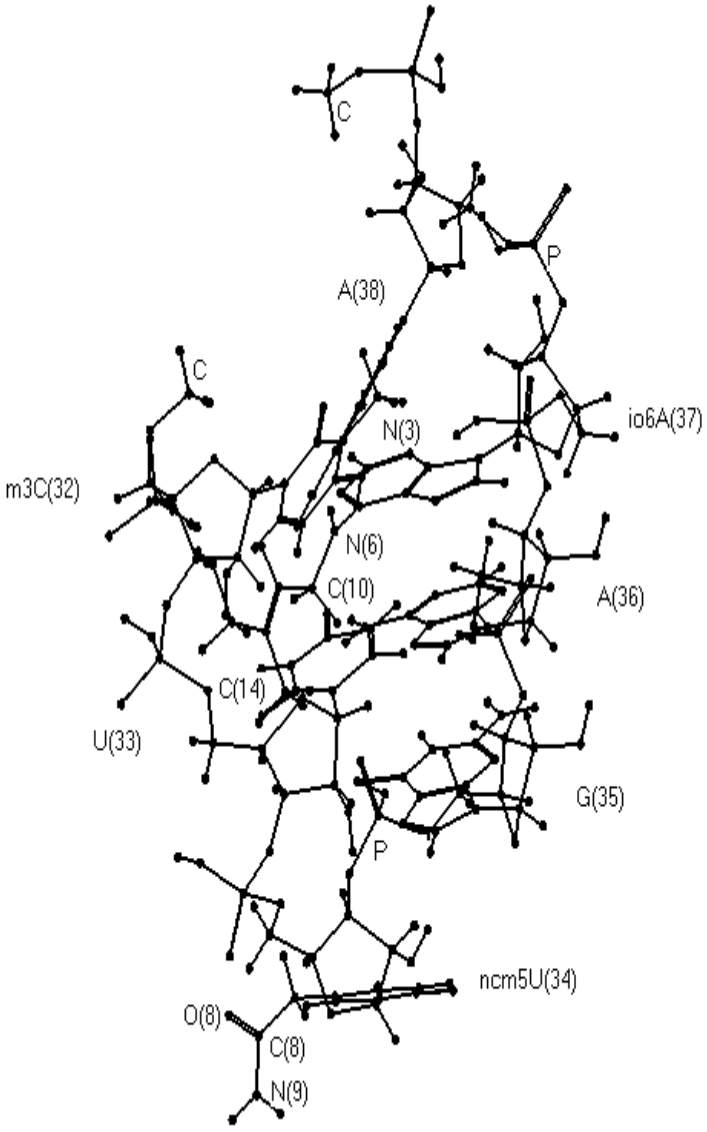


Fig. 4.4: Predicted PCILO most stable structure for *trans*-io^{6}A at 37th position and ncm^{5}U at 34th position in model anticodon loop segment, preferred glycosyl torsion angle values are ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°).

The 5-carbamoylmethyl uridine orientation is stabilized by intramolecular interactions between O(4)$_{(34)}$…HC(7) and O(8)…HC(6)$_{(34)}$. The 5-carbamoylmethyl substituent also gets stabilized due to interactions like pO$_{(33\text{-}34)}$…HC(7) and pO$_{(33\text{-}34)}$…HC(6)$_{(34)}$ with ribose-phosphate backbone.

The favored orientation of hydroxyisopentenyl substituent for (*trans*-io^{6}A) in anticodon loop segment, with the preferred glycosyl torsion angles ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°) is similar to that of isolated hypermodified base [8] *trans*-io^{6}Ade or *trans*-zeatin. The intramolecular interaction between N(1)$_{(37)}$…HC(11) is present in this model anticodon loop segment as also predicted for isolated hypermodified base [8] *trans*-io^{6}Ade or *trans*-zeatin. However, the torsion angle values favored in the loop segment (α=60°, β=300°, γ=150°) are significantly different from the favored values for isolated hypermodified base (α=0°, β=0°, γ=±120°).

The orientation of 5-carbamoylmethyl substituent (ncm^{5}U) retains similar conformation, although glycosyl torsion angles may take the preferred values ($\chi_{(32)}$=91°, $\chi_{(33)}$=234°, $\chi_{(34)}$=3°, $\chi_{(35)}$=8°, $\chi_{(36)}$=1°, $\chi_{(37)}$=22°, $\chi_{(38)}$=33°) or retain values specified in Holbrook tRNA^{Phe} model. The torsion angle values for the favored 5-carbamoylmethyl uridine orientation (ncm^{5}U$_{(34)}$) in both the cases are (η=150°, ε=270°, ϕ=180°, $\chi_{(34)}$=3°). The glycosyl torsion angle for (ncm^{5}U$_{(34)}$) specified in Holbrook tRNA^{Phe} model ($\chi_{(34)}$=3°), also turns out to be the preferred value and thus remains unchanged.

Higher energy (0.4 kcal/mol) PCILO stable alternative conformation presented in Table 4.4, is arrived at by taking (θ=60°). The intramolecular interactions between N(1)$_{(37)}$…HC(11), O(13)…HC(11), O1'$_{(33)}$…HC(10) and O1'$_{(33)}$…HC(14)$_{(37)}$ contribute to stability and are included in Table 4.3. However, this structure lacks the interaction between O2'$_{(32)}$…HO(13). Next higher energy (0.7 kcal/mol) alternative conformation may be arrived at by flipping of (θ=180°). This structure also lacks the interaction between O2'$_{(32)}$…HO(13). No other interesting interactions are noticeable. Another higher energy (0.8 kcal/mol) alternative conformation is reached by flipping of (ψ1=180°) along with (θ=60°). The interaction between O(13)…HC(11) is no longer possible. Interactions between N(1)$_{(37)}$…HC(11), O1'$_{(33)}$…HC(10) and O1'$_{(33)}$…HC(14) may contribute to stability.

Results of automated geometry optimization through molecular mechanics force field (MMFF) method are reasonably close to PCILO data and are shown in Table 4.5. The orientation of hydroxyisopentenyl adenosine and 5-carbamoylmethyl substituent is similar to PCILO preferred structure (fig. 4.4). The torsion angles α, and

β differ by 30° from PCILO values, where as torsion angles γ, δ, Ψ1, Ψ2, and θ have nearly similar values. The torsion angle η shows larger deviation from PCILO most stable values, but ε and ϕ agree with PCILO results.

Table 4.5: Optimized values of the torsion angles by automated geometry optimization using MMFF method for *cis*- and *trans*-io^{6}A$_{(37)}$ along with ncm^{5}U$_{(34)}$, starting from the PCILO preferred values.

Torsion angles	Fig. ref.
α=47°, β=-104°, γ=175°, δ=4°, ψ1=-82°, ψ2=-6°, θ=173°, η=52°, ε=-126°, ϕ=171°, $\chi_{(32)}$=87°, $\chi_{(33)}$=-96°, $\chi_{(34)}$=2°, $\chi_{(35)}$=10°, $\chi_{(36)}$=-30°, $\chi_{(37)}$=29°, $\chi_{(38)}$=55°. Rel. Energy:= 20.2 kcal/mol	4.3
α=32°, β=-91°, γ=-151°, δ=175°, ψ1=-9°, ψ2=14°, θ=-52°, η=58°,ε=-103°, ϕ=173°, $\chi_{(32)}$=99°, $\chi_{(33)}$=-100°, $\chi_{(34)}$=8°, $\chi_{(35)}$=20°, $\chi_{(36)}$=-25°, $\chi_{(37)}$=27°, $\chi_{(38)}$=65°. Rel. Energy:= 0.0 kcal/mol	4.4

4.3.2.3 Comparison between anticodon loop segments with *cis*-io^{6}A and *trans*-io^{6}A:

The hydroxyisopentenyl substituent spreads away from five membered imidazole moiety of adenosine in *cis*-io^{6}A as well as in *trans*-io^{6}A. The hydroxyl oxygen O(13) participates in hydrogen bonding with HC(11) in case of *trans*-io^{6}A, but this interaction is absent in *cis*-io^{6}A. The O2' of 32nd nucleoside interacts with HC(14) of *cis*-io^{6}A, where as in *trans*-io^{6}A it interacts with hydroxyl oxygen HO(13) as evident from, Table 4.3. The O(6) of 35th guanosine and O1'of 33rd ribose ring interact with HC(13) of *cis*-io^{6}A, where as these interact with HC(14) in case of *trans*-io^{6}A as evident from Table 4.3. The O(13) of *cis*-io^{6}A is involved in hydrogen bonding with HN(6) of 36th adenosine, but such kind of interaction is absent in *trans*-io^{6}A. The PCILO preferred *trans*-io^{6}A structure (fig. 4.4) is more stable (1.4 kcal/mol) than PCILO preferred *cis*-io^{6}A structure (fig. 4.3). The full geometry optimization through molecular mechanics force field (MMFF) method also supports these results. The results of (MMFF) method presented in Table 4.5 show that *trans*-io^{6}A is 20.2

kcal/mol more stable than *cis*-io^{6}A. The MMFF geometry optimization calculations utilized PCILO preferred structures as initial structures for geometry optimization.

4.4 Conclusion:

The base substituent orientation of hypermodified nucleoside 5-carbamoylmethyl uridine is found to remain unaffected whether glycosyl torsion angle value as specified in Holbrook tRNA^{Phe} model [3] ($\chi_{(34)}$=3°) is retained or the preferred value ($\chi_{(34)}$=333°) is used instead. The 5-carbamoylmethyl substituent orientation does not obstruct codon-anticodon interactions.

In the seven nucleotide anticodon loop structure of tRNA^{Ser}(UGA), the 5-carbamoylmethyl substituent takes extended '*trans*' orientation. The –NH2 group of ncm^{5}U$_{34}$ in seven nucleotide anticodon loop structure is away from O(4) of U$_{(34)}$. The orientation of 5-carbamoylmethyl substituent is found similar irrespective of whether glycosyl torsion angle values specified in Holbrook tRNA^{Phe} model [3] are used or the glycosyl torsion angle preferred values are utilized. The hydroxyisopentenyl adenosine (*cis*- or *trans*-io^{6}A) is accommodated well in seven nucleotide anticodon loop structure. After allowing glycosyl torsion angles to change, the orientation of hydroxyisopentenyl substituents in (*cis*- or *trans*-io^{6}A) anticodon loop segment is found similar to that for isolated hypermodified base [8] *cis*- or *trans*-zeatin. In these two modified nucleosides the C(11)H is favorably placed to participate in intramolecular hydrogen bonding with N(1). Thus the N(1) side remains shielded in these two hypermodified nucleosides (*cis*- or *trans*- io^{6}A) in anticodon loop structure. Such orientation of hydroxyisopentenyl substituents in (*cis*- or *trans*- io^{6}A) may be helpful for maintaining the reading frame during translation and prevent misreading of mRNA.

The 5-carbamoylmethyl substituent (ncm^{5}U$_{(34)}$) and hydroxyisopentenyl substituents in (*cis*- or *trans*- io^{6}A) are not interacting with each other.

4.5 References:

1. Teichmann, T.; Urban, C.; Beier, H.; Plant Mol. Biol. 1994, 24, 889-901.
2. Sprinzl, M.; Horn, C.; Brown M.; Ioudovitch, A.; Steinberg, S.; Nucleic Acid Res. 1998, 26, 148-153.
3. Holbrook, S. R.; Sussman, J. L.; Warrant, R.W. and Kim, S.H. J. Mol. Biol. 1978, 123, 631-660.

4. Diner, S.; Malrieu, J. P.; Claverie, P. Theoret. Chim. Acta. 1969, 15, 100-110.
5. Malrieu, J. P. in Semiempirical Methods of Electronic Structure Calculations, Part A, Techniques, Segal, G. A.; Ed.; Plenum, New York, 1977, p. 69-103.
6. Diner, S.; Malrieu, J. P.; Jordan, F.; Gilbert, M. Theoret. Chim. Acta. 1969, 13, 1-17.
7. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A.; in *Ab Initio* Molecular Orbital Theory, Wiley, New York, 1986.
8. Sonawane, K. D.; Sonavane, U. B.; Tewari, R. J. Biomol. Struct. Dyn. 2002, 19, (4), 637-648.

CHAPTER-V

STUDIES ON OTHER INTERESTING MODIFIED NUCLEOSIDS

PART-A: N6-(N-glycylcarbonyl) adenine (gc^6Ade)

5.1 Introduction

The hypermodified nucleic acid base N6-(N-glycylcarbonyl) adenine, gc^6Ade, occurs at anticodon 3'-adjacent position in specific tRNAs with uridine (U) terminated anticodon [1, 2]. The N(1), N(3) and N(7) are the possible sites of protonation in adenine (fig. 5.1). The effects of intermolecular hydrogen bond donor - acceptor interactions, as present in crystal structure packing environment, could be simulated through protonation of the acceptor site in gc^6Ade [3, 4].

The effects of the single protonation at each of the N(3), N(7), or N(1) sites in gc^6Ade have been reported earlier [4-6]. Conformational flipping induced by diprotonation of gc^6Ade has been reported from crystal structure investigations [7], however, multiple protonation theoretical studies on hypermodified bases have not been reported so far. The present work gives more information about the preferred conformations of (N(1), N(3)), (N(3), N(7)), and (N(1), N(7)) diprotonated gc^6Ade. Unlike the N(1) or N(7) site towards which O(10) may orient alternatively, the N(3) site is remote and does not interact directly. Thus it is of interest to inquire what influence N(3) protonation has with accompanying N(1) or N(7) protonation. It is also very interesting to inquire which way O(10) orients preferably when both the N(1) and N(7) sites are protonated.

5.2 Nomenclature, Convention, and procedure:

Fig. 5.1 shows the numbering of the atoms in the gc^6Ade molecule along with the identification of the various torsion angles, which describe rotation around the respective singly bonded acyclic chemical bonds and are required to describe the molecular conformation. The torsion angle α [N(1)C(6)N(6)C(10)] describes rotation around the bond C(6)N(6) and measures the orientation of the bond N(6)C(10) with respect to the C(6)N(1) from the cis (eclipsed, 0°) position in the right-hand sence of rotation. Likewise, the torsion angles β [C(6)N(6)C(10)N(11)], γ [N(6)C(10)N(11)C(12)], δ[C(10)N(11)C(12)C(13)], ϵ [N(11)C(12)C(13)O(13a)], θ [C(12)C(13)O(13a)H] define rotation of the succesive chemical bonds along the main extension of the substituent. The diagram also depicts N(1), N(3), and N(7) protonation sites in gc^6Ade.

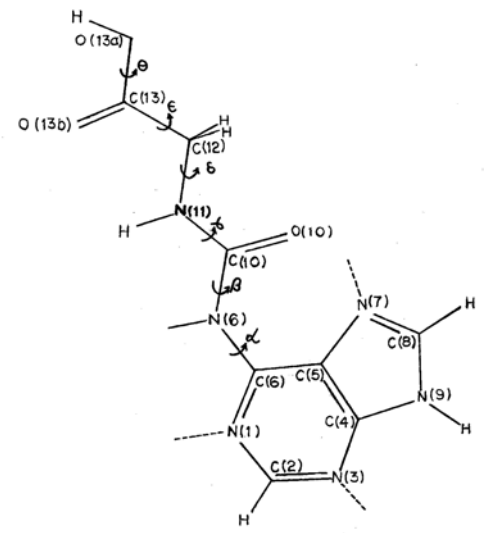


Fig. 5.1 Atoms numbering and torsion angles identification in gc^{6}Ade molecule. An all-*trans* ($\alpha = \beta = \gamma = \delta = \epsilon = \theta = 180°$), fully extended but proximal conformation is shown here. The N(1), N(3), and N(7) protonation sites are indicated.

For all the energy calculations of various molecular conformations the quantum chemical perbturbative configuration interaction with localized orbitals (PCILO) method has been used [8, 9].

The most stable structure and the alternative stable conformations have been searched in the multidimensional conformational space using the logical selection of grid points approach [10]. For N(1), N(3), and N(7) protonation sites alike, the hydrogen nucleus is placed 1.0 A° apart along the direction of the external angle bifurcator. The total number of electrons in the molecule, however, remain unchanged consequent to protonation. In order to compare PCILO results with alternative molecular orbital methods, full geometry optimization calculations in which all bond distances and angles are also allowed to change, besides all the torsion angles, have been made using Tripos's Sybyl 6.5 Version implementation of PM3, and MNDO methods [11-13] for a few selected conformations.

5.3 Results and Discussions:

The predicted most stable structure based on PCILO calculations, for the ((N(1), N(3)) diprotonated gc^6Ade is depicted in (fig.5.2). The preferred values of the torsion angles are α=0°, β=180°, γ=0°, δ=300°, ϵ=240°, and θ=180°. The acceptor carbonyl oxygen O(10) is suitably placed, on the same side, as the donor HN(1) for associating through the six membered (O(10)C(10)N(6)C(6)N(1)H) hydrogen bonding.

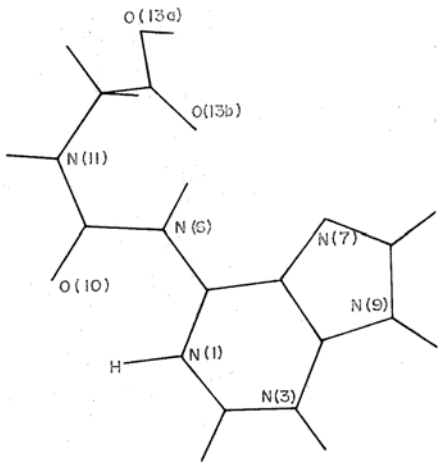


Fig. 5.2 Preferred conformation of (N(1), N(3)) - diprotonated gc^6Ade.

Additionally, the carboxyl oxygen O(13b) is favorably placed for participating in the seven membered (O(13b)C(13)C(12)N(11)C(10)N(6)H) hydrogen bonding interaction with the HN(6). The geometrical parameters for the likely N(6)H…O(13b) and N(1)H…O(10) hydrogen bonding are shown in Table 5.1. These geometrical parameters indicate that the N(6)H…O(13b) hydrogen bond is shorter than the N(1)H…O(10) bond and provides stability to the structure in (fig. 5.2). The corresponding full geometry optimized torsion angle values using PM3, and MNDO methods are also quite similar, and are shown in Table 5.2.

The predicted most stable structure for the (N(3), N(7)) diprotonated gc^6Ade is shown in (fig. 5.3). The acceptor, carbonyl oxygen O(10) is placed here on the same side as the donor HN(7) and stabilize the structure through the seven-membered (O(10)C(10)N(6)C(6)C(5)N(7)H) hydrogen bond. The hydrogen bonding between N(6)H…O(13b) provides the stabilization. The torsion angles for the structure in (fig. 5.3) are α=180°, β=180°, γ=0°, δ=300°, ϵ=210°, and θ=180°. The hydrogen bonding

parameters are shown in Table 5.1. The results of full geometry optimization using PM3 and MNDO method are shown in Table 5.2.

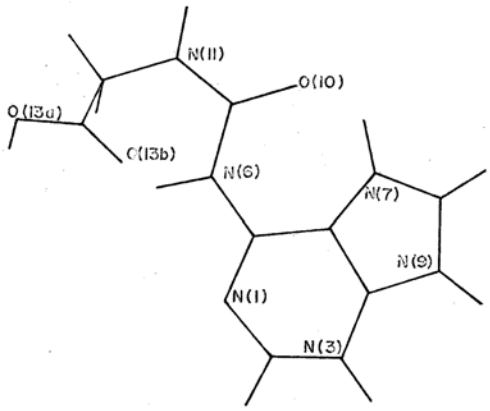
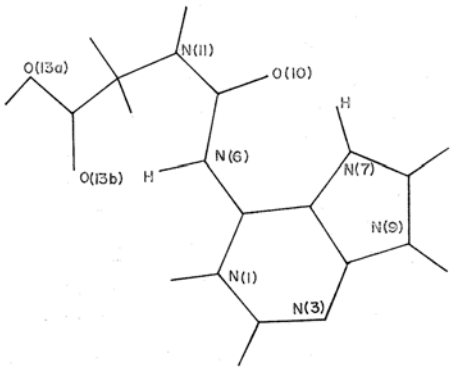


Fig. 5.3 Preferred conformation of (N(3), N(7)-diprotonated gc^{6}Ade.

Table 5.1:Geometrical parameters for hydrogen bonding in the predicted preferred conformations for (N1, N3), (N3, N7), and (N1, N7) - diprotonated gc^{6}Ade.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (Å)	Distance Atom pair 2-3 (Å)	Angle 1-2-3 (deg.)	Figure references
O(10)…H-N(1)	2.09	1.01	123	5.2
O(13b)…H-N(6)	1.77	1.01	121	5.2
O(10)…H-N(7)	1.41	1.01	129.5	5.3
O(13b)…H-N(6)	1.98	1.01	107.6	5.3
O(10)…H-N(7)	1.64	1.01	126	5.4
O(13b)…H-N(6)	1.69	1.01	171	5.4
O(10)…H-N(1)	2.25	1.01	119.3	5.5
O(13b)…H-N(6)	1.69	1.01	171	5.5

The predicted most stable structure for (N(1), N(7))-diprotonated gc^{6}Ade is shown in (fig. 5.4). The set of torsion angles for (fig. 5.4) are α=180°, β=150°, γ=0°, δ=60°, ϵ=90°, and θ=180°. The hydrogen bonding between N(6)H…O(13b) and N(7)H…O(10) contribute to stability. Unlike N(7)H, N(1)H is not involved in hydrogen bonding.



. Fig. 5.4 Preferred conformation of (N(1), N(7))-diprotonated gc^{6}Ade.

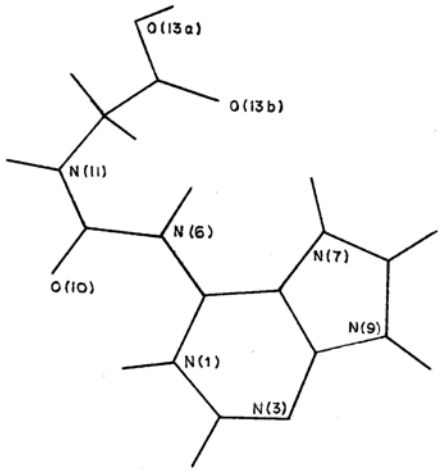


Fig. 5.5 Alternative distal conformation of (N(1), N(7)) - diprotonated gc^{6}Ade.

As in the case of the (N(1), N(3)) and (N(3), N(7)) diprotonation results, (N(1), N(7)) diprotonation also leads to participation of the N(6)H in hydrogen bonding with the O(13b). The hydrogen bonding parameters are shown in Table 5.1. The corresponding torsion angle values with full geometry optimization using PM3 method are broadly similar, and are shown in Table 5.2. Alternative distal (α=0°) conformation having 6.5 kcal/mol higher energy is shown in (fig. 5.5). Thus O(10)…HN(1) interaction, in this case, is much weaker and can scarcely replace the O(10)…HN(7) hydrogen bonding of (fig.5.4). Results of full geometry optimization using PM3 and MNDO support the PCILO results and are shown in Table5.2.

Table 5.2: Relative stability of the variously diprotonated gc^{6}Ade structures in Figures 5.2–5.5 by PCILO and full geometry optimization using PM3 and MNDO methods.

Specifics Probed	PCILO	PM3	MNDO
(a) Molecule: gc^{6}Ade(N(1), N(3)) diprotonated; Reference Figure 5.2			
	α=0°, β=180°, γ=0°, δ=300°, ϵ=240°,θ=180°	α=1°, β=177°, γ=14°, δ=300°,ϵ=229°,θ=183°	α=-5°,β=169°,γ=-8°, δ=292°,ϵ=226°,θ=181°
Relative energy	36.6 kcal/mol	0.0 kcal/mol	0.0 kcal/mol
(b) Molecule: gc^{6}Ade(N(3), N(7)) diprotonated; Reference Figure 5.3			
	α=180°,β=180°, γ=0°, δ=300°, ϵ=210°,θ=180°	α=180°, β=177°, γ=-5°, δ=294°, ϵ=224°, θ=182°	α=172°,β=171°,γ=-8°, δ=290°,ϵ=221°,θ=181°
Relative energy	44.7 kcal/mol	9.9 kcal/mol	15.8 kcal/mol
(c) Molecule: gc^{6}Ade(N(1), N(7)) diprotonated; Reference Figure 5.4			
	α=180°,β=150°,γ=0°, δ=60°, ϵ=90°, θ=180°	α=175°,β=185°,γ=346°, δ=60°, ϵ=129°, θ=176°	α=151°,β=193°, γ=2° δ=71°,ϵ=123°,θ=178°
Relative energy	0.0 kcal/mol	12.3 kcal/mol	15.5 kcal/mol
(d) Molecule: gc^{6}Ade(N(1), N(7)) diprotonated distal; Reference Figure 5.5			
	α=0°,β=150°,γ=0°, δ=60°, ϵ=90°, θ=180°	α=-8°,β=188°,γ=342°, δ=62°, ϵ=125°, θ=177°	α=330°,β=190°,γ=4°, δ=74°,ϵ=113°,θ=178°
Relative energy	6.5 kcal/mol	11.7 kcal/mol	14.0 kcal/mol

The PCILO conformational energy calculations show clear preferences Table 5.2, for the gc^{6}Ade (N(1), N(7)) diprotonation. The structure in (fig.5.4) is the most stable. The gc^{6}Ade (N(1), N(3)) - diprotonated structure in (fig.5.2) has over 36.6 kcal/mol higher energy than gc^{6}Ade (N(1), N(7)) diprotonation structure in (fig.5.4).

The gc^6Ade (N(3), N(7)) diprotonated structure in (fig.5.3) has still higher, 44.7 kcal/mol energy over the structure in (fig.5.4). The comparison of PCILO results with PM3 and MNDO geometry - optimized values in Table 5.2 shows that the gc^6Ade (N(1), N(3) diprotonation is preferred by PM3 and MNDO methods alike. The complete geometry optimization using MNDO and PM3 methods results in somewhat different values for the torsion angles than the PCILO preferred values, also the energy differences between the compared structures decrease from PCILO to MNDO to PM3.

Although direct experimental evidence for comparison with the present investigations is limited, nevertheless the crystal structures of anionic gc^6Ade salts having metal cations [14], dicationic gc^6Ade salts with anionic halides [7], or related nonionic ribonucleoside gc^6Ade [3] provide valuable data on interactions under such a wide range of conditions.

The geometrical parameters for the predicted intramolecular hydrogen bonds in (fig.5.3 and fig. 5.4) suggest these interactions to be fairly strong and viable. Instances of seven-membered intramolecular hydrogen bonds having the OCNCCOH motif have been observed in the crystal structures [15, 16] of antibacterial novobiocin and rifamycin. This indicates that seven-membered O(10)CN(6)CCN(7)H and O(13b)CCN(11)CN(6)H intramolecular hydrogen bonds as predicted in the present study may also be realized occasionally in condensed phases.

5.4 Conclusion:

The proximal conformation preferred for (N(3), N(7)) diprotonated as well as for (N(1), N(7) diprotonated gc^6Ade (figs.5.3 and 5.4) has N(6)H accessible for participation in Watson-Crick base pairing. This feature may allow extended codon - anticodon interactions involving the anticodon 3'-adjacent nucleosides and thus may result in the altered definition of the reading frame during protein biosynthesis, whenever N(7) is protonated. The common feature in (figures 5.2-5.5) is the N(6)H…O(13b) intramolecular hydrogen bonding, which is invariably predicted for diprotonated gc^6Ade.

5.5 References:

1. Yarus, M. Science. 1982, 218, 646-652.

2. Tsang, T. H.; Ames, B. N.; Buck, M. Biochim. Biophys. Acta. 1983, 741, 180- 196.

3. Parthasarathy, R.; Ohrt, J. M.; Chheda, G. B. Biochemistry 1977, 16, 4999–5008.
4. Tewari, R. Int J Quantum Chem 1994, 51, 105–112.
5. Tewari, R. Int J Quantum Chem 1997, 62, 551–556.
6. Tewari, R. Chem. Phys. Lett. 1995, 238, 365-370.
7. Parthasarathy, R.; Soriano-Garcia, M.; Chheda, G. B. Nature, 1976, 260, 807–808.
8. Diner, S.; Malrieu, J. P.; Jordan, F.; Gilbert, M. Theor Chim Acta 1969, 15, 100–110.
9. Malrieu, J. P. In Semiempirical Methods of Electronic Structure Calculations, Part A, Techniques, Segal, G. A., Ed.; Plenum: New York, 1977, p. 69–103.
10. Tewari, R. Int J Quantum Chem. 1987, 31, 611–624.
11. Stewart, J. J. P. J Comp Chem 1991, 12, 320–341.
12. Dewar, M. J. S.; Thiel, W. J Am Chem Soc 1977, 99, 4899– 4907.
13. Sybyl 6.5, Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144.
14. Parthasarathy, R.; Ohrt, J.M.; Chheda, G. B. Acta Cryst, 1976, B32, 2648–2653.
15. Boles, M. O.; Taylor, D. J. Acta Cryst B. 1975, 31, 1400–1406.
16. Brufani, M.; Cerrini, S.; Fedeli, W.; Vaciago, A. J Mol Biol. 1974, 87, 409–435.

PART-B: LYSIDINE (k^2C)

5.6 Introduction

The hypermodified nucleoside, 4-amino-2-(N6-lysino)-1-(β-D-ribofuranosyl) pyrimidinium ("lysidine"), naturally occurs in the first 'wobble' position of anticodon loop of isoleucine tRNA ($tRNA^{Ile}$) from *Mycoplasma capricolum* [1]. It is a modification of cytosine through replacement of O(2) and is designated as k^2C.

The lysidine modification prevents mistranslation of AUG as isoleucine and that of AUA as a methionine. The modification of C to k^2C also switches the aminoacylation identity of the tRNA from methionine to isoleucine [2]. Thus strict recognition of A and no misrecognition of G by lysidine plays very important role. The base pairing specificity from C.G to k^2C.A can be achieved by the modification of C to k^2C. The long lysine moiety may fold back towards the ribose ring, thus allowing the carbonyl group (or the amino group) of the side chain to form a hydrogen bond with the 2'-hydroxyl group [3].

It has also been shown that codon and amino-acid specificity of $tRNA^{Ile}$ are both converted by a single post-transcriptional modification, from cytidine to lysidine [2].

5.7 Nomenclature, Convention and Procedure:

The atom numbering and identification of the torsion angles (which specify the internal rotation around the various acyclic chemical bonds) are depicted in (fig.5.6). The oxygen atom in position 2 of cytidine is replaced by the ε-nitrogen atom of lysine amino acid. The torsion angle α [N(1)C(2)N(2)C(7)] denotes the rotation of C(7) around bond C(2)-N(2) and is measured in the right hand sense of rotation, with reference to the eclipsed orientation of C(2)N(1) and N(2)C(7) bonds. Likewise, the following chemical bonds sequences along the main extension of the substituent define the subsequent torsion angles β [C(2)N(2)C(7)C(8)], γ [N(2)C(7)C(8)C(9)], δ [C(7)C(8)C(9)C(10)], Ψ [C(8)C(9)C(10)C(11)], ϕ [C(9)C(10)C(11)C(12)], ξ [C(10)C(11)N(11)H], θ [C(10)C(11)C(12)O(12b)], η [C(11)C(12)O(12b)H].

The ribose - backbone torsion angles take values as specified in Holbrook $tRNA^{Phe}$ model [4]. The torsion angles in the ribose - phosphate backbone are distinguished by the subscript b to refer to the backbone. These backbone torsion angles values retain the nomenclature as in the tRNA model [4] referring likewise to the right hand sense of rotation around the central bond, measured from the eclipsed

position of the outer bonds γ_b[H-O5'-C5'-C4'], δ_b[O5'-C5'-C4'-C3'], ε_b[C5'-C4'-C3'-O3'] and ζ_b[C4'-C3'-O3'-H]. The glycosyl torsion angle is held anti χ=3° and the ribose ring puckering is C3'-endo, similar to the wobble nucleoside at 34th position in tRNA model.

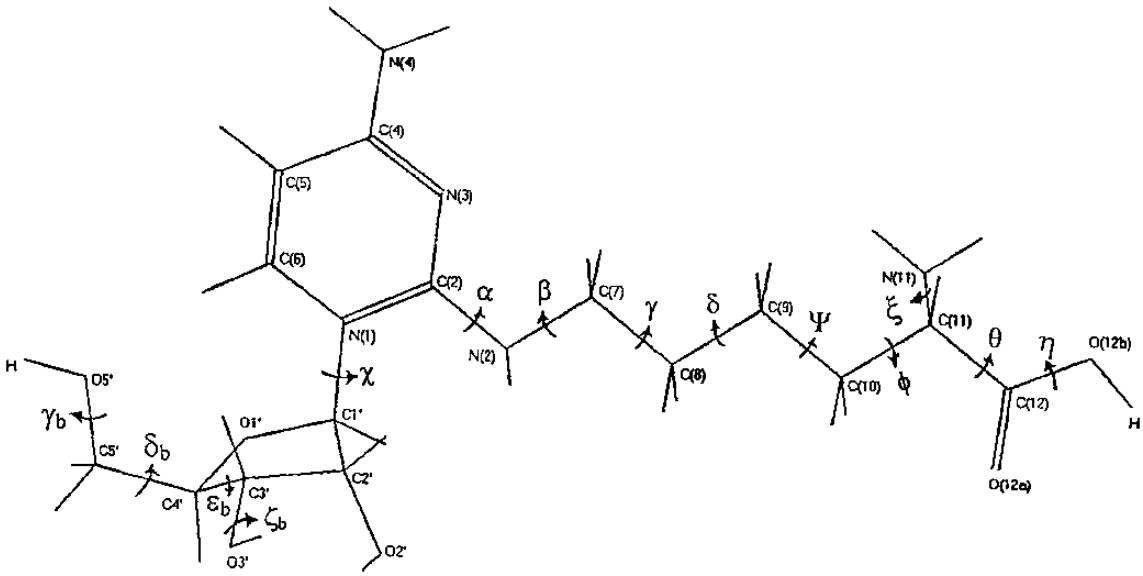


Fig.5.6 Atom numbering and various torsion angles in lysidine (nonzwiteroinic form).

The quantum chemical perturbative configuration interaction with localized orbitals PCILO method [5-7] has been utilized for the energy calculations of the various molecular conformations. The appropriately selected bond lengths and bond angle values for various forms of lysidine molecule have been used in the conformational energy calculations. In the case of hydrogen atoms, the standard values from the reference source data [8], suited for the particular type of bonding environment have been used. Full geometry optimization calculations using molecular mechanics force field (MMFF) method [9] have been utilized to compare the salient features.

The full geometry optimization has also been carried out to compare the salient features using *ab initio* (molecular orbital Hartree-Fock SCF) quantum mechanical energy calculations using 3-21G* basis set (PC Spartan Pro version 6.0.6). Appropriate values of the torsion angles in the ribose - phosphate backbone corresponding to the nucleoside at 34th position in Holbrook model of tRNA^{Phe} [4] has been adopted.

5.8 Results and Discussion

5.8.1 Lysidine in Zwitterion form:

The PCILO preferred most stable base substituent orientation for zwitterionic form of hypermodified nucleoside lysidine is shown in fig.5.7. The positive charge is present on the N(1) of cytidine and N(11) of lysine moiety. The torsion angles describing the base substituent orientation are (α=180°, β=180°, γ=30°, δ=60°, Ψ=180°, ϕ=60°, ξ=330°, and θ=150°).

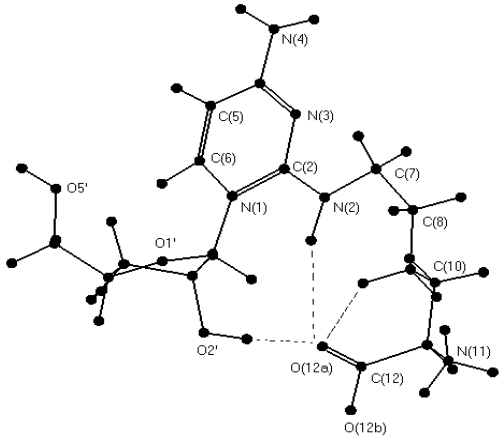


Fig. 5.7: PCILO most stable structure for Zwiterionic form of lysidine

The preferred orientation of lysine moiety is such that the amino acid carboxyl end folds back towards the ribose ring. The carboxyl oxygen O(12a) of lysine moiety forms hydrogen bond with the 2'-hydroxyl group of ribose sugar ring. The carboxyl oxygen O(12a) also interacts with N(2) and C(9) of lysine moiety. So the preferred structure is stabilized by intramolecular interactions between O(12a)…HO2', O(12a)…HN(2), N(2)…HC(9), O(12a)…HC(9), O(12b)…HN(11), O5'…HC(6) and O(12a)…HC1' as indicated in Table 5.3. The interaction between N(2)…HC(10) is the additional stabilizing factor. The preferred glycosyl torsion angle for lysidine also turns out to be χ=3° as in the Holbrook model. The orientation of lysine moiety is 'trans' to the N(1) of cytidine. The modified nucleoside lysidine does not allow Watson-Crick base pairing with G, because O(2) of cytidine is replaced by the bulky lysine group.

Table 5.3: Geometrical parameters for hydrogen bonding in the preferred conformations of various forms of lysidine.

Atoms Involved (1-2-3)	Distance atom pair 1-2 (A°)	Distance atom pair 2-3 (A°)	Angle 1-2-3 (deg.)	Figure Reference
O(12a)…H-O2’	1.67	0.96	173.37	5.7, 5.9, 5.10
O(12a)…H-N(2)	2.59	1.09	157.84	5.7, 5.9, 5.10, 5.11
N(2)…H-C(9)	2.43	1.09	94.40	5.7, 5.9, 5.10, 5.11
N(2)…H-C(10)	2.66	1.09	110.44	5.7, 5.9, 5.10
O(12a)…H-C(9)	2.02	1.09	131.77	5.7, 5.9, 5.10
O(12b)…H-N(11)	1.83	1.01	118.48	5.7
O5’…H-C(6)	2.64	1.09	118.23	5.7, 5.9, 5.10, 5.11
O(12a)…H-C1’	2.86	1.09	122.52	5.7, 5.9, 5.10
O(12b)…H-N(11)	1.95	1.01	109.31	5.9, 5.10
O(12a)…H-O2’	1.58	0.96	149.22	5.11
O(12a)…H-C(9)	1.89	1.09	147.64	5.11
O(12b)…H-N(11)	1.73	1.09	121.80	5.11
O(12a)…H-N(3)	1.38	1.09	163.59	5.12
O(12a)…H-C(7)	1.63	1.09	143.00	5.12
O(12b)…H-N(3)	2.03	1.09	109.81	5.12
O(12b)…H-C(8)	2.29	1.09	140.84	5.12
O(12b)…H-N(11)	1.99	1.01	112.9	5.12
O1’…H-N(2)	1.76	1.09	137.50	5.12
O5’…H-N(2)	2.39	1.09	98.30	5.12
O5’…H-C3’	2.13	1.09	110.47	5.12

Higher energy (3.0 kcal/mol) alternative conformation Table 5.4 is arrived by flipping of θ=330°. This structure is stabilized by intramolecular interactions between O(12b)…HO2’, O(12b)…HN(2), O(12b)…HC(9), O(12a)…HN(11). Interaction between O5’…HC(6) may also contribute to stability. In this alternative conformation one of the interesting feature is that O(12b) interacts with O2’H, N(2) and C(9), instead of O(12a) in the most stable conformation (fig.5.7). Results of geometry optimization through molecular mechanics force field (MMFF) method are

α=165°, β=171°, γ=58°, δ=54°, ψ=-140°, ϕ=72°, and χ=-3°. All these optimized torsion angle values have differences within the range of (±20°), except for ψ which differs by 40° from the PCILO preferred value. The full geometry optimization carried out by HF-SCF method results in torsion angles α=175°, β=101°, γ=59°, δ=67°, ψ=-142°, ϕ=75°, and χ=20°. The torsion angles α and δ have minor differences from the preferred PCILO values. The torsion angles β, γ, ψ, ϕ, ξ, θ and χ show larger variations from values for the PCILO most stable structure.

5.8.1.1 The specific recognition of A by Lysidine (k^2C):

The $tRNA^{Ile}_{minor}$ specifically recognizes the codon AUA, two models of k^2C.A pairs are possible. In one model fig. 5.8(a), the nucleoside k^2C is in a zwiterionic form with -NHR group in position 2. This model may not be stable, because of the steric hindrance between the bulky R-Group of the lysine moiety and the ribose ring. In this model the lysine moiety (R) is located '*cis*' to the N(1) atom of cytidine. However, in the preferred PCILO most stable structure (fig.5.7), the lysine moiety (R) is located '*trans*' to the N(1) of cytidine, as in model fig.5.8(b). As the lysine moiety (R) prefers the '*trans*' orientation, the hydrogen of –NHR group in position 2 may not be available for Watson- Crick pairing with 'A'

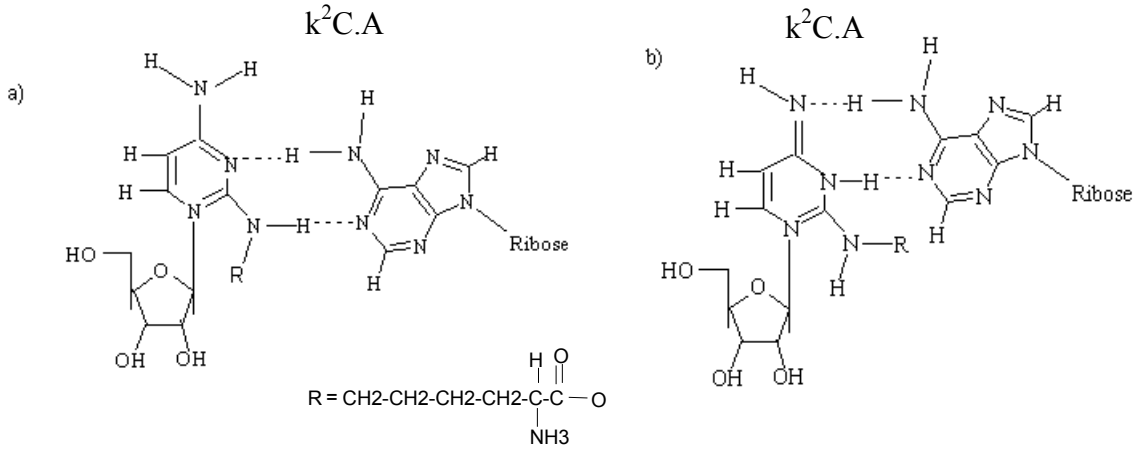


Fig. 5.8: Model diagram a) k^2C.A pairing showing '*cis*' conformation of R, where as b) shows 'trans' orientation of R group.

Table 5.4: Alternative stable conformations for various forms of lysidine:

Trosion Angles	Rel. Eng. (kcal/mol)	Fig. Ref.
α=180°,β=180°,γ=30°,δ=60°,Ψ=180°,ϕ=60°,ξ=330°,θ=150°,χ=3°	0.0	5.7
α=180°,β=180°,γ=30°,δ=60°,Ψ=180°,ϕ=60°,ξ=330°,θ=330°,χ=3°	3.0	
α=180°,β=180°,γ=30°,δ=60°,Ψ=180°,ϕ=60°,ξ=30°,θ=150°,η=180°,χ=3°	0.0	5.10
α=180°,β=180°,γ=180°,δ=60°,Ψ=60°,ϕ=60°,ξ=30°,θ=60°,η=180°,χ=3°	3.7	
α=180°,β=180°,γ=180°,δ=60°,Ψ=180°,ϕ=300°,ξ=30°,θ=60°,η=180°,χ=3°	4.4	
α=180°, β=180°, γ=30°, δ=180°, Ψ=150°, ϕ=0°, ξ=0°, θ=120°,χ=3°	0.0	5.11
α=180°, β=90°, γ=180°, δ=180°, Ψ=300°, ϕ=0°, ξ=0°, θ=120°,χ=3°	2.9	
α=180°, β=180°, γ=30°, δ=180°, Ψ=150°, ϕ=0°, ξ=0°, θ=300°,χ=3°	4.6	

5.8.2 Nonzwitterionic Lysidine :

Fig. 5.9 depicts the PCILO predicted most stable conformation for hypermodified nucleoside nonzwitterionic lysidine. The molecule has positive charge on N(1) of cytidine. The set of torsion angles for the preferred orientation of the base substituent are (α=180°, β=180°, γ=30°, δ=60°, Ψ=180°, ϕ=60°, ξ=30°, θ=150°, η=150°). The structure is stabilized by intramolecular interactions between O(12a)…HO2', O(12a)…HN(2), O(12a)…HC(9), N(2)…HC(9), O(12a)…HC1', O(12b)…HN(11), O5'…HC(6) and N(2)…HC(10), which are included in Table 5.3. The lysine moiety folds back towards the ribose ring. The carboxyl oxygen O(12a) of lysine moiety forms hydrogen bond with O2'H of ribose sugar ring. The carboxyl oxygen O(12a) also forms hydrogen bond with N(2) and C(9) of lysine moiety. The lysine moiety has orientation similar to shown in fig.5.7. Starting from the PCILO preferred conformation (fig.5.9), results of automated geometry optimization through molecular mechanics (MMFF) method are α=170°, β=170°, γ=48°, δ=54°, ψ=-168°, ϕ=67°, η=--179°, and χ=5°. The MMFF optimized values are closely related (±15°) to the PCILO preferred torsion angle values. The results of full geometry optimization by HF-SCF method result in torsion angle values α=173°, β=94°, γ=56°, δ=63°, ψ=-164°, ϕ=70°, η=-175°, and χ=20°. The torsion angle β shows larger difference but γ, ψ, ϕ, ξ, θ and χ differ by about 20° from the PCILO values. The other torsion angles differ much less.

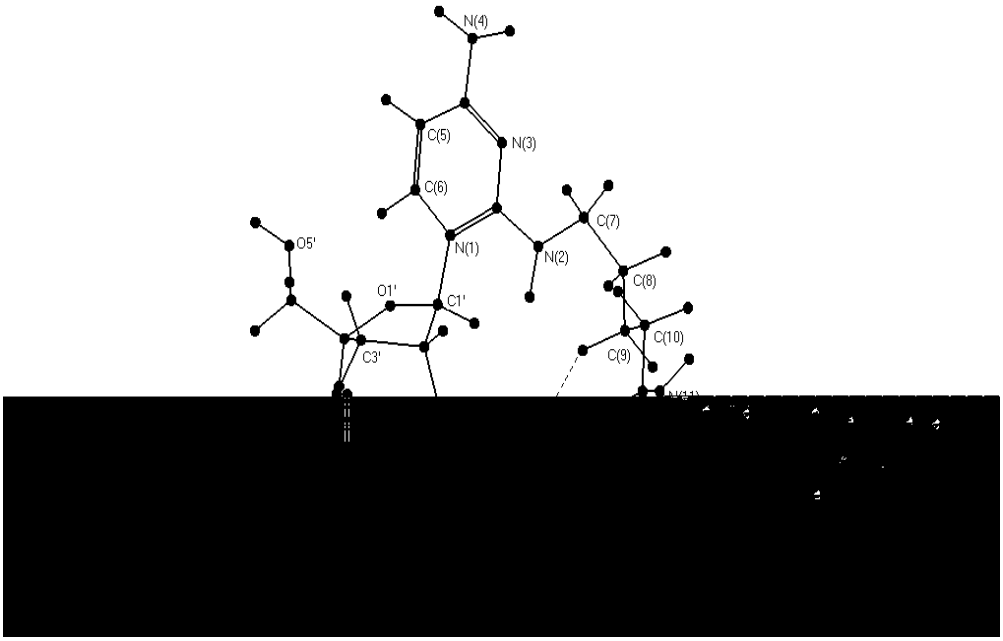


Fig. 5.9: PCILO most stable structure for nonzwitterionic Lysidine

When the glycosyl torsion angle is allowed to change freely, still the glycosyl torsion angle choice remains unchanged. Thus, the orientation of the base substituent (nonzwitterionic lysidine) is unchanged. The orientation of the lysine moiety is '*trans*' to the N(1) of cytidine. Positively charged nonzwitterionic lysidine molecule does not form Watson-Crick base pairing with 'G'. In contrast to model in fig. 5.8(a), this form of lysidine is not able to form WC base pairing with 'A', due to the '*trans*' orientation of lysine moiety (R).

5.8.3 Lysidine in Neutral form:

The PCILO predicted most stable structure, of neutral lysidine is depicted in fig.5.10. The preferred torsion angle values describing the base substituent orientation in lysidine are (α=180°, β=180°, γ=30°, δ=60°, Ψ=180°, ϕ=60°, ξ=30°, θ=150°, η=180°). The structure is stabilized by intramolecular interaction between O(12a)…HO2', O(12a)…HN(2), N(2)…HC(9), N(2)…HC(10), O(12a)…HC(9), O5'…HC(6), O(12a)…HC1' and O(12b)…HN(11), which are included in Table 5.3. The lysine moiety folds back towards the ribose ring. The carboxyl oxygen O(12a) of lysine moiety forms hydrogen bond with O2'H of ribose sugar ring. The carboxyl oxygen O(12a) is also involved in hydrogen bonding with N(2) and C(9) of lysine moiety. The orientation of lysine moiety is similar to the preferred structures in figures 5.7 and 5.9.

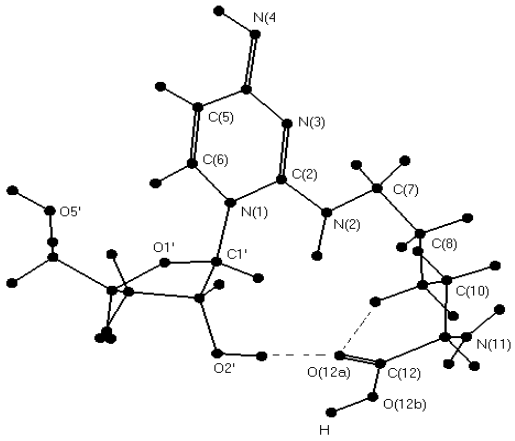


Fig. 5.10: PCILO most stable structure for neutral lysidine

Results of molecular mechanics force field (MMFF) full geometry optimization of the PCILO preferred structure in fig.5.10 result in α=172°, β=163°, γ=51°, δ=54°, ψ=-173°, ϕ=71°, η=-179°, and χ=14°. The optimized torsion angle values are nearly equal (±10°) to the preferred structure fig. 5.10. The geometry optimized torsion angles obtained by HF-SCF method are α=173°, β=83°, γ=53°, δ=64°, ψ=-167°, ϕ=72°, η=-176°, and χ=24°. The torsion angle values for α, δ, ψ, ϕ, ξ, η and χ are within 20° of the PCILO most stable values, torsion angle β shows larger difference where as γ, θ and χ differ by 30°. Higher energy (3.7 kcal/mol) PCILO alternative stable conformation, in Table 5.4, may be realized by flipping to γ=180°, Ψ=60°, θ=60°. Interaction between ribose - phosphate backbone and the substituent O5'…HC(11),may contribute to stability. However, interaction of carboxyl ogygen O(12a) with O2'H is absent in this alternative structure. Another higher energy (4.4 kcal/mol) alternative conformation may be arrived by flipping of the indicated torsion angles to γ=180°, ϕ=300°, and θ=60°. Intramolecular interactions between O(12b)…HC(9) may provide stability to the lysidine amino acid terminal. Interaction between base substituent and ribose-backbone O5'…HC(11) is also a likely factor contributing to stability.

When glycosyl torsion angle is allowed to change freely, the glycosyl torsion angle retains anti (χ=3°) conformation, as specified for wobble nucleoside in $tRNA^{Phe}$ model. Thus the preferred orientation of hypermodified nucleoside lysidine (neutral form) remains unaffected. The orientation of lysine substituent remains *trans* to the

N(1) atom of cytidine. Unlike the model fig. 5.8(b), in neutral lysidine position 3rd of the pyrimidine ring lacks the hydrogen atom. Thus neutral form of lysidine molecule can not form Watson-Crick base pairing with A.

5.8.4 Lysidine (Tautomer form):

The PCILO predicted most stable structure of lysidine (Tautomer form) with zwitterionic form and positive charge on N(1) and N(11) atoms of lysidine is shown in fig. 5.11. The lysidine is in the tautomer form having –NHR group in position 2, -NH group in position 3, and =NH group in position 4. The set of preferred torsion angle values for the favored substituent orientation are (α=180°, β=180°, γ=30°, δ=180°, Ψ=150°, ϕ=0°, ξ=0°, θ=120°). The structure is stabilized by intramolecular interactions between O(12a)…HO2', O(12a)…HN(2), O(12a)…HC1', N(2)…HC(9), O(12a)…HC(9), O(12b)…HN(11) and O5'…HC(6), which are included in Table 5.3. The lysine moiety folds back towards the ribose ring. The carboxyl oxygen O(12a) of lysine moiety interacts with O2'H of ribose ring. The carboxyl oxygen O(12a) may also have hydrogen bonding with N(2) and C(9).

The preferred orientation of lysine substituent may be related to that of figs. 5.7, 5.9 and 5.10. As compared to lysidine (zwitterion form) in fig.5.7 (α=180°, β=180°, γ=30°, δ=60°, Ψ=180°, ϕ=60°, ξ=330°, θ=150°) with that of lysidine in tautomer form fig.5.11, the torsion angles instead take the indicated value (α=180°, β=180°, γ=30°, δ=180°, Ψ=150°, ϕ=0°, ξ=0°, θ=120°). Among these torsion angles δ, ψ, ϕ, ξ, and θ have different values, where as other torsion angles like α, β, and γ have similar values. Notably the torsion angles (δ=180°, θ=120°) have different values in fig. 5.11, in contrast to (δ=60°, θ=150°) in fig.5.7. Consequently stronger interaction occurs between O(12a)…HC(9) and O(12a)…HO2' but interaction between N(2)…HC(10) is lacking in fig.5.11, also see Table 5.3. It may also be noted that different values of (ϕ=0°, ξ=0°, θ=120° in fig.5.11) are found as compared to (ϕ=60°, ξ=330°, and θ=150°) in fig. 5.7. These changes result in stronger interaction between O(12b)…HN(11) in (fig.5.11) as evident from Table 5.3.

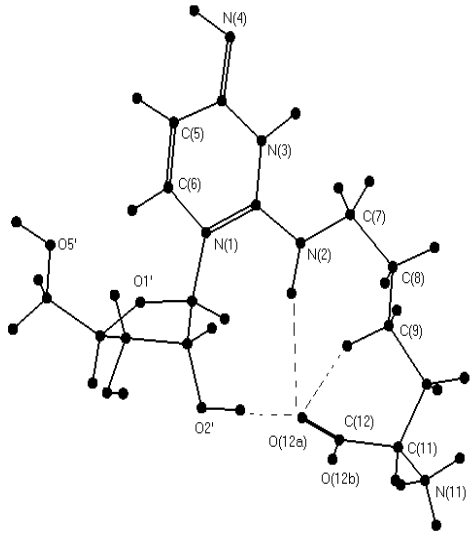


Fig. 5.11: PCILO most stable structure for lysidine (Tautomeric form) at χ=3°.

Starting from PCILO most stable structure (fig.5.11), results of full geometry optimization by using molecular mechanics force field (MMFF) method are α=-175°, β=177, γ=82°, δ=-142°, ψ=103°, ϕ=-35°, ξ=-27°, θ=130°, and χ=9°. The MMFF optimized values differ by more than 50° for γ, by more than 40° for δ, and by more than 77° for Ψ, but the other torsion angles have smaller differences. The geometry optimization by using HF-SCF results are α=-178°, β=-177°, γ=76°, δ=-150°, ψ=106°, ϕ=-35°, ξ=2°, θ=116°, and χ=16°. The torsion angle γ differs by more than 46°, δ differs by 30°, Ψ differs by 44° and ϕ differs by more than 35°. The other torsion angles have minor differences.

Higher energy (2.9 kcal/mol) PCILO alternative conformation included in Table 5.4 is reached with torsion angles β=90°, γ=180°. Interactions between O(12a)…HC(9), O(12b)…HN(11) and O5'…HC(6) may provide stabilization. The intramolecular hydrogen bonding between O(12a)…HC(8), O(12a)…HN(3), N(3)…HC(7) and O(12b)…HC(8) may provide additional stability. Hydrogen bonding between O(12a)…HO2' and O(12a)…HN(2) is not present. The next higher energy conformation 4.6 kcal/mol may be realized by flipping of (θ=300°). The structure is stabilized by O(12b)…HO2', O(12b)…HC(9), O(12b)…HN(2) and O(12a)…HN(11) interactions. Atom O(12b) is involved, instead of O(12a) as in the preferred conformation (fig. 5.11). Interaction between the base and ribose - phosphate backbone O5'…HC(6) is another factor for stability.

In the most stable conformation of lysidine (Tautomer form) (fig. 5.11), the orientation of lysine moiety is '*trans*' to the N(1) atom of cytidine. So the most stable conformation is similar to the model diagram fig.5.8(b). Lysidine tautomer form (fig.5.11) has an –NHR group as the substituent in position 2, >NH group in position 3, and exocyclic =NH group in position 4. In this structure NHR in position 2, >NH in position 3, and =NH in position 4 may not pair with sites in G, but pairing with 'A' may be possible. Thus tautomer form of lysidine can recognize codon AUA instead of AUG.

When glycosyl torsion angle for nucleoside is varied freely, the resulting preferred value is (χ=183°). Adoption of this preferred value results in new set of torsion angle values for the preferred orientation of the lysidine substituent (fig.5.12), (α=180°, β=90°, γ=180°, δ=120°, Ψ=300°, ϕ=30°, ξ=150°, θ=90°).

The structure is stabilized by intramolecular interactions between O(12a)…HN(3), O(12a)…HC(7), O(12b)…HN(3), O(12b)…HC(8), O(12b)…HN(11), O1'…HN(2) and O5'…HN(2). Geometrical parameters for these hydrogen bonding interactions are included in Table5.3.

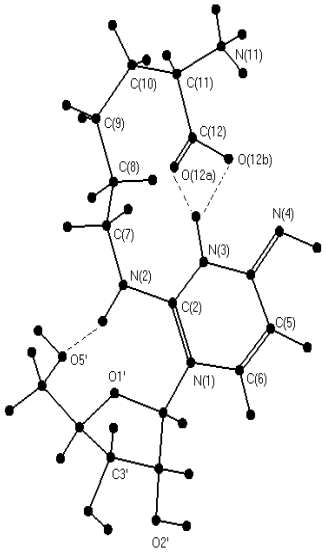


Fig.5.12: PCILO predicted structure for tautomeric form of lysidine with χ=183°.

The other oxygen O(12b) of carboxyl group of lysine moiety also forms reasonable hydrogen bonding with N(3)H of cytidine. When glycosyl torsion angle takes preferred value (χ=183°), the hydrogen bonding between O(12a)…HO2' and O(12a)…HN(2) (fig.5.11) is no longer possible, instead of this the structure is stabilized by O5'…HN(2) (fig.5.12). In this structure (fig.5.12) the glycosyl torsion

angle (χ=183°) has the 'syn' conformation. With the preferred value for glycosyl torsion angle (χ=183°) the orientation of lysine moiety turns towards the cytidine base (opposite to the direction of 2'-OH group of ribose ring). The carboxyl oxygen's O(12a) and O(12b) form hydrogen bonds with N(3)H, see Table 5.3. Thus the Watson-Crick base pairing site N(3)H in position 3, and =NH in position 4 are completely blocked. Therefore with 'syn' glycosyl orientation, the hypermodified nucleoside lysidine in tautomeric form may not participate in codon-anticodon interactions.

The automated geometry optimized torsion angles by using MMFF method are α=173°, β=109°, γ=-155°, δ=141°, ψ=-86°, ϕ=50°, ξ=151°, θ=105°, and χ=-164°. The full geometry optimization through computationally expensive Hatree-fock Self Consistence Field (HF-SCF) method have also been carried out. The torsion angles are α=170°, β=97°, γ=-171°, δ=139°, ψ=-89°, ϕ=61°, ξ=135°, θ=106°, and χ=-136°. These agree with the PCILO results except for some minor differences.

5.9 Conclusion:

The conformational preferences of hypermodified nucleoside lysidine with its various tauromeric and ionic forms have been studied using PCILO, MMFF and HF-SCF methods. The orientation of bulky lysine moiety (R) is located '*trans*' to the N(1) atom of cytidine in these various forms of lysidine. The carboxyl oxygen O(12a) of lysine moiety forms hydrogen bonding with O2'H of ribose sugar ring. The lysidine with or without zwitterionic form (with positive charge on N(1) of cytidine) and neutral form of lysidine may be unable to recognize 'A' from codons. Because to recognize A there should be a hydrogen bond donor group at position 3, and one hydrogen bond acceptor group at position 4. These hydrogen bond donor-acceptor situations are not fulfilled by the above discussed three forms of lysidine i.e. (fig. 5.7, 5.9, and 5.10). These various forms of lysidine avoid the misrecognition of 'G', but only tautomer form of lysidine fig.5.11, can form Watson-Crick base pairing with 'A'. Because tautomeric form of lysidine has one hydrogen bond donor group at position 3 and acceptor group at position 4. So according to model fig. 5.8(b) tautomeric form of lysidine can form Watson-Crick base pairing with 'A' and may recognize the codon AUA instead of AUG. The glycosyl torsion angle is held anti (χ=3°) in these various lysidine forms. The 'syn' conformation with glycosyl torsion angle χ=183° for tautomeric form (fig. 5.12) blocks one of the Watson-Crick base pairing site N(3)H

due to O(12a)…HN(3) and O(12b)…HN(3) interactions. The interactions between O5'…HN(2) may provide additional stability to the structure (fig. 5.12).

5.10 Reference:

1. Andachi, T.; Yamao, F.; Muto, A and Osawa, S. J. Mol. Biol. 1989, 209, 37-54.
2. Muramatsu, T.; Nishikawa, K; Nemoto, F; Kuchino, Y.; Nishimura, S.; Miyazawa, T and Yokoyama, S. Nature (London) 1988, 336, 179-181.
3. Yokoyama, S. and Nishimura, S.; Chapter 12, tRNA: Structure Biosynthesis and function, Edited by D.Soll and U. L. Rajbhandary, page 216-223.
4. Holbrook, S. R.; Sussman, J. L.; Warrant, R.W. and Kim, S.H. J. Mol. Biol. (1978), 123, 631-660.
5. Diner, S.; Malrieu, J. P and Claverie, P. Theoret.Chim.Acta, 1969, 13, 1-17.
6. Diner, S.; Malrieu, J. P.; Jordan, F.; Gilbert, M. Theoret.Chim.Acta 1969, 15, 100-110.
7. Malrieu, J. P. in Semiempirical Methods of Electronic Structure Calculations, Part A, Techniques, Segal, G. A.; Ed.; Plenum, New York, 1977, p. 69-103.
8. Kennard, O in CRC Handbook of Chemistry and Physics, 61st ed. Weast. R.C Astle, M. J. Eds. (CRC Press, Boca Raton, 1980-81), pp.f208-f211.
9. Halgren, T. A. J. Comp.Chem. 1996, 17, 490-519.

CHAPTER-VI

ANTICODON LOOP STRUCTURE WITH MODIFIED COMPONENTS

6.1 Introduction

Molecular dynamics (MD) simulation of anticodon stem-loop is discussed in this chapter. In $tRNA^{Ser}$(UGA) from Nicotiana rustica, modified nucleosides, 5-carbamoylmethyl uridine (ncm^5U) occurs at 34^{th} 'wobble' position and hydroxyisopentenyl adenosine (io^6A) occurs at 37^{th} position [1, 2]. MD studies on anticodon stem loop of $tRNA^{Ser}$(UGA) are reported here. MD studies on E.Coli $tRNA^{Ser}$(GGA) which also contains io^6A at 37^{th} position, are also included. The hypermodified nucleoside hydroxyisopentenyl adenosine (io^6A) occurs at anticodon 3'-adjacent position in those tRNAs which recognize the codons starting with U. Anticodon 'wobble' nucleoside as well as the anticodon 3'-adjacent nucleoside are most often extensively modified [3-5]. While wobble base modifications may extend or restrict the selection of synonymous codons to be read by the tRNA, the anticodon 3'-adjacent modifications may help define the reading frame for codon-anticodon interactions [6-8]. These modifications may also optimize the codon-anticodon interactions to enable nearly equal residence times for tRNA anticodon-mRNA codon interactions irrespective of diverse base sequences, to allow smooth inphase protein biosynthesis [9, 10].

In order to understand the structural and functional significance of hypermodified nucleosides, *trans*-io^6A or *trans*-ribosylzeatin at 37^{th} position and ncm^5U at 34^{th} position, a molecular dynamics (MD) simulation study has been undertaken.

6.2 Computational procedure

The starting co-ordinates were extracted from $tRNA^{Phe}$ crystal structure (Holbrook, et. al., 1978) [11] (PDB code 5TRA). The ribose-phosphate backbone was modelled on $tRNA^{Phe}$ [11] and the desired base sequences of $tRNA^{Ser}$(UGA) from Nicotiana rustica (fig. 6.1) and E.Coli $tRNA^{Ser}$(GGA) (fig. 6.2) were utilized to build models of the respective anticodon stem loops. For this the commercially available, Sybyl-Biopolymer modeling software package from Tripos and Associates, Inc. was used. The anticodon stem loop part consisted of 17 nucleotides of $tRNA^{Ser}$(UGA) from Nicotiana rustica (fig.6.1) and E.Coli $tRNA^{Ser}$(GGA) (fig. 6.2). To truncate the anticodon stem-loop model methyl phosphate groups (Me-p-) were placed at both 5' and 3' terminals. The molecules are neutralized by adding a hydrogen atom at O1p of $pO4^-$ group of each nucleotide. The system is partially solvated by 600 TIP3P model

water molecules. Cubic periodic box of length 39.79 A° has been considered. The Kollman-All-Atom force field [12] with Gasteiger- Marsilli charges, 8 A° non-bonded cutoff, and 'constant' dielectric function with dielectric constant value set at unity (=1) was used. The other options used for MD simulations include canonical ensemble (NTV), initial velocity considered as a Boltzman distribution, 10fs of non-bonded update with scaled velocities, 5ps of temperature coupling time, 1fs time step, data recorded at 1000fs intervals for MD calculations. Shake [13] algorithm was used to constrain all the X-H bonds of anticodon stem part and a constraint condition applied to the last base pair of anticodon stem in order to prevent disturbance of the stem part (figures 6.1 and 6.2). Density of partially solvated system is 0.43.

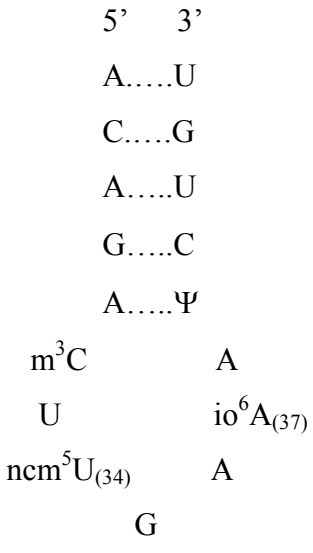


Fig. 6.1 $tRNA^{Ser}$(UGA)

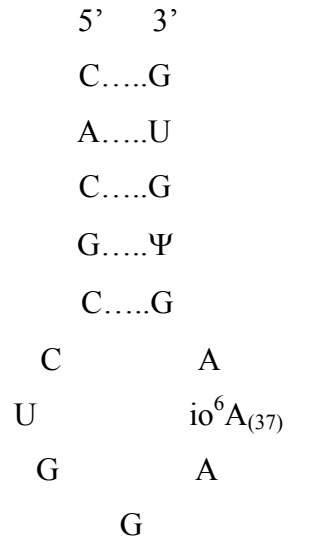


Fig. 6.2 E. Coli $tRNA^{Ser}$(GGA)

The equilibration protocol consisted of 2000 steps of steepest descent minimization applied to the whole system in order to relax possible steric clashes between solute and solvent. This was followed by 5ps of molecular dunamics at 300K, where only the water molecules were allowed to reorient around the solute atoms. After the water relaxation, again 1000 steps of steepest descent minimization was applied to the system.

In these two MD runs, the temperature was progressively increased to 300K in steps of 50k, 100k, 150k, 200k, 225k, 250k, 275k, 300k with 1ps of MD at each time step. After this a MD for longer duration of 208 ps for $tRNA^{Ser}$(UGA) (fig.6.1) and

158ps simulation for $tRNA^{Ser}$(GGA) (fig. 6.2) were made at 300k temperature and 1 atmosphere pressure. Up to 8ps data collection (snap shot) done at 250fs and from 9ps onwards data collected at 1000fs for $tRNA^{Ser}$(UGA) and $tRNA^{Ser}$(GGA) molecule. All the calculations were performed on silicon Graphics power o2 workstation having (R5000 MIPS) processor and 512 MB RAM.

6.3 Results and Discussion

6.3.1 $tRNA^{Ser}$(UGA) from Nicotiana rustica:

6.3.1.1 Anticodon stem/loop analysis:

The (fig. 6.3) shows anticodon stem/loop of $tRNA^{Ser}$ (UGA) at 208ps. The base pairs in the tRNA stems are well maintained during the 208ps of MD trajectory. The anticodon loop bases are found to be well stacked. The bases 32^{nd}, 33^{rd}, and 34^{th}, 35^{th}, 36^{th}, 37^{th}, 38^{th} are stacked.

The anticodon loop is stabilized by interactions between $N(7)_{(36)}\ldots HN(3)_{(33)}$, $O(2)_{(33)}\ldots HO2'_{(33)}$, and $N(7)_{(35)}\ldots HO2'_{(33)}$, shown in (fig. 6.4). The $O2'H_{(33)}$ shows dual hydrogen bonding pattern by interacting with $N(7)_{(35)}$ and $O(2)_{(33)}$, which may help in maintaining the anticodon loop structure. A sharp curve (fig. 6.5) is observed between 33^{rd}, 34^{th}, and 35^{th} bases, that is like a 'U' turn in anticodon loop. The interaction of $N(3)H_{(33)}$ with that of $N(7)_{(36)}$ may enable the 'U' turn in anticodon loop.

The C3'-endo ribose ring puckering is maintained for 27^{th}, 28^{th}, 29^{th}, 30^{th}, 31^{st}, 32^{nd}, 34^{th}, 35^{th}, 36^{th}, 37^{th}, 38^{th}, 39^{th}, 40^{th}, 41^{st}, and 42^{nd} nucleotides. In case of 33^{rd} and 43^{rd} nucleotides the ribose ring fluctuates between c3'-endo and c2'-endo puckerings.

The orientation of glycosyl torsion angle is 'anti' for all 17 bases. Thus the base pairing is maintained at stem region, also base stacking in anticodon loop region is well maintained.

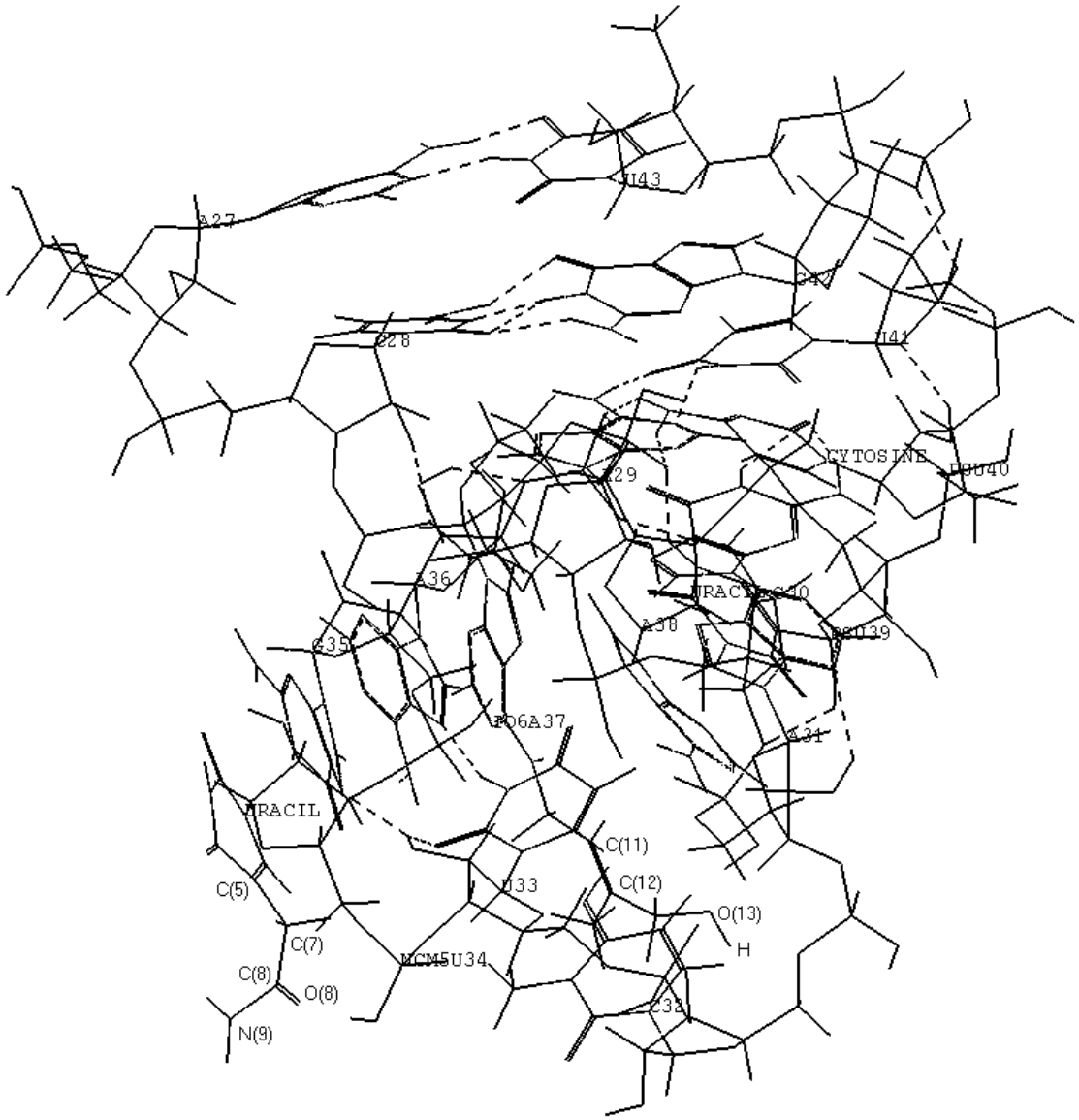


Fig. 6.3: Anticodon stem loop of $tRNA^{Ser}$(UGA) at 208ps.

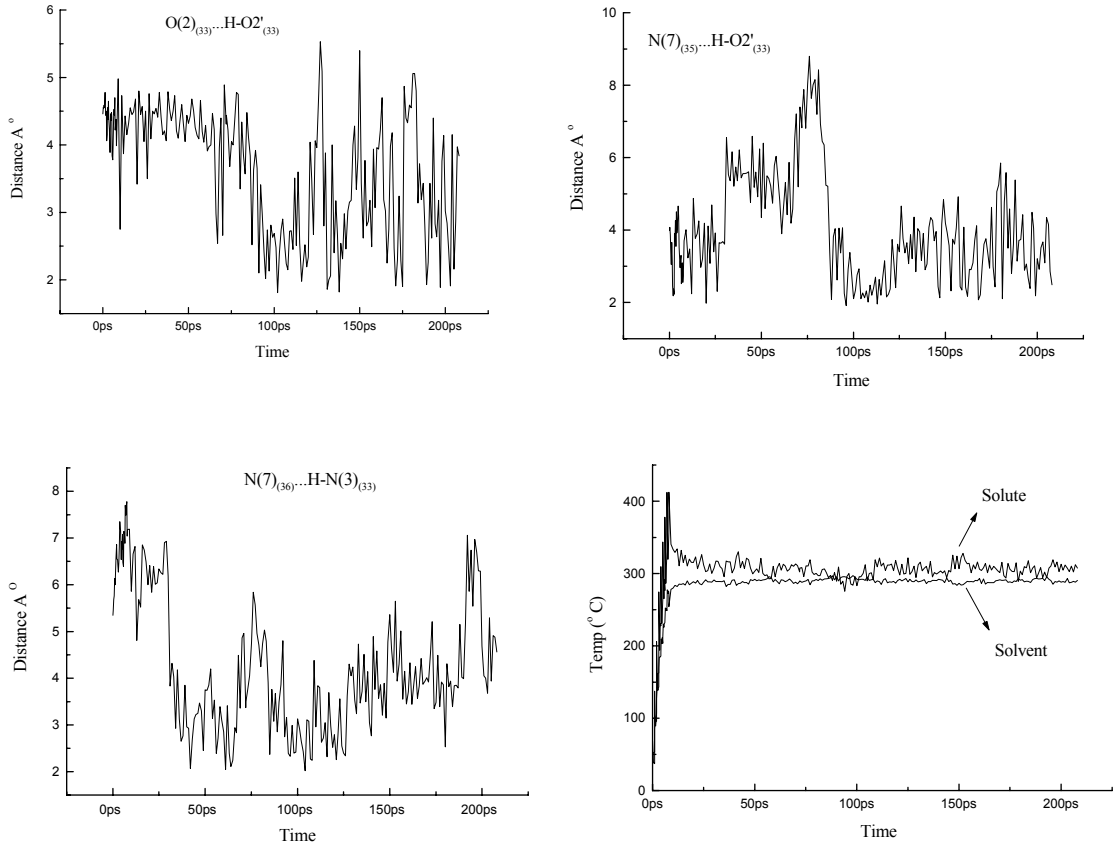


Fig. 6.4: Intramolecular interactions stabilizing the anticodon loop structure and solute-solvent temperature equilibrium graph for $tRNA^{Ser}$(UGA) during MD.

6.3.1.2 The conformation of hydroxyisopentenyl substituent:

The orientation of hydroxyisopentenyl substituent in anticodon loop is discussed here. The torsion angle α maintains nearly 0° value, 'distal' conformation, throughout the 208ps MD simulation run (α ≅ 0°). The range of β is in between –50° to -100° for initial 40ps, from 50ps onwards it prefers 180° value for maximum time. The torsion angle γ mostly remains steady at ±120° value, some times it takes 180° value also. From 180ps onwards γ prefers 90° to 120° values. Through out 208ps MD simulation the torsion angle δ retains trans (δ=180°) conformation. The torsion angle Ψ1 prefers nearly ±60° value. It also takes in between ±120° values. The torsion angle Ψ2 takes all alternative values like 180°, ±60°. The graphical representation of torsion angles α, β, γ, δ, Ψ1 and Ψ2 are shown in (fig. 6.6). The torsion angle θ shows fluctuation in between 180°, ±60° as shown in (fig. 6.7). No intramolecular

interactions involving hydroxyl oxygen O(13) are observed in anticodon loop of $tRNA^{Ser}$(UGA).

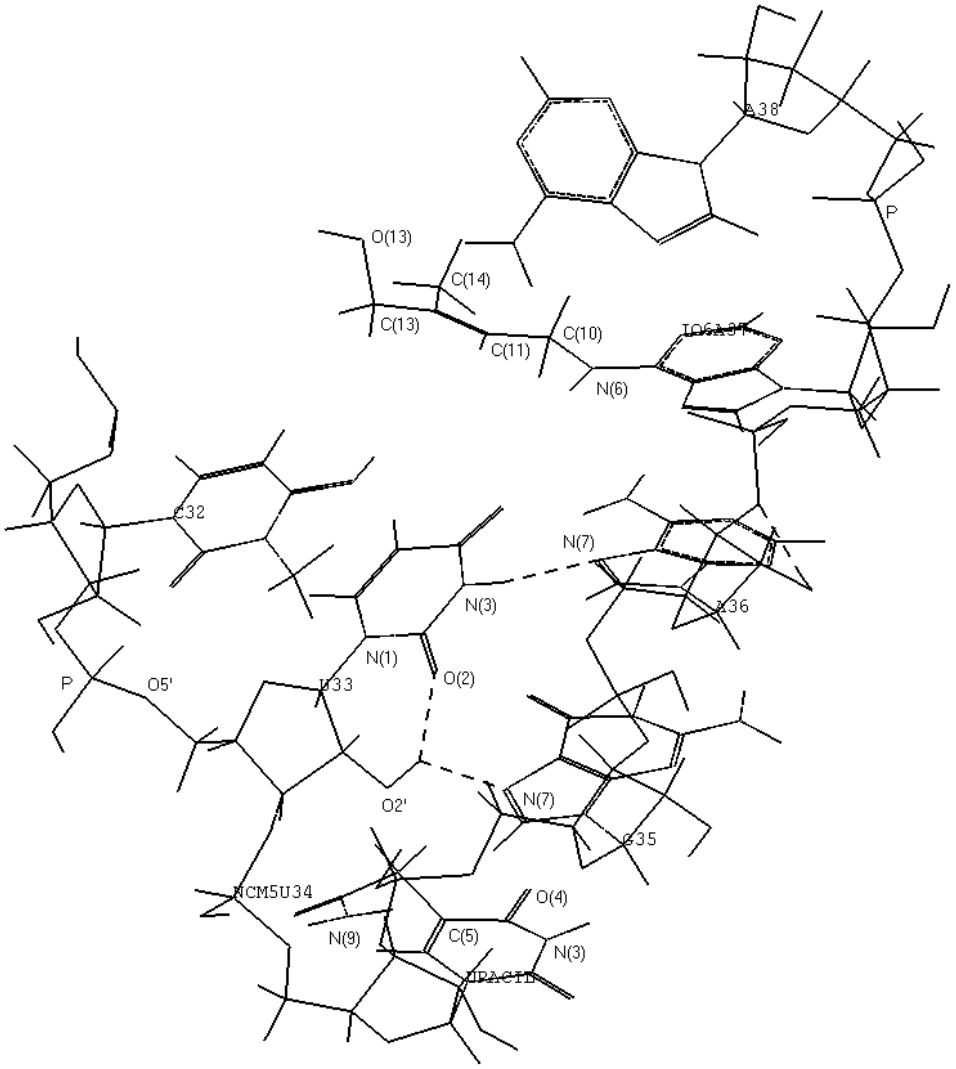


Fig. 6.5: Anticodon loop structure showing io^6A at 37^{th} and ncm^5U at 34^{th} position, MD snap shot taken at 208ps.

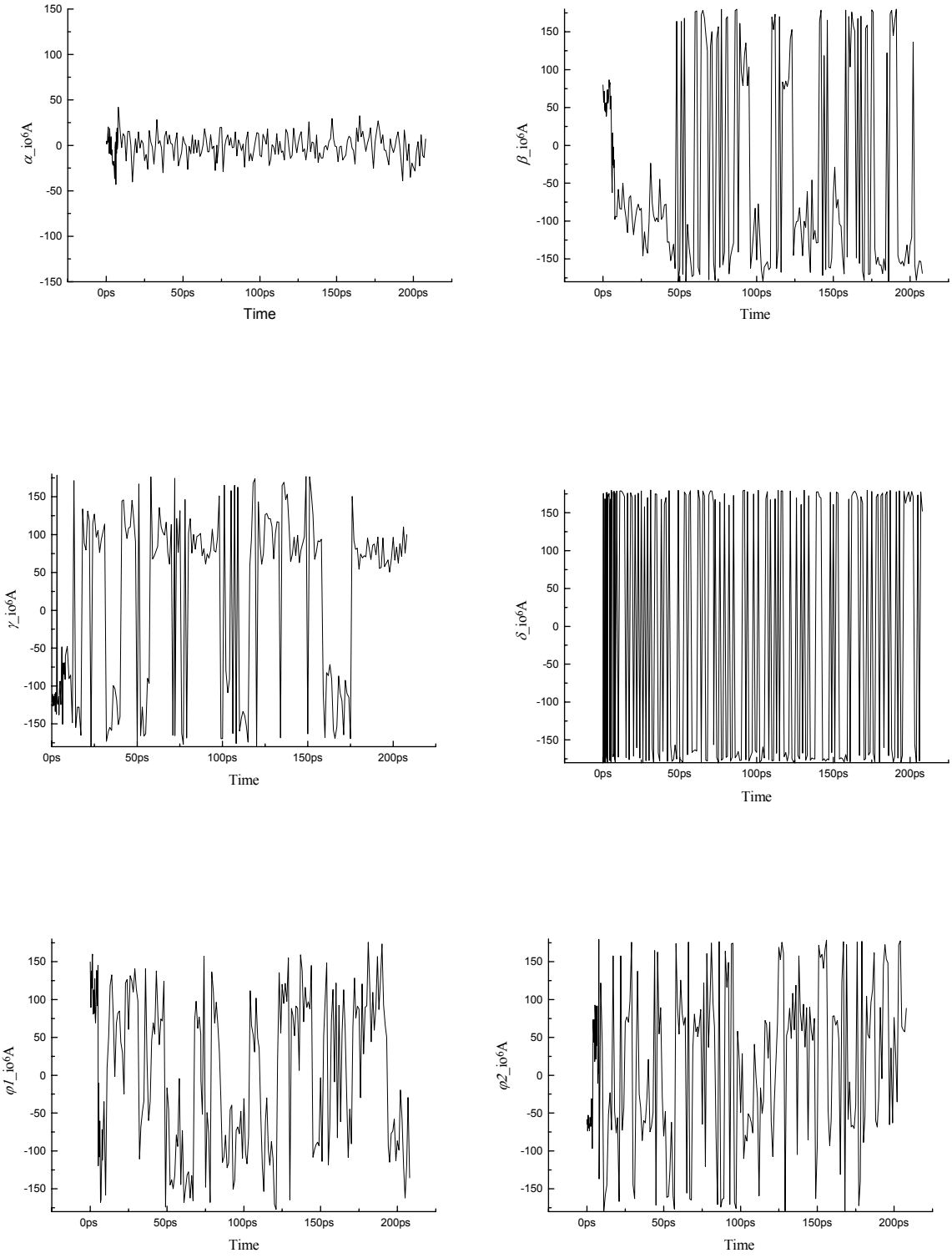


Fig. 6.6: Graphical representation of fluctuations in torsion angles α, β, γ, δ, Ψ1 and Ψ2 of hydroxyisopentenyl substituent at 37th position during MD simulation.

6.3.1.3 The conformation of ncm^{5}U(34):

The 5-carbamoylmethyl substituent (ncm^{5}U), which occurs at 34th position in anticodon loop, shows trans conformation. The (fig. 6.7) shows fluctuations of the various torsion angles in ncm^{5}U during MD simulation.

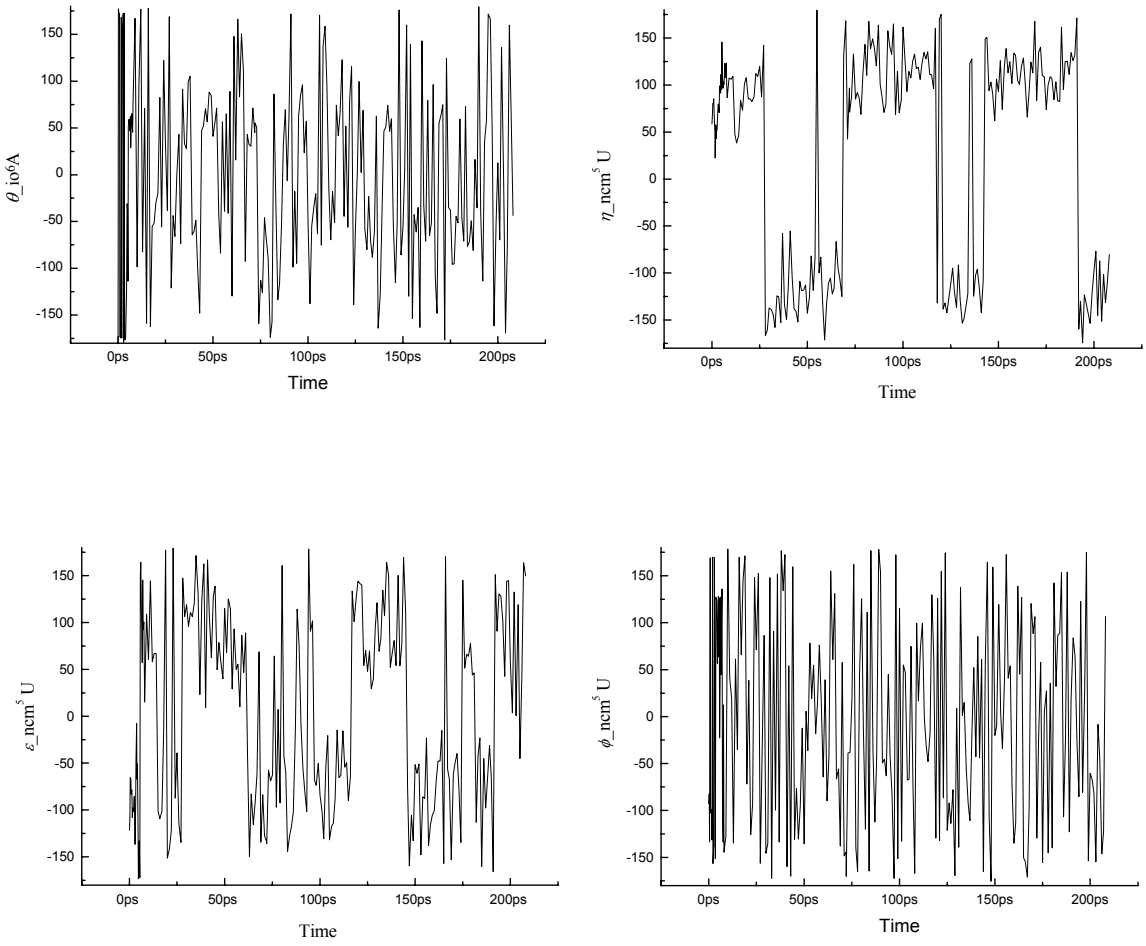


Fig. 6.7: Torsion Angles Fluctuations: θ of io^{6}A at 37th position and η, ε, and ϕ of ncm^{5}U at 34th position.

The torsion angle η shows values in between ±120° for maximum time. It also shows 150° value. The torsion angle ε shows alternative values, but most of the time it is in between –50° to –120° and between 50° to 120° and the torsion angle ϕ takes values in between 0° and 180° as shown in (fig. 6.7).

The ncm^{5}U substituent does not obstruct access to Watson-Crick base pairing sites. No interactions are observed between io^{6}A$_{(37)}$ and ncm^{5}U$_{(34)}$ hypermodified bases.

In tRNA^{Ser}(UGA), the solute-solvent temperature equilibrium is shown in (fig. 6.4). An equilibrated system should give the same internal temperature, for solvent and the solute [14]. Thus this solvated molecule system has reached equilibrium.

6.3.2 E. Coli tRNA^{Ser}(GGA):

6.3.2.1 Anticodon stem/loop analysis:

The anticodon stem/loop of E.Coli tRNA^{Ser}(GGA) contains io^{6}A at 37th position, a snap shot taken at 158ps is illustrated in fig.6.8(a). Base pairing in the stem is maintained, also the stacking of bases in anticodon loop is well maintained. The anticodon loop gets stability due to intramolecular interactions between pO$_{(35-36)}$…HN(3)$_{(33)}$ which is also a 'U' turn feature, N(7)$_{(35)}$…HO2'$_{(33)}$, pO$_{(35-36)}$…HN(6)$_{(32)}$, N(3)$_{(37)}$…HO2'$_{(37)}$ and N(7)$_{(36)}$…HN(6)$_{(32)}$ which are shown in fig. 6.8(b). The graphical representation of these interaction features is in fig. 6.9(b). The N(6)H$_{(32)}$ hydrogen bonds with N(7)$_{(36)}$ and pO$_{(35-36)}$. Besides 'U' turn feature hydrogen bonding between pO$_{(35-36)}$…HN(3)$_{(33)}$ and a sharp curve between 33rd , 34th and 35th is observed in this molecule fig.6.8(b). The base C(32) is close to the phosphate (p) between A$_{(36)}$-A$_{(37)}$.

The ribose ring puckering of 27th, 28th , 29th , 30th , 32nd , 33rd, 34th, 35th, 36th, 37th, 38th, 39th, 40th, 41st, 42nd, and 43rd nucleotides have c3'-endo form. The 31st ribose sugar ring shows c2'-endo conformation. All glycosyl torsion angles retain 'anti' conformation.

6.3.2.2 Hydroxyisopentenyl substituent orientation:

The anticodon loop structure containing io^{6}A at 37th position is shown in fig. 6.8(b). The torsion angle α retains 'distal' orientation, as $\alpha \cong 0°$ in fig. 6.9(a), during the course of MD simulation. The torsion angle β takes value 180° from 20ps to 90ps, then β value is between 60° to 90° up to 120ps and subsequently β is trans 180° again from 121ps to 158ps. The orientation of γ takes all values between 180°, ±60°. It shows 180° up to first 83ps, then from 84ps to 158ps show values nearly ±60°. The torsion angle δ mostly takes *trans* value, 180°. The torsion angle Ψ1 fluctuates between ±120°, some times values are also in between ±60°. The torsion angle Ψ2 up

to first 120ps fluctuates in between ±60°, then from 124ps to 158ps value in between 60° to 90° is seen. The torsion angle θ may vary in the range of 180°, ±60° values. The graphical representation of α, β, γ, δ, Ψ1 and Ψ2 are shown in fig 6.9(a), where as torsion angle θ is shown in fig 6.9(b).

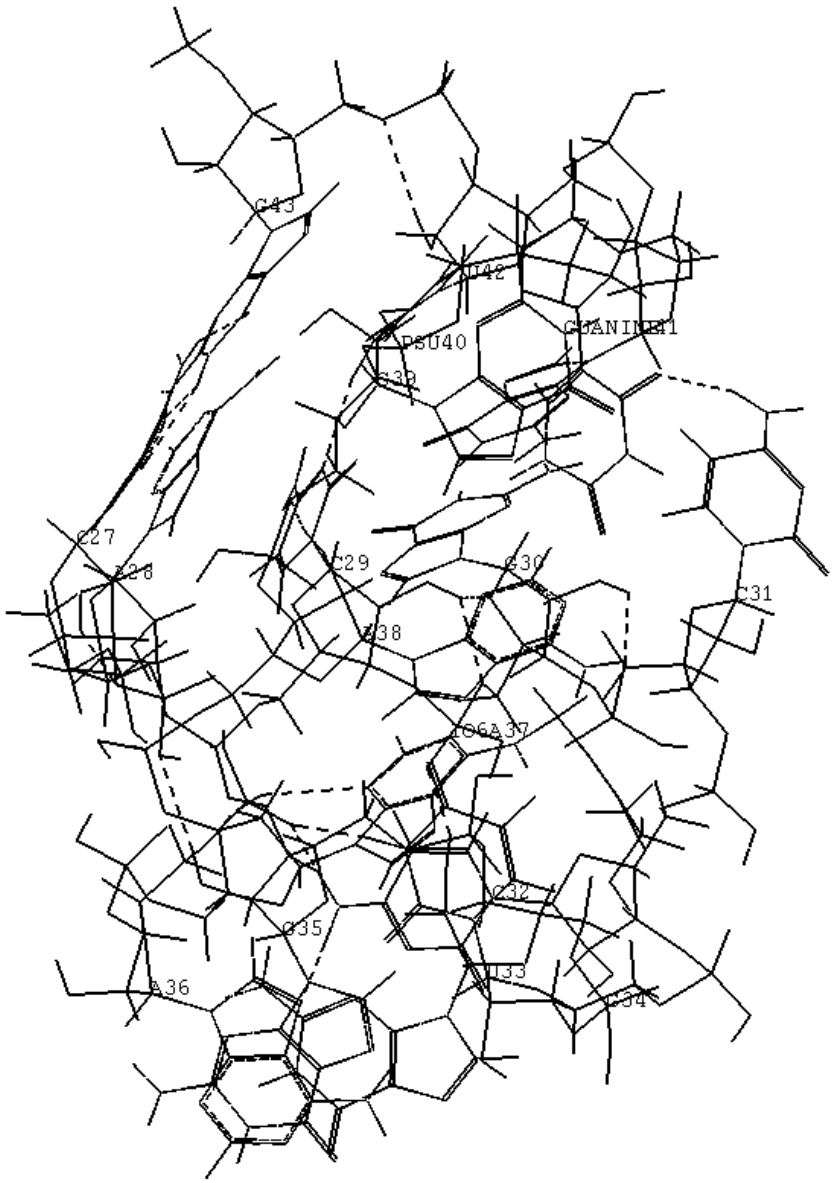


Fig. 6.8(a): The $tRNA^{Ser}$(GGA) anticodon stem/loop structure at 158ps.

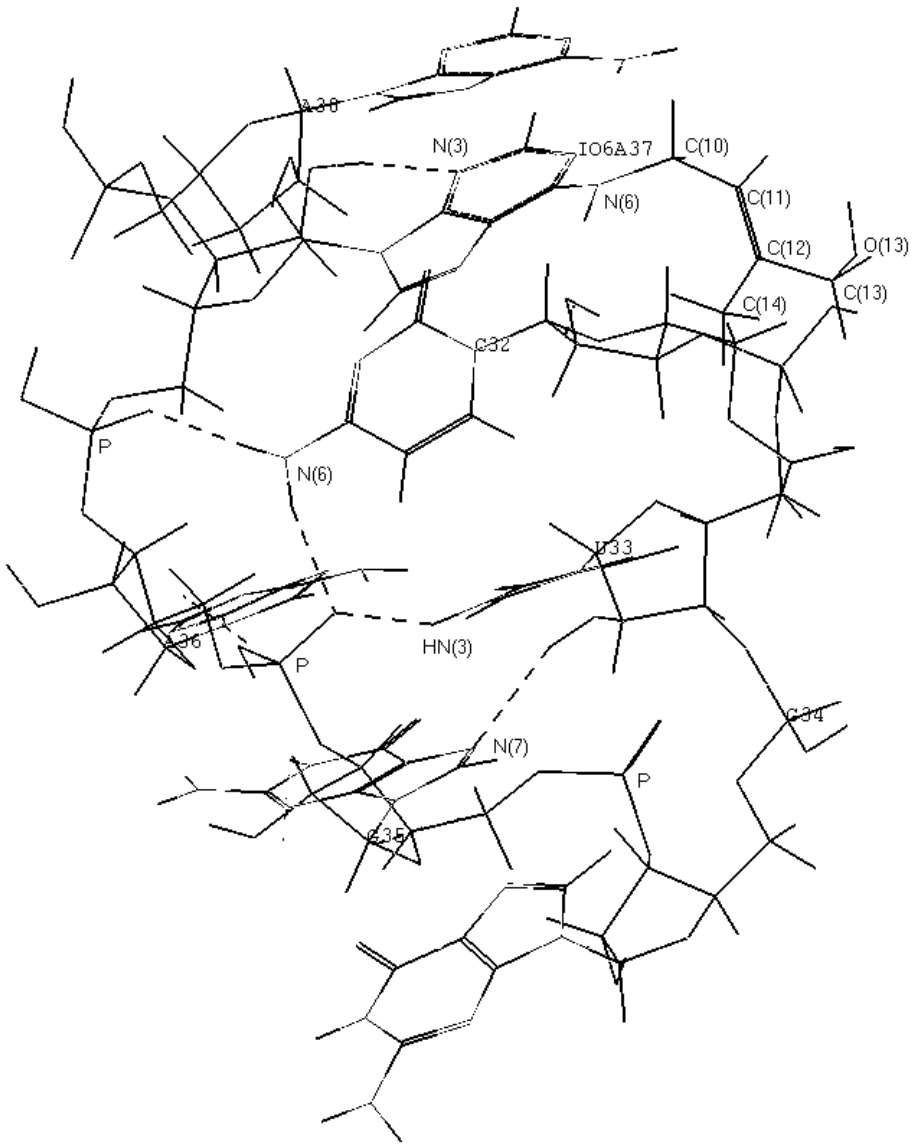


Fig. 6.8(b): Anticodon loop structure of $tRNA^{Ser}$(GGA) at 158ps with 'U' turn feature.

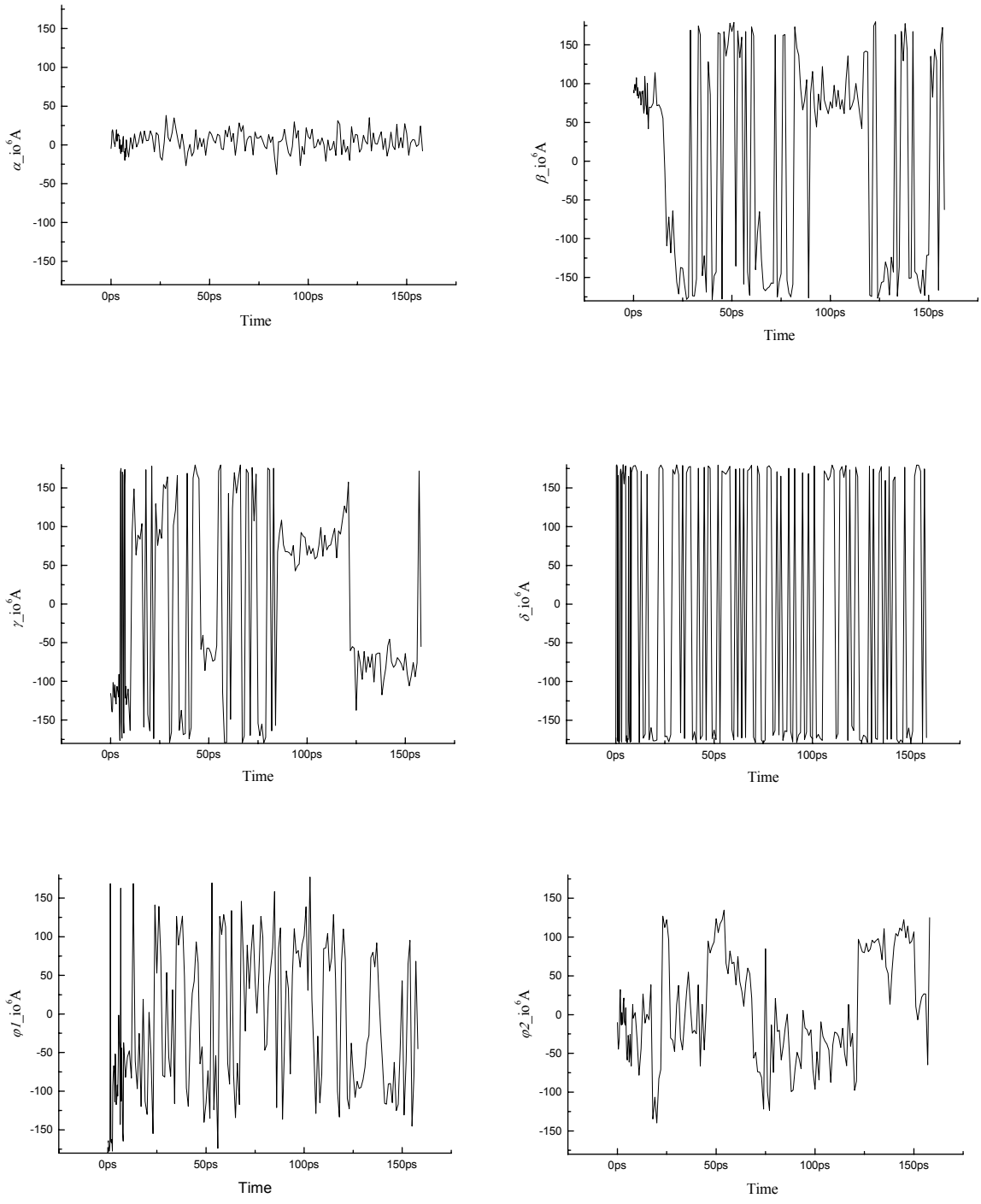


Fig. 6.9(a): Fluctuations during MD simulation of Torsion angles α, β, γ, δ, Ψ1, and Ψ2 for hydroxyisopentenyl substituent in $tRNA^{Ser}$(GGA)

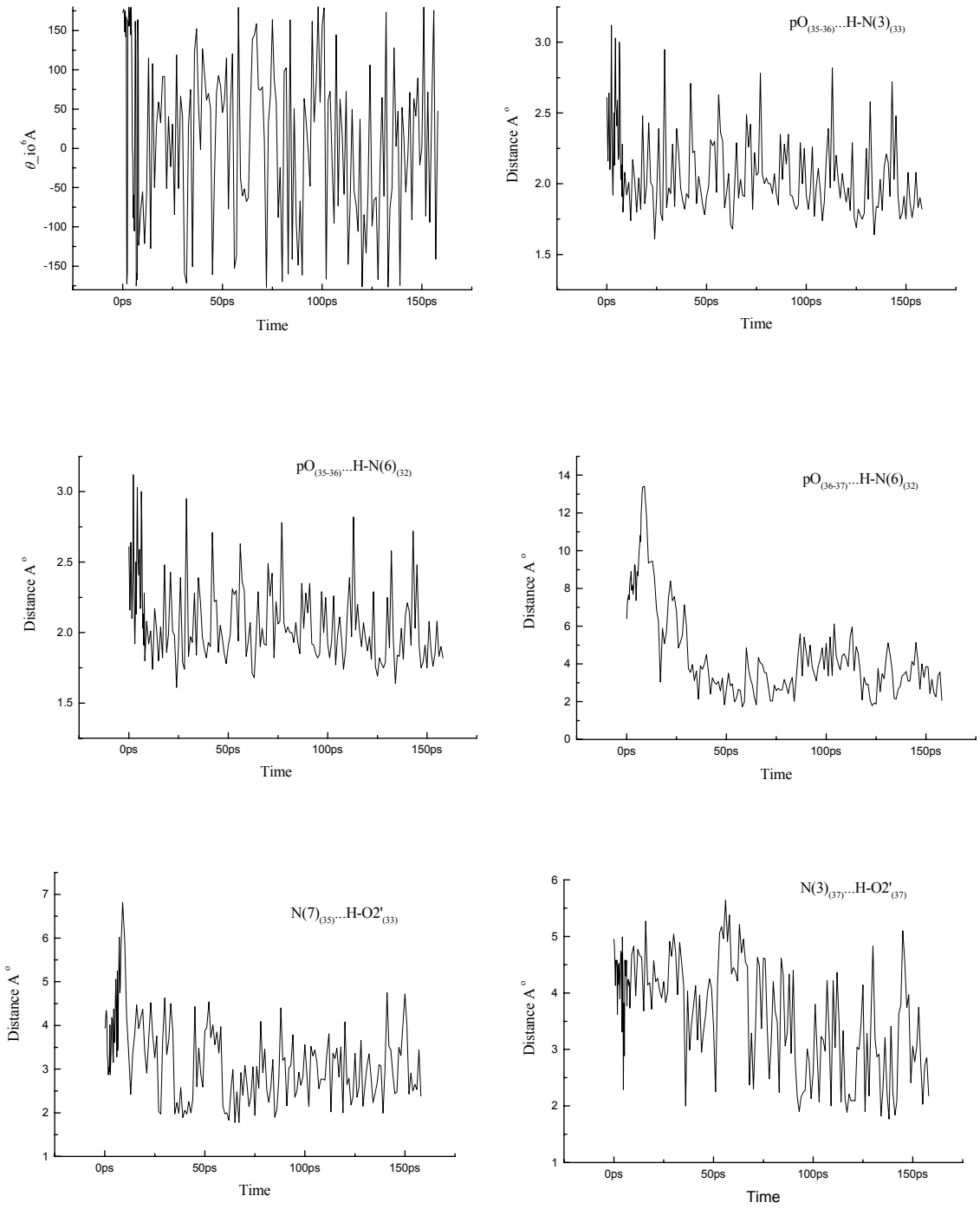


Fig. 6.9(b): Fluctuation of torsion angle θ, in $io^6A_{37,}$ and fluctuation of key distances for intramolecular interactions in anticodon loop of $tRNA^{Ser}$(GGA).

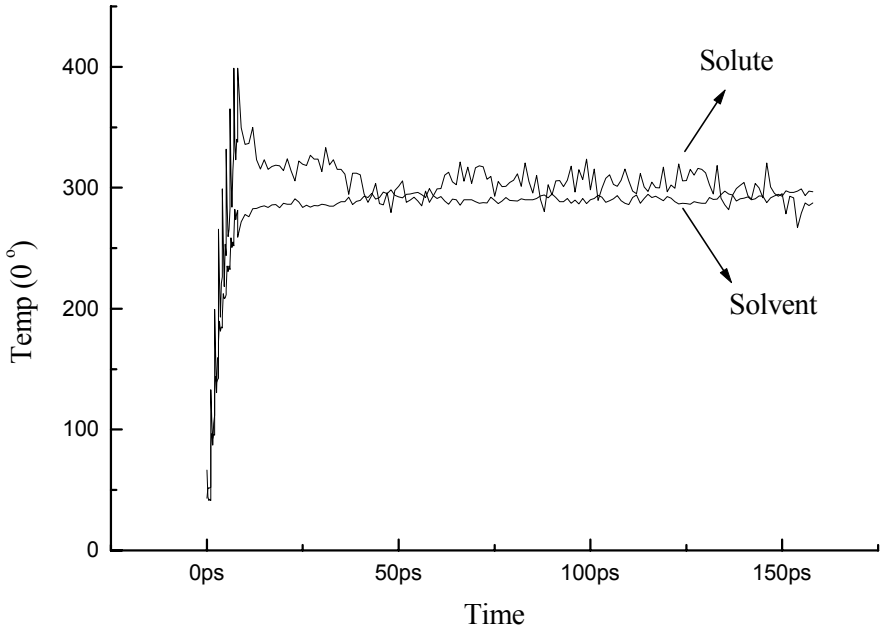


Fig. 6.10: Progress of the solute-solvent temperature equilibration.

Fig. 6.10 depicts the solute-solvent temperature equilibration graph. This molecular system has got thermal-equilibrium. The condition of thermal equilibrated system that the same internal temperature for the solvent and solute [14] is met.

6.4 Conclusion:

The hypermodified nucleoside io^6A at 37^{th} position in $tRNA^{Ser}$(UGA) and $tRNA^{Ser}$(GGA) is well accommodated in the anticodon loop. The hydroxyisopentenyl substituent spreads away from the five membered imidazole moiety (has 'distal' conformation), in agreement with the results of semi-empirical and molecular mechanics methods. The N(1) and N(6)H sites of 37^{th} adenine are not accessible for Watson-Crick base pairing. The anticodon loop is stabilized by various intramolecular interactions. The 5-carbamoylmethyl uridine (ncm^5U) at 34^{th} position does not restrict Watson-Crick base pairing. No interaction is found between io^6A at 37^{th} position and ncm^5U at 34^{th} position in this MD simulation also.

6.5 References:

1. Teichmann, T.; Urban, C.; Beier, H. Plant Mol. Biol.1994, 24, 889-901.
2. Sprinzl, M.; Horn, C.; Brown M.; Ioudovitch, A.; Steinberg, S. Nucleic Acid Res. 1998, 26, 148-153.

3. Adamiak R. W. and Gornicki, P. Prog. Nucleic Acids Res. Mol. Biol. 1985, 32, 27-74.
4. Motorin, y.; Bec, G.; Tewari, R.; Grosjean, H. RNA 1997, 3, 721-733.
5. Morin, A.; Auxilien, S.; Senger, B.; Tewari, R.; Grosjean, H. RNA 1998, 4, 24-37.
6. Agris, P. F. Prog. Nucleic Acids Res. Mol. Biol. 1996, 53, 79-129.
7. Persson, B. C. Molec. Microbiol. 1993, 8, 1011-1016
8. Limbach, P. A.; Crain, P.F. and McCloskey, J. A. Nucleic Acids Res. 1994, 22, 2183-2196.
9. Watts, M. T.; Tinoco, I. Biochemistry 1978, 17, 2455-2463.
10. Grosjean, H.; deHenau, S.; Crothers, D. M. Proc. Natl. Acad. Sci. (USA) 1978, 75, 610-614.
11. Holbrook, S. R.; Sussman, J. L.; Warrant, R.W. and Kim, S.H. J. Mol. Biol. 1978, 123, 631-660.
12. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. J. Am. Chem. Soc. 1984, 106, 765-785.
13. Ryckaert, J. P.; Ciccotti, G.; Berendson, H. J. C. J. Comput. Phys. 1977, 23, 327-336.
14. Sybyl Force field manual 1996, varsion 6.3, page 331.

List of Publications

1. N(7)-protonation induced conformational flipping in hypermodified nucleic acid bases N6 - (N-threonyl carbonyl) adenine and its 2-methylthio- or N(6)-methyl- derivatives.
   U.B. Sonavane, K.D. Sonawane, Annie Morin, Henri Grosjean and R. Tewari
   Int. J. Quantum Chem. 75 (3), 223-229 (1999).
2. Kailas D. Sonawane Henri Grosjean and Ravindra Tewari
   Journal of Biosciences Vol. 24 (Supplement 1) September 1999, P183.
3. Conformational flipping of the N(6) substituent in diprotonated N6 - (N- glycyl carbonyl) adenines: the role of N(6) H in purine-ring-Protonated ureido adenines.
   K. D. Sonawane, U. B. Sonavane and R. Tewari
   Int. J. Quantum Chem. 78, 398-405 (2000).
4. Conformational Preferences of Anticodon 3'-adjacent Hypermodified Nucleic Acid Base *cis*-or *trans*-Zeatin and its 2-methylthio Derivative, *cis*-or *trans*-ms^2Zeatin.
   K. D. Sonawane, U. B. Sonavane and R. Tewari
   J. Biomol. Struct. Dynamics**,** 19 (4), 637-648, (2002).
5. Conformational preferences of the base substituent in hypermodified nucleotide Queuosine 5`-monophosphate 'pQ' and protonated variant `pQH+'
   U.B. Sonavane, K.D. Sonawane and R. Tewari
   (Communicated)