Restricted backbone preference in the conformational landscape of amino acids:

Do they have a role to play in the peptide structure?

A thesis submitted in partial fulfillment of requirements for the Degree of Doctor of Philosophy

By

Pankaj Dubey
(MP12008)

Research Advisor
Prof. K. S. Viswanathan

Department of Chemical Sciences

Indian Institute of Science Education & Research (IISER) Mohali

Knowledge City, Sector 81, SAS Nagar, Manauli P.O. 140306,

Punjab, India

April 2019

In dedication to my family and loved ones

CERTIFICATE

I hereby certify that the thesis titled ‘Restricted backbone preference in the conformational landscape of amino acids: Do they have a role to play in the peptide structure?’ submitted for the degree of Doctor of Philosophy by Mr. Pankaj Dubey is the record of research work carried out by him during the period August 2014 to April 2019 under my guidance and supervision and further state that this work has not formed the basis for the award of any degree, diploma or a fellowship to any other university or institute. I further state that the work in this thesis was carried out independently by Mr. Pankaj Dubey.

In my capacity as the supervisor of the candidate’s Ph.D. thesis work, I certify that the above statements are true to the best of my knowledge.

Prof. K. S. Viswanathan

(Research Supervisor)

Place: Mohali

Date:

DECLARATION

The work presented in this thesis has been carried out by me under the guidance of Prof. K. S. Viswanathan at the Indian Institute of Science Education and Research Mohali. This work has not been submitted in part or in full for a degree, diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort has been made to indicate this clearly, with due acknowledgment of collaborative research and discussions. This thesis is a bonafide record of the original work done by me and all the sources listed within have been detailed in the bibliography.

Pankaj Dubey

Acknowledgments

To my supervisor **Prof. K. S. Viswanathan**: Sir, thank you for everything. Thank you for your trust, faith, freedom, support, guidance, care and love. It is an honor and privilege to be the last member of your academic journey. All these years you have nurtured me in every aspect of life, and I owe it all to you. If I become only half the thinker, half the teacher, half the person that you are, it will surely be one of my most significant accomplishments.

I extend my sincere gratitude to **Dr. Sugumar Venkataramani** and **Dr. P. Balanarayan** for their constant support throughout my Ph.D. work, for the regular assessment of research progress and their valuable suggestions.

I thank **Prof. Arvind**, Director IISER Mohali, **Prof. D. P. Sarkar** and **Prof. N. Satyamurthy**, former Directors, for the institutional supports and facilities.

I am thankful to **Dr. S. Arulananda Babu**, Head, Department of Chemical Sciences, IISER Mohali, for the departmental facilities and his support.

Special gratitude goes out to **Dr. Angshuman Roy Choudhury**, for helping and providing the computational resources.

I am grateful to my colleagues, **Dr. Ginny Karir, Jyoti, Dr. Kanupriya, Dr. Anamika, Srijit, Divita, Amala, Priyanka, Jai, Himanshi, Dipali**, who provided me with a positive and learning lab environment. **Dr. Anamika, Srijit and Jai**, collaborated on some of the work, and their effort is highly appreciated.

Special thanks to **Chitra Ma'am** for her love, care, support, and delicious foods.

I am grateful to my friends **Ashish, Deepu, Neeraj, Mayank, Shivam, Pritwish, Panda, Promit, Manoaj, Chitranjan, Naveen** and others for the beautiful moments and all the support.

I want to acknowledge the tremendous sacrifices that my parents made to ensure that I can pursue this journey. For this and much more, I am forever in their debt. It is to them that I dedicate this dissertation. I am grateful to my siblings, who have provided me through moral and emotional support in my life. I am also thankful to my other family members who have supported me along the way. And finally, I would like to thank Gargi for her endless support throughout this journey.

The financial support from MHRD and CSIR is highly acknowledged.

Pankaj Dubey

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

UV	Ultraviolet
GM	Gifford-McMohan
KBr	Potassium Bromide
HF	Hartree-Fock
MP	Møller-Plesset perturbation theory
CBS	Complete basis set
DFT	Density functional theorem
SCF	Self-consistent field
M06	Minnesota functional
FWHM	Full width at half maximum
ZPE	Zero point vibrational energy
AIM	Atoms in molecules
BCP	Bond critical points
NBO	Natural bond orbitals
NBODEL	NBO deletion
PES	Potential energy surface
IRC	Intrinsic reaction coordinates
TS	Transition states
CCC	cis-cis-cis or type I backbone geometry
$\mathbf{\Delta E_{ZPC}}$	Zero point corrected relative energy
ΔG	Relative free energy
$\mathbf{P_{conformer}}$	Normalized Boltzmann distribution of a conformer
$\mathbf{P_{\ backbone\ type}}$	Normalized Boltzmann distribution of a backbone type
$\mathbf{P_{\ I}}$	Normalized Boltzmann distribution of a backbone type I
$\mathbf{P_{\ II}}$	Normalized Boltzmann distribution of a backbone type II
$\mathbf{P_{\ (I+II)}}$	Sum of Boltzmann distribution of a backbone type I and type II
$\mathbf{\Delta E_{II-I}}$	Average stability of backbone type II relative to type I
$\mathbf{P_{\ P-I}}$	Normalized Boltzmann distribution of a backbone type P-I

$\mathbf{P}_{P\text{-}IV}$	Normalized Boltzmann distribution of a backbone type P-IV
$\mathbf{P}_{(P\text{-}I + P\text{-}IV)}$	Sum of Boltzmann distribution of a backbone type P-I and type P-IV
$\mathbf{\Delta E_{P\text{-}IV - P\text{-}I}}$	Average stability of backbone type P-IV relative to type P-I
$\mathbf{\Delta E_{linear}}$ **(amino acid)**	Average stability of linear-form backbone relative to the turn-like backbone in amino acids
$\mathbf{\Delta E_{linear}}$ **(dipeptides)**	Average stability of linear-form backbone relative to the turn-like backbone in dipeptides
$\mathbf{\Delta\Delta E_{linear}}$	Change in the relative stability of linear form backbone structure going from amino acids to dipeptides
H ($\alpha_R + \alpha'_R$)	Helix region
E (***ext + ppII***)	Extended strand region
$\mathbf{P}_{(H)}$	Helix propensity
$\mathbf{P}_{(E)}$	Extended-strand propensity
$\mathbf{\Delta E}_{(E\text{-}H)}$***(amino acid/dipeptide)***	Average stability of extended strand relative to helix structure in amino acid/dipeptide
$\mathbf{\Delta\Delta E}_{(E\text{-}H)}$	Change in the relative stability of extended-strand going from amino acid to dipeptide

List of Publications

Dubey, P.; Mukhopadhyay, A.; Viswanathan, K. S. Do Amino Acids Prefer Only Certain Backbone Structures? Steering through the Conformational Maze of L-Threonine Using Matrix Isolation Infrared Spectroscopy and *Ab initio* Studies. *J. Mol. Struct.*, **2019** 1175 (5), 117–129.

Manuscript in preparation:

Dubey, P.; Viswanathan, K. S, Conformational picture of L-glutamic acid. A matrix isolation FTIR and *ab initio* studies (in preparation).

Dubey, P.; Viswanathan, K. S, Matix isolation FTIR and *ab initio* studies of conformations of L-methionine (in preparation).

Journal publication based on the work not included in this thesis:

Mukhopadhyay, A; **Dubey, P.** Investigation of solvation of ammonium salts: A Raman spectroscopy and *Ab initio* study; J. Raman. Spectrosc. **2017**, 48, 1-11

Dubey, P.; Mukherjee, S.; Choudhury A. R.; Viswanathan, K. S. Fluorescence Enhancement of Tb^{3+} in the $Tb3^{+}$ -TrimesicAcid-Gd^{3+} Complex: Role of Polynuclear Structures; *ChemistrySelect*, **2019**, 4, 2747-2752.

Book Chapter:

Dubey, P.; Saini, J.; Verma, K.; Karir, G.; Mukhopadhyay, A.; Viswanathan, K.S. Matrix Isolation Spectroscopy-A Window to Molecular Processes A book chapter in Molecular and Laser Spectroscopy, Edited by V.P. Gupta; *Elsevier*, **2017**, ISBN: 9780128498835.

Synopsis

S.1 Introduction

The word protein was first coined by J.J. Berzelius in 1838 A.D. and was derived from the Greek word "proteios," which means "of the first rank." Proteins are the most abundant biological macromolecules that facilitate almost every process that takes place in a cell. The process by which protein adopts its functional native structure is called "protein folding" and is one of the most complex reactions known. All the efforts to comprehend the protein folding process mostly rely upon Anfinsen's theory, which suggests that the primary amino acid sequences govern the folded structure of peptides and proteins.

The association between the amino acid composition and secondary structure is rationalized in term of *conformational propensity*, which is defined as the quantitative measure of the occurrence of a given amino acid in the helical, sheet and turn region of folded proteins and peptides. The *conformational propensity* values obtained, however, reflects both the *intrinsic propensity* of the corresponding amino acid residue and the environmental influence. Therefore, the conformational propensity values of amino acids are 'context dependent' and vary from one protein to another. Furthermore, these *conformational propensity* values also do not replicate the unfolded state of a protein or the initial steps of secondary structure formation.

The traditional understanding of the unfolded state of the protein indicates that all natural amino acids residues uniformly occupy the sterically allowed region of Ramachandran plot. Likewise, the thermodynamical perspective of maximum entropy suggests that the unfolded proteins do not exhibit any order or residual structure and can, therefore, be considered as a completely random structure.[6] However, several recent research evidence contradicts above understanding of the unfolded state of the protein and indicate about the presence of residual or

ordered structure, which depends upon the amino acid composition of unfolded protein or peptide.

Therefore, a reliable understanding of the partially ordered structure of the unfolded state of protein and peptides requires the knowledge of *intrinsic conformational propensities*, which is the quantitative measure of the occurrence of respective amino acids in the helical, sheet and turn region; however, independent of environmental effect. In the past, several attempts have been made to study and measure the intrinsic conformational propensity of amino acid, both experimentally and computationally, employing several model systems. However, the basis of the physical origin of the *intrinsic conformational propensity* of amino acid and factor determining it is not well understood and is still a matter of debate.

The recently published work by Vymetal *et al.*, which focuses on the *conformational intrinsic propensity* measurement using various force fields, further bolsters this argument. In this work, the author expresses his view upon the origin of *intrinsic conformational propensity* and to quote him "The physical origin of this distinct propensities remains still unclear, and it is questionable whether it can be investigated by computed stimulation using the current methods and models."

An alternative approach to understanding the physical origin of *intrinsic propensity* is presented in this work and is based on comprehending the conformational preferences of free amino acids and its effect upon the structure of dipeptides. Amino acids show a rich conformational landscape owing to the presence of multiple rotatable bonds. Glycine, the simplest amino acids, is a three-rotor molecule and has been found to show various conformations. In amino acids other than glycine, the presence of the bulky side-chains results in

a further complex but rich conformational landscapes. A comprehension of the conformational landscape of naturally occurring α-amino acids is therefore required.

Similarly, the conformational landscape analysis of dipeptides can reveal the structural preferences after peptide bond formation. An understanding of the conformational preferences of the amino acids before peptide bond formation and those of the dipeptides after peptide bond formation will, therefore, help rationalize the bridging link to understand the origin of intrinsic propensity.

Amino acids have been widely studied in the solid and solution phase. The gas phase study is useful in providing the conformational landscape of neutral or canonical form of amino acids. Investigation of low energy conformers of amino acids is challenging both experimentally and computationally. Several spectroscopic techniques such as microwave spectroscopy, infrared spectroscopy, multiphoton ionization spectroscopy are widely used for the conformational study of amino acids. Infrared spectroscopy, coupled with the matrix isolation technique, provides spectra with small enough linewidths to resolve the features due to individual conformers. Apart from the structural determination, infrared spectroscopy is also useful in probing intramolecular hydrogen bonding interactions that may be present in certain conformers of amino acid. Thus, the matrix isolation infrared spectroscopy allows for a good understanding of the conformations of amino acids.

S.2 Experimental and Computational Details

Experimental study of amino acids was done using matrix isolation infrared spectroscopy. The amino acid sample was deposited employing a hot nozzle, placed near the cold substrate window. Ar was used as matrix gas in all the experiments was allowed to flow over the heated amino acid sample, and the matrix gas/amino acid mixture was then deposited on the 12 K

substrate. A KBr window was used as the substrate for deposition. Typical temperatures of the hot nozzle ranged from 100º C to 150º C. Care was taken to avoid decomposition of the amino acids during the deposition. The flow rate of the matrix gas was typically three mmol/h. After deposition that usually lasted for 1 to 2 hours was complete, an FTIR spectrometer (Bruker Tensor 27) operating at a resolution of 0.5 cm^{-1} was used to record the spectra of trapped species over the region of 4000-400 cm^{-1}. The matrix was then annealed at 32 K, and the spectra were recorded after recooling the matrix back to 12 K. Annealing was done to observe conformational cooling, if any.

The experimental results were corroborated with the *ab initio* quantum mechanical computations. The computations were carried out with Gaussian-09 suite of programs. The potential energy surface (PES) of amino acids and dipeptides were scanned to obtain possible conformers. The PES scan was carried out by optimizing a large number of initial geometries of the concerned amino acid at HF/6-31G level of theory. The optimized geometries obtained were then refined at the higher level of theory such as MP2 and M06-2X, employing 6-311G++(d,p) basis set. Factors determining conformational stability were explored by performing atoms in molecules (AIM) and natural bond orbital (NBO) analysis.

S.3 Conformational Landscape of L-Threonine

Herein, we investigated the conformational landscape of L-threonine, both experimentally and computationally. L-threonine is a polar amino acid containing a hydroxyl group as the side-chain and is one of the two naturally occurring α-amino acids with two chiral centers. The conformations of L-threonine using *ab initio* computations have been studied by Zhang *et al.* and Szidarovszky *et al.* In the experimental study using rotational spectroscopy, Alonso and

coworker observed seven different conformers L-threonine and suggested that stability of low energy conformers arises due to two different types of intramolecular hydrogen bonding.

In molecules where conformers are stabilized by hydrogen bonding, infrared spectroscopy often provides unequivocal evidence for the presence of such non-covalent interactions. Hence, the conformational study of L-threonine was revisited to understand the role of intramolecular hydrogen bonding in conformational stability. Apart from the hydrogen bonding interaction, the role of the hydroxyl side-chain group and orbital delocalization interactions in defining the conformational stability were also addressed.

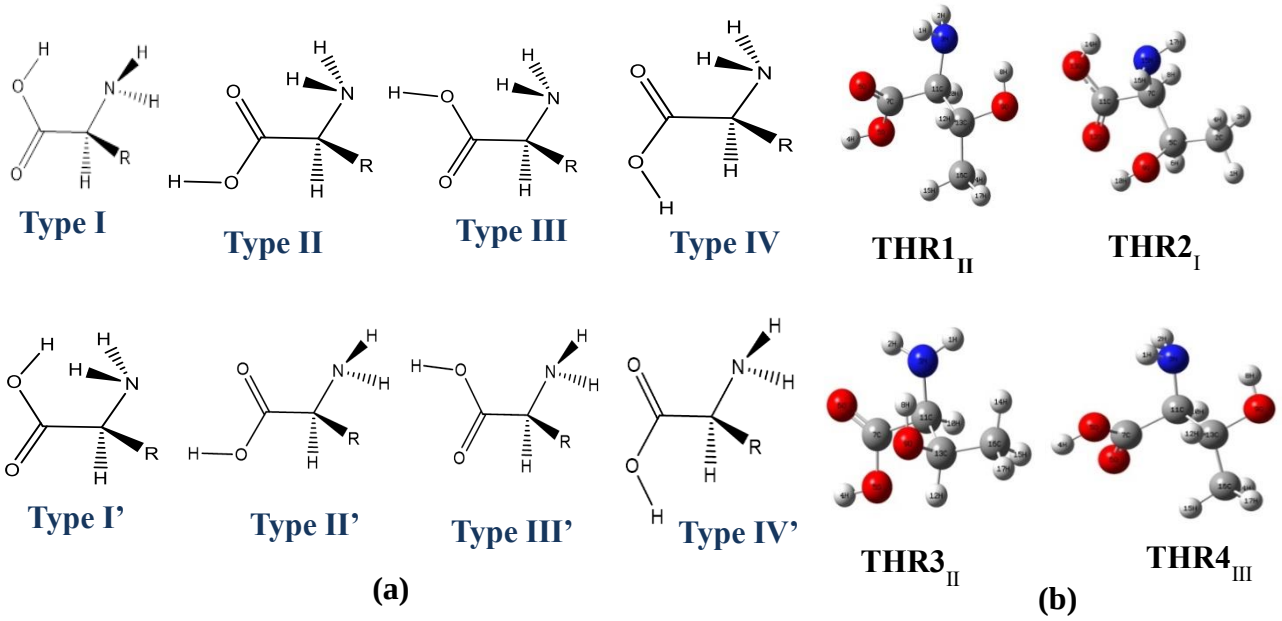


Fig. S1 a) Eight backbone structure of amino acid; b) Four lowest energy conformers of L-threonine

The potential energy surface scan yielded 38 conformers of L-threonine, which were then classified into eight classes based on the backbone structure (Fig. S1(a)). The conformational nomenclature based upon conformer relative energy and backbone type was proposed. As shown in Fig. S1(b), the Roman numeral subscript in the nomenclature for each conformer indicates the

backbone type to which a given conformer belongs too. For example, $THR1_{II}$ suggest that conformer THR1 belongs to backbone type II.

Matrix isolation study of L-threonine was carried out by heating sample at two different temperatures of 403 K and 423 K. The experimental spectra were found to be in good agreement with the computed scaled spectra of four lowest energy conformers (shown in Fig. S1(b)), whose populations were calculated to be in excess of ~10 %. The O-H spectral feature of backbone type I conformer, i.e., $THR2_I$ was found to be red-shifted by approximately 119 cm^{-1} from the feature corresponding to the same mode of $THR1_{II}$. The significant shift is due to intramolecular hydrogen bonding, which is also corroborated by AIM analysis. In the earlier studies, the stability of backbone type II conformers (i.e., $THR1_{II}$ and $THR3_{II}$) was attributed to the presence of bifurcated $C{=}O^{\cdots}H_2N$ hydrogen bond. However, our AIM analysis does not show any bond critical points and refutes the existence of bifurcated $C{=}O^{\cdots}H_2N$ hydrogen bond in backbone type II, and also is in agreement with the experimental result.

The NBO deletion study indicated that vicinal orbital delocalization interaction plays a significant role in stabilizing backbone type II. The role of the side-chain O-H group in defining conformational stability was understood by comparing the conformational preference of L-threonine with glycine (where just a proton is a 'side-chain'). This comparison suggests that the side-chain of L-threonine is unlikely to play a defining role in the conformational preferences, and at best, it may be supportive.

We then proposed a pictorial representation to depict the energetics of the various conformers, which is shown in Fig. S2. We call this "conformational dartboard." Each track on the dartboard represents a given backbone structure and all the conformations belonging to this backbone fall on that particular track. Based on the energy of a given conformer corresponding

to this backbone structure, the conformers is located on this track. Beginning from the X-axis which corresponds to the lowest energy conformer, which is arbitrarily set as zero, the complete 360º is taken to represent 10.0 kcal mol^{-1}. A conformer, therefore, can be positioned at a θ, which is given by relative energy ΔE, where θ = ΔE x (360/10). The first octant of the dartboard (Fig. S2) which represents 1.25 kcal mol^{-1} is populated predominantly by conformer lying on the track corresponding to backbone type I and II and a few to III. This observation suggests that of eight backbone structures, low energy conformers of L-threonine prefer predominantly type I and II and rather rarely type III.

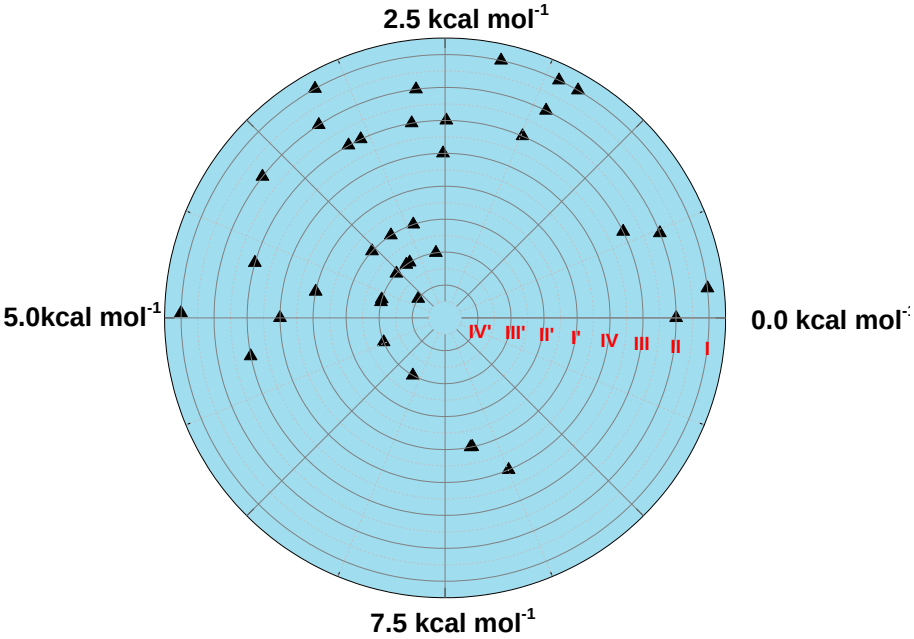


Fig. S2 Conformational dartboard of L-threonine

S.4 Conformational Picture of L-Glutamic Acid

L-glutamic acid is one of principal neurotransmitter of the vertebrate central nervous system and is known to play a significant role in the neural differentiation of a developing brain. In the past, the conformational study of L-glutamic acids has been attempted both computationally and experimentally. Alonso and co-worker in their microwave spectroscopy of L-glutamic acid

reported five conformers of L-glutamic acid. Furthermore, the stability of all the conformers was attributed to the presence of two types of intramolecular hydrogen bonding, i.e. (O-H···N) and bifurcated (C=O···H_2N). However, as mentioned earlier in our work on conformations of L-threonine, the experiments and computations clearly do not show the presence of any bifurcated (C=O···H_2N) hydrogen bond. Therefore, to understand the role of intramolecular hydrogen bonding and the influence of acidic side-chain group and on the conformational landscape of amino acids, we performed the matrix isolation study of L-glutamic acid.

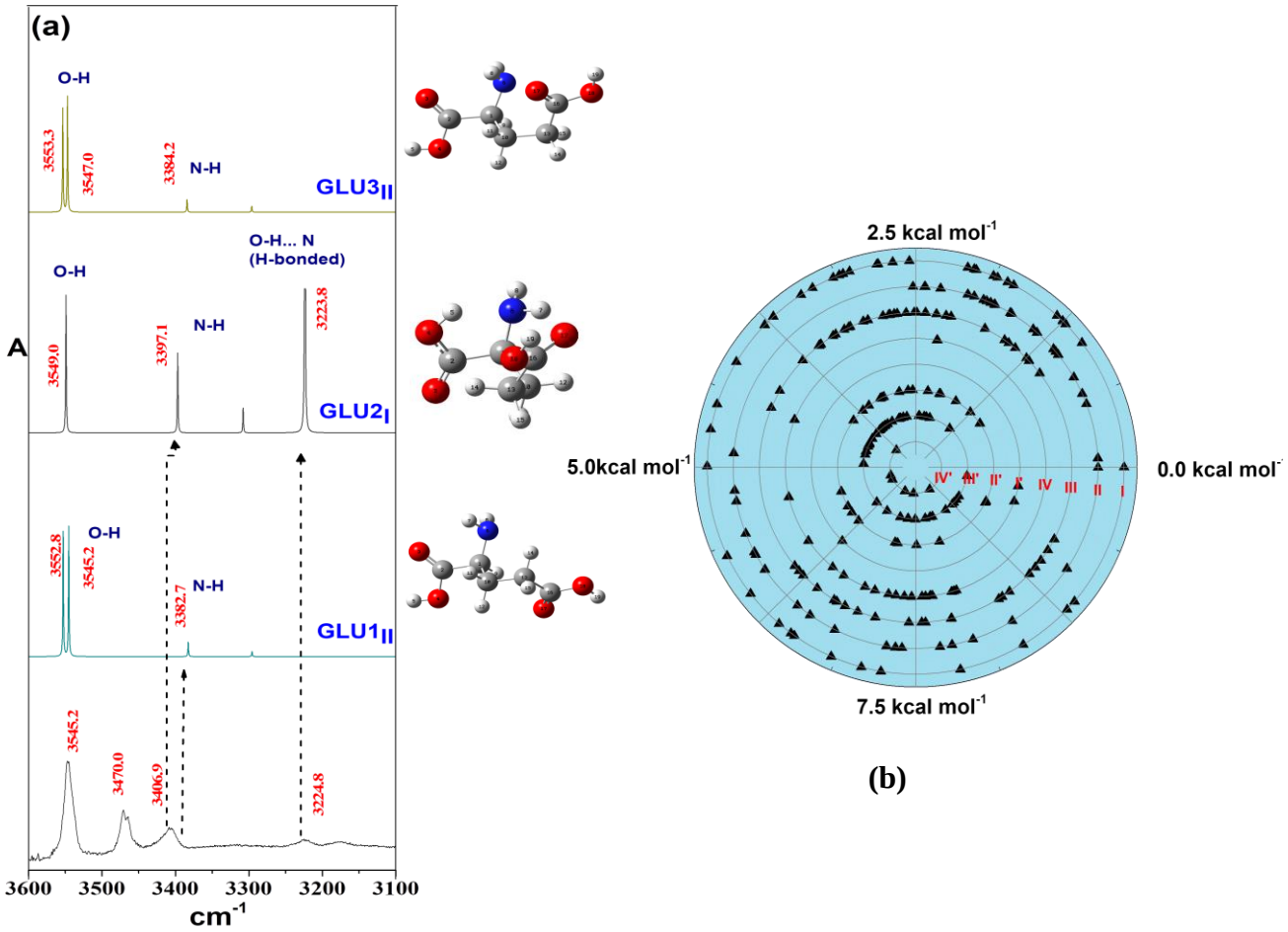


Fig. S3 Comparison of experimental and computed spectra of L-glutamic acid. The spectra span the region 3600-3100 cm^{-1} (grid a) i) IR spectra of Matrix isolated L-glutamic acid deposited at 388 K; (ii-iv) Computed spectra showing the spectral features of the three lowest energy conformers, calculated at the MP2/6-311++G(d,p) level of theory; (b) Conformational dartboard of L-glutamic acid at MP2/6-311++G(d,p) level of theory.

The matrix isolation FTIR study of L-glutamic acid was carried out by heating the sample at 388K and 398 K. The PES scan yielded 267 conformers of L-glutamic acid, at MP2/6-311++G(d,p) level of theory. The backbone type classification method, which was used earlier in the case of L-threonine, was employed for the conformational classification. All the features in the experimental spectra were assigned to the computed scaled wavenumbers of conformer $GLU1_{II}$, $GLU2_{I,}$ and $GLU3_{II,}$ as shown in Fig. S3(a). The O-H stretching mode of backbone type I conformer of glutamic acid, i.e., $GLU2_I$ shows a sizeable spectral shift of ~ 320 cm^{-1}, thus suggesting the presence of intramolecular (O-H$^{\cdots}$N) hydrogen bonding. However, the O-H stretching mode of conformer $GLU1_{II}$ and $GLU3_{II}$ does not show any shift and are very closely spaced. The AIM analysis of backbone type II conformer does not locate any bond critical point along bifurcated (C=O$^{\cdots}H_2$N) bond path, thus eliminating the stability due to the hydrogen bond.

The observation is also consistent with our previous study of L-threonine, where the backbone type II conformers were found to be stabilized by vicinal delocalization interaction and not due to bifurcated (C=O$^{\cdots}H_2$N) hydrogen bonding. The conformational dartboard plot of L-glutamic acid conformer, like L-threonine, also demonstrates the tendency of low energy conformers to prefer only type I and II structures. As shown in Fig. S3(b), the first octant, which represents conformers with relative energy of 1.25 kcal mol^{-1} is predominantly populated by backbone type I and II conformers.

S.5 Conformations of L-Methionine

In the amino acids discussed earlier the low energy conformers were found to prefer backbone type I and II, and the side-chain groups, i.e., polar (L-threonine) and acidic (L-glutamic acid) were found to play an only supportive role, but not a defining factor for conformational stability. After rationalizing the effect of polar and acidic amino acids on the conformational preferences,

the effect of the hydrophobic side-chain on the conformational landscape was studied. Therefore, the matrix isolation FTIR and *ab initio* computational study of L-methionine was performed. Apart from being non-polar amino acids, L-methionine is also sulfur-containing amino acids, and hence the conformational study of L-methionine is essential. The experiments were carried out by the heating sample at 408 K and 423 K and depositing it on the cold substrate window together with the matrix gas. In the computational study, the PES scan of L-methionine yielded 165 conformers at MP2/6-311++G(d,p) level of theory.

The conformers were classified based on their backbone structure, as discussed in the earlier section. The experimental features of L-methionine were corroborated with the computed scaled wavenumbers of lower energy conformer. The assignment between experimental and calculated wavenumbers shows the presence of three lowest energy conformers, i.e., $MET1_{II}$, $MET2_{I}$, and $MET3_{III}$ in the experiments. As observed previously in L-threonine and L-glutamic acid, the backbone type I conformer of methionine, i.e., $MET2_{I}$ also shows a shift in the O-H spectral feature and thus suggesting the presence of intramolecular (O-H$^{\cdots}$N) hydrogen bond. The observation was further supported by AIM analysis, which shows bond critical points along (O-H$^{\cdots}$N) bond path.

Similarly, the backbone type II conformer, i.e., $MET1_{II}$ does not show bifurcated (C=O$^{\cdots}$H$_2$N) hydrogen bond both experimentally and computationally. The role of the non-polar sulfur-containing side-chain on the conformational preference was found to be supportive, but not decisive. The conformational dartboard plot of L-methionine also demonstrates the preferred choices for backbone structure I and II to the lower energy forms.

S.6 α-Amino Acids, Tale of Two Structures

The tendency of low energy conformers of a non-polar (L-methionine), polar (L-threonine) and acidic (L-glutamic acid) amino acids to prefer only specific backbone structure indicates about the generality in the trend over all the amino acid. Furthermore, it also suggests that the nature of the side-chain group does not play a definitive role in conformational preference, and at best, they may be supportive. Therefore, in this chapter, we have looked for the generality of the above trend of choice for specific backbone structures, over the entire naturally occurring α-amino acid. The conformational manifold of alpha-amino acids, other than previously discussed, was studied only by *ab initio* computations. The PES scan was performed for all the amino acids, except arginine and the conformers obtained were classified, based on the backbone structure (Fig. S1(a)).

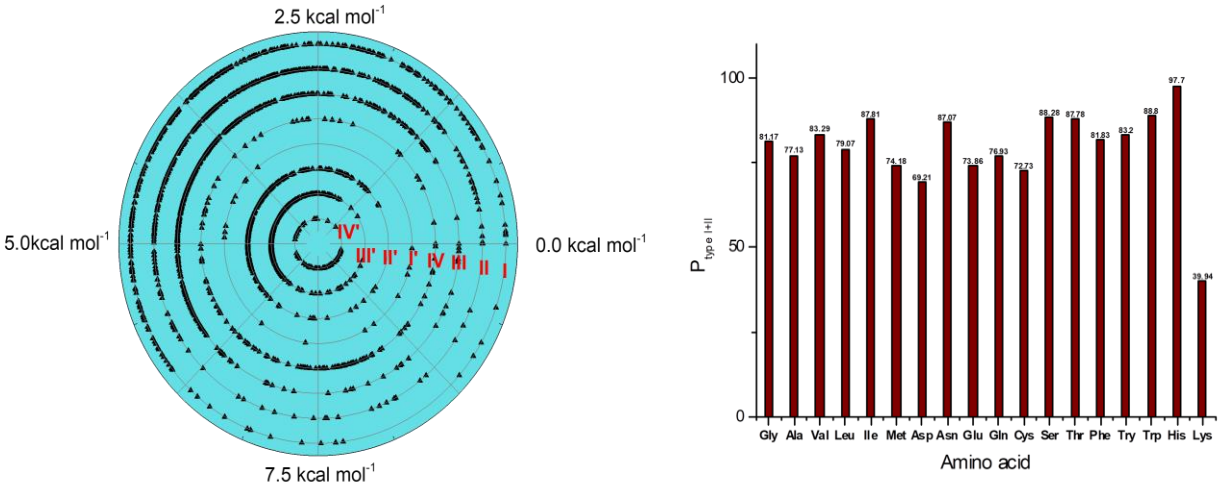


Fig. S4 (a) Conformational dartboard of naturally occurring α-amino acid, except arginine; (b) The sum of type I and II backbone propensities of α-amino acid, except arginine.

The conformational dartboard comprising of all the conformers of naturally occurring amino acid is shown in Fig. S4(a). The first octant of the dartboard, which represents conformer energies up

to 1.25 kcal mol^{-1} is populated by conformers corresponding to backbone type I and II. The backbone preference tendency was then rationalized quantitatively by backbone propensity. The backbone propensity is defined as the quantitative measure of amino acid tendency to adopt a given backbone geometry. The sum over the normalized Boltzmann distribution of all the conformers, which belongs to a particular backbone type, gives backbone propensity. Fig. S4(b) shows the sum of backbone propensity of type I and II (Y-axis) for various α-amino acids (X-axis). The average overall amino acids indicate that ~ 80% of the total population of alpha-amino acid exists in the form of either backbone type I or II. The restricted structural preference observed in α-amino acid can be useful in understanding the physical origin of intrinsic conformational propensity and rationalizing the conformational choices of dipeptides.

S.7 What is the Role of the Amino Acid Conformational Dichotomy on the Conformational Landscape of the Dipeptides?

The amino acids are the fundamental units of proteins and peptides, and therefore, the limited backbone preference of α-amino acid conformers can play a pivotal role in defining the conformational preference of dipeptides. The relationship between the conformers of amino acids and the structure of peptides was examined using dipeptides (Fig. S5(a)) as models. The model dipeptide is an amino acid residue which is methyl protected at both C-terminal and N-terminal, as shown in Fig. S5(a). This model dipeptide mimics the structure of an amino acid residue in a polypeptide chain, but with no neighboring residue. Therefore, the model dipeptides are widely used in understanding the intrinsic conformational propensity of amino acid residue and are a good system to rationalize the effect of restricted structural preference of amino acids. The PES scan, as was done for the amino acids, was carried out to obtain the conformers of model dipeptides. Furthermore, a classification of the conformers for these dipeptides, based on

backbone geometry, very much similar to that employed for the amino acid classification, was done. The conformational dartboard plot containing the conformers of all model dipeptides, (except diarginine), reflects the preference of backbone type P-I and P-IV as low energy conformer, as shown in Fig. S5(b). The backbone propensity computed based on the method discussed in the previous section shows that 98% of the population of model dipeptides adopts either backbone structure I or IV (Fig. S5(c)). The amino acid preference for the backbone type I and II, and model dipeptides inclination to take backbone type P-I and P-IV were quantitatively correlated.

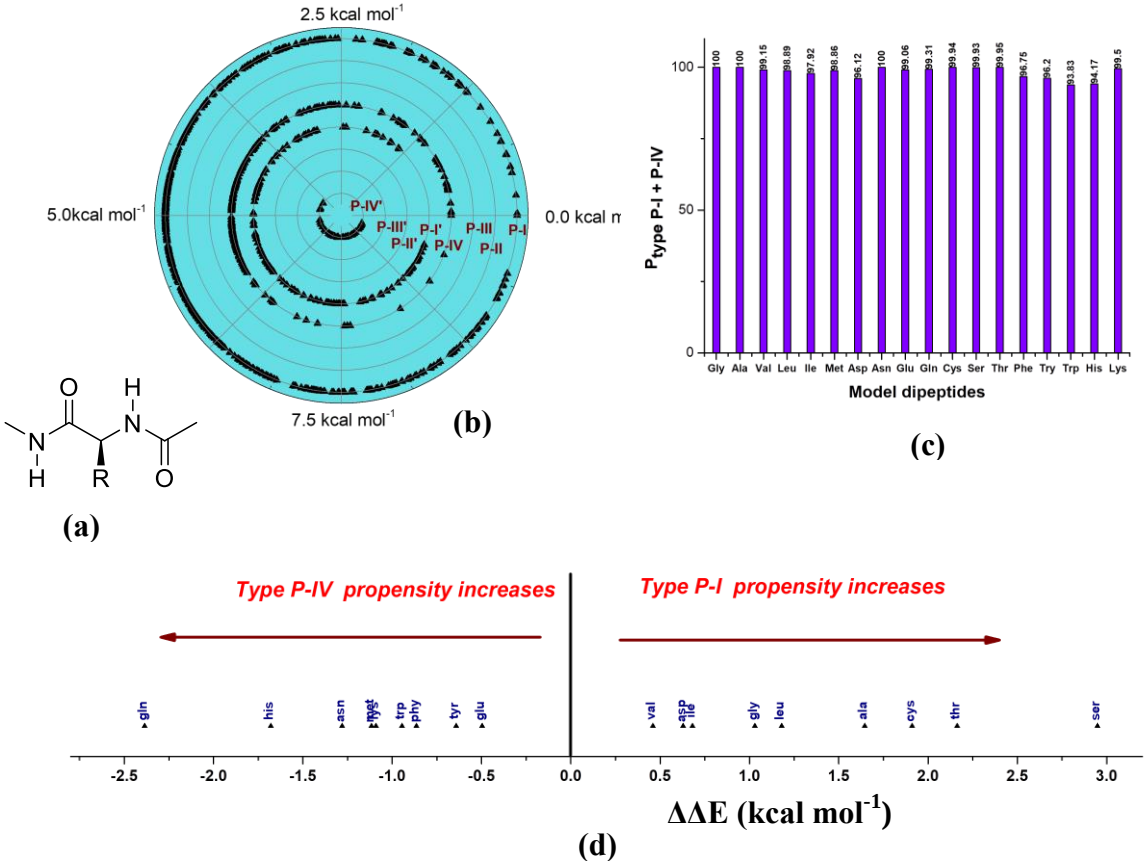


Fig. S5 (a) Chemical structure of model dipeptide; (b) Conformational dartboard of model dipeptides; (c) The sum of type P-I and P-IV backbone propensities of model dipeptdies; (d) ΔΔE scale demonstrating the relative preference for backbone structure type P-I and P-IV after peptide bond formation.

Based on the relative propensity of backbone type II and I (P_{II} - P_I), average stability ($\Delta E_{II - I}$) of backbone type II (relative to type I) was calculated. Similarly, for model dipeptide, the average stability ($\Delta E_{P\text{-}IV - P\text{-}I}$) of backbone P-IV (relative to type P-I) was computed based on the relative propensity of backbone type IV and I ($P_{P\text{-}IV}$ - $P_{P\text{-}I}$). After that, a ΔΔE, i.e. ($\Delta E_{P\text{-}IV - P\text{-}I}$) - ($\Delta E_{II - I}$) stability scale, as shown in Fig. S5(d), was used to demonstrate the relative tendency of any given amino acid to adopt backbone type P-I or P-IV, after peptide bond formation. The amino acid falling on the positive side of the ΔΔE scale shows a relatively higher tendency for backbone type P-I after peptide bond formation and vice-versa.

The backbone type I and II of amino acids were found to adopt the Ramachandran space(φ,ψ) similar to helix and strand. Likewise, the (φ,ψ) area adopted by backbone type P-I and P-IV of model dipeptides is found identical to helix and strand, respectively. Therefore, the correlation between the conformational preference of amino acids and model dipeptides was also rationalized based on two vital secondary structure of the protein, i.e., helix and strand. The rationalization thus helped us in understanding the relationship between the structural preference of amino acids both before and after peptide bond formation. Furthermore, it also suggests that the intrinsic conformational propensities of amino acid residues originate from the amino acid conformer, which exists majorly in two backbone geometries.

S.8 Conclusions

The thesis presents an in-depth study of conformational preference of amino acids and the implication of limited structural preference of amino acid on the structure of peptides. The IR spectra of matrix isolated L-threonine, L-glutamic acid, and L-methionine shows the presence of more than one conformer trapped in experiments. Based on the spectral shift observed both in experiments and computations, the backbone type I conformer were found to show an

intramolecular hydrogen bonding interactions. However, the backbone type II structures, which were suggested to be stabilized by bifurcated hydrogen bonding interactions, does not show hydrogen bonding interaction. The NBO deletion study suggests that vicinal delocalization interactions stabilize backbone type II structures.

The low energy conformers of L-threonine, L-glutamic acid, and L-methionine observed in experiments and computation were found to adopt backbone only type I and II. The observation was pictorially represented in the form of a conformational dartboard. The *ab initio* conformational studies of other α-amino acids demonstrate the existence of generality in the above trend. The low energy conformers of α-amino acids prefer backbone type I or II, irrespective of the chemical nature of the side-chain group it possesses. A quantitative analysis of conformer population indicates that ~80 % of the amino acid populations remain as backbone type I and II structures. The effect of this limited structure on the structure of peptides was rationalized based on the comparative study of model dipeptides.

In summary, the thesis presents a detailed study of conformation landscape of α-amino acids, together with the factors defining conformational stability. The preference for two backbone structure was used to rationalize the bridging link to understand the origin of intrinsic propensity the structural preferences in peptides.

Chapter 1

Introduction

1.1 Preface

Proteins are biopolymers, comprising of long chains of amino acid residues, and perform several functions in the living organisms, such as DNA replication, metabolic reaction catalysis, molecular transport, and several others. Theoretically, a protein can yield a large number of conformations. However, only a few of the conformation predominates under biological conditions.[1-6] The structures of proteins have been rationalized into four distinct levels, i.e., primary, secondary, tertiary, and quaternary. The primary structure of protein represents its amino acids composition in a polypeptide chain, while the secondary structure of a protein is the highly regular sub-local structure of a polypeptide chain. The two major types of the secondary structures of proteins are α-helix and β-strand or β-sheet, which are stabilized by hydrogen bonding network between the peptide groups.[7] The secondary structures folded into compact three-dimensional form resulting in the tertiary structure of the protein. The quaternary structure is the three-dimensional form, containing two or more polypeptide chains, which are stabilized by non-covalent interactions and disulfide linkages.

The physical process by which a polypeptide chain adopts its native and functional three-dimensional structure is called protein folding and is a very complex reaction.[5] Various efforts to understand the protein folding event is based on the Anfinsen's theory, which suggests that the primary amino acid sequence governs the native form of proteins and polypeptides.[8] The relationship between amino acid sequence and secondary structure of a protein is elucidated using the idea of conformational propensity, which is the quantitative measure of the occurrence of a given amino acid residue in the different secondary structure regions of proteins, i.e., helix,

strand or turn.[9] This tendency of amino acid residue to attain specific secondary structure depends on the intrinsic conformational propensity of amino acid residue and the environmental effects (effect of neighboring residue). Therefore, the conformational propensities, estimated for amino acid residue vary from one protein to others, and, thus, are context dependent.[10-12] Moreover, the conformational propensity of amino acid residue estimated in the native form of protein does not replicate the unfolded structure of the protein.

The unfolded or denatured structures of proteins are the initial states for thermodynamic and mechanistic studies of the protein folding process. The classical understanding of the denatured state of protein denotes that the entire amino acid residue, except glycine and proline, nearly uniformly occurs in the sterically allowed regions of Ramachandran plot, i.e., extended region and right-handed helical region.[13,14] The perspective of maximum entropy of denatured structure suggests that the unfolded protein in water do not show any residual structure and therefore can be considered as a random coil.[15,16] However, Tiffany and co-workers, using far-UV circular dichroism studies, suggested that denatured proteins are not entirely random, but exhibit some local order.[17,18] Likewise, several other pieces of evidence indicated the presence of a certain residual structure in the unfolded form, which is governed by the amino acid composition of the unfolded state.[19-25] Therefore, for comprehending the denatured state of protein or peptide, the knowledge of intrinsic conformational propensity of amino acid residues is essential.

The intrinsic conformational propensity is the quantitative estimate of the occurrence of a concerned amino acid residue in the helical, sheet, or turn region of Ramachandran plot, in the absence of environmental influence. In the past, the intrinsic conformational propensity of amino acid residue has been extensively studied employing various experimental and computational

methods. Peptides of varying length such as pentapeptides, tripeptides or dipeptides have been used to estimate the preference of amino acid residues to attain specific regions of the Ramachandran Plot.[26-36] The physical origin of the intrinsic conformational preferences of peptides or the tendency to adopt only certain φ,ψ space, is attributed to the steric factor[37,38] and is still a matter of debate. In the recent past, Vymatal *et al.*[36] reported intrinsic conformational propensity of amino acid residue employing force field computations, and suggested: "The physical origin of this intrinsic conformational propensity remains unclear, and it is questionable whether it can be investigated by computer simulation using the current methods and model."

The free amino acids are chemically different from the corresponding amino acid residues in the peptide chain only by the absence of peptide bonds. Exploring the conformational landscape of free amino acids and dipeptides and rationalizing the structural similarity, if any, can serve as a potential link to understand the origin of the intrinsic preferences of peptides.

1.2 Conformations of α-Amino Acid

Amino acids are class of biomolecules that contains both carboxyl (COOH) as well as amine group (NH_2) groups, along with the side-chain of different chemical nature and bulkiness. The most common form of amino acids found in nature is α-amino acid, having L-isomeric form. Except for glycine, all other α-amino acids exist in two-enantiomeric form, i.e., D and L amino acids, where L-form constitutes all the proteins during the translation in the ribosome. The D-forms are an abundant component of the bacterial cell wall and some proteins produced after post-translational modifications.[39,40] Amino acids are classified into four categories viz. non-polar, polar, acidic, and basic, based on the chemical nature of the side-chain group. Due to the presence of both acidic and amine groups, amino acids are amphoteric. In the solution phase,

amino acids can exist in negative, positive, or Zwitter-ionic form, depending on the pH of the solution. However, in gas-phase, amino acids remain in neutral or canonical form.

The conformational landscape of amino acids is observed to be complex as compared with other organic molecules, because of the presence of multiple freely rotatable single bonds. The simplest amino acid, Glycine, is a three-rotor molecule, whereas, L-lysine, which contain butyl-amine side-chain group is an eight-rotor molecule. The side-chain groups with varying chemical nature can form intramolecular non-covalent interactions such as hydrogen bonding with COOH and NH_2 moieties present in the backbone. Furthermore, the amino acids containing aromatic side-chain groups can show the interaction between aromatic pi electron cloud and backbone COOH and NH_2 moieties. Apart from the above possible interactions, factors such as orbital delocalization interactions can also play a vital role in determining the conformational landscape of α-amino acids.

Conformations of α-amino acids in the solid and solution phase have been widely studied in the literature using different experimental and computational techniques.[41-50] In the solid and solution phase, amino acids exist in charged or Zwitter-ionic form, and the interactions between these charged species govern the conformational stability. Furthermore, the intermolecular interactions present in solid and solution phase also play a crucial role in defining the conformational preferences. In contrast, to condense phase (solid and solution), amino acids in the gas phase are present in neutral or canonical form and are generally free from the intermolecular interactions. Therefore, the conformational studies in gas-phase present the inherent molecular structure of amino acids. Moreover, the gas phase studies appear to be quite appealing compared to the analysis carried out in the solution phase, as the spectral congestion in the gas phase is very minimal relative to that in the condensed phase.

The conformations of α-amino acids in the gas phase have been extensively studied, in the past, employing different experimental methods, such as double resonance spectroscopy, microwave spectroscopy, and infrared spectroscopy. The solid amino acids were evaporated using a resistive heating method or laser ablation method to attain optimal sample concentrations. The experimental results in several of these studies were also corroborated by the *ab initio* quantum computational results. Alonso and co-workers have studied the conformations of several α-amino acids, using molecular beam Fourier transform microwave spectroscopy and *ab initio* computations.[51-65] Research works discussing the conformational landscape of various α-amino acids by *ab initio* computations, at different levels of electronic structure calculations, have also been reported. [66-77]

Infrared spectroscopy coupled with matrix isolation techniques traps the molecule of interest in a matrix of inert gas at very low temperature, and hence resulting in sample molecules free from intermolecular interaction, and thus, sharp spectral features are observed. The spectra recorded are also free from rotational congestion, as the sample molecules are trapped in a rigid-inert gas cage, thus prohibiting any rotational motion. As discussed earlier, the conformations of α-amino acids can manifest hydrogen binding interaction, involving backbone and side-chain moieties. For a molecular system, where the conformers are believed to be stabilized by such non-covalent interactions, the matrix isolation infrared spectroscopy can provide unequivocal evidence for the presence of such interactions, through shifts in the vibrational frequencies. The matrix isolation IR spectroscopy has also been widely used in the past to study the conformations of several α-amino acids such as glycine, alanine, phenylalanine, serine, valine, proline, isoleucine, leucine, cysteine, and tyrosine.[78-92] The conformers of various α-amino acids were

suggested to be stabilized by intramolecular hydrogen bonding interactions involving backbone and side-chain moieties.

1.3 Motivation of Thesis

The primary objectives of the research work presented in this thesis are three-fold, which are as follow

a) Exploration of the conformational landscape of α-amino acids using the matrix isolation infrared spectroscopy and *ab initio* quantum computation methods. Earlier works of conformational studies are reported only for some amino acids, and no general rationalization of the conformational picture has been presented. Here in this thesis, an effort to rationalize the structures of all 20 naturally occurring α-amino acids is presented.

b) Apart from presenting the conformational landscape of α-amino acids, we also address the factors that govern the conformational preferences. Presence of non-covalent interactions such as hydrogen bonding, in different conformers of α-amino acids, was examined based on shift observed in experimental spectra and atoms in molecule (AIM) analysis. The role of orbital delocalization interactions was also explored by natural bond orbital (NBO) analysis. As follows earlier, the side-chain groups of various α-amino acids can participate in hydrogen bonding interactions with the backbone moieties, and thus can influence the conformational preferences of concerned amino acids. Therefore, the roles of side-chain groups were also examined.

c) To understand the relationship between the preferred conformers of α-amino acids, with that of dipeptides, the conformational studies of dipeptides were also conducted, using electronic structure calculations. The association between conformers of α-amino acids and dipeptides were rationalized based on the geometrical similarity of the conformers of both species. Furthermore, the relationship was also understood based on the Ramachandran region (φ, ψ space) adopted by

the conformers of both the species, i.e., α-amino acids and dipeptides. The conformer preferences of dipeptides are also referred to as intrinsic conformational propensity, and therefore, examining the relationship between conformers of α-amino acids and those of dipeptides, can serve as a vital link for understanding the physical origin of intrinsic conformational propensities in dipeptides.

1.4 Conformational Landscape of L-Threonine

The conformations of polar amino acid L-threonine was studied both experimentally and computationally. The 38 conformers of L-threonine, obtained after scanning the potential energy surface (PES), at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theories, were classified into different classes based on the backbone geometry of conformers. The vibrational assignment of matrix isolated L-threonine was found to be in good agreement with the computed spectra of four lowest energy conformers viz. $THR1_{II}$, $THR2_{I}$, $THR3_{I}$, and $THR4_{III}$. The roman subscripts in the nomenclature of conformers represent the backbone structure of conformer. The low energy conformers observed in matrix isolation IR study were the same as those reported in microwave study, by Alonso and co-workers. The stability of backbone type II conformer ($THR1_{II}$) was attributed to the presence of bifurcated hydrogen bonds, involving C=O (backbone COOH group) and NH_2 group, whereas that of backbone type I conformer ($THR2_{I}$) was suggested to be due to the hydrogen bond between O-H (COOH group) and NH_2 group. The presences of such non-covalent interactions in different conformers were verified experimentally as well as computationally. The role of the polar side-chain of L-threonine in defining the conformational preferences was also investigated. The structure and energetics of all conformers were presented in a dartboard fashion, which we called as a conformational dartboard. The

association between structures and energetics of low energy conformers were examined to understand the preference of low energy conformers to adopt only specific geometries.

1.5 Conformational Picture of L-Glutamic Acid

To understand the role of the acidic side-chain and intramolecular hydrogen bonding interactions, on the conformational landscape of α-amino acid, we performed matrix isolation study of L-glutamic acid. The potential energy surface scan exercise yielded 267 conformers of L-glutamic acid at MP2/6-311++G(d,p) level of theory, and the conformers were classified based on their backbone geometries. The assignment between experimental and computed wavenumbers suggested the presence of three low energy conformers in our experiments viz. $GLU1_{II}$, $GLU2_{I}$, $GLU3_{II}$. The lower energy conformers of L-glutamic show the same backbone structure, as observed in the case of L-threonine. The presence of hydrogen bonding interactions in backbone type II and type I conformers of L-glutamic acid was investigated both experimentally and computationally. The importance of the acidic side-chain of L-glutamic acid in defining the conformational preferences was also explored. The association between the structure and energetics of various conformers was rationalized based on the conformational dartboard representation, the same as that done in case of L-threonine.

1.6 Conformations of L-Methionine

After rationalizing the role of the polar and acidic side-chain in the conformational preferences of amino acid, the effect of the non-polar side-chain in defining the conformational preferences of amino acid was studied by exploring the conformational landscape of L-methionine. The 165 conformers of L-methionine, obtained after scanning the potential energy surface, at MP2/6-311++G(d,p) level of theory, were classified based on the method used for L-threonine and L-glutamic acid. The IR spectra of matrix isolated L-methionine were found to be in good

agreement with the computed wavenumbers of the three lowest energy conformers viz. $MET1_{II}$, $MET2_{I,}$ and $MET3_{III}$. The lowest energy conformers of L-methionine show same backbone structure (backbone type II) as that of L-threonine and L-glutamic acid, and similarly, the 2nd lowest energy conformers also share common backbone structure relative to that observed in previous two amino acids. The presence of hydrogen bonding interactions in different conformers and the role of the non-polar side-chain in the conformational preferences were also analyzed. Like, L-threonine and L-glutamic acid, the conformers of L-methionine were also represented together in the form of the conformational dartboard, to comprehend the relation between different conformers.

1.7 α-Amino Acids, Tale of Two Structures

The conformers of the remaining α-amino acids were studies computationally, and the potential energy surface was scanned to obtain possible conformers. Matrix isolation IR studies of a few other amino acids were also carried out, but due to sample decomposition during the deposition, the vibrational assignments were not performed. The conformers of various α-amino acids, obtained in PES scan, were classified using the method discussed in the above section. All the conformers of α-amino acids were represented together in the form of the conformational dartboard. The relation between the low energy conformers of α-amino acids was rationalized based on their structure and energy. Based on the relative abundance of conformers and population of backbone types, an estimate of quantitative preferences of a given conformer and backbone type in various α-amino acids was established. The quantitative analysis suggested that on an average ~81 % of the total population of α-amino acids, at room temperature, existed in two backbone structures, i.e., type I and II. The restricted conformational preferences observed in α-amino acids can be vital for understanding the intrinsic conformational propensity of

dipeptides. The relative preferences between two preferred backbone structures, i.e., type I and II in different α-amino acids were also rationalized. The preferred orientations of side-chain groups in conformers of α-amino acids were also explored.

1.8 What is the Role of the Amino Acid Conformational Dichotomy on the Conformational Landscape of Dipeptides?

The preference of only two backbone types as lower energy conformers in α-amino acids can have crucial importance in the structural preferences of dipeptides. The conformational preferences or intrinsic conformation propensity of peptides have been studied in the past. The conformations of natural dipeptides, which have two amino acids bonded together by peptide bonds and terminal COOH (C- terminal) and NH_2 (C- terminal) groups present, have attracted much attention since last one-two decades. The C and N-terminal functional groups play a significant role in defining the conformational landscape of such dipeptides. However, regarding the intrinsic conformational propensity of peptides, the effect of terminal groups has to be avoided, as, in a regular peptide chain; these terminal groups have little influence on the conformations of central residues. For this purpose, molecules resembling the central segments of peptides are considered, i.e., capped peptides or model peptides, chemically protected at both the terminals. Model peptides of varying length have been employed to investigate the intrinsic conformational propensity.

In this thesis, the conformational landscape and intrinsic conformational preferences of capped dipeptides were studied, using *ab initio* quantum mechanical computational studies. The conformers of model dipeptides were obtained by scanning the potential energy surface and were classified into different classes, based on their backbone structure. The classification method employed for dipeptides conformers were equivalent to that used in amino acids. The

conformational preferences of dipeptides were analyzed based on the structure and energetics of conformers, presented in conformational dartboard form. The quantitative analysis, the same as that used for amino acids, showed that ~98% of the total population of dipeptides existed in two backbone structures.

The preferred backbone structures of dipeptides were geometrically equivalent to that of amino acids, and the relationship between these geometrically similar backbone structures was explored. Furthermore, the association between amino acids and dipeptides conformers was also analyzed based on the Ramachandran regions (φ, ψ space) adopted by these conformers.

In summary, the thesis presents a detailed and systematic study of conformers of amino acids and dipeptides. The factors determining the conformational preferences and the role of the side-chain were also explored. The relationship between the structural choices of amino acids and dipeptides was carried out to understand any existing variation in the conformational preferences going from amino acids to dipeptides form. The important issue of the physical origin of intrinsic conformational propensity or conformational preference of dipeptides is also discussed.

Chapter 2

Experimental and Computation Methods

2.1 Introduction

The technique of matrix isolation was first introduced by Pimental and co-workers in 1954, primarily to study reactive species by trapping them in a solid matrix of inert gas.[93-98] The method has since then been widely used to study reactive species, exotic molecules, reaction intermediates, molecular conformations, and non-covalent interactions. In this technique, the sample of interest is first diluted with a large excess of an inert gas such as Ar or N_2 and less frequently with Ne, Kr and Xe and then deposited as a solid film, on a cold substrate, maintained at ~10 K. When Ne is used as a matrix a temperature of ~4 K is used. Due to the dilution of the sample with a large excess of inert gas, the sample molecules are predominantly isolated from each other and are surrounded by the inert gas atoms. This situation minimizes intermolecular interactions between the sample molecules. Fig. 2.1, depicts a cartoon representation of the analyte molecule trapped in such a cryogenic inert matrix.[99,100] For comparison, the situation prevailing in a solid or liquid phase sample is also shown, where intermolecular interactions can be seen to be present. The molecules trapped in the inert gas matrix are studied using spectroscopic tools such as infrared, UV-visible [101], or ESR.[102]

The technique was primarily developed to study short-lived reactive species, which are challenging to study owing to their very brief lifetime. However, once frozen in the matrix of inert gas, these reactive intermediates do not undergo reactive collision, which results in an increased lifetime; thus enabling one to probe these species. The possibilities of these reactive species engaging in unimolecular reactions are also prohibited by the activation barriers, which are difficult to surmount at the low temperature (~10K) of the matrix.

The experimental procedure involves mixing of the sample molecules of interest with a large excess of inert gas, typical matrix gas to sample ratios being 1000:1 or higher. The matrix gas/sample mixture is then passed through a nozzle, into the vacuum chamber housing a cryostat. The pressure in the vacuum chamber is typically maintained at 1 x 10^{-6} mbar. A KBr substrate is mounted on the cryotip onto which the matrix gas/sample mixture is deposited to form the thin homogeneous film. After a sufficient amount of the sample is deposited, the sample molecules are probed using different spectroscopic methods. The cold substrate is usually maintained at ~12 K, (or ~4K if Ne is used as the matrix). At these temperatures, the sample mixture deposited on the cryotip is immobilized, thus resulting in the significant minimization of Doppler and collision broadening. The low cryogenic temperature in the matrix also results in a substantial reduction of spectral congestion by ensuring that only the lowest rovibrational states of the analyte are populated.

2.2 Advantages of the Matrix Isolation Technique

The high dilutions of the sample in the inert gas results in the analyte molecules to be surrounded only by inert gas atoms and predominantly isolated from other analyte molecules. This picture is represented in the form of the cartoon in Fig. 2.1. The spectral feature in matrix isolated spectra shows line width of typically 2-5 cm^{-1} as compared with the 40-50 cm^{-1} line widths seen in condensed phase or gas phase spectra at room temperature. The uncongested spectral features, the small line widths, absent of Doppler and collision broadening and the isolation allows this technique to be used for the studying molecular conformations, weak non-covalent interaction, and reactive species.

The matrix isolation setup involves a combination of several interdependent techniques, to produce the required rigid matrix sample film at low temperature. The cryogenic technology

provides a low temperature of 10 K or 4 K, which is essential for the rigid inert gas matrix. High vacuum conditions are required for efficient working of the cryogenic system and also for inert gas matrix film to be prepared with a negligible concentration of impurities. The thin films formed on the cold substrate under high vacuum condition are probed employing the different spectroscopic technique.

The matrix isolated species can be probed and characterized by employing various spectroscopic methods such as infrared, UV/Visible,[101] fluorescence[103], and ESR spectroscopy[102]. Also, other spectroscopic techniques such as Raman, circular dichroism are also gaining application.[104] The material to be used as a substrate depends upon the probe used for analysis. For mid-infrared spectroscopy ranging from 4000-400 cm^{-1}, KBr or NaCl windows are used, whereas CsBr or CsI windows are used for the study in the far IR region, i.e., below 400 cm^{-1}. For fluorescence and UV/visible spectroscopy, quartz window are employed, and sapphire tipped copper rods are used as matrix materials for ESR.

The choice of material used as matrix gas determines the success of the technique. A chemically inert material is preferred as the matrix to avoid any reaction between the trapped species and matrix molecules. The matrix must be highly pure and should be transparent in the spectral absorption range of trapped species, to avoid any spectral interference. To prevent analyte diffusion in the matrix, the rigidity of the matrix at the temperatures of the matrix is essential. To ensure the rigidity of the matrix, the matrix should be kept at a temperature which is at least 30% below the melting point of the matrix material used. The materials which fulfill the required conditions of chemical inertness and spectral transparency in the working region are rare gases such as Ar, Kr, Ne, Xe and the inert gas such as N_2, which are therefore commonly used as matrix materials. For measurement using fluorescence spectroscopy, alkanes are used as

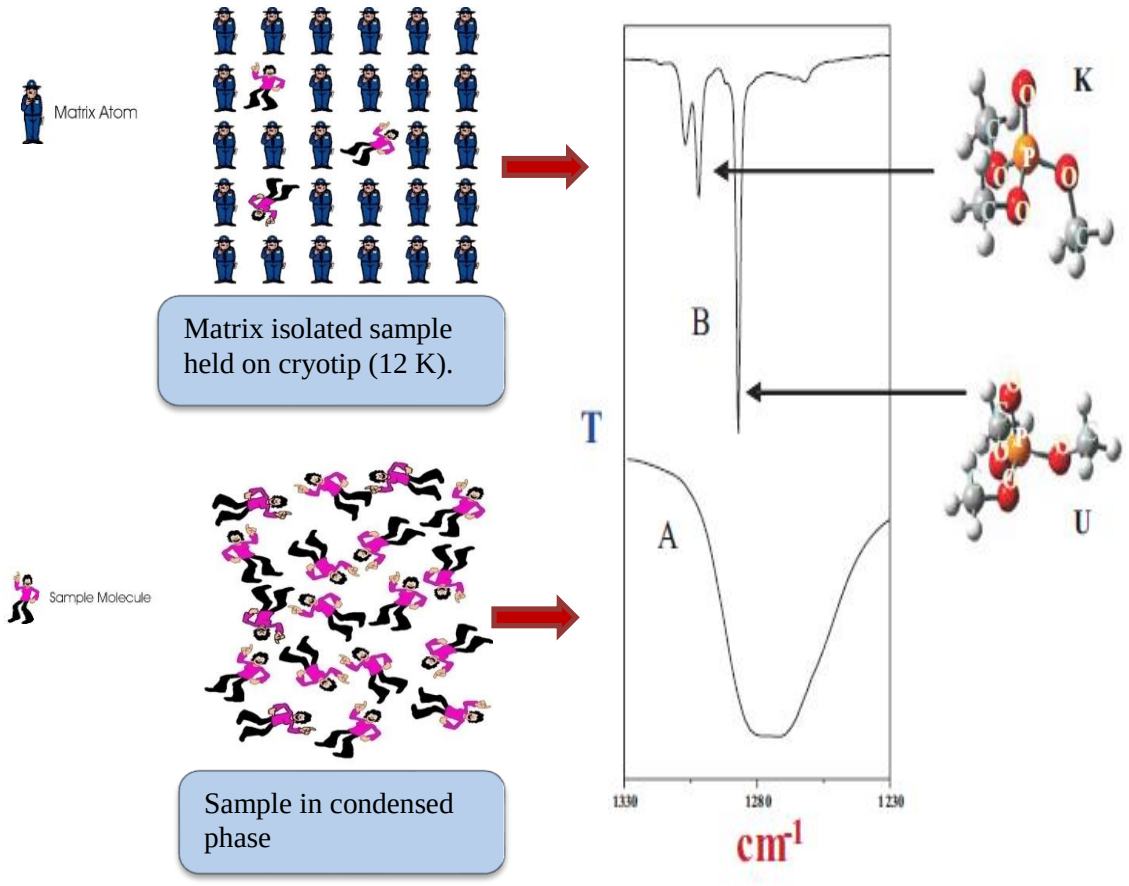


Fig.2.1 Matrix isolation; The procedure and the outcome.

Figure source: Viswanathan and co-workers, *J.Ind. Inst. Sci. 85* 2005, 403; Viswanathan and co-workers, J. *Phys. Chem. A*,.102, 1998, 2944

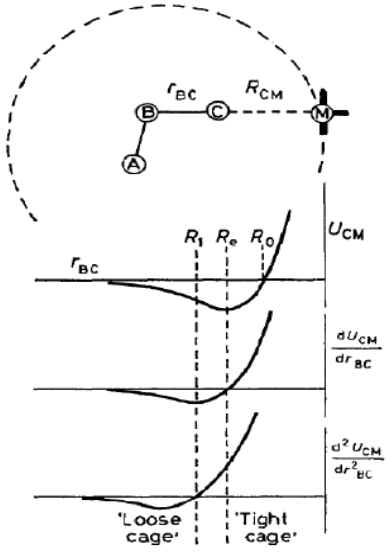


Fig.2.2 Plot showing the dependence of U, U' and U" on matrix cage size

a matrix material and are referred to as Shpolskii materials. Other than the above-discussed materials, CH_4, O_2, SF_6, CO, CO_2, etc. have also been used as matrix materials; however, all these materials show disadvantages such as chemical reactivity and spectral interference in the range of interest.

2.3 Challenges in the Matrix Isolation Technique

For a sample to be studied using matrix isolation technique, a sufficient vapor pressure at easily attainable temperature is needed, so that the sample can be easily deposited on the cold substrate and detectable signal can be recorded. A sample having negligible vapor pressure is difficult to be deposited in sufficient amount at room temperature condition, and therefore, a heating furnace is required to raise the vapor pressure of the sample. However, the use of high temperature is associated with the risk of sample decomposition, and thus non-volatile sample such as amino acids pose problems when studied using the matrix isolation technique.

2.4 Solvation Effects of the Matrix

In reality, there does not exist a perfectly inert matrix; the assumption of a "non-interacting matrix" is therefore practically unrealizable. In actual practice, the matrix and the species isolated in the matrix are not entirely devoid of any intermolecular interaction. While this interaction is minimal, it certainly cannot be ignored, and this aspect will be addressed in the following section.

2.4 (a) Matrix shifts

The effect of matrix gases interaction with the sample molecules can be observed by comparing the variation seen in the vibrational frequencies of a molecule trapped in the matrix, with that of gas phase frequencies. These frequency shifts of sample molecules reflect the extent of

perturbation on the vibrational levels of the trapped species by matrix molecules. The magnitude of this interaction can be estimated from the shift in the vibrational frequency ($\Delta\nu$) of the trapped species relative to the gas phase value.

The interaction of matrix gases with the trapped species results from various factors, such as electrostatic ($\Delta\nu_{elec}$), inductive($\Delta\nu_{ind}$), repulsive($\Delta\nu_{rep}$) and dispersion($\Delta\nu_{disp}$) interactions. The interaction of rare-gas matrix molecules with the trapped species is majorly driven by long-range London dispersion($\Delta\nu_{disp}$) and short-range repulsive forces($\Delta\nu_{rep}$). Sample molecule trapped in the matrix can show a red or blue shift with respect to gas phase frequencies, and the observation was explained by Pimental and Charles, using the Buckingham equation.[105] The equation predicts the effect of matrix cage cavity size on the direction of frequency shift, i.e., blue or red.

$$\mathbf{(\Delta\nu) = (\nu_{matrix} - \nu_{gas}) = (B_e/hc\omega_e)(U'' - 3AU'/\omega_e)}$$

B_e = is the rotational constant ($h/8\pi^2 c\mu r_e$)
a = anharmonicity constant,
U = Energy due to solute-solvent interaction,
U' = $\{\partial U/\partial r_{BC}\}$ and U"= $\{\partial^2 U/\partial^2 r_{BC}\}$, r_{BC} = equilibrium distance between atom B and C of the molecule ABC as shown in Fig.2.2
ω_e = Harmonic oscillator frequency

The effect of matrix cage size on U, U' and U" is shown in Fig. 2.2. For a molecule ABC placed in a tight cage, the distance between matrix atom M and atom C, i.e., R_{CM} will be smaller than R_e, equilibrium bond distance between atom B and C and therefore U' and U" will be positive. Such tight cage interaction will result in a positive value of $\Delta\nu$, and hence, a blue shift will be observed. A loose cage would imply R_{CM} value larger than R_e, i.e., host atom placed at a larger distance from matrix atoms, and thus resultant negative $\Delta\nu$ and hence redshift will be observed. For polyatomic molecules, the high-frequency modes such as stretching vibration mimic loose cage effect owing to the small vibrational amplitudes and hence show red shift upon

matrix perturbation. The low-frequency modes such as bending and rocking show large amplitude motion and thus depict a tight cage condition in a matrix environment, and hence a blue shift in frequency is observed.

Thus the blue shift observed for low-frequency modes such as bending and rocking modes of polyatomic molecules and red shift seen for high-frequency modes is rationalized in term of matrix perturbation and cage effect on the normal modes of the molecule. Therefore it is important to recognize the matrix effect to infer the spectra of matrix isolated species accurately.

2.4 (b) Multiple trapping sites

The rare and inert matrix gases upon solidification on cold substrate adopt either hexagonal close-packed (h.c.p.) or face-centered cubic crystal (f.c.c.) structure, depending upon the gas used. The analyte molecules deposited on the cold substrate along with matrix gas, get trapped in substantial holes and less frequently in interstitial sites. Thus the sample molecules, trapped in different sites, experience matrix environments that are different. This phenomenon possibly induces variable shifts, which result in either spectral broadening or splitting of single vibrational feature into multiplet structure referred to as site effects. The site effect can be identified by depositing the sample molecule in the different matrix as it is highly unlikely that the two different matrix gases will result in exactly similar sites. The intensity of spectral features at different sites will be determined by the stability of each trapping site and annealing the matrix at elevated temperature will result in the unstable sites to disappear, thus sharpening the observed spectral features. The problem of the effect of multiple sites on the spectral features can be studied by varying the distribution of different sites, which can be achieved by altering the rate of matrix-sample deposition.

2.4 (c) Molecular rotation in the trapping sites

The rigid and unsymmetrical matrix site, in general, prohibits the rotational motion of immobilized trapped species. However, small molecules such as water, methane, ammonia and alkali halides, show rotations in some matrices. The rotational lines of trapped species can be identified by temperature cycling, which results in reversible intensity changes. The reversible intensity variation arises due to change in the rotational level of the population during the temperature cycling process.

2.4 (d) Aggregations

Isolation of sample molecules in the matrix isolation method can be assured only at large dilution condition, i.e., matrix/analyte (M/A) ratio greater than 1000:1. At lower dilution, complete isolation is not achieved and chances of trapping molecular aggregates such as a dimer, trimer, and multimer along with the monomer increase significantly. Therefore the concentration-dependent experiment should be performed to distinguish the spectral features resulting from the sample aggregates. The probability of isolation of small molecule such as carbon monoxide, which occupies a single substantial site, can be calculated as follows. The probability of interaction is simply defined as the chance of finding another sample molecule occupying one of the 12 nearest neighbor sites of the cage. Likewise, the probability of the absence of the second CO molecule can be found out by the formula $P = (1 - r)^{12}$, where r is the reciprocal of the matrix ratio. For the value of r tending to zero, the expression becomes $P = 1 – 12r$. For a matrix ratio of 1000:1, the P calculated will be 0.988 and thus ensure sample isolation nearly up to 99%. However, the experimental studies of carbon monoxide at (M/A) ratio of 1000:1 shows a higher percentage of dimer and higher aggregate formation as compared to the above-calculated value of 1% only. Similarly, lithium atoms show complete dimerization even with (M/A) ratio of

10,000:1 owing to the fast aggregation kinetics of lithium as compared to the rigidity kinetics of matrix gas. A careful selection of concentration window of the sample without compromising on the intensity of monomer features should be used to eliminate the formation of the aggregates.[106]

2.4 (e) Lifting the degeneracy of vibrational levels

The symmetry of the trapping site determines the matrix effect on the sample molecule. As discussed earlier, the different vibrational modes of trapped molecules experience different perturbations in the different sites. In case of degenerate normal modes of vibration, the asymmetry of trapping site can lift the vibrational degeneracy and hence result in splitting of the degenerate vibrational modes, as observed in CO_2 and $C_2 H_2$ in N_2 matrix.

2.5 Matrix Isolation FTIR Set up

The matrix isolation FTIR method is composed of the following components:

a) Cryogenic assembly

b) High vacuum system

c) Sample introduction

d) FTIR spectrometer

a) Cryogenic assembly

The low temperature of 12 K in matrix isolation studies produced using a closed cycle helium compressor cooled cryostat CH202w/HC4E1 model (Sumitomo Heavy Industries Ltd.) Such cryostats are commercially available and work on the principle of Gifford-McMohan (GM) cycle, with helium as working fluid. The essential components of GM closed cycle cryostat are as follow:

1) Compressor 2) Expander 3) Vacuum Shroud 4) Radiation Shield

The expander, also commonly known as cold head or cold finger, is connected to the compressor via two gas lines for closed-loop supply of helium. One of the gas line supplies compressed helium from the compressor unit to expander while other gas line returns the expanded helium from expander to compressor, for the next cycle. The flow and pressure of compressed helium for GM refrigeration are controlled by the compressor (specification). The heat load on the expander, due to conduction and convection, is limited and controlled by the vacuum shroud. Likewise, the radiation shield reduces the heat load on the second stage of cooling by insulating the second stage from room temperature thermal radiation, emitted from vacuum shroud. Other than the above major component, the cryogenic unit is also accompanied by instrumentation skirt to provide electrical feed-through, vacuum port and a temperature controller to measure and control the cold window temperature. An electronically controlled unit (Lakeshore Instrument; Model 335P) is employed to monitor and maintain the cold window temperature, ranging from 12 K to 300 K. Regulation of cold window temperature is attained by using a heater unit, which is mounted near the cold window and couple with closed loop Proportional-Integral-Derivative (PID) controlled. A silicon diode sensor is also placed near a cold window to measure temperature. Temperature more than 12 K is a prerequisite condition for annealing the matrix, to enhance the diffusion of trapped species, thus resulting in complex formation. Temperature variation based studies are also in understanding the effects of multiple sites on the vibrational features of trapped species. The compressor is cooled by 10 kW water chiller (Werner Finley) that keeps flowing water at approximately 15 ºC. The cryogenic setup, along with its associated components, is shown in Fig. 2.3, and 2.4.

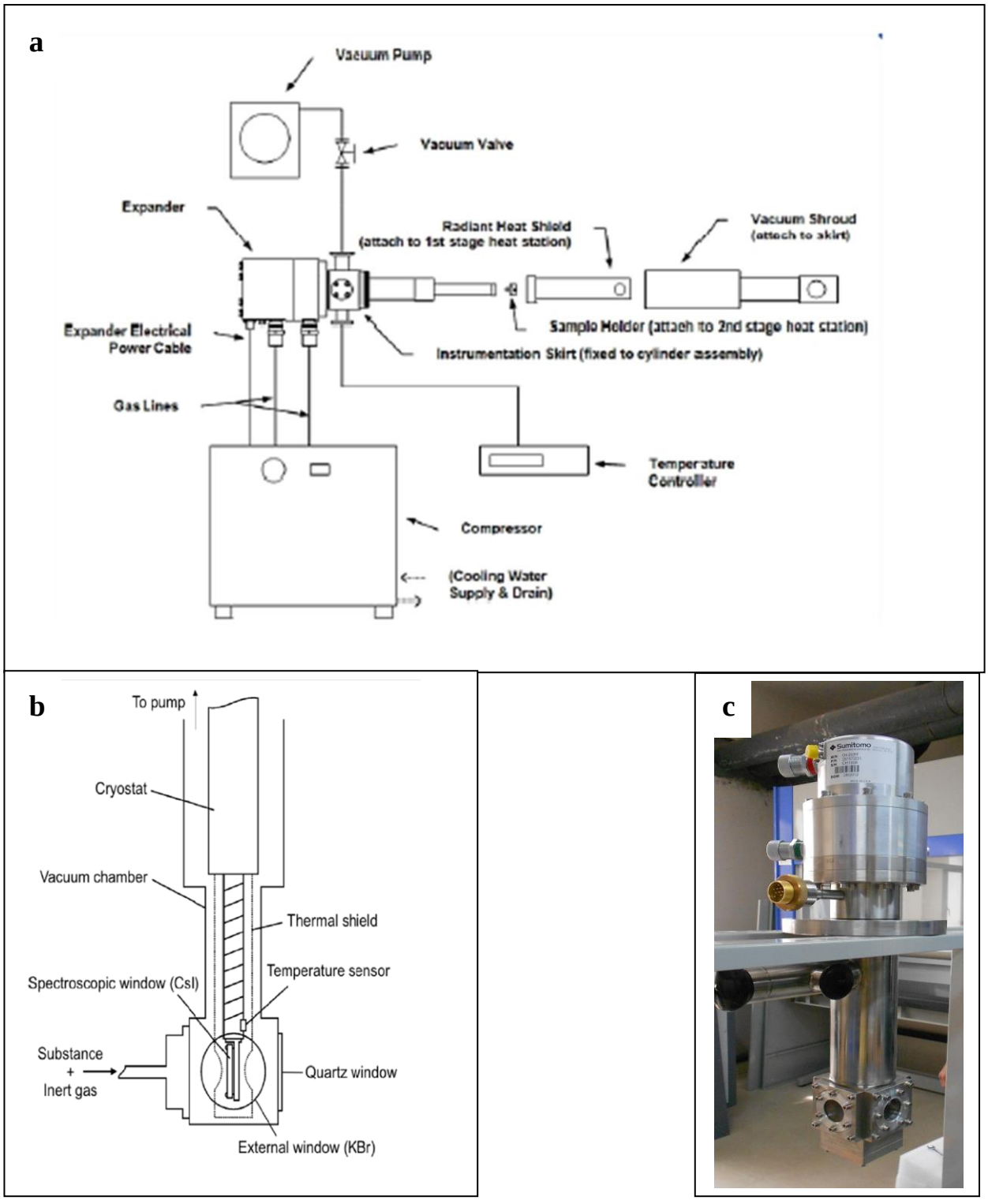


Fig. 2.3 (a) The general schematic diagram for the cryosystem assembly; (b) The schematic for the cryostat head mounted on the FTIR Spectrophotometer; (c) Picture of the cryostat mounted in Matrix Isolation set up at IISER Mohali

(d): Cryostat without radiation shield

(e): Cryostat with radiation shield

(f): Temperature Controller

(g): Helium Hoses

(h): Helium compressor

Fig. 2.4 Cryostat and its associated units

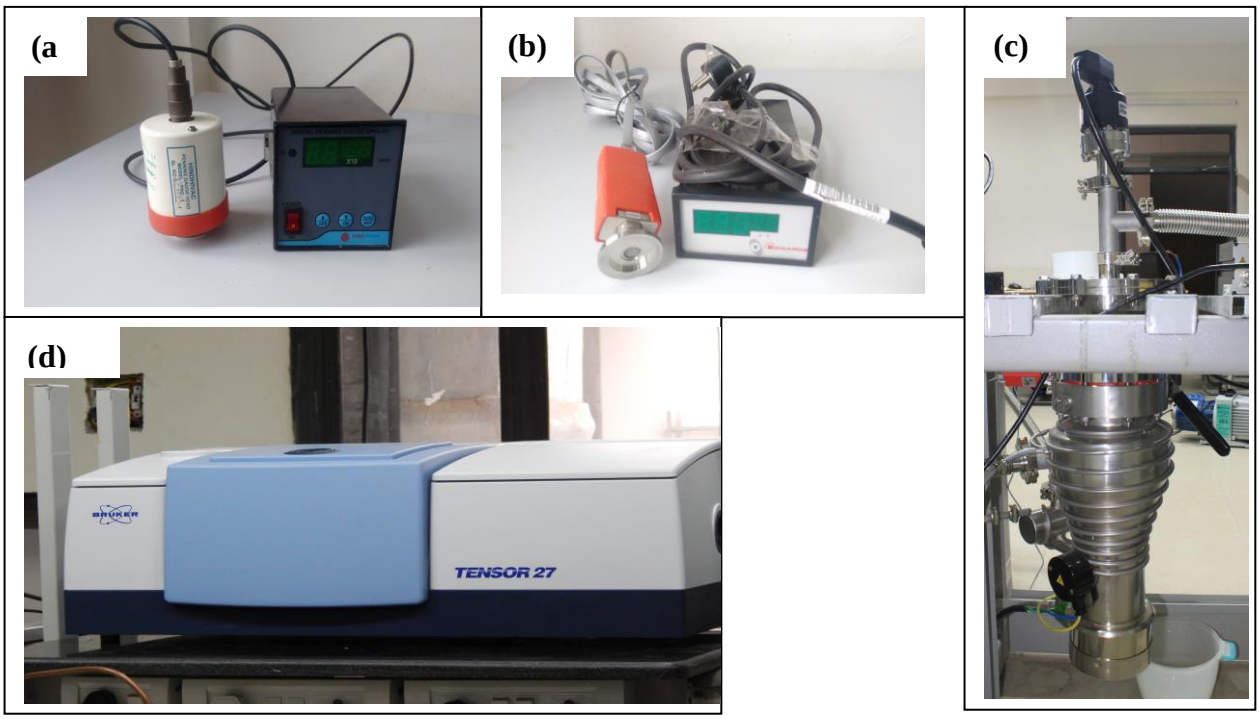


Fig. 2.5 (a) Penning gauge; (b) Pirani gauge; (c) Diffusion pump; (d) FTIR spectrometer

b) Vacuum system

A high vacuum condition (~ 10^{-6} mbar) is an essential prerequisite for matrix isolation method both in terms of supporting the cryogenic setup to attain a low temperature and to ensure that species of interest are trapped without impurities. The high vacuum condition in our experiments is achieved by employing an oil diffusion pump backed with a rotary pump (Hind Hivac; ED6). First a rough vacuum of 10^{-2} mbar is accomplished by mechanical rotary pump operating at pumping speed of 200 L/min, which is then followed by pumping through oil diffusion pump (Edwards, Diffstak MK2 series 100/300) running at a pumping speed of 300 L/sec, to attain a vacuum condition of 10^{-6} mbar or lower. The diffusion pump work on the principle of momentum transfer from the directional oil jet to the gases molecules. The oil vapors from the boiler are first directed upward through a tower to an array of nozzles pointing outward and downward toward the diffusion pump wall. The oil vapors upon passing through these array of nozzles results in the formation of directional jet flow, which in turn transfer its momentum to the gases present in the chamber. The diffusion pump walls are maintained at a temperature of ~15^{o}C, to condense the working fluid vapors. The cold water supply to the diffusion pump walls is maintained by the chiller (Werner Finley). Digital gauges such as Penning gauge (Hind Hivac) and Pirani Gauge 26 (Edwards APG 100 Active Pirani Gauge), both shown in Fig. 2.5 are used to measure the vacuum conditions.

c) Fourier transform infrared spectroscopy

The trapped species were studied using a Bruker-Tensor 27 Fourier Transform Infrared (FTIR) spectrometer (Fig. 2.5d). The instrument was operated at a resolution of 0.5 cm^{-1}. The FTIR machine is provided with silicon carbide (Globar) MIR source, KBr beam splitter, and a DLaTGS detector. The following FT parameters were used in the data acquisition: triangular

apodization function, a zero-filling factor of 8, Mertz phase correction, respectively. The number of scans co-added to obtain spectra with good signal to noise ratio ranged from 8 to 64.

d) Sample introduction

The sample/inert gas mixture is prepared in a mixing chamber of 1 L capacity and made of stainless steel. The chamber contains multiple ports, which can be attached to a sample holder via glass stopcocks. The mixing chamber is also connected to the vacuum chamber housing of cryostat through a ¼ "diameter metal tube (Fig. 2.6 (a) and (b)). The mixing chamber in all experiments is maintained at a high vacuum condition and is pumped by the oil diffusion pump. For the sample with sufficient or high vapor pressure, the sample is allowed into the mixing chamber by maintaining the sample in a temperature bath. The temperature bath keeps the volatile sample at desired vapor pressure and thus controls the amount of sample introduced in the mixing chamber. Following the sample introduction, the mixing chamber is topped up to about 1000 mbar with matrix gas (Ar, N_2), to eventually obtain the required sample dilution condition (M/A ratio). The matrix-analyte mixture is deposited to the cold substrate window, and the rate of deposition is regulated by a needle valve (Model: EVN 116, Pfeiffer Vacuum), shown in Fig. 2.6 (c). The typical rate of deposition of matrix gas in most experiments was from 1.5-3.0 mmol/hr.

The matrix isolation study of non-volatile samples is challenging, owing to the low vapor pressure of the sample. The desired sample concentration for a non-volatile analyte such as amino acids is achieved by increasing the vapor pressure employing a home built heating furnace, placed near the cold substrate window. The heating furnace consists of three major parts, which are shown in Fig. 2.6(d) and is mentioned below

1) Sample holder: A T-shaped stainless steel tube, which was brazed to the inner side of the vacuum flange. The outside of the vacuum flange was connected to the mixing chamber via a metal tube, as discussed previously.

2) Heater: The T-shaped metal tube containing a non-volatile sample was heated by using a thin and insulated nichrome wire (Lakeshore, 0.5mm diameter), which were wound tightly around the metal tube. The current through the heater was controlled by a D.C. power supply (Aplab, 16V-6A). Typical currents used were ~1 A.

3) Sensor: A bead sized thermocouple sensor (company and model) was placed at the tip of the T-shaped tube to measure the temperature of the heated nozzle. The thermocouple wires were coupled to the outside of the vacuum system through a nozzle in the vacuum flange and were connected to a multimeter (specification) for temperature readout. A thin layer of coated copper wire was placed in between the thermocouple wires and the heater wire (nichrome) to avoid circuit shorting. Typical temperatures of the nozzle were between 100 to 150˚ C.

The experiments start with the pumping of the cryostat and mixing chamber to a vacuum of nearly ~1 x 10^{-6} mbar. After the vacuum is achieved, the mixing chamber is isolated from the oil diffusion pump and cryostat chamber through a bellows valve. The compressor is turned on for cooling the cryostat from room temperature to 12 K and mixing chamber is filled with the matrix gas. The matrix gas from the mixing chamber was introduced into the cryostat chamber via a needle valve (Pfeiffer gas dosing valve; EVN 116), connected to an effusive beam nozzle. In all the experiments discussed in this thesis, the amino acids sample was mixed with the matrix gas inside the cryostat chamber. The desired vapor pressure of amino acids was achieved by maintaining the sample containing hot furnace at the required temperature, and the vapor of

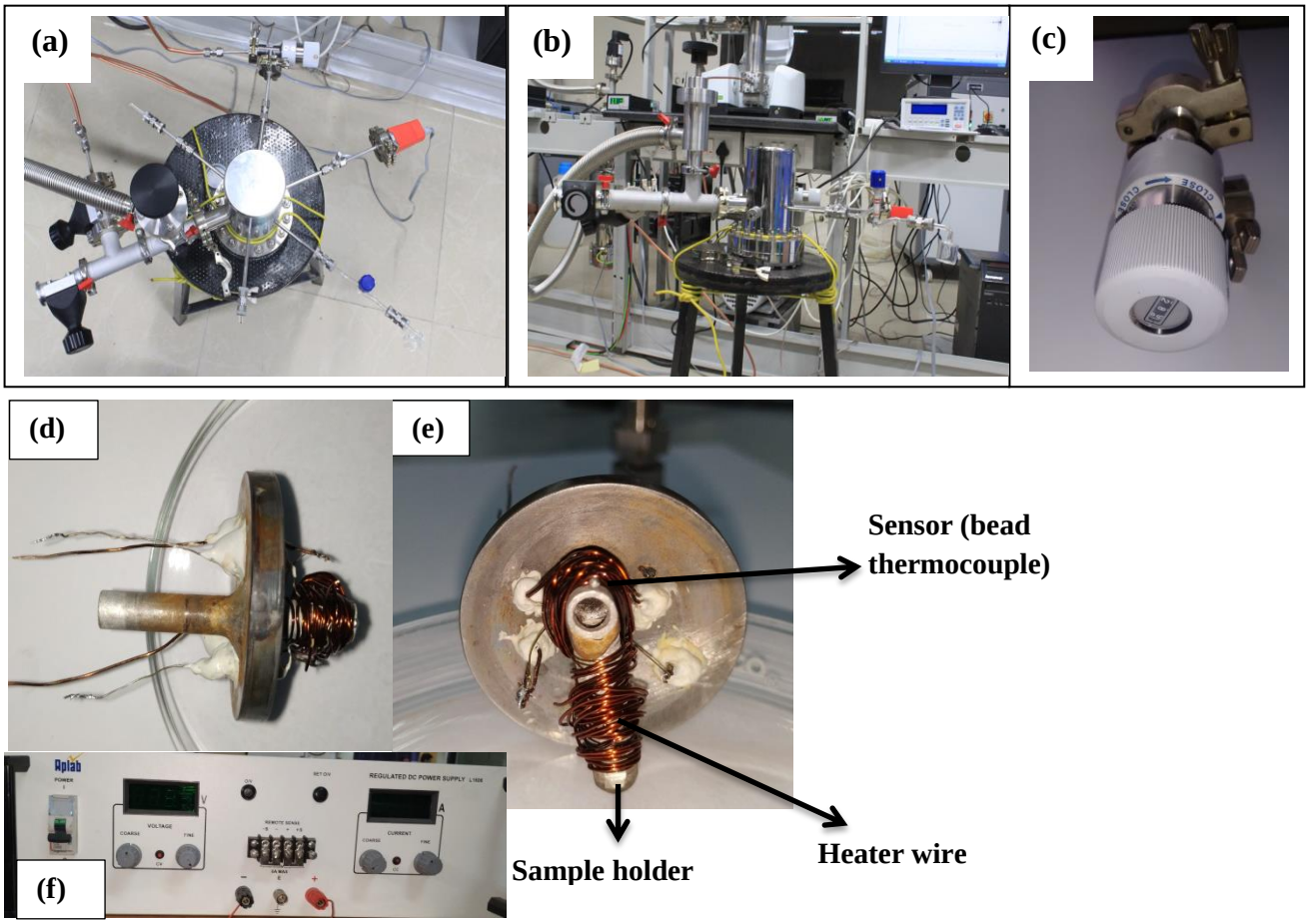
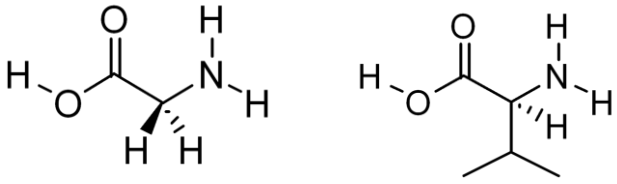
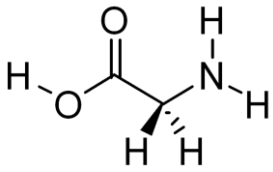


Fig. 2.6 Sample introduction assembly; (a) Mixing chamber (top view); (b) Mixing chamber (side view); (c) Flow controller (needle valve) ; (d) Hot nozzle (top view); (e) Hot nozzle side view; (f) D.C. power supply .



Glycine (3-rotor molecule) **Valine (4-rotor molecule)**

Fig. 2.7 Structure of glycine and L-valine



(O-H bond) =2 fold rotation;
(C-Cα bond)= 6 fold rotation;
(Cα –N bond)=3 fold rotation;
Total 36 initial geometries

HF/6-31G

11 stationary points were obtained

B3LYP/6-311G

8 stationary points obtained

M06-2X/6-311++G(d,p)

7 Conformers were obtained

Unique stationary points were obtained by comparing the energy and rotational constants of the optimized geometry.

Fig. 2.8 Schematic of potential energy surface (PES) scan of glycine

amino acids were then swept on the cold window by the inert gas. The different component of sample introduction set up along with the hot furnace is shown in Fig. 2.6.

After the deposition of sample and matrix gas onto the cold window, the spectra were recorded. The matrix film is then annealed to induce molecular diffusion of trapped species to enable the homogeneous site formation and conformational interconversion if possible. During the annealing process, the matrix film is raised from 12 K to 27 K for N_2 and 30 K for Ar and kept for an hour or so and then re-cooled to 12 K. The spectra of the annealed matrix are then recorded. An increase or decrease in the intensity of spectral features after annealing is attributed to the site effect and conformational interconversion during annealing. The site effects are usually verified by performing the same experiment in a different matrix.

2.6 *Ab Initio* Computations

The quantum mechanical computational study of the conformers of amino acids and model dipeptides were carried out using Gaussian-09 package.[107] The following protocol was followed in the computational work.

2.6.1 Potential energy surface (PES) scan

All the naturally occurring α-amino acids are chiral, except glycine, and exhibit a large number of low energy conformers and rich potential energy surface (PES). A large number of conformers of amino acids stems from the possible rotations about various bonds. For example, glycine the simplest α-amino acid is a three-rotor system, and, valine a four-rotor molecule, as shown in Fig. 2.7. The presence of bulky side-chains in other α-amino acids contributes to the number of rotational bonds present in α-amino acids and hence results in a large number of conformers. One of the primary objectives of the research presented in this thesis was to find out various local minima on the respective potential energy surfaces of amino acids and structural

characterization of all the conformers. Furthermore, we were also interested in understanding the structural correlation between the low energy conformers of α-amino acids with that of the structures of dipeptides.

The maximum possible local minima from PESs of all α-amino acids and model dipeptides were obtained by scanning the PES surface of respective α-amino acids and model dipeptides. A schematic of the PES scan methodology is shown in Fig. 2.8. Total of 36 initial geometries of glycine is generated by systematically rotating the three bonds. The three dihedral angles H-O-C-C_α, O-C-C_α-N and C-C_α-N-H, were turned by two-fold, six-fold, and threefold respectively. The 36 geometries yielded 11 stationary points at HF/6-31G level. The unique stationary points from 36 optimized file were extracted by comparing the rotational constants of optimized geometries. These 11 stationary points were then used as starting geometry for optimization at the B3LYP/6-311G level. The rotational constants screening of 11 files resulted in 8 stationary points, which were later refined at M06-2X/6-311++G(d,p). A total of seven stationary points were obtained, for which the frequency calculations were also done, using an analytical gradient. The optimization at M06-2X method was performed at tight optimization and ultrafine grid condition.[108]

A similar PES scan exercise was conducted for all α-amino acids and model dipeptides, except arginine. The details of the PES scan of all α-amino acids and model dipeptides along with the stationary points obtained at various stages of PES scan will be discussed in the coming chapters.

2.6.2 Atoms in molecules (AIM) analysis

The presence of intramolecular hydrogen bonding interactions in amino acids and dipeptides conformers were computationally identified and analyzed using topological analysis of electron

density ρ (r). Bader and co-worker based on their 'atoms-in-molecules' theory, proposed charge density based criteria for hydrogen bonding interactions.[109] A diverse range of other molecular properties can be deduced from the electron density ρ (r) topological analysis. In the electronic charge density ρ (r) topological framework the spatial coordinates with the vanishing gradient of ρ (r), i.e. ($\nabla\rho(r_c) = 0$), are referred as critical points. The critical points after being located are characterized by the second derivative of charge density, i.e., $\nabla^2\rho(r_c)$ and Hessian of ρ(r). The Eigenvalue of Hessian, which is a (3x3) symmetric matrix of partial double derivative, λ_1, λ_2, and λ_3 are also called as principal axis of curvature. The sum of Eigenvalue ($\sum_{i=1}^{3} = \lambda i$) at bond critical points is referred to as laplacian of charge density ρ(r) and is of fundamental importance. The critical point of electron density topology is characterized in terms of rank and signature of the Hessian matrix. The rank of the critical point is symbolized by ω, and is defined as the number of non-zero eigenvalues or non-zero curvature of electron density ρ at the critical point. The signature of a critical point is the algebraic sum of the sign of the eigenvalues of the Hessian matrix and is denoted by σ. The critical points are represented in the form of rank and signature i.e. (ω, σ). For example, a (3,-1) represents a critical point with two negative and one positive curvature of electron density and corresponds to the bond between two atoms. Likewise, a (3, 1) CP denotes to ring, a (3, 3) CP to a cage and a (3,-3) CP corresponds to a maximum. The number of critical points of the various type which can co-exist in a system with the finite number of nuclei is defined by the Poincare-Hopf relationship.

$$\boldsymbol{n - b + r - c = 1}$$

Where n is the number of nuclei, b is the number of bond CP, r is the number of ring CP and c is the number of cage critical points.

The AIM analysis of amino acid conformers was performed using AIM2000 package.[110] The electron density topology and critical points from the wave function input file were constructed using AIM2000 package. The nature of interactions was characterized by the topological criteria at the critical points, as proposed by Koch and Popelier.[111,112] For an interaction to qualify as hydrogen bonding contact, the $\rho(r)$ and $\nabla^2\rho(r_c)$ at the bond critical point should be within the range of 0.002-0.034 a.u., and 0.024-0.139 a.u., respectively. The sign of $\nabla^2\rho(r_c)$ at BCP, differentiate the closed shell electronic interaction as found in ionic, hydrogen bond and van der Waals interaction with the shared shell interactions, as seen in a covalent bond. A negative $\nabla^2\rho(r_c)$ at BCP implies increased electronic charge density between the atom pair and thus refer to shared shell interaction. Similarly, a positive $\nabla^2\rho(r_c)$ at BCP involves depleted charge density between the atom pair and is suggestive of closed-shell interactions.

The preferential charge accumulation on a given plane of BCP is characterized in term of ellipticity at the BCP and is presented by $[(\lambda_1/ \lambda_2) - 1]$, where λ_1 and λ_2 are the negative eigenvalues of Hessian. Employing the topological parameter such as $\rho(r)$ and $\nabla^2\rho(r_c)$, the local kinetic energy, the local potential energy and total interaction energy at BCP were also calculated, using the method described by Espinosa *et. al*.[113]

2.6.3 Natural bond orbital (NBO) analysis

Natural bond orbitals (NBO) theory was developed by Weinhold and co-workers and is widely used to understand and quantify the role of various orbital interactions in the conformational stability of the concerned molecule.[114] Natural bond orbitals (NBO) are an orthogonal set of localized orbitals that describes the Lewis-like bonding pattern of electron pairs. NBO analysis is based on the process of transforming a delocalized wave function into the localized one or two center orbitals. In the NBO analysis, the concerned atomic orbital basis set is transformed into

natural atomic orbital (NAO), which in turn forms natural hybrid orbital and natural bonding orbital. The Natural Lewis structure (NLS) is determined by Lewis type NBO and the remaining non-Lewis NBO complete the span of the basis and describe the delocalization effect.

The NBO analysis was carried out with Gaussian 09 package using a keyword of pop=nbo. The stability achieved due to orbital delocalization interaction is given by second-order perturbation energy E(2). The magnitude of E(2) of a given donor-acceptor orbital pair interactions is directly proportional to the overlap coefficient between donor (i) and acceptor (j) orbitals, and the overlap coefficient is denoted by F(i,j), and is inversely related to the energy difference between donor and acceptor orbitals [E(j)-E(i)]. The role of specific orbital interaction was analyzed by employing the NBO deletion (NBODEL) method. The NBODEL method involves deletion of interaction concerned from the function space and examines the effect of deleting such interaction on the stability and energy of the structure.

Chapter 3

Conformational Landscape of L-Threonine

3.1 Introduction

L-threonine (2S,3R) is one of the four diastereomers of threonine and is an essential amino acid and constitutes a significant part of connective tissues and plays a vital role in bone and liver health. The L-threonine residues in the GABA receptors of brain cell display a critical role in molecular docking and receptions of the anesthetic agent, such as propofol.[115] The structure and functions of enzymes and proteins rely heavily on the nature of amino acid residues, and it has also been demonstrated that the structure of the protein is determined by the side-chain orientations of amino acid residues.[116] Therefore, an in-depth understanding of the conformations of amino acids can serve as a basis for rationalizing the structure and functions of enzyme and proteins.

The gas phase conformational study of several amino acids such as glycine, alanine, leucine, isoleucine, valine, phenylalanine, cysteine, serine, lysine and others, employing experiments and *ab initio* computations has been addressed in the past.[78-92] The *ab initio* conformational study of L-threonine has also been reported.[120-124] Zhang *et al.*, in their systematic analysis of 1296 geometries at B3LYP/6-311*G level of theory, found a total of 71 conformers of L-threonine.[123] Likewise, Szidarovszky *et al.*, reported a total of 56 conformers of L-threonine, following their work with 7776 initial geometries at HF/3-21G level of theory, followed by optimization at B3LYP and MP2 methods.[124] Alonso and co-workers employing Fourier Transform Microwave spectroscopy reported seven different structures of L-threonine.[64] In the solution phase, the conformations of L-threonine were studied using circular dichroism.[125] However, the conformational study of L-threonine employing vibrational spectroscopy has not

been reported and is done in this work. In the molecular systems, where conformations are thought to be stabilized by hydrogen bonding interactions, the infrared spectroscopy can provide unambiguous evidence for the presence of such weak non-covalent interactions. Therefore, the conformational study of amino acids using vibrational spectroscopy is essential.

Apart from understanding the conformations of amino acid, the factors that determine the conformational preference has not been addressed in the earlier work, which this study address. The presences of intramolecular hydrogen bonding interactions in various conformers of L-threonine were analyzed by atom in molecule (AIM) analysis. Likewise, the conformational stability resulting from the electron delocalization interaction was understood by Natural Bond Orbital (NBO) analysis.

The matrix isolation infrared spectroscopy has been widely applied in the conformational studies of several amino acids such as glycine,[78,79] alanine,[80,81] phenylalanine,[82] serine,[83-85], leucine,[89], isoleucine,[87] lysine,[88], leucine,[89] cysteine,[90,126], tyrosine[92], valine,[117] and proline,[118,119]. The conformational study of L-threonine using matrix isolation FTIR, in the present work, is carried out by employing a hot nozzle placed in front of the cold substrate (KBr) window. The L-threonine vapors, produced at an elevated temperature, were swept onto the cold window, by Ar, flowing through an effusive molecular beam nozzle. The Ar also served as the matrix for trapping threonine.

3.2 Experimental Details

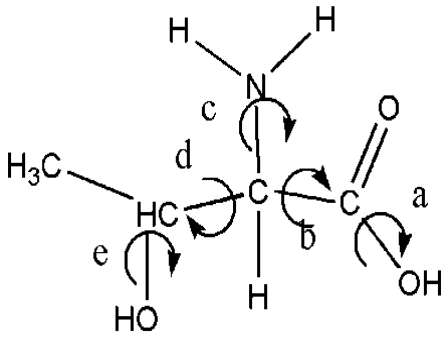
The detail of matrix isolation setup used for the conformational study is described in chapter 2, and only operational details are given here. A closed cycle helium compressor-cooled cryostat (HC-4E1) was used to achieve a low temperature of ~10 K in the matrix isolation experiments. Grade I Argon gas (purity of 99.999%) was used as the matrix gas. The non-volatile L-threonine

was deposited onto the cold window by heating the amino acid, in a hot nozzle effusive beam source at two different temperatures; ~423 K and ~407 K. The vapors of L-threonine thus produced in the hot nozzle source were swept away on to the cold window by a flow of Ar. The typical deposition rate of matrix gas in all the experiments was maintained at ~3 mmol/hour, and a typical deposition lasted for 2 hours. The vibrational spectra of matrix isolated L-threonine were recorded using an FTIR spectrometer (Bruker, Tensor-27), operating at a resolution of 0.5 cm^{-1}. After the IR spectrum of trapped species was recorded at 12 K, the temperature of the cold substrate was raised to 30 K for an hour to allow for any conformational relaxation. The substrate is then brought back to 12 K. This process is referred to as annealing. After annealing the matrix, the infrared spectra of L-threonine were again recorded.

3.3 Computational Details

Ab initio computations were performed using Gaussian 09 suite of programs.[107] The initial geometries for the potential energy surface (PES) scan were obtained through systematic variations of important dihedrals angles of L-threonine. The schematic of the initial geometries generation through the dihedral angle variation is shown in Fig. 3.1. The 1296 different structures obtained were used as starting geometries for the optimization at HF/6-311G level of theory. The optimization process yielded 130 stationary points, which were then re-optimized at the B3LYP/6-311G level of theory. This exercise resulted in the 52 stationary points of L-threonine, and the structure obtained were then refined at B3LYP, M06-2X and MP2 level of theory employing 6-311++G(d,p) basis set. The optimization at M06-2X method was performed using opt=tight and int=ultrafine options. Frequencies of the optimized geometries were calculated to examine if these structures were minima on the potential energy surface.[108] The frequency was calculated using analytical gradients, at the B3LYP, M06-2X and MP2 level of

theory employing 6-311++G(d,p) basis set, and the exercise resulted in the final 38 conformers of L-threonine, each conformer showing real frequency for all the normal modes. The frequency calculations were also used in the assignment of the experimentally observed features of L-threonine.



a = 2 fold rotation ; b = 6 fold rotation
c = 3 fold rotation ; d = 6 fold rotation
e = 6 fold rotation
Total 2x6x3x6x6 = 1296 initial geometries

Fig 3.1 Systematic dihedral rotation of L-threonine for initial geometry generation

The relative free energies of the conformers were calculated at M06-2X and MP2 methods employing 6-311++G(d,p) basis set. The free energies were estimated at 298 K (room temperature) and at the temperature at which the L-threonine was heated during the deposition (423 K), and the relative abundance of conformers was calculated using the relative free energies. Furthermore, the conformer energy at MP2/CBS limit was also computed employing the extrapolation method suggested by Helgaker et. al.[127] The relative energy of the conformers was obtained using the formula shown below.

$$\Delta E_{MP2/CBS} = \frac{4^3 \times \Delta E(pVQZ) - 3^3 \times \Delta E(pVTZ)}{4^3 - 3^3}$$

Where ΔE(pVQZ) and ΔE(pVTZ) are the relative energies of the conformers at the MP2/aug-cc-pVQZ and MP2/aug-cc-pVTZ levels of theory, respectively. The single point

energies of L-threonine conformers were also calculated at the CCSD(T)/6-311++G(d,p) level of theory.

The energy barrier and transition state (TS) structures for conformational interconversion between lowest energy structures were computed at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theories. The transition state geometries obtained were also checked to see the minima on the surface that they connected, by performing intrinsic reaction coordinates (IRC) calculation.

The vibrational assignment for the lower energy conformers was done by scaling the computed wavenumbers. A scaling factor of 0.9298 was used to scale the computed features in the range of 3800-2000 cm^{-1}. This scaling factor was derived by assigning the strong feature observed in the experiments at 3551.8 cm^{-1} to the strongest computed wavenumber in the region at 3820.1 cm^{-1}. Likewise, the calculated wavenumbers in the spectral range between 2000-400 cm^{-1} were scaled by a factor of 0.9483, obtained by assigning the computed feature at 1867.0 cm^{-1} to the experimental feature at 1770.4 cm^{-1}. The mode-by-mode scaling was used in our work, as has been done in many of our and other previous work[128-138] because the matrix perturbation on the vibrational modes varies from one spectral region to that in other.[105]

The hydrogen bonding interactions in the conformers of L-threonine were examined by electron density topology analysis, using AIM2000 package.[110] The wave functions used in this analysis were computed at the MP2/6-311++G(d,p) level of theory. The strength of hydrogen bonding interaction, wherever present was assessed using the method suggested by Espinosa et. al.[113] The importance of orbital interactions in determining the conformational preferences of L-threonine was analyzed by performing Natural Bond Orbital (NBO) analysis.[114] The role of specific orbital interactions was further examined by NBO deletion study, which involves the

deletion of particular orbital interactions from the function space and rationalizes the effect of such deletion on the energetics of various conformers.

3.4 Results

3.4.1 Experimental

The infrared spectra of matrix isolated L-threonine in the argon matrix is shown in Fig. 3.2. The spectra of the different spectral region are shown separately, i.e., 3800-3350 cm^{-1} in the grid a, 1800-1750 cm^{-1} in grid b. The spectral region is shown in the grid (a) corresponding to the O-H stretching region, while that in the grid (b) represent the C=O stretching region. The spectral region corresponding to the O-H bends/skeleton vibrations is shown in Fig. 3.3.The O-H stretching region (grid a) of L-threonine shows spectral features at 3623.8, 3570.0, 3554.1, 3551.8, 3549.2, 3537.5, 3534.0, 3508.5 and 3432.6 cm^{-1}. Likewise, the C=O stretching region (grid b) exhibit vibrational features at 1785.8, 1784.1, 1779.9 and 1770.4 cm^{-1}. The vibrational features observed in the O-H bend and skeleton region are shown in Fig. 3.3 and are listed in Table 3.4.

The matrix isolated IR spectra of L-threonine show weak spectral features due to H_2O, which is a ubiquitous impurity observed in matrix isolation experiments. However, the intensities of features due to H_2O were not increased when the amino acids were deposited, relative to that found when matrix gas was deposited alone. The onset temperature for the thermal decomposition of L-threonine is reported to be ~280ºC, which is much higher than the heating temperature used in our experiments.[139] Furthermore, NH_3 and H_2O which are said to be the thermal decomposition products of L-threonine were not observed in our matrix isolation infrared spectra; thus, confirming that the amino acid did not undergo any decomposition during our deposition.

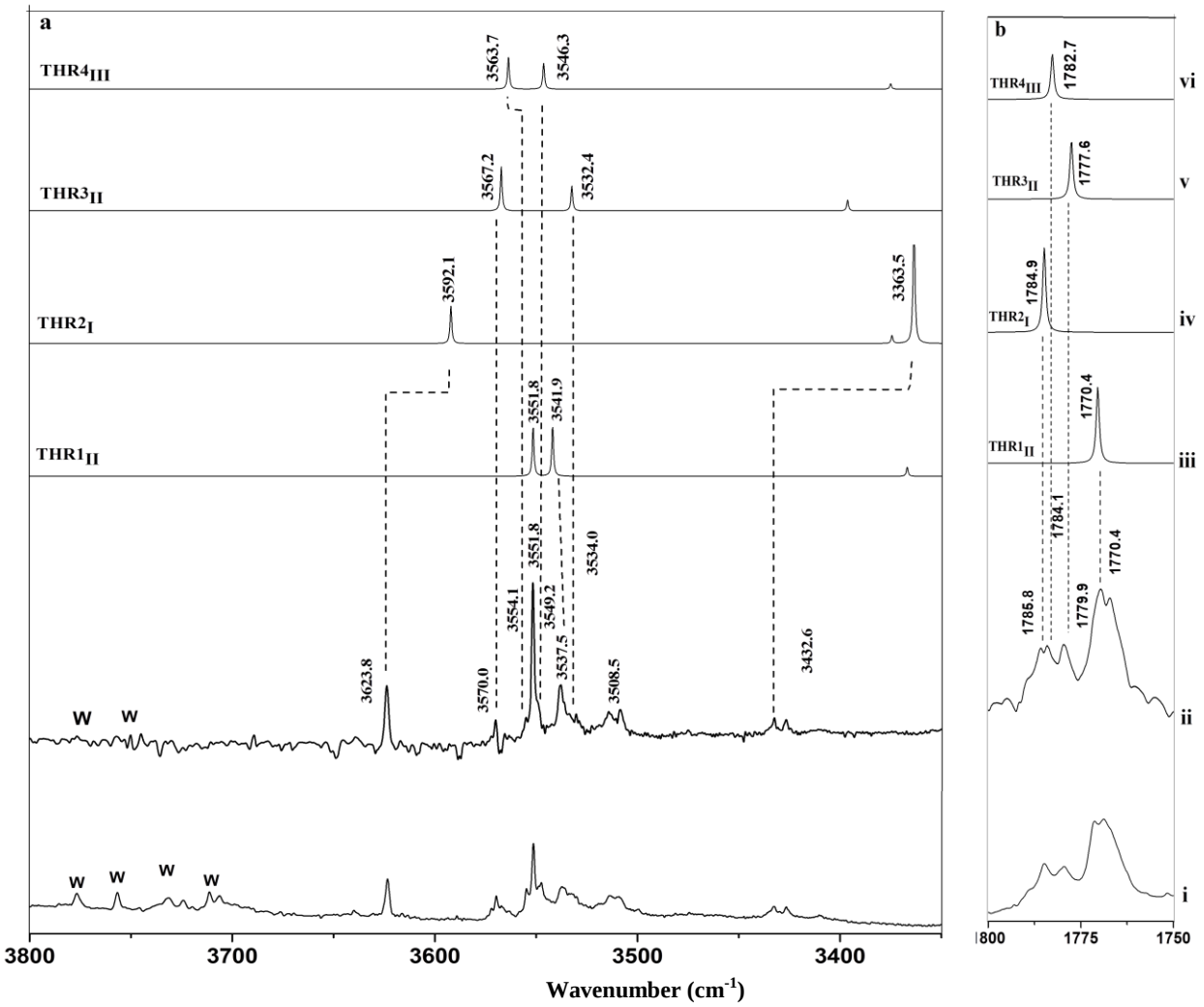


Fig. 3.2 Comparison of the experimental and computed spectra of L-threonine; The spectra span the region 3800-3350 cm^{-1} (grid a) and 1800-1750 cm^{-1} (grid b); (i-ii) IR spectra of matrix isolated L-threonine in an argon matrix deposited at 407 K (i) and 423 K(ii); Features marked as 'w' at 3777.2, 3757.6, 3731.9 and 3711.2 cm^{-1} is due to the water in an argon matrix; (iii-vi) Computed and scaled spectra showing the spectral features of the four lowest energy conformers, calculated at the M06-2X/6-311++G(d,p) level of theory. The computed wavenumbers in grid 'a' and 'b' are scaled using a scaling factor of 0.9298 and 0.9483 respectively.

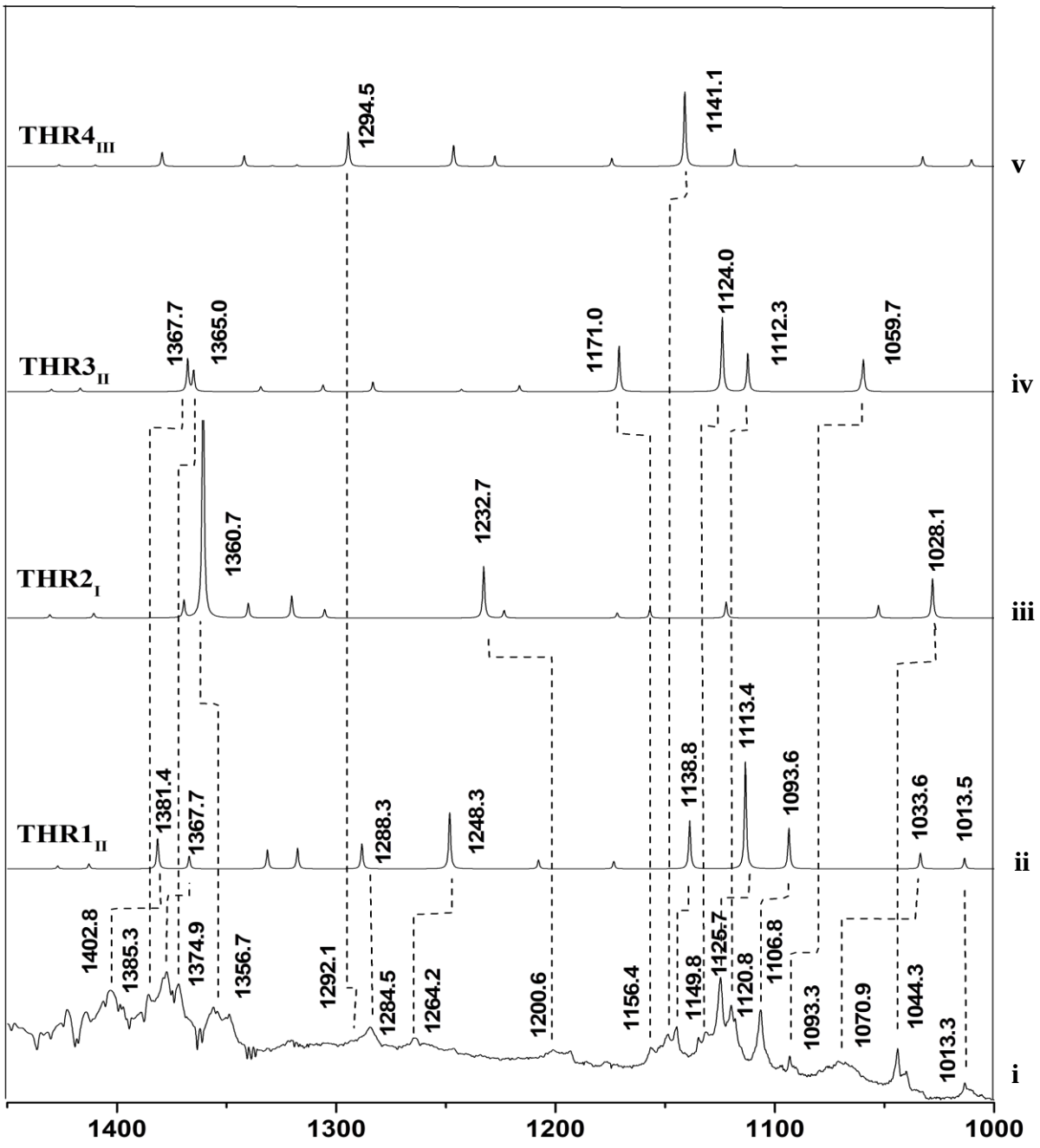
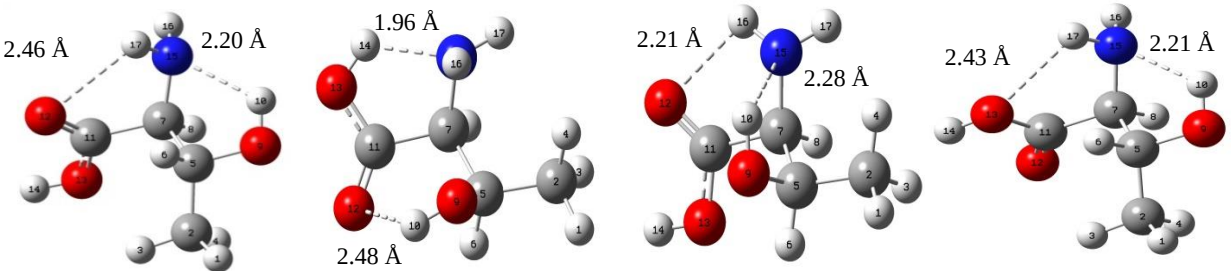


Fig. 3.3 Comparison of experimental and computed spectra of L-threonine, the spectra span the region 1450-1000 cm^{-1}. Trace (i) is the IR spectrum of matrix isolated L-threonine deposited at 423 K; traces (ii-v) show the computed spectra of the four lowest energy conformers, calculated at the M06-2X/6-311++G(d,p) level of theory

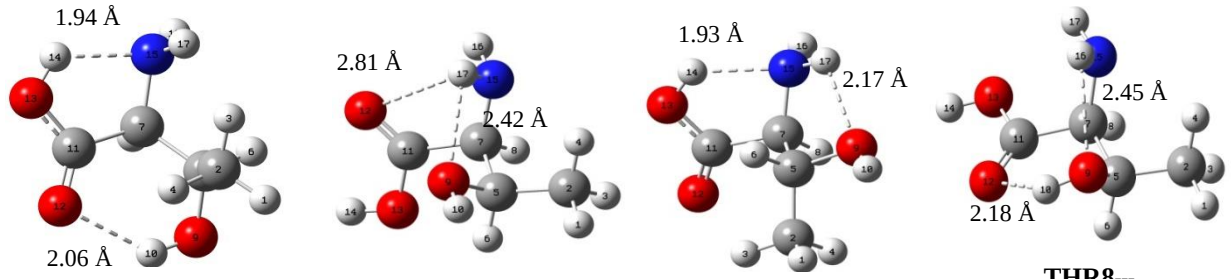


$THR1_{II}$
ΔE_{M06} =0.00 kcal mol^{-1}
ΔE_{MP2} =0.00 kcal mol^{-1}

$THR2_{I}$
ΔE_{M06} =-0.51 kcal mol^{-1}
ΔE_{MP2} =0.18 kcal mol^{-1}

$THR3_{II}$
ΔE_{M06} =0.08 kcal mol^{-1}
ΔE_{MP2} =0.60 kcal mol^{-1}

$THR4_{III}$
ΔE_{M06} =0.81 kcal mol^{-1}
ΔE_{MP2} =0.72 kcal mol^{-1}

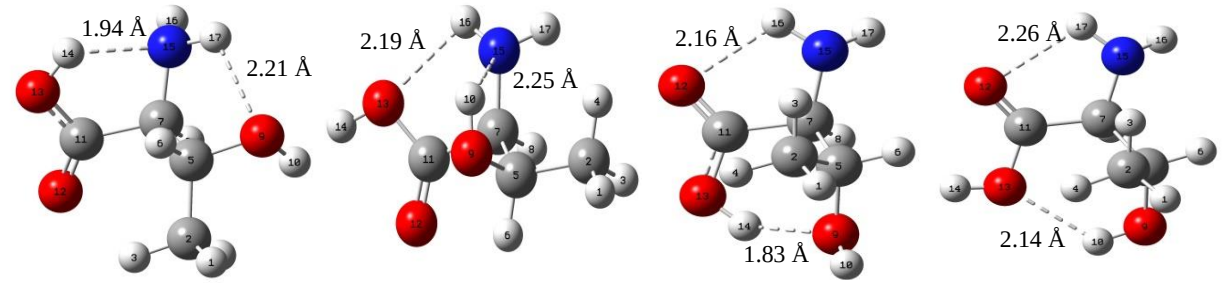


$THR5_{I}$
ΔE_{M06} =0.91 kcal mol^{-1}
ΔE_{MP2} =1.66 kcal mol^{-1}

$THR6_{II}$
ΔE_{M06} =1.40 kcal mol^{-1}
ΔE_{MP2} =1.78 kcal mol^{-1}

$THR7_{I}$
ΔE_{M06} =1.64 kcal mol^{-1}
ΔE_{MP2} =1.79 kcal mol^{-1}

$THR8_{III}$
ΔE_{M06} =0.94 kcal mol^{-1}
ΔE_{MP2} =1.86 kcal mol^{-1}



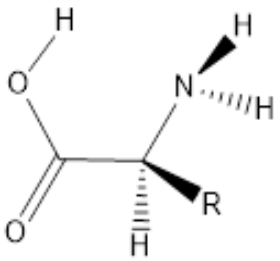
$THR9_{I}$
ΔE_{M06} =1.82 kcal mol^{-1}
ΔE_{MP2} =2.16 kcal mol^{-1}

$THR10_{III}$
ΔE_{M06} =2.05 kcal mol^{-1}
ΔE_{MP2} =2.49 kcal mol^{-1}

$THR11_{IV}$
ΔE_{M06} =1.80 kcal mol^{-1}
ΔE_{MP2} =2.52 kcal mol^{-1}

$THR12_{II}$
ΔE_{M06} =1.84 kcal mol^{-1}
ΔE_{MP2} =2.70 kcal mol^{-1}

Fig. 3.4 Optimized structures of the 12 lowest energy conformers of L-threonine at M06-2X/6-311++G(d,p) level of theory. The relative energies of the conformers, at the M06-2X/6-311++G(d,p) and (MP2/6-311++G(d,p) levels, are also given, the energies being relative to the energy of $THR1_{II}$

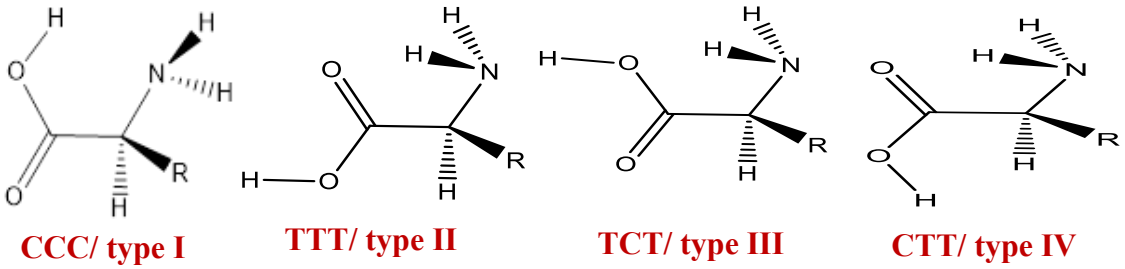


Three key Dihedral angles

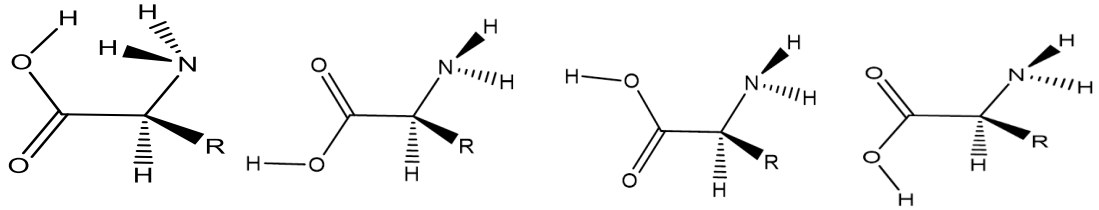
1. ***H-O-C-C ; 0.0 (cis configuration)***
2. ***O-C-C-N ; 0.0 (cis configuration)***
3. ***C-C-N-lp(N); 0.0 (cis configuration)***

CCC or Type I

Fig.3.5 Backbone definition of α-amino acid



CCC/ type I **TTT/ type II** **TCT/ type III** **CTT/ type IV**



CCT/ type **TTC/ type II'** **TCC/ type III'** **CTC/ type IV'**

Fig. 3.6 Eight backbone structure of α-amino acid

3.4.2 Computations

Fig. 3.4 shows the geometry of 12 lowest energy conformers of L-threonine, computed at M06-2X/6-311++G(d,p) level of theory. The energy of conformers relative to THR1, at M06-2X/6-311++G(d,p) and MP2/6-311++G(d,p) level are also shown. The energy of all 38 conformers of L-threonine at various level of theory is listed in Table 3.1. The dipole moment of all 38 structures of L-threonine, computed at M06-2X/6-311++G(d,p) are shown in the last column of Table 1. The four lowest energy conformers, i.e., THR1, THR2, THR3, and THR4, of the 38 conformers, lie within relative energy of ~1 kcal mol^{-1}, as shown in Fig. 3.4.

The conformers were classified based on the orientation of backbone moieties. The backbone of α-amino acids contains a carboxylic group (COOH), an amino group (NH_2) and α-carbon (Cα), as shown in Fig. 3.5. The orientation of the backbone could be defined in terms of three dihedral angle, i.e., H-O-C- Cα, O-C- Cα-N and C- Cα-N-lp, where lp is the lone pair of the nitrogen atom. For the geometry shown in Fig. 3.5, all these three dihedral angles adopt a cis configuration, i.e., dihedral angle value of zero degrees, and therefore the backbone geometry is called as 'CCC (cis-cis-cis) or type I' geometry. Since the backbone orientation is defined based on the three dihedral angles and in case of each dihedral angle adopting two possible orientations, i.e., cis or trans, a total of 8 different backbone (2^3) orientation is possible, as shown in Fig. 3.6. All eight backbone structures, i.e., I, II, III, IV, I', II', III' and IV' differ from each other concerning at least one or more dihedral angle orientation. The backbone geometry with prime notation varies with their respective non-prime backbone structure in term of NH_2 direction, i.e., third dihedral angle. The classification method based on backbone structure is inclusive for all the α-amino acids, and any conformer of α-amino acids could be classified into one the eight backbone geometries.

Table 3.1 Relative zero-point corrected energies of the conformers of L-threonine computed at B3LYP/6-311++G(d,p)[a], M06-2X/6-311++G(d,p)[b] and MP2/6-311++G(d,p)[c] levels. Uncorrected relative energies at the CCSD(T)/6-311++G(d,p)[d] and MP2/CBS[e] levels are also shown. All energies are in kcal mol^{-1} and are relative to the energies of the conformer THR1$_{II}$. Dipole moments (D) of all the conformers at M06-2X/6-311++G(d,p) level of theory shown in the last column.[f]

Conformer	ΔE[a]	ΔE[b]	ΔE[c]	ΔE[d]	ΔE[e]	Dipole moment[f]
THR1$_{II}$	0.00	0.00	0.00	0.00	0.00	2.3
THR2$_{I}$	-0.22	-0.51	0.18	0.05	-0.70	4.4
THR3$_{II}$	0.56	0.08	0.60	0.73	0.35	3.0
THR4$_{III}$	0.88	0.81	0.72	0.85	1.03	3.0
THR5$_{I}$	1.09	0.91	1.66	1.33	0.37	5.0
THR6$_{II}$	1.90	1.40	1.78	2.10	2.26	2.6
THR7$_{I}$	1.62	1.64	1.79	1.97	1.14	4.5
THR8$_{III}$	1.43	0.94	1.86	1.59	1.36	2.5
THR9$_{I}$	1.72	1.82	2.16	2.24	1.31	4.2
THR10$_{III}$	2.42	2.05	2.49	2.67	2.46	3.6
THR11$_{IV}$	2.21	1.80	2.52	2.59	1.79	4.7
THR12$_{II}$	2.65	1.84	2.70	2.97	2.73	1.1
THR13$_{III}$,	2.22	1.97	2.72	2.33	2.28	2.2
THR14$_{III}$	2.68	2.17	2.77	3.07	3.09	0.7
THR15$_{II}$,	3.22	3.03	3.02	3.36	3.37	3.1
THR16$_{III}$	2.95	2.73	3.20	3.28	3.40	1.5
THR17$_{III}$	2.92	2.23	3.31	3.22	2.83	3.4
THR18$_{I}$	2.38	2.45	3.32	3.24	2.90	4.3
THR19$_{III}$,	3.12	2.35	3.40	3.25	3.00	2.9
THR20$_{II}$	3.66	3.08	3.42	3.49	3.68	2.1
THR21$_{II}$,	3.32	3.20	3.42	3.57	3.47	2.9
THR22$_{III}$,	3.82	3.49	3.50	3.94	4.19	3.6
THR23$_{III}$,	3.81	3.31	3.82	4.10	4.05	1.2
THR24$_{II}$,	5.32	4.98	3.82	4.10	4.05	2.4
THR25$_{II}$	4.13	3.74	3.95	4.01	4.08	0.7
THR26$_{IV}$,	3.78	3.61	4.01	3.55	2.72	4.9
THR27$_{III}$,	4.39	4.08	4.54	4.87	4.87	2.9
THR28$_{III}$	4.79	4.16	4.55	4.68	4.86	2.1
THR29$_{III}$,	4.37	4.00	4.61	4.75	4.64	3.0
THR30$_{I}$,	4.16	4.20	4.68	4.53	4.31	5.2
THR31$_{I}$	4.83	4.52	4.97	4.88	4.45	6.3
THR32$_{IV}$	5.00	4.73	5.00	4.97	4.52	2.3
THR33$_{III}$	5.57	5.29	5.31	5.45	5.83	2.8
THR34$_{III}$,	5.85	3.20	5.60	5.85	6.18	2.8
THR35$_{III}$,	6.94	6.66	6.69	7.00	7.44	4.3
THR36$_{I}$,	7.61	7.82	7.81	7.90	7.70	3.0
THR37$_{I}$,	7.18	6.99	7.83	7.63	6.74	3.5
THR38$_{IV}$	8.07	7.73	8.13	8.49	8.22	5.1

All energies in kcal mol^{-1}

All 38 conformers that are listed in Table 3.1 can be associated with one these eight backbone structures, i.e., I, II, III, IV, I', II', III' and IV'. Superimposed on this eight backbone structure are small variations in the side-chain orientation, which yields a large number of distinct conformers of L-threonine. The side-chain O-H group referred to as O-H$_{sc}$, participates

in hydrogen bonding interaction with backbone COOH and NH_2 moiety, in some conformers, either as a proton donor or an acceptor.

The Roman numeral subscript in the nomenclature of a conformer infers the backbone type to which the corresponding conformer belongs. For example, in the conformer $THR1_{II}$ subscript II indicate that THR1 belong to backbone type II. Interestingly, of the 12 lowest energy conformers of L-threonine shown in Fig. 3.4, all, except one, belongs to backbone type I, II or III. The lone exception is the conformer $THR11_{IV}$ which adopt backbone type IV geometry.

The thermodynamic parameters of the conformers, whose relative zero points corrected energy is under 2.50 kcal mol^{-1}, were computed using M06-2X and MP2 methods, and are listed in Table 3.2. The relative abundance of conformers based on the relative free energy at 298 K and 423 K (experimental temperature), were also calculated, as shown in Table 3.2. While the four lowest energy conformers, i.e., $THR1_{II,}$ $THR_{I,}$ $THR3_{II,}$ and $THR4_{III}$ continue to dominate population distribution at both 298 K and 423 K, the other high energy conformers, as anticipated, begin to have increased population at a higher temperature. As can be seen, at 423 K (experimental temperature) the gas phase population distribution of four lowest energy conformers $THR1_{II,}$ $THR_{I,}$ $THR3_{II,}$ and $THR4_{III}$ are 18.74%, 19.73%, 13.74%, and 10.73% respectively, computed at M06-2X/6-311++G(d,p) level of theory. The relative abundance of these four conformers thus accounts for ~65 % of the total population. Likewise, at the MP2 level, these four conformers account for ~70 % of the total population. These four conformers, which have relative abundances more than 10 %, are therefore the probable candidates for observation in our experiments.

Table 3.2 Computed thermodynamical parameters of L-threonine at 298 K and 423 K, at the M06-2X/6-311++G(d,p) and MP2/6-311++G(d,p) levels. Zero point corrected energies (ΔE/kcal mol^{-1}), formation enthalpies (ΔH/kcal mol^{-1}), entropy contribution (TΔS/kcal mol^{-1}), and Gibbs free energies (ΔG/kcal mol^{-1}) relative to the values for THR1$_{II}$ are given. Population (%)[a] of the various conformations computed using the relative free energies are also shown for the first 10 conformers.

Conformer	ΔE	ΔH	TΔS	ΔG	(%)[a]	ΔH	TΔS	ΔG	(%)[a]
M06-2X									
298K						**423K**			
THR1$_{II}$	**0.00**	**0.00**	**0.00**	**0.00**	**22.82**	**0.00**	**0.00**	**0.00**	**18.74**
THR2$_{I}$	**-0.51**	**-0.65**	**-0.42**	**-0.23**	**33.86**	**-0.71**	**-0.66**	**-0.05**	**19.83**
THR3$_{II}$	**0.08**	**0.08**	**-0.13**	**0.21**	**15.99**	**0.10**	**-0.16**	**0.26**	**13.74**
THR4$_{III}$	**0.81**	**0.88**	**0.29**	**0.59**	**8.39**	**0.90**	**0.43**	**0.47**	**10.73**
THR5$_{I}$	**0.91**	**0.75**	**-0.48**	**1.23**	**2.87**	**0.70**	**-0.74**	**1.44**	**3.38**
THR6$_{II}$	**1.40**	**1.61**	**0.34**	**1.28**	**2.65**	**1.69**	**0.57**	**1.12**	**4.94**
THR7$_{I}$	**1.64**	**1.63**	**-0.09**	**1.71**	**1.27**	**1.61**	**-0.14**	**1.75**	**2.33**
THR8$_{III}$	**0.94**	**0.93**	**-0.05**	**0.98**	**4.37**	**0.95**	**-0.05**	**1.00**	**5.71**
THR9$_{I}$	**1.82**	**1.77**	**-0.16**	**1.93**	**0.88**	**1.75**	**-0.25**	**2.00**	**1.74**
THR10$_{III}$	**2.05**	**2.09**	**0.09**	**2.00**	**0.79**	**2.12**	**0.16**	**1.96**	**1.83**
MP2									
298K						**423K**			
THR1$_{II}$	**0.00**	**0.00**	**0.00**	**0.00**	**38.60**	**0.00**	**0.00**	**0.00**	**27.26**
THR2$_{I}$	**0.18**	**0.09**	**-0.34**	**0.33**	**22.21**	**0.05**	**-0.39**	**0.43**	**16.26**
THR3$_{II}$	**0.60**	**0.61**	**-0.13**	**0.70**	**11.82**	**0.64**	**-0.09**	**0.73**	**11.40**
THR4$_{III}$	**0.72**	**0.77**	**0.29**	**0.57**	**14.83**	**0.79**	**0.32**	**0.48**	**15.45**
THR5$_{I}$	**1.66**	**1.49**	**-0.56**	**1.89**	**1.60**	**1.43**	**-0.65**	**2.07**	**2.33**
THR6$_{II}$	**1.78**	**2.00**	**0.51**	**1.65**	**2.40**	**2.09**	**0.62**	**1.48**	**4.70**
THR7$_{I}$	**1.79**	**1.77**	**-0.15**	**1.88**	**1.62**	**1.76**	**-0.17**	**1.93**	**2.75**
THR8$_{III}$	**1.86**	**1.89**	**0.10**	**1.82**	**1.81**	**1.91**	**0.13**	**1.78**	**3.27**
THR9$_{I}$	**2.16**	**2.10**	**-0.27**	**2.29**	**0.81**	**2.07**	**-0.31**	**2.38**	**1.62**
THR10$_{III}$	**2.49**	**2.53**	**0.16**	**2.42**	**0.65**	**2.57**	**0.20**	**2.37**	**1.63**

[a]Population distributions (%) were calculated considering all 38 THR conformers.

3.4.3 Vibrational assignments

The four conformers discussed above which show different backbone types and different side-chain (O-H$_{sc}$) orientations have different O-H and C=O stretching frequencies. These wavenumbers serve as potential signatures for identifying the various conformers, trapped in the matrix. The assignments of the features in the experimental spectra were performed by comparison with synthetic spectra generated using the computed scaled wavenumbers of the four lower energy conformers. Synthetic spectra corresponding to each of the four lowest energy conformers were generated, as shown in Fig. 3.2 and 3.3 (trace iii-vi). The synthetic spectra were

computed with SYNSPEC package and by using a Lorentzian line shape and line width of 1.0 cm^{-1}. The intensities of the computed features in the synthetic spectra were obtained by using the computed intensities for each vibrational modes of a given conformer, followed by intensity scaling for the population of particular conformer at 423 K (nozzle temperature during sample deposition). The assignment between experimental features and the computed wavenumbers of these four lowest energy conformers are listed in Table 3.3 and 3.4.

3.4.3 (a) 3800 - 3350 cm^{-1} region

Fig. 3.2 (grid a) show experimental (trace i-ii) and the computed spectra (iii-vi) of the four lowest energy conformers of L-threonine, in the O-H stretching spectral region. The feature in this spectral region at 3551.8 cm^{-1} (trace ii), is the most intense and is, therefore, most likely due to the υ(O-H) of the lowest energy conformer. However, our *ab initio* computations do not assertively identify the lowest energy structure, as the conformer $THR1_{II}$ is computed to be the most stable structure at MP2/6-311++G(d,p) and CCSD(T)/6-311++G(d,p) level of theories, while at B3LYP, M06-2X and MP2/CBS levels conformer $THR2_{I}$ is computed at lowest energy conformer. Therefore, this intense experimental feature at 3551.8 cm^{-1} could be assigned to either $THR1_{II}$ or $THR2_{I}$. If the computed O-H stretch of the conformer $THR2_{I}$ is attributed to the above discussed experimental feature and a scaling factor calculated based on that assignment (work out to be 0.9817), the scaled computed wavenumbers of other conformers show an unsatisfactory assignment with other experimental features. However, if the experimental feature at 3551.8 cm^{-1} is assigned to the strongest computed wavenumber of the conformer $THR1_{II}$ and a scale factor is calculated based on this assignment, the scaled computed features of other conformers provides a convincing assignment with all the feature in the experiment thus making this assignment acceptable. This argument, therefore, can be considered to offer unambiguous experimental

evidence that conformer $THR1_{II}$ is the lowest energy conformer of L-threonine. As discussed in the earlier section, the scaling factor based on this assignment works out to be 0.9298.

The above observation of conformer $THR1_{II}$ as the most stable structure and conformer $THR2_{I}$ as the second lowest energy conformer is consistent with conformational ordering observed in the rotational spectroscopy study by Alonso and co-workers.[64] A similar backbone preference, i.e., backbone type II being the most stable and type I being the 2nd lowest energy conformer has also been reported in the conformational study of serine [61,84] as well as several non-polar amino acids.[79,80,87,89,117] Though, the frequency calculations at both MP2 and M06-2X level, with 6-311++G(d,p) basis set, present nearly similar vibrational assignment, the mean square deviation of the experimental and the scaled computed features at the MP2 level was observed to be twice than that calculated when computed wavenumbers at M06-2X level were used. Therefore, the assignment using the computed scaled wavenumber at M06-2X level provides better agreement with the experimental features and hence has been used.

The assignment of other experimental features with the computed wavenumbers of the four lowest energy conformers at M06-2X/6-311++G(d,p) level of theory is shown in Table 3.3. The feature at 3537.5 cm^{-1} (trace ii) can be assigned to the $\upsilon(O\text{-}H_{sc})$ of conformer $THR1_{II}$, which is computed to be observed at 3541.9 cm^{-1} (trace iii). Likewise, the experimental wavenumber at 3432.6 cm^{-1} (trace ii) can be assigned to the $\upsilon(O\text{-}H)$ of conformer $THR2_{I}$. The O-H stretching vibrations of conformer $THR2_{I}$ appear red-shifted by ~119 cm^{-1}, relative to the feature corresponding to the same mode of the conformer $THR1_{II}$. This large redshift arises due to an intramolecular hydrogen bonding interaction, which is also corroborated by our calculations of the vibrational frequency of the O-H for this conformer. The computed wavenumber of side-chain O-H stretching, i.e., $\upsilon(O\text{-}H_{sc})$ of $THR2_{I}$ (trace iv) agrees well with the experimental feature

Table 3.3 Experimental (in argon matrix) and computed scaled vibrational wavenumbers (cm^{-1}) at the M06-2X/6-311++G(d,p) level for the different conformers of L-threonine. See text for details on scaling

Experimental	Computed scaled[a]				IR intensity[b]	Vibrational modes
	$THR1_{II}$	$THR2_{I}$	$THR3_{II}$	$THR4_{III}$		
3623.8		3592.1			67.7	ν ($O-H_{sc}$)
3570.0			3567.2		113.7	ν (O-H)
3554.1				3563.7	107.8	ν (O-H)
3551.8	3551.8				93.1	ν (O-H)
3549.2				3546.3	87.9	ν ($O-H_{sc}$)
3537.5	3541.9				93.3	ν ($O-H_{sc}$)
3534.0			3532.4		64.0	ν ($O-H_{sc}$)
3508.5	3441.8					2ν (C=O)
3432.6		3363.5			247.0	ν (O-H)
1785.8		1784.9			335.2	ν (C=O)
1784.1				1782.7	333.0	ν (C=O)
1779.9			1777.6		335.7	ν (C=O)
1770.4	1770.4				319.1	ν (C=O)

[a] The scaling factors for the computed vibrational wavenumbers of L-threonine conformers: 0.9298 for the O−H stretch region (3800-2000 cm^{-1}) and 0.9483 for the C=O stretch and skeleton vibration (2000-400 cm^{-1}). See text for details.
[b] The computed IR intensity(km mol^{-1}) corresponding to each normal mode.

Table 3.4 Experimental (argon matrix) and computed scaled vibrational wavenumbers (cm^{-1}) at the M06-2X/6-311++G(d,p) level of theory, of vibrational modes in the region 1450 -1000 cm^{-1}, for the four conformers of L-threonine

Experimental	Computed scaled[a]				IR intensity[b]
	$THR1_{II}$	$THR2_{I}$	$THR3_{II}$	$THR4_{III}$	
1402.8	1381.4				46.3
1385.3			1367.7		67.1
1377.0	1367.1				19.7
1374.9			1365.0		43.5
1356.7		1360.7			413.8
1292.1				1294.5	91.4
1284.5	1288.3				38.4
1264.2	1248.3				85.9
1200.6		1232.7			73.6
1156.4			1171.0		93.6
1152.3				1141.1	197.3
1149.8	1138.8				72.3
1134.8			1124.0		153.7
1125.7	1113.4				160.2
1120.8			1112.3		80.1
1106.8	1093.6				61.4
1093.3			1059.7		63.6
1070.9	1033.6				23.5
1044.3		1028.1			56.2
1013.3	1013.5				16.2

[a] Scaling factors for computed vibrational frequencies: 0.9483. See text for details.
[b] Computed infrared intensity(km mol^{-1}) corresponding to each mode.

at 3623.8 cm^{-1} (trace ii). A comparison with the computed wavenumber suggests that features observed at 3570.0 and 3534.0 cm^{-1} (trace ii) can be assigned to the conformer $THR3_{II}$ for the υ(O-H) and υ($O-H_{sc}$) mode respectively (trace v). Similarly, the experimental features at 3554.1 and a shoulder at 3549.2 are in agreement with the computed wavenumbers of $THR4_{III}$ for the υ(O-H) and υ($O-H_{sc}$) mode respectively (trace vi). It must be highlighted that conformers with backbone geometry II, i.e., $THR1_{II}$ and $THR3_{II}$, show closely lying wavenumbers for the O-H stretch of the carboxylic and the side-chain OH groups. However, conformer $THR2_{I}$ that possess backbone type I shows a significant difference between the wavenumbers of the O-H stretch of the groups mentioned above, due to the operation of the hydrogen bonding interaction involving the carboxylic O-H. The feature at 3508.5 cm^{-1}, for now, is left unassigned and the assignment will be discussed later while addressing the carbonyl stretching region.

To summarize the spectral behavior in the O-H stretching region of L-threonine, the backbone type I conformer ($THR1_{II}$) shows the large difference in the carboxyl O-H and side-chain O-H stretching vibrational wavenumbers, both experimentally and computationally. However, conformers with backbone type II and III have comparatively closely spaced wavenumbers for the same two modes, both in our experiments and computations. The above observation is rationalized based on the presence of hydrogen bonding interaction involving carboxylic O-H in the backbone type I conformer, which is absent in backbone type II and III.

The N-H vibration of conformer $THR1_{II,}$ $THR_{I,}$ $THR3_{II,}$ and $THR4_{III}$ are computed to occur at 3366.9, 3374.3, 3396.3 and 3375.1 cm^{-1} respectively. However, due to poor infrared absorption cross sections, these N-H vibrations were not seen in our experiments. Similarly, the computed wavenumbers of the C-H stretching vibrations of four lowest energy conformers are grouped around 2927.7 to 2915.6 cm^{-1}, whereas the asymmetric C-H vibrations were clustered

around 2937.5 to 2930.9 cm^{-1}. These modes also show very poor infrared absorption cross-section, and thus were not observed in our IR spectra of L-threonine.

3.4.3 (b) 1800-1750 cm^{-1} region

The carbonyl stretching absorption region of amino acid occurs in the range of 1800-1750 cm^{-1}, is shown in Fig. 3.2 (grid b) and Table 3.4. The experimental wavenumbers in the C=O stretching region are seen at 1770.4, 1785.8, 1779.9 and 1784.1 cm^{-1} (trace ii). The computed scaled wavenumbers corresponding to the C=O stretching vibrations of conformer $THR1_{II,}$ $THR_{I,}$ $THR3_{II,}$ and $THR4_{III}$ occurs at 1770.4 (trace iii), 1784.9 (trace iv), 1777.6 (trace v) and 1782.7 cm^{-1} (trace vi) respectively and are in good agreement with experimental features mention above.

The overtone mode of C=O stretching vibration was computed employing the anharmonic frequency calculations, and the computed scaled wavenumbers of C=O overtone mode were found to occur in the region of 3450 cm^{-1}. The unassigned feature at 3508.5 cm^{-1}, in the O-H stretching region, could be assigned to the 2υ(C=O) mode of conformer $THR1_{II}$, whose computed scale wavenumbers for overtone vibration was observed at 3441.5 cm^{-1}. The 2υ(C=O) mode of conformer $THR_{I,}$ $THR3_{II}$ and $THR4_{III}$ were found to occur at 3472.8, 3470.4 and 3463.4 cm^{-1} respectively. The experimental features corresponding to the overtone mode of these conformers were not observed, possibly due to the smaller relative abundance. Furthermore, it is likely that, even if the overtone features of these conformers had occurred with smaller intensity, they were probably unresolved from the O-H stretching features occurring to the higher wavenumber side of 3508 feature.

3.4.3 (c) 1450-1000 cm^{-1} region

The computed and experimental wavenumbers of the spectral region ranging in 1450-1000 cm^{-1} are shown in Fig. 3.3 and Table 3.4. The normal modes in this spectral region demonstrate

vibrations in the form of deformation, torsional motions and atomic excursion involving bending motion. The prominent experimental features in this region were observed at 1402.8, 1377.0, 1284.5, 1264.2, 1149.8, 1125.7, 1106.8, 1070.9 and 1013.3 cm^{-1}, as shown in Fig. 3.3 (trace i). These experimental features are found in good agreement with the computed wavenumber of conformer $THR1_{II}$, computed to occur at 1381.4, 1367.1, 1288.3, 1248.3, 1138.8, 1113.4, 1093.6, 1033.6 and 1013.5 cm^{-1} respectively (trace ii). Likewise, the experimental wavenumbers observed at 1356.7, 1200.6 and 1044.3 cm^{-1} (trace iii) corroborate with the computed scaled wavenumber of conformer $THR2_{I}$, which are computed to occur at 1360.7, 1232.7 and 1028.1 cm^{-1} respectively (trace iii).

The comparison between the computed values suggests that the experimental features observed at 1385.3, 1374.9, 1156.4, 1134.8, 1120.8 and 1093.3 cm^{-1}, could be assigned to the conformer $THR3_{II}$. The feature of $THR3_{II}$ corresponding to these experimental wavenumbers are computed at 1367.7, 1365.0, 1171.0, 1124.0, 1112.3 and 1059.7 cm^{-1} respectively (trace iv). The other experimental features observed in this region at 1292.1 and 1152.3 cm^{-1} are assigned to the conformer $THR4_{III}$ which was computed to occur at 1294.5 and 1141.1 cm^{-1} respectively (trace v). Even though the assignments of experimental features with the computed wavenumber of the four lowest energy conformers are conclusive, but the spectral complexity of the features in this region makes the assignments tentative. Therefore clearly, based on the spectral assignment of O-H and C=O stretching region, we observed the presence of four lowest energy conformer of L-threonine, trapped in our matrix isolation experiments.

3.4.4 Conformer interconversion

The matrix isolation experiments of L-threonine were performed using an effusive beam hot nozzle, as mentioned in the earlier sections. Therefore, the application of an effusive beam

nozzle eliminates the possibilities of conformation cooling during the process of sample deposition. Hence, the conformational abundance in the gas phase before sample deposition is likely to be frozen in the matrix phase. However, the conformational relaxation to the lowest energy conformers, after the deposition, is still possible during the annealing of the matrix at 30 K. Based on the earlier studies of matrix isolation, it has been established that conformational interconversion and relaxation to the lowest energy structure is feasible during the process of annealing of matrix at 30 K, provided the energy barrier of relaxation is less than ~5 kJ mol^{-1} (~1.2 kcal mol^{-1}).[80]

Therefore, to verify the possibility of conformation cooling to lowest energy conformers during the annealing, the energy barrier and transition structure separating low energy conformers were computed at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theories. The transition state geometries obtained were further verified by intrinsic reaction coordinates (IRC) calculation. The energy barrier separating the relaxation of conformer $THR3_{II}$ and $THR4_{III}$ to the lowest energy conformer $THR1_{II}$ were calculated to be 3.7 and 2.7 kcal mol^{-1} respectively, at the MP2/6-311++G(d,p) and 3.8 and 1.9 respectively, at the M06-2X/6-311++G(d,p) level of theory. The relaxation of conformer $THR3_{II}$ and $THR4_{III}$ to $THR2_{I}$ and conformer $THR2_{I}$ to $THR1_{II}$ proceeds via rotation around multiple dihedral angles such as H-N-Cα-C, N-Cα-C=O, Cα-C-O-H and others, and thus are likely to have a larger barrier for interconversion. Hence, these relaxations are unlikely to occur in the matrix, during annealing. It can be observed that the energy barrier in all the cases of larger than 5 kJ mol^{-1}, thus rules out the possibilities of any conformational relaxation in the matrix, after the deposition of the matrix. Therefore, it is established that the conformational distribution present at the time of deposition, in the gas phase, are trapped in the matrix with no further relaxation.

3.4.5 Atoms in molecule (AIM) analysis of amino acid backbone types

Non-covalent interaction, such as hydrogen bonding, plays a vital role in determining the conformational preferences of amino acids. The presences of such intramolecular hydrogen bonding in various conformers of L-threonine were analyzed by charge density topology study using AIM2000 program.[110] The nature of such weak interactions were characterized based on topological parameters such as charge density ($\rho(r_c)$) concentrated around the nucleus and the Laplacian of electron density ($\nabla^2\rho(r_c)$), as proposed by Koch and Popelier.[111,112] In hydrogen bonded interactions, the value of $\rho(r_c)$ and $\nabla^2\rho(r_c)$ ranges from 0.002-0.034 and 0.024-0.139, respectively. The stability due to hydrogen bonded interaction was also computed based on the topological parameters at bond critical points, using the method suggested by Espinosa et. al.[113]

For closed-shell interactions, Espinosa et al. proposed a method to estimate the hydrogen bond (E_{HB}) dissociation energy using electron density ($\rho(r_c)$) and Laplacian of electron density ($\nabla^2\rho(r_c)$) at bond critical points. The local electronic kinetic energy density G(r) was evaluated from the electron density $\rho(r_c)$ at the bond critical point (equation i). The local electronic potential energy ($V(r_c)$) at bond critical points was estimated using the virial theorem(equation ii). The hydrogen bonding energy (E_{HB}) for closed shell interactions, can be estimated using the proportionality relation shown below in equation iii.

$$\mathbf{G(r_c) = 3/10\ (3\pi^2)^{2/3}\ \rho^{5/3}(r_c) + 1/6\ (\nabla^2\rho(r_c))} \qquad \mathbf{\ldots.\ (i)}$$

$$\mathbf{V(r_c) = \hbar^2/4m\ (\nabla^2\rho(r_c)) - 2G(r_c)} \qquad \mathbf{\ldots.\ (ii)}$$

$$\mathbf{E_{HB} = \tfrac{1}{2}\ V(r_c)} \qquad \mathbf{\ldots.(iii)}$$

The AIM analysis and observed bond critical points of various backbone type conformer of L-threonine are shown in Fig. 3.7 (a-e). The values of topological parameters such as electron

density, Laplacian of electron density and others, at bond critical points of different conformers, computed at MP2/6-311++G(d,p) level of theory, are listed in Table 3.5.

The backbone type I conformer $THR2_I$ shows a bond critical point along the (O-H$^{\cdots}$N) bond path, thus indicating the presence of an intramolecular hydrogen bonding interactions, involving the NH_2 group and hydroxyl moiety of COOH group. The bond critical point observed is in agreement with the vibrational assignment of the O-H stretching vibrations of this conformer made earlier, which was observed significantly red-shifted (at 3432.6 cm^{-1}) compared to the similar features of other conformers. Such type of hydrogen bonding interactions is found to be unique for the backbone type I conformers, thus providing stability to these conformers. The hydrogen bonding interaction (E_{HB}) at bond critical point computed using Espinosa method was found to be 10.7 kcal mol^{-1}.

In the previously reported work of amino acids conformations, backbone type II, III and IV were suggested to be stabilized due to a bifurcated (C=O$^{\cdots}NH_2$), (C-O$^{\cdots}NH_2$) and (C=O$^{\cdots}NH_2$) intramolecular hydrogen bonded interaction respectively.[82,87,88] However, AIM analysis performed in this study does not locate bond critical points corresponding to such interaction in the backbone type II ($THR1_{II}$ and $THR3_{II}$), type III ($THR4_{III}$) and type IV ($THR11_{IV}$), as shown in Fig. 3.7. Nevertheless, the conformer $THR1_{II}$ shows hydrogen bonded interaction involving the side-chain hydroxyl group (O-H_{sc}), as shown in Fig. 3.7(b). Similarly, the conformer $THR4_{III}$ and $THR11_{IV}$ also appears to have an intramolecular hydrogen bonding contact involving side-chain hydroxyl group (O-H_{sc}), as shown in Fig. 3.7(d) and (e). The AIM analysis of conformer $THR3_{II}$ does not indicate hydrogen bonding interaction due to the side-chain hydroxyl group (O-H_{sc}).

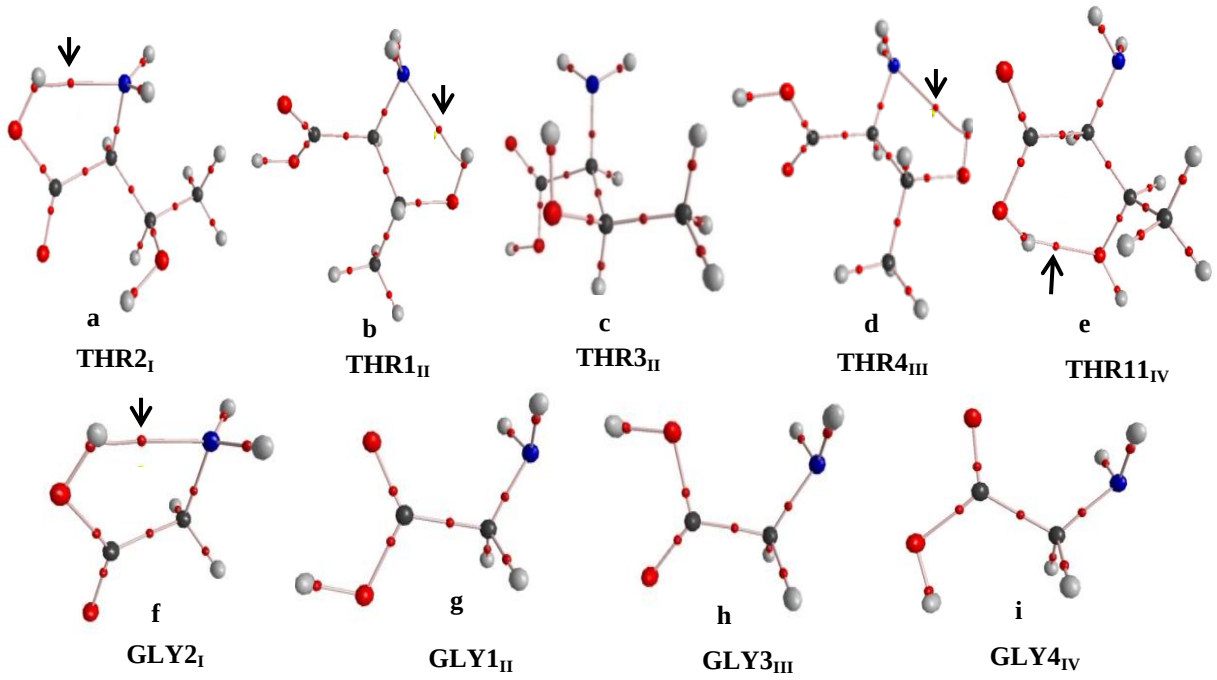


Fig. 3.7 AIM analysis of the different conformers of L-threonine (top) and glycine (bottom) at MP2/6-311++G(d,p). Bond critical points corresponding to the intramolecular hydrogen bonds are indicated by the black arrow

Table 3.5 Electron density [$\rho(r_c)$], Laplacian of electron density [$\nabla^2\rho(r_c)$], local electronic kinetic energy density [$G(r_c)$], local electronic potential energy density [$V(r_c)$] and hydrogen bond energy (E_{HB}) values for the hydrogen bonded interactions in the different conformers of L-threonine and glycine, at the MP2/6-311++G(d,p) level. The values for $\rho(r_c)$, $\nabla^2\rho(r_c)$, $G(r_c)$ and $V(r_c)$ are given in atomic units and for E_{HB}, in kcal mol^{-1}

Conformer	Interactions	$\rho(r_c)$	$\nabla^2\rho(r_c)$	$G(r_c)$	$V(r_c)$	E_{HB}
L-Threonine						
$THR1_{II}$	$O\text{-}H_{SC}\cdots N$	0.0224	0.0824	0.0189	-0.0171	-5.36
$THR2_{I}$	$O\text{-}H\cdots N$	0.0381	0.1114	0.0309	-0.0340	-10.67
$THR4_{III}$	$O\text{-}H_{SC}\cdots N$	0.0213	0.0799	0.0180	-0.0161	-5.04
$THR11_{IV}$	$O\text{-}H_{SC}\cdots O$	0.0366	0.1272	0.0328	-0.0338	-10.60
Glycine						
$GLY2_{I}$	$O\text{-}H\cdots N$	0.0378	0.1096	0.0305	-0.0336	-10.54

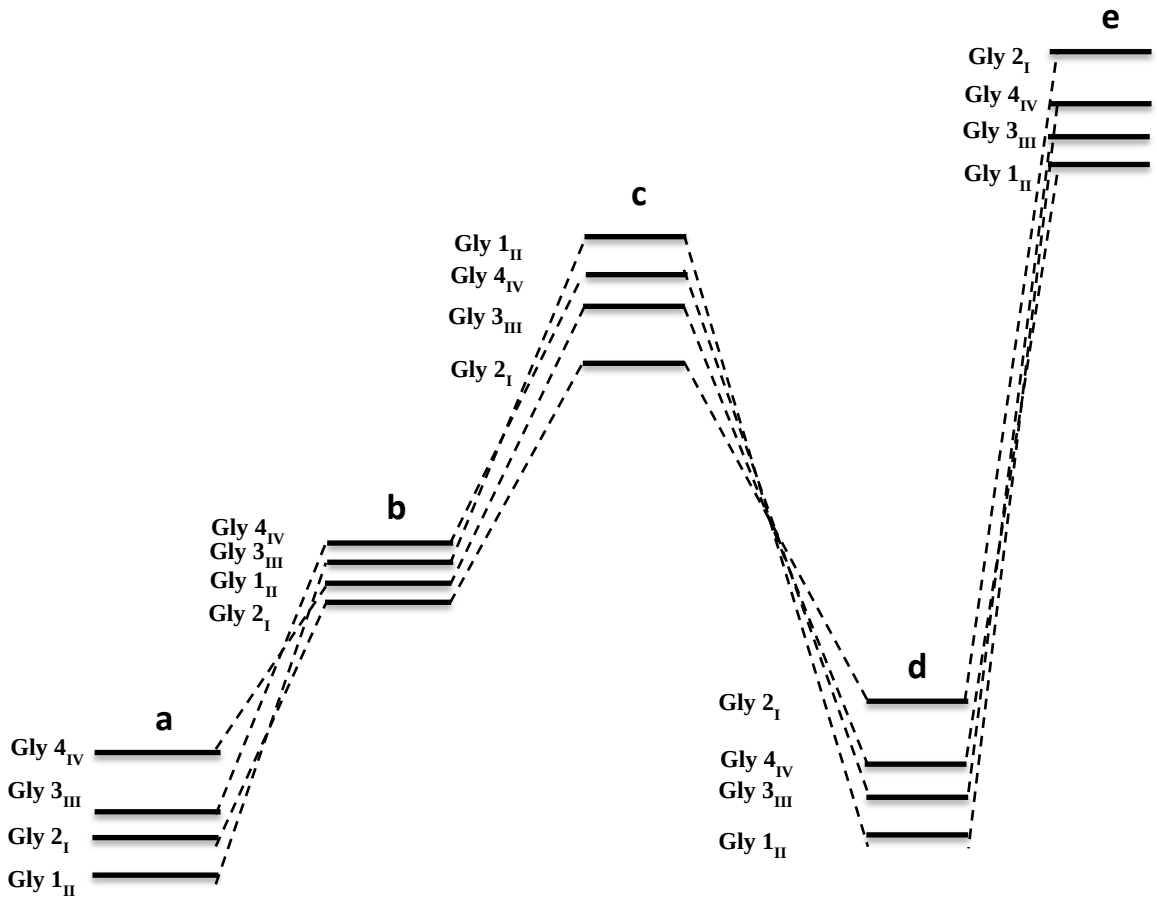


Fig. 3.8 Diagram showing the correlation between the different backbone type conformers in glycine; (a) with all the hyperconjugative interactions present; (b) with the geminal interactions deleted; (c) with the vicinal interactions deleted; (d) with the remote interactions deleted, and (e) with all the hyperconjugative interactions deleted

Table 3.6 NBO deletion analysis showing the energies obtained after deletion of various hyperconjugative interactions in different conformers of glycine computed at the M06-2X/6-311++G(d,p) level of theory

Glycine NBO DEL				Change in energy	
Conformer	Energy (Hartree)	Interaction Deleted	Energy after deletion (Hartree)	Hartree	kcal mol^{-1}
$GLY2_I$		Geminal	-284.36301	0.04465	28.02
	-284.40766	Vicinal	-283.94695	0.46071	289.10
		Remote	-284.37918	0.02848	17.87
		All	-283.87023	0.53743	337.24
$GLY1_{II}$		Geminal	-284.36243	0.04676	29.34
	-284.40919	Vicinal	-283.92878	0.48041	301.46
		Remote	-284.40789	0.00130	0.82
		All	-283.88078	0.52841	331.58
$GLY3_{III}$		Geminal	-284.35844	0.04820	30.24
	-284.40663	Vicinal	-283.93173	0.47491	298.01
		Remote	-284.40446	0.00217	1.36
		All	-283.87971	0.52693	330.65
$GLY4_{IV}$		Geminal	-284.35435	0.04570	28.67
	-284.40005	Vicinal	-283.92996	0.47009	294.98
		Remote	-284.39716	0.00289	1.81
		All	-283.87484	0.52521	329.57

The above-discussed AIM analysis results are in agreement with our infrared spectra. However, the issue that needs to be explored is whether the hydrogen bonding interactions involving the side-chains in backbone type II and III conformers play a decisive role in the energy ordering of such conformer. The issue can be rationalized by comparison with glycine, which has no side-chain. The AIM analysis of different backbone type conformers of glycine is shown in Fig. 3.7 (f-i). The relative energy ordering of glycine conformers, i.e., $GLY1_{II}$, $GLY2_{I}$, $GLY3_{III}$ and $GLY4_{IV}$, computed at MP2/6-311++G(d,p) level of theory, is found to be 0.00, 0.59, 1.59 and 5.70 kcal mol^{-1} respectively. The energy ordering of glycine conformers is similar to that observed for L-threonine, i.e., backbone type II conformer ($GLY1_{II}$) is more stable than type I conformer ($GLY2_{I}$). The AIM analysis of type II conformer of glycine ($GLY1_{II}$) neither shows any bifurcated($C=O^{\cdots}NH_2$) hydrogen bonding nor can it even have any hydrogen bond due to the side-chain, as it is devoid of any side-chain.

Likewise, in the conformers having backbone type III ($GLY3_{III}$) and type IV ($GLY4_{IV}$), bifurcated hydrogen bonding interactions were not seen, as shown in Fig. 3.7 (h) and (i). Therefore, the side-chain hydroxyl group ($O\text{-}H_{sc}$) involved hydrogen bonded interactions in L-threonine are unlikely to play a defining role in conformational preference, and at best they may be only supportive. It is possible that backbone stability of glycine and L-threonine are directed by orbital delocalization interactions, which will be discussed next.

3.4.6 NBO analysis

Natural bond orbital (NBO) analysis, along with the use of the orbital interaction deletion method (NBODEL) was carried out to comprehend the importance of orbital delocalization interactions in the stability of various backbone types of glycine and L-threonine. All the three hyper-conjugation interactions viz. geminal, vicinal and remote are operative in glycine and L-

threonine. The magnitude of stabilization because of geminal, vicinal and remote interactions was quantified by understanding the change in the relative energy ordering of $GLY1_{II}$, $GLY2_{I}$, $GLY3_{III,}$ and $GLY4_{IV}$, by systematic deletion of geminal, vicinal and remote and all interactions. The relative energy ordering of glycine conformers with the different backbone type, with all the interactions operative and with the systematic deletion of geminal, vicinal, and remote and all interactions are shown in Fig. 3.8 and Table 3.6. The "all interaction deletion" implies the deletion of all the acceptor orbitals, thus resulting in the operations of no hyperconjugation interactions in the molecule. The conformer ordering one obtains following the deletion of all interaction may be considered to result purely from steric interactions.

With all the interactions present, the energy ordering of various conformers observed is $GLY1_{II}$, $GLY2_{I}$, $GLY3_{III,}$ and $GLY4_{IV}$, presented in the increasing order of energy; thus suggesting that backbone type II conformers is most preferred orientation. The deletion of only geminal interactions (vicinal and remote retained), swaps the energy ordering between type I and II conformers, resulting in type I conformer to be the most stable. Backbone type III and IV continue to remain higher in energy compared with type I and II, as shown in Fig. 3.8 (b). After the deletion of vicinal interaction (geminal and remote retained) the conformer order changes and backbone type II structure (which was most stable, to begin with) becomes the least stable and backbone type I gets the most stable. Hence, the vicinal interactions have an essential role to play in determining the stability of backbone type II conformers. It must be mentioned here that stability of backbone type II conformers in case of L-threonine and few other amino acids was attributed to the presence of an intramolecular bifurcated($C=O^{\cdots}NH_2$) hydrogen bonded interaction. [64,82,87,88,140] However, our NBO results suggest that the stability of backbone type II conformers arises due to vicinal orbital delocalization interactions and not due to

bifurcated($C=O^{\cdots}NH_2$) hydrogen bonding. The deletion of remote orbital interactions also changes the relative energy ordering and the backbone type I conformers becomes the highest energy conformer while maintaining the backbone type II as the most stable, as shown in Fig. 3.8 (d). As discussed in the earlier section that the backbone type I conformer exhibit an intramolecular hydrogen bond involving the O-H and NH_2 moiety. Therefore, it is reassuringly consistent, that the deletions of remote orbital interactions erase the hydrogen bonded interactions, thus resulting in the destabilization of type I structure. After the removal of all the interactions, the backbone type II remains the most stable structure, as shown in Fig. 3.8 (e).

Hence, the backbone type II conformers are stabilized due to vicinal orbital delocalization interaction, whereas the type I conformers are stable due to remote orbital delocalization interactions. A similar NBO study was also carried out with the lowest energy conformers of each backbone type of L-threonine, i.e., $THR1_{II}$, $THR2_{I}$, $THR4_{III,}$ and $THR11_{IV}$. The role of orbital delocalization interactions in the stability of various backbone type conformer of L-threonine was found to be the same as observed in glycine. Thus, one may conclude that the conformational preferences in these amino acids are governed by a complex interplay of hydrogen bonding interaction and vicinal orbital delocalization interactions.

3.5 Discussions

Alonso and co-worker have reported the conformational study of L-threonine using Molecular Beam Fourier Transform Microwave Spectroscopy (MB-FTMW), and seven different conformers were observed in the study.[64] The study was conducted by laser ablating the solid sample of L-threonine, and the laser ablated species formed were introduced to the microwave cavity by mixing it with a carrier gas. The seven conformers seen in their study were Ia, IIb, I'b, $III_{\alpha}a$, IIc, IIa, and $III_{\beta}b$, mentioned in the decreasing order of energy. These conformers match

with our structures for $THR1_{II}$, $THR2_{I}$, $THR3_{II}$, $THR4_{III}$, $THR5_{I}$, $THR7_{I}$, and $THR8_{II}$ respectively, both about the structure as well as energies. The structural agreement of our conformers with those reported by Alonso and co-worker is corroborated by the excellent agreement of the computed rotational constant of conformers at MP2/6-311++G(d,p) level of theory, between our values and those reported by Alonso.

In their microwave study, Alonso and coworker suggested the stability of the conformers Ia ($THR1_{II}$) and I'b ($THR3_{II}$) to the presence of two intramolecular hydrogen bonded interactions, one between COOH and NH_2 moiety (bifurcated $C=O^{\cdots}NH_2$ hydrogen bond) and other involving the side-chain hydroxyl group and nitrogen lone pair ($O\text{-}H_{sc}^{\cdots}N$ hydrogen bond). But the AIM analysis performed in this study, we did not find any bond critical points along bifurcated $C=O^{\cdots}NH_2$ hydrogen bond interactions, as discussed in an earlier section. However, in the conformer $THR1_{II}$ hydrogen bonded interactions due to the side-chain hydroxyl group and nitrogen lone pair was seen, but no interaction due to $O\text{-}H_{sc}$ was identified in case of conformer $THR3_{II}$. The stability of conformers IIb ($THR2_{I}$), IIc ($THR5_{I}$) and IIa ($THR7_{I}$) were suggested to be due to the presence of an $O\text{-}H^{\cdots}N$ hydrogen bond (between hydroxyl of COOH and nitrogen lone pair), which we also observed in our study. The stability of conformers $III_{\alpha}a$ ($THR4_{III}$)and $III_{\beta}b$ ($THR8_{III}$) was attributed to the presence of two hydrogen-bonded interaction, one involving COOH and NH_2 moiety (bifurcated $C\text{-}O^{\cdots}NH_2$ hydrogen bond) and another due to $O\text{-}H_{sc}$. While the above suggestions appear to be plausible, but AIM analysis for these conformers does not show bond critical point along bifurcated $C\text{-}O^{\cdots}NH_2$ H-bond path. The conformer stability in both the structure results due to electronic orbital delocalization interaction, as discussed in the earlier sections of our work.

While we observed four lowest energy conformers of L-threonine in our study, Alonso and co-worker had seen seven conformers in their study, in which the sample was laser ablated before introducing to the microwave cavity. It is probable that the transient temperature in the laser ablation process may have resulted in a significant population to pool in the higher energy conformers as compared to the resistive heating method employed in our study, where the computed populations of next three conformers were found to be less than 6 % (Table 3.2).

To pictorially summarize the 38 conformers of L-threonine obtained in our study based on their structure and energetics, we presented all the conformers in a dartboard fashion, as shown in Fig. 3.9, which we called as *"conformational dartboard."* In the dartboard plot, eight tracks at varying radii represent the eight backbone structures and all the conformers belonging to same backbone type fall on the corresponding track. The angular position (θ) of conformers in this plot depends on the energy of the conformers. Starting from the *X-axis*, which represents the zero of energy, the full 360º is taken to represent the 10.0 kcal mol^{-1}, which represent the limit for the highest energy conformer for L-threonine. Thus the angular position (θ) corresponding to any conformer is given by [**ΔE** (energy of conformer relative to the lowest energy conformer) **x** (360/10)]. All the conformers with the same backbone structure but with varying side-chain orientations will lie on the same track but at different angular locations. As can be seen from the conformational dartboard, shown in Fig. 3.9, the first quadrant which represents the conformers with the relative energy up to 2.5 kcal mol^{-1} from the lowest energy form (which is located on X-axis; θ= 0.0), is populated by structures lying on the track corresponding to backbone type I, II and III. These backbone structures filled only the first and second quadrants, thus suggesting that these conformers correspond to the lower energy forms. The quadrant 3 and 4 are populated by

the tracks corresponding to the backbone type IV, I' II', III' and IV', which therefore represent the higher energy forms.

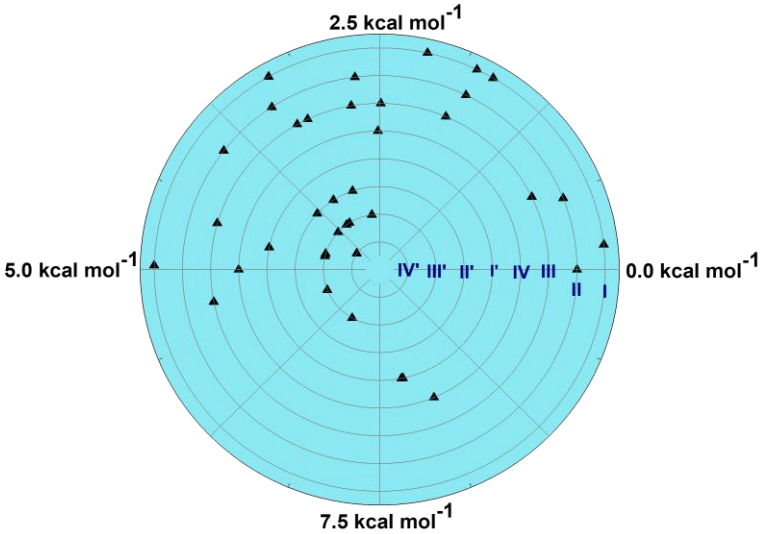


Fig. 3.9 The conformational dartboard for 38 conformers of L-threonine at MP2/6-311++G(d,p) level of theory

3.6 Conclusions

The conformational landscape of amino acid L-threonine was understood employing matrix isolation FTIR spectroscopy and *ab initio* QM computations. The L-threonine vapor produced by heating the sample at 407 and 423 K were swept by flowing argon on to the KBr substrate, maintained at 12 K. The potential energy surface scan yielded 38 different conformers of L-threonine, at M06-2X/6-311++G(d,p) and MP2/6-311++G(d,p) level of theories. These 38 conformers were classified into eight classes based on their backbone structure. The lowest energy conformers were found to adopt backbone type I, II and III structure, while the high energy forms adopted backbone type IV, I', II', III' and IV' structures. The spectral features

observed in experiments were found to be in good agreement with the computed scaled wavenumbers of four lowest energy conformers viz. $THR1_{II}$, $THR2_{I}$, $THR3_{III}$ and$THR4_{III}$.

The backbone type I conformers manifested intramolecular hydrogen bonded interaction, whereas the type II, III, and IV conformers only showed side-chain and not the N-H···C=O hydrogen bonding interactions, as observed in our AIM analysis. The NBO analysis suggested that orbital delocalization interactions play an essential role in determining the conformational landscape of L-threonine. In summary, the conformational preferences were observed to be governed by the complex interplay of hydrogen bonding interactions and orbital delocalization interactions, particularly the vicinal and remote orbital delocalization interactions.

We have also demonstrated that there exists a pattern in the backbone structure adopted by the low energy conformers of L-threonine, wherein the backbone type I, II and to a lesser extent III are preferred over other backbone types. The data has been represented in the form of a conformational dartboard plot. A general rationalization of the conformational landscape of amino acids is useful in understanding its role in the structure of proteins.

Chapter 4

Conformational Picture of L-Glutamic Acid

4.1 Introduction

L-glutamic acid is the most abundant excitatory neurotransmitter in the vertebrate central nervous system and plays a significant role in the neural differentiation of developing the brain. Glutamic acid also serves as the precursor for the synthesis of inhibitory gamma-aminobutyric acid (GABA) in neurons.[141] Furthermore, the structure and function of proteins depend upon the nature of amino acid residue and the conformation of side-chain groups.[116] Therefore, the knowledge of low energy conformers of glutamic acid can be useful to understand the nature of interactions between glutamic acid and specific receptors.

The structure of amino acids, in solid and solution phase, has been extensively studied in the past.[41-48] The Zwitterionic or doubly charged form plays a vital role in the stability of conformers in the solid and solution phase. Gas phase conformation analysis is useful in understanding the molecular geometry in a neutral condition and in the absence of intermolecular interactions occurring in a condensed medium. The conformational landscape of α-amino acids, in the gas phase, has been explored using various experimental techniques and *ab initio* quantum computational method. Matrix isolation infrared spectroscopy, owing to its advantage of providing sharp spectral features, have been widely used in the past, to study conformations of α-amino acids, such as glycine[78,79], alanine,[80,81] serine,[83,84,126] lysine,[88] isoleucine,[87] leucine,[89] cysteine,[90,142] tyrosine,[92] proline[118,119] valine,[117] and threonine,[143].

Navarrete *et al.*, reported conformations of L-glutamic acid using semi-empirical and *ab initio* computation at HF/4-31G level of theory.[144] Gas phase acidity and basicity of glutamic acid, using *ab initio* methods, has also been discussed in the past.[145,146] In the experimental study,

Alonso and co-workers observed five conformers of L-glutamic acid, using rotational spectroscopy.[140] The stability of low energy conformers was attributed to the presence of different types of intramolecular hydrogen bonded interactions. For molecular conformations, which are believed to be stabilized by non-covalent interactions such as hydrogen bonding, infrared spectroscopy can provide evidence for the presence of such interactions. The conformational study of glutamic acid using infrared spectroscopy has not been reported in the past, and therefore, is presented in this chapter.

Furthermore, as discussed in the previous chapter, the lower energy conformers of L-threonine adopted only specific backbone structures, as seen in conformational dartboard analysis, and the polar side-chain group of L-threonine was found to play only a supportive role in determining the conformational preferences.[143] Therefore, to explore the generality to above trend, i.e., low energy conformers preferring only specific backbone structures and to understand the role of the acidic side-chain in deciding the conformational preferences in α-amino acid, the conformational study of L-glutamic acid was done.

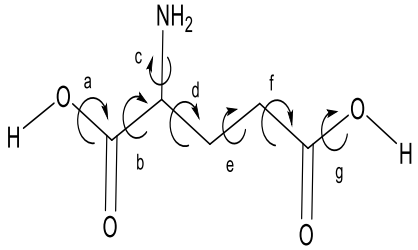
4.2 Experimental Details

The matrix isolation FTIR studies were carried out employing a closed cycle helium compressor-cooled cryostat (HC-4E1). The experimental setup and vacuum system have been described in details in chapter 2. A heated nozzle kept close to the cold substrate and maintained at ~388 K, and ~378 K was employed for sample deposition. The vapors of L-glutamic acids, thus produced were swept by a stream of argon gas (Grade-I, Sigma Gases, and Service, 99.999%), and was deposited on to the cold substrate (KBr window) which was maintained at 12 K. The standard flow rate for deposition in all the experiments was ~1.5-2.0 mmol/hr. The FTIR spectra of matrix isolated L-glutamic acids were recorded using IR spectrometer (Bruker Tensor-27), operating at

a resolution of 0.5 cm^{-1}. After spectrum recording at 12 K, the matrix was annealed by raising the temperature to 30 K for an hour and then cooled back to 12 K. The spectra of annealed spectra were also recorded.

4.3 Computational Details

All the calculations were performed using Gaussian 09 package.[107] The potential energy surface (PES) of L-glutamic acid was scanned to obtain possible conformers. The initial geometries for PES scan were generated by the systematic rotation of backbone and side-chain dihedral angles of L-glutamic acid, as shown in Fig. 4.1. The 2592 initial geometries of L-glutamic acid were then optimized at HF/6-31G level of theory, and the exercise yielded 394 stationary points. These 394 stationary points were then used as starting geometry for optimization at B3LYP/6-311G level of theory. The 317 stationary points, obtained at B3LYP/6-311G level, were then refined at M06-2X and MP2 level of theory, employing a 6-311++G(d,p) basis function.



a = 2 fold rotation ; b = 6 fold rotation
c = 3 fold rotation ; d = 3 fold rotation
e = 3 fold rotation ; f = 4 fold rotation
g = 2 fold rotation
Total 2x6x3x3x3x4x2 = 2592 initial geometries

Fig. 4.1 Systematic dihedral rotation of L-glutamic acid for initial geometry generation

The frequency calculation was done at M06-2X and MP2 level, using analytical gradients. At M06-2X and MP2 level, based on the frequency analysis of stationary points, 299 and 267 conformers, respectively, of L-glutamic acid were obtained, and each conformer shows real frequency for all the normal modes. The unique stationary points at every stage of PES scan calculations were extracted by comparing rotational constants of optimized geometries and excluding the redundant files. The various conformers of L-glutamic acids were then classified

by their backbone structure into eight types, i.e., I, II, III, IV, I', II', III' and IV' as discussed in the last chapter and shown in Fig. 4.2. Superimposed on these eight backbone structures are the slight differences in side-chain orientations, which results in a large number of conformers. The optimized geometry of 10 lowest energy conformers of L-glutamic acid computed at MP2/6-311G++G(d,p) level is shown in Fig. 4.3. The relative energies of these conformers at MP2/6-311G++G(d,p) and M06-2X/6-311G++G(d,p) are also shown.

The Roman numeral subscript in the nomenclature of each conformer indicates the backbone type the conformer belongs. For example, $GLU1_{II}$ suggest that conformer GLU1 belongs to backbone type II. The relative energies of 20 low energy conformers, comparable to that of $GLU1_{II}$, population distribution, and dipole moment are listed in Table 4.1. The free energy of the conformers relative to the lowest energy conformers was also computed at the temperature of sample vaporization (388 K), using MP2/6-311G++G(d,p) and are shown in Table 4.1. The barriers for conformer interconversion between the low energy conformers were computed at the MP2/6-311G++G(d,p) level of theory. The transition states corresponding to a conformer pair were also verified with intrinsic reaction coordinates calculation.

The computed vibrational wavenumbers of low energy conformer were scaled before assigning to the experimental features. Since the matrix perturbs the different spectral mode differently,[105] a mode-by-mode scaling factor was used, as has been done in many of our previous works.[128-138,143] The computed wavenumbers in the range of 2000-400 cm^{-1} were scaled using a factor of 0.9643. This scaling factor was obtained by assigning the experimental feature observed at 1751.9 cm^{-1} to the computed feature observed at 1816.8 cm^{-1}. Similarly, a scaling factor of 0.9347, which was computed by assigning the experimental feature at 3545.2 cm^{-1} to

the calculated feature at 3792.8 cm^{-1}, was used to scale the computed wavenumbers in the region of 4000-2000 cm^{-1}.

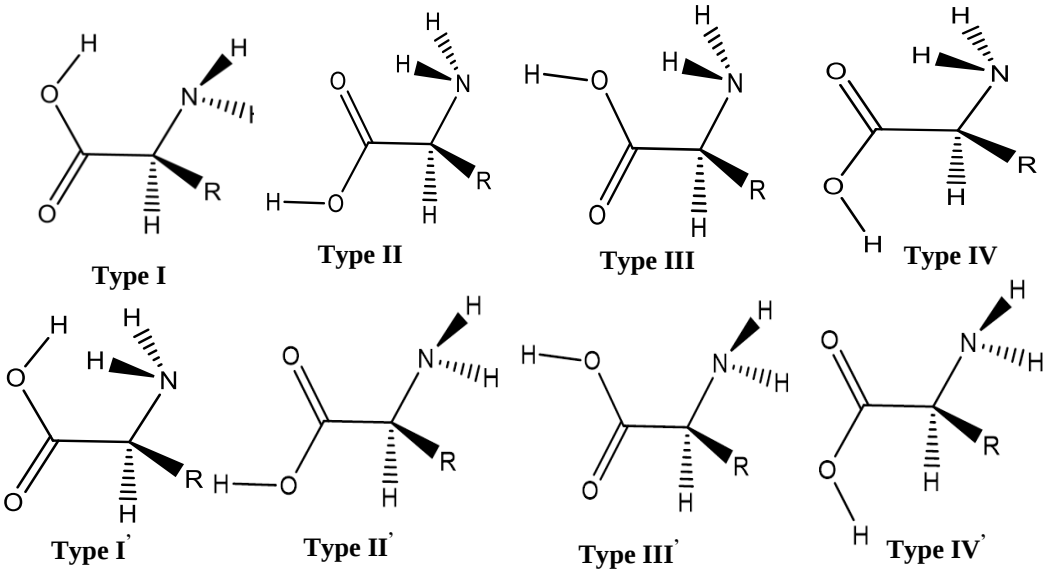


Fig. 4.2 Eight backbone structures of α-amino acid

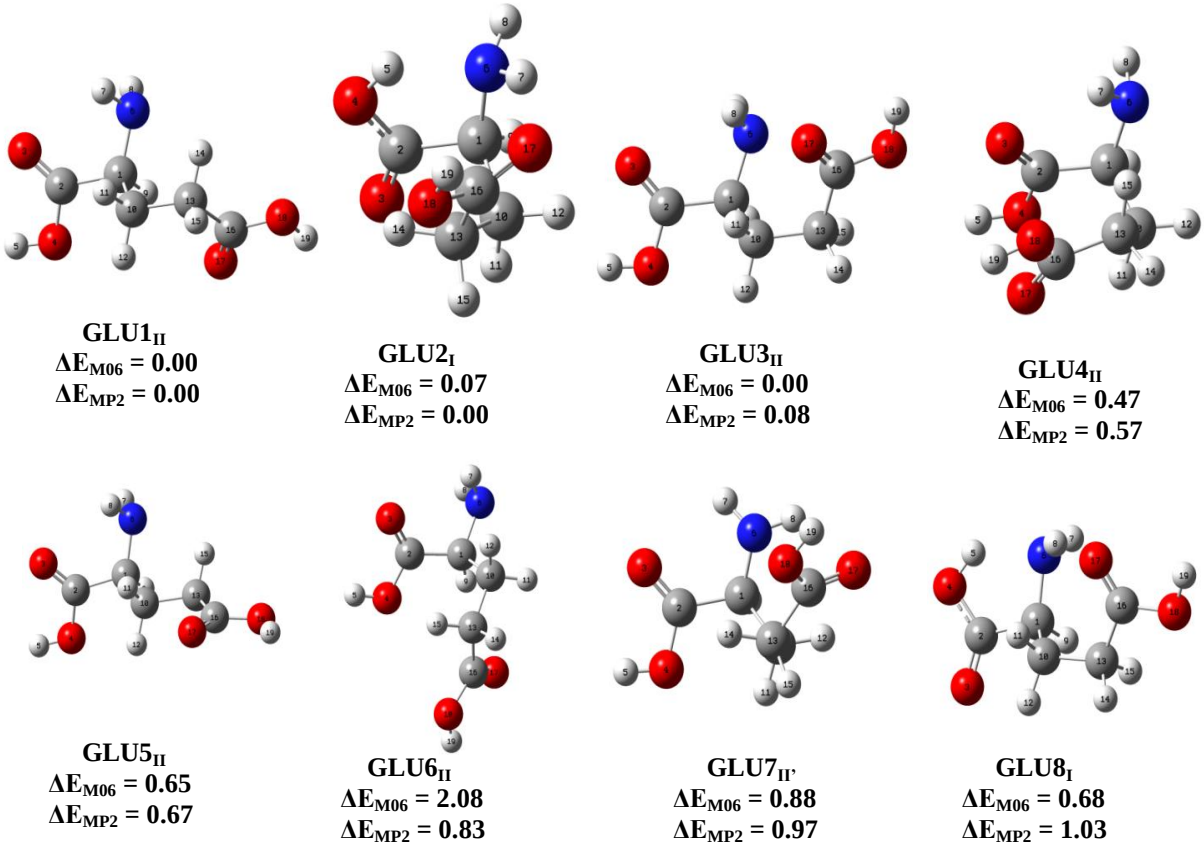


Fig. 4.3 Optimized geometry of 8 lowest energy conformers of L-glutamic acid at MP2/6-311++G(d,p) level of theory. The relative energy of conformers at M06 and MP2 level with 6-311++G(d,p) basis set, are also shown. The energies are relative to conformer $GLU1_{II}$ and are in kcal mol^{-1}

Table 4.1 Relative zero-point corrected energies of the conformers of L-glutamic acid computed at M06-2X/6-311++G(d,p) ($\Delta E_{M06\text{-}2X}$) and MP2/6-311++G(d,p)(ΔE_{MP2}) levels. The relative free energies at 388 K (sample deposition temperature) at the MP2/6-311++G(d,p)(ΔE_{MP2}) levels are also shown. All energies are in kcal mol^{-1} and are relative to the energies of the conformer $GLU1_{II}$. Dipole moments (D) of all the conformers at MP2/6-311++G(d,p) level of theory shown in the last column[f]

Conformer	$\Delta E_{M06\text{-}2X}$	ΔE_{MP2}	ΔG_{MP2}	%P $_{(\Delta EMP2)}$	%P $_{(\Delta GMP2)}$	Dipole moment[f]
$GLU1_{II}$	0.00	0.00	0.00	11.21	22.67	4.5
$GLU2_{I}$	0.07	0.00	1.41	11.17	3.18	2.8
$GLU3_{II}$	0.00	0.08	0.99	9.99	5.74	0.9
$GLU4_{II}$	0.47	0.57	1.73	5.06	2.05	5.2
$GLU5_{II}$	0.65	0.67	0.57	4.41	10.23	1.0
$GLU6_{II}$	2.08	0.83	1.59	3.52	2.47	5.6
$GLU7_{II'}$	0.88	0.97	2.17	2.92	1.10	5.2
$GLU8_{I}$	0.68	1.03	1.74	2.67	2.00	0.8
$GLU9_{III}$	1.16	1.04	1.92	2.63	1.56	6.5
$GLU10_{II}$	1.16	1.07	1.91	2.52	1.58	1.9
$GLU11_{II}$	3.11	1.11	0.72	2.40	8.27	6.5
$GLU12_{II}$	3.88	1.14	2.46	2.30	0.74	0.7
$GLU13_{III}$	1.15	1.21	2.69	2.09	0.54	2.1
$GLU14_{I}$	1.35	1.22	2.48	2.04	0.72	2.4
$GLU15_{III}$	0.92	1.25	3.28	1.98	0.23	4.6
$GLU16_{I}$	0.93	1.25	3.64	1.95	0.14	2.8
$GLU17_{II'}$	1.23	1.27	1.62	1.91	2.38	2.3
$GLU18_{I}$	0.95	1.31	2.03	1.80	1.33	6.4
$GLU19_{II}$	1.82	1.34	1.15	1.73	4.60	0.7
$GLU20_{III'}$	0.92	1.37	2.58	1.67	0.62	2.8

All energies in kcal mol^{-1}

The electron density topology analysis, using AIM 2000 program suite[110] was performed to locate the intramolecular hydrogen bonding interactions in different conformers of L-glutamic acid. The wave function computed at MP2/6-311G++G(d,p) level of theory was used to construct the electron density topology. The strength of hydrogen bonding interaction was estimated using the topological parameters such as electron density and Laplacian of electron density, at bond critical points, as explained by Espinosa *et al.*[113]

4.4 Results

4.4.1 Conformer's relative energy and population

The *ab initio* computed low energy conformers of L-glutamic acid at the MP2/6-311G++G(d,p) level of theory are shown in Fig. 4.3. The conformer $GLU1_{II}$, $GLU2_{I}$, and $GLU3_{II}$ are nearly similar zero-point corrected energy (ΔE_{ZPC}), both at MP2 and M06-2X level of theory as shown in Table 4.1. Interestingly, 14 out of 20 low energy conformers of L-glutamic acid, listed in Table 4.1, adopt backbone type II or I structure. The tendency of low energy conformers of L-glutamic acid to prefer only specific backbone structures was also observed in the case of L-threonine, as discussed in the last chapter.

Our experiments were performed by maintaining the sample at 388 K to obtain the sufficient vapor pressure of the L-glutamic acid. Therefore the thermodynamic parameters of all the conformers were also computed at 388 K (sample heating temperature) employing MP2/6-311G++G(d,p) level of theory. The relative free energy of low energy conformers, at 388 K (ΔG_{MP2}) are listed in Table 4.1. The addition of a thermal correction factor resulted in significant variation in the relative energy of the conformers, as seen, especially in the case of conformer $GLU2_{I}$. The relative free energy ΔG_{MP2} of the conformer $GLU1_{II}$ and $GLU2_{I}$, which are isoergic based on ΔE_{ZPC}, however after thermal correction, ΔG_{MP2} between two increases to 1.4 kcal mol^{-1}, as shown in Table 4.1. Based on the ΔE_{ZPC} and $\Delta G_{MP2,}$ the gas phase populations of low energy conformers were estimated and are listed in Table 4.1. The relative abundances of the three lowest energy (ΔE_{ZPC}) conformers $GLU1_{II}$, $GLU2_{I}$ and $GLU3_{II,}$ are 11.21%, 11.17 %, and 9.99%. These three conformers, which have populations more than ~10 % are therefore possible candidates to be observed in experiments.

4.4.2 Conformer interconversion

The matrix isolation study of L-glutamic acid was done using an effusive beam, and thus, no conformational cooling is expected during the process of sample deposition. Hence, the conformational population in the gas phase before the deposition is expected to be frozen on the matrix. However, it can be argued that even after the deposition, conformer may relax to the low energy conformers, during the process of matrix annealing. It has been recognized that the conformational relaxation to the lower energy structure is possible during the annealing of the matrix at 30 K, provided the barrier for interconversion is less than ~5 kJ mol^{-1} (<1.2 kcal mol^{-1}). [80,143]

The conformer $GLU1_{II}$ and $GLU3_{II}$ has similar backbone geometry and varies only in terms of the side-chain orientation. The geometry of the side-chain group of L-glutamic acid can be defined based on the relative orientation of three dihedral angles, i.e. (N-C_α-C_β-C_γ), (C_α-C_β-C_γ-C_δ) and (C_β-C_γ-C_δ-O_{sc}), where O_{sc} is oxygen atom of side-chain carbonyl group, as shown in Fig.4.4. The side-chain dihedral angles of conformer $GLU1_{II}$ are -56.2 (gauche - or G^-), -67.0 (gauche - or G^-) and 174.4 (trans or *T*) respectively and hence, the side-chain orientation of $GLU1_{II}$ is "G^-G^-T." Likewise, in the case of $GLU3_{II,}$ the side-chain is in "G^-GT" orientation. The conformer $GLU1_{II}$ and $GLU3_{II}$, which have the same backbone geometry of type II, differ only by the relative orientation of C_α-C_β-C_γ-C_δ dihedral angle. The interconversion of the conformer $GLU3_{II}$ to $GLU1_{II}$ occur via rotation around the C_α-C_β-C_γ-C_δ dihedral angles, and the energy barriers separating these two structures, computed at MP2/6-311++G(d,p) level of theory, was found to be 5.5 kcal mol^{-1}.

The relaxation of the $GLU3_{II}$ to $GLU2_I$ requires rotation around multiple bonds, as both the conformers have different backbone geometry as well as side-chain orientation. For an

interconversion to occur via rotation around multiple bonds ($GLU3_{II}$ to $GLU2_{I}$), the energy barrier for the conversion is likely to be very high and thus, are improbable to happen in the matrix, during annealing. The energy barriers of interconversion between different conformer pairs, which differ from each other by a single dihedral angle orientation, are shown in Fig. 4.5. As it can be, the barriers separating every case are larger than 5kJ mol^{-1} (1.2 kcal mol^{-1}). Therefore, conformers relaxation to lower energy form does not occur, during the matrix annealing at 30 K, and it is the conformational distribution present in the gas phase that is being trapped are present in our spectra.

4.4.3 IR spectra of matrix-isolated L-glutamic acid

The matrix isolated IR spectra of L-glutamic acid trapped in an argon matrix is shown in Fig. 4.6 and 4.7. The spectra have been shown over the region of 3600-3100 cm^{-1} (Fig. 4.6, grid a) and 1850-1700 cm^{-1} (Fig. 4.6, grid c) corresponding to the O-H stretching region and C=O stretching region respectively. The spectral features in the O-H stretching region were observed at 3545.2, 3470.0, 3406.9 and 3224.8 cm^{-1}. The strongest spectral feature in the O-H region at 3545.2 cm^{-1} is unusually broader as compared to the spectral features usually observed in the matrix isolation experiments. However, a closer analysis of this spectral feature shows the presence of doublet and shoulders, as shown in Fig. 4.6 (grid b). The C=O stretching region shows the spectral wavenumbers at 1779.7, 1751.9, 1744.6, 1727.8 and 1719.8 cm^{-1}. The spectral features corresponding to the skeletons vibration region, i.e., 1450-1000 cm^{-1} are shown in Fig. 4.7.

The onset temperature for thermal decomposition of L-glutamic acid is reported to be 473 K, and NH_3 and H_2O were found to be volatile decomposition products, along with some non-volatile decomposition products.[147] In the spectra of L-glutamic acid the features due to H_2O, a ubiquitous impurity was observed, but the intensities of features due to H_2O were not increased

when amino acid was deposited relative to that observed when the matrix gas alone was deposited. The features of matrix isolated NH_3 in an argon matrix is reported to be observed at 3435.0 cm^{-1},[148] no such features are seen in the matrix isolated spectra of L-glutamic acid presented in this study, as shown in Fig. 4.6. The absence of NH_3 feature suggests that no thermal decomposition of L-glutamic acid was occurring in our experiment. Moreover, H^1 NMR, C^{13} NMR, and ESI-MS analysis of the sample, before and after heating, was done to investigate the presence of non-volatile decomposition products. As shown in Fig. 4.8, the H^1 NMR, C^{13} NMR and ESI-MS spectra of L-glutamic acid before and after heating (388 K) are nearly identical; hence, rules out the possibility of sample decomposition, during the experiments.

4.4.4 Vibrational assignments

The O-H and C=O spectral regions of L-glutamic acid have contributions due to the backbone carboxylic group (COOH) as well as due to carboxylic moiety present in the side-chain (CH_2-CH_2-COOH). The three lowest energy conformers of L-glutamic acid, i.e., $GLU1_{II}$, $GLU2_{I,}$ and $GLU3_{II}$ have different backbone type and side-chain orientation and thus varying O-H and C=O stretching vibrations. These vibrational features are important signatures to identify the different conformers trapped in the matrix. The vibrational assignment of experimental features was assisted by generating synthetic spectra, using the scaled vibrational wavenumbers of low energy conformers. The SYNSPEC package was used to generate the synthetic spectra and Lorentzian line shape, and a line width of 1.0 cm^{-1} was employed. The intensity of computed features of various conformers scaled for the relative abundance of the individual conformer. The experimentally seen vibrational wavenumbers along with the computed wavenumbers of three lowest energy conformers together with a spectral assignment are listed in Table 4.2 and 4.3.

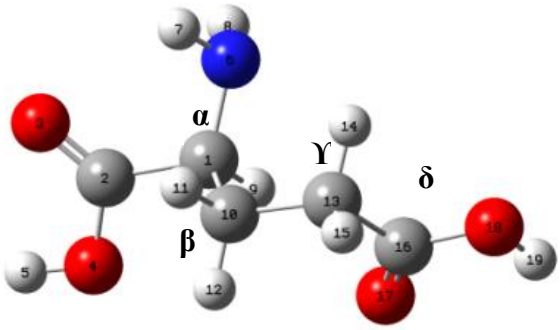


Fig. 4.4 Side-chain orientation of $GLU1_{II}$ in terms of N-C_α-C_β-C_γ = -56.2 (G^-), C_α-C_β-C_γ-C_δ = -67.0 (G^-) and C_β-C_γ-C_δ-O_{sc} = 174.4 (T) dihedral angles. Thus, the side-chain group is in *G^-G^-T* orientation.

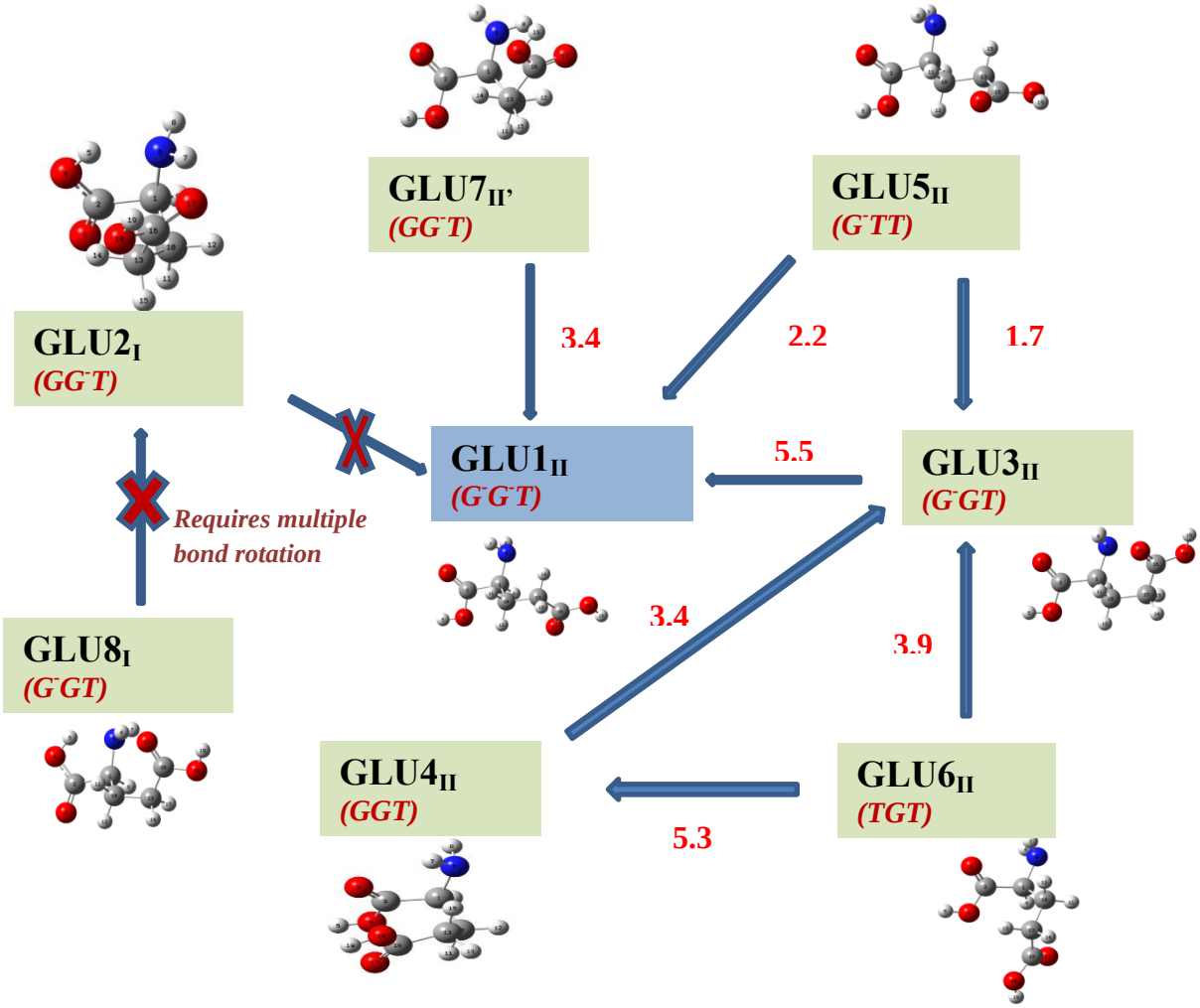


Fig. 4.5 The interconversion barrier for conformers of L-glutamic acid; The three *italic* letters in each conformer (for ex. *(G^-G^-T)* in $GLU1_{II}$) represent the orientation of side-chain dihedral angles, i.e. (N-C_α-C_β-C_γ), (C_α-C_β-C_γ-C_δ) and (C_β-C_γ-C_δ-O_{sc}). The energy barrier separating conformers are (shown in red) calculated at MP2/6-311++G(d,p) level of theory

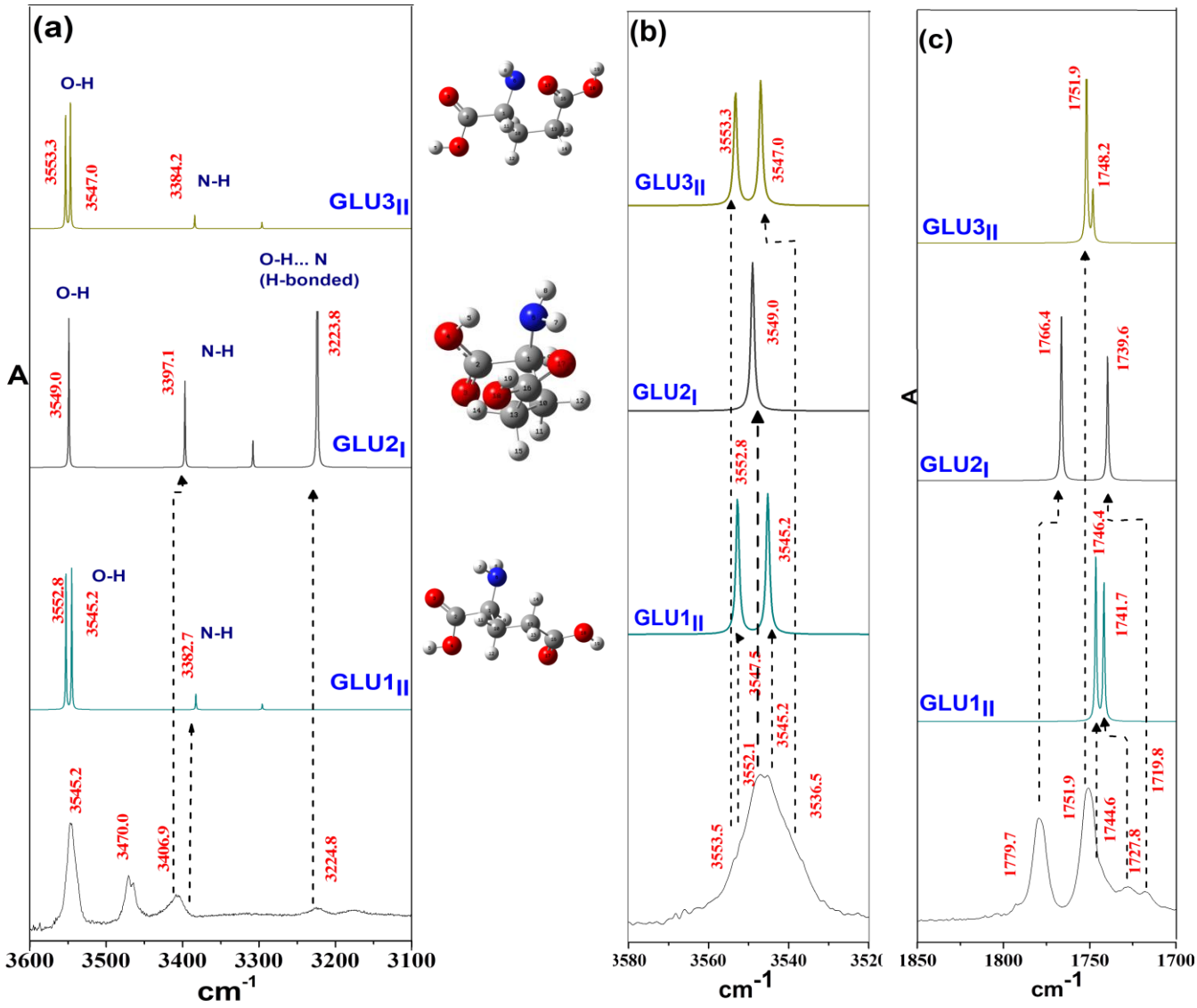


Fig. 4.6 Comparison of experimental and computed spectra of L-glutamic acid. The spectra span the region 3600-3100 cm^{-1} (grid a), 3580-3520 cm^{-1} (grid b) and 1850-1700 cm^{-1} (grid c). i) IR spectra of matrix isolated L-glutamic acid deposited at 388 K; (ii-iv) Computed spectra showing the spectral features of the three lowest energy conformers, calculated at the MP2/6-311++G(d,p) level of theory. The Computed wavenumbers in the grid 'a', 'b' and 'c' are scaled using a scaling factor of 0.9347, 0.9347 and 0.9643 respectively

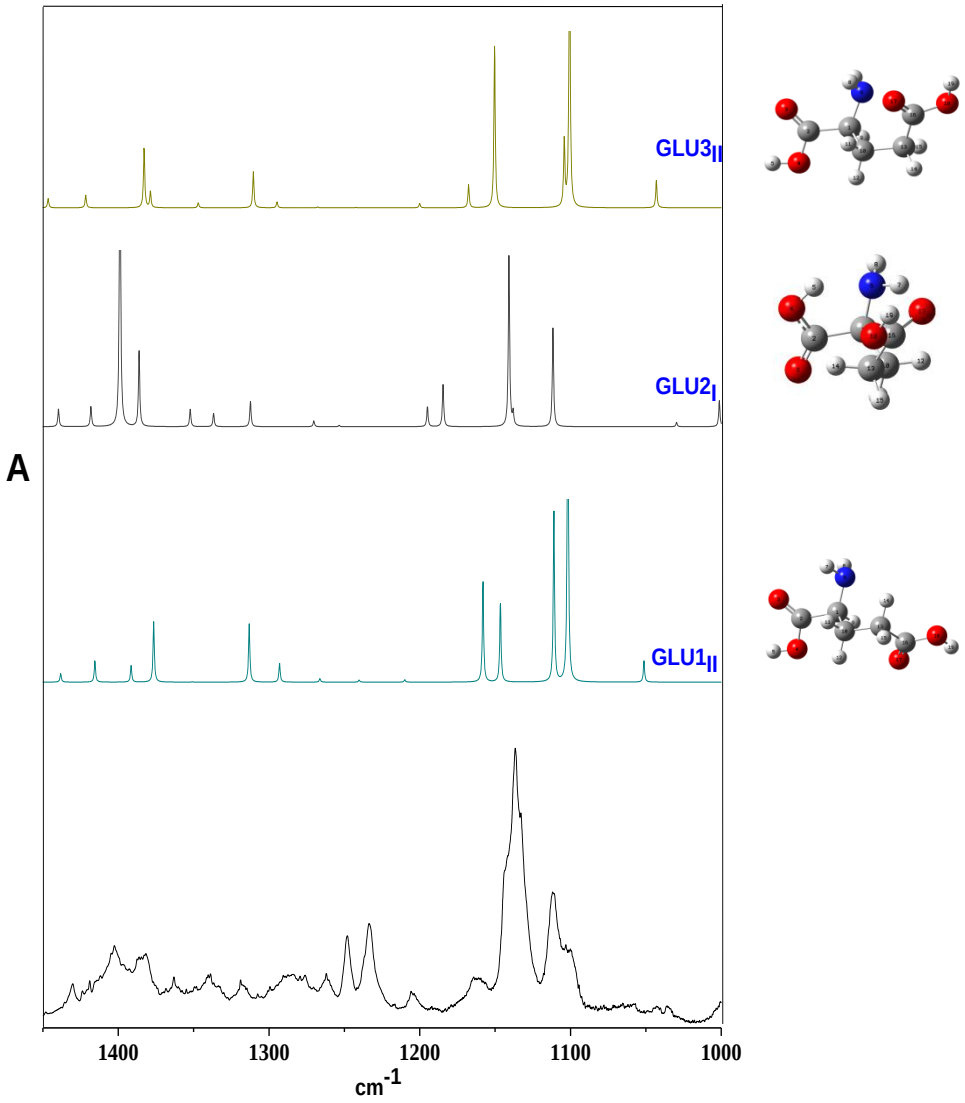
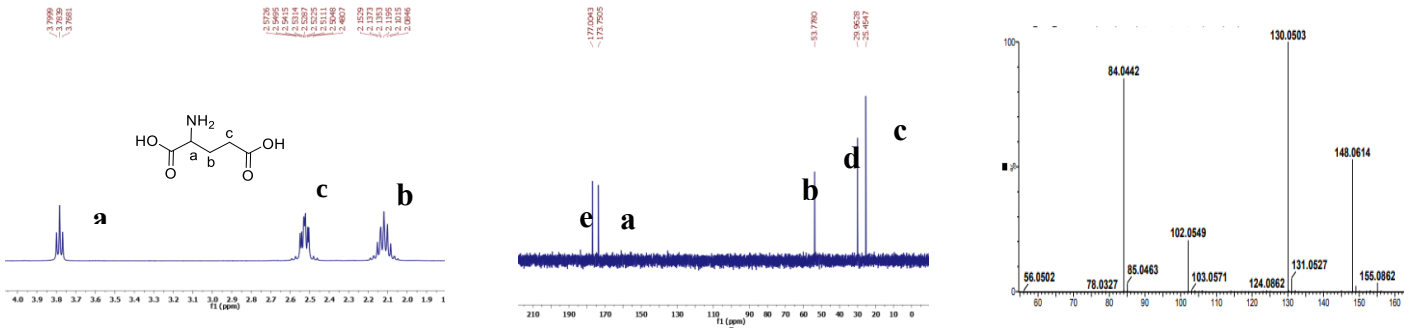


Fig. 4.7 Comparison of experimental and computed spectra of L-glutamic acid. The spectra span the region 1450-1000 cm^{-1}. i) IR spectra of matrix isolated L-glutamic acid deposited at 388 K; (ii-iv) Computed spectra showing the spectral features of the three lowest energy conformers, calculated at the MP2/6-311++G(d,p) level of theory. The computed wavenumbers are scaled using a scaling factor of 0.9643.

Pure sample

H^1 NMR **C^{13} NMR** **ESI-MS**



Sample after heating at 388 K

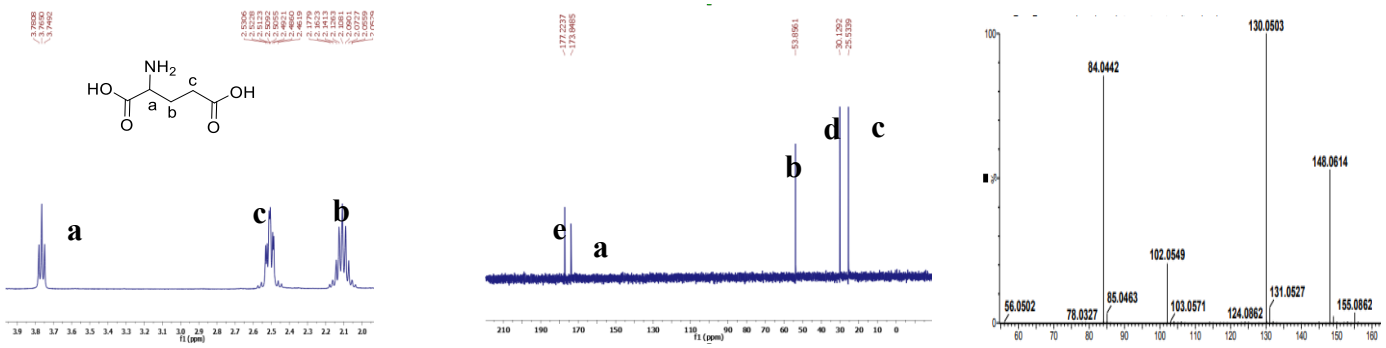


Fig. 4.8 H^1 NMR, C^{13} NMR and ESI-MS spectra of L-glutamic acid, before (top) and after (bottom) sample heating

4.4.4 (a) 3600-3100 cm^{-1} region

The experimental (trace i) and computed spectra of the three lower energy conformers (trace ii-iv) in the 3600-3100 cm^{-1} region are shown in Fig. 4.6 (grid a). The spectral region from 3600-3100 cm^{-1} contain features due to backbone O-H, side-chain O-H and backbone N-H stretching vibrations. The intense feature observed in the experiment occurs at 3545.2 cm^{-1} and shows full-width at half maximum (FWHM) of ~20 cm^{-1}. In the computational study, $GLU1_{II}$, $GLU2_{I,}$ and $GLU3_{II}$ are nearly isoergic, both at MP2/6-311G++G(d,p) and M06-2X/6-311G++G(d,p) level of theory, as shown in Table 4.1. Therefore, the strongest feature seen in the experiment could be assigned to any of the three conformers, i.e., $GLU1_{II}$, $GLU2_{I,}$ or $GLU3_{II}$. When the strongest computed feature of $GLU2_I$ is assigned the strongest experimental feature at 3545.2 cm^{-1}, and a scaling factor based on this assignment is calculated (which work out to be 1.028), the scaled computed wavenumbers of other conformers shows unsatisfactory assignments with the experimental features. Likewise, the scaling factor obtained after assigning the strongest computed feature of conformer $GLU3_{II}$ to the experimental feature at 3545.2 cm^{-1} yields inadequate assignment for the other conformers. However, when the strongest computed feature of the conformer $GLU1_{II}$ is assigned to the experimental wavenumber at 3545.2 cm^{-1}, and a scale factor is calculated based on it, the scaled computed features of other conformers provide agreeable assignments for all the experimentally observed features. The scaling factor based on the above assignment turned out to be 0.9347 and was used to scale all the computed features in the range of 3800-2000 cm^{-1}.

While, the structure and frequency computations both at M06-2X and MP2 level, employing the 6-311G++G(d,p) basis set, shows a nearly similar assignment, the mean square deviation between the experimental and computed features at M06-2X level was found to be

higher than that at MP2 level. The vibrational assignment between the experimental features and the computed wavenumbers, based the conformational abundances obtained with relative free energy conformers ($\%P_{(\Delta GMP2)}$) was also carried out. However, the assignment between the experimental and computed features was found to be inadequate. Hence, calculated frequencies at MP2 level and the relative abundance of conformer obtained with ΔE_{ZPC} provides a better assignment with the experimental features and therefore, has been used.

Based on the assignment of the feature at 3545.2 cm^{-1} to the O-H stretch of $GLU1_{II}$, the following assignments were made, using the computed values at the MP2/6-311G++G(d,p) level of theory, shown in Table 4.2. The computed scaled spectral features of υ(O-H) and $\upsilon(O\text{-}H_{sc})$ for $GLU1_{II}$ (3545.2 and 3552.1 cm^{-1}) and $GLU3_{II}$ (3547.0 and 3553.5 cm^{-1}) are found to be very closely placed, as shown in Fig. 4.6 (grid 1, trace ii, and iv). However, in the case of $GLU2_{I,}$ the υ(O-H) and $\upsilon(O\text{-}H_{sc})$ was computed to be 3223.8 and 3549.0 $cm^{-1,}$ respectively (trace iii). The large red-shift of nearly 320 cm^{-1} observed in the backbone O-H stretching, is due to the presence of intramolecular hydrogen bonding. This red-shifted computed feature is in excellent agreement with the experimental feature at wavenumber 3224.8 cm^{-1} (trace i).

The experimental feature at 3545.2 cm^{-1} is observed to be unusually board, with an FWHM of ~ 20 cm^{-1}. A closer analysis of this feature shows the presence of doublet features at 3547.5, 3445.2 cm^{-1}, along with few shoulders at 3553.5, 3552.1 and 3536.5 cm^{-1}, as shown in (grid 2, trace i). The scaled computed vibrational wavenumber of υ(O-H) and $\upsilon(O\text{-}H_{sc})$ mode of $GLU1_{II}$, and $GLU3_{II}$ are very found to be closely spaced, as mentioned earlier. Therefore, the doublet and multiple shoulders features observed in the above experimental feature can be due to these closely spaced computed features. The feature at 3545.2 and shoulder at 3552.1 cm^{-1} (grid b, trace i) can be assigned to the computed scaled υ(O-H), and $\upsilon(O\text{-}H_{sc})$ features of $GLU1_{II}$ at

3545.2 and 3552.8 cm^{-1} respectively (grid b, trace ii). Similarly, the υ(O-H) and υ($O-H_{sc}$) mode of $GLU3_{II}$ computed at 3547.0, and 3553.3 cm^{-1} (grid b, trace iv) are in agreement with the experimental feature at 3536.5 and 3553.5 cm^{-1} respectively (grid b, trace ii). The υ($O-H_{sc}$) of $GLU2_I$ is computed to be at 3549.0 cm^{-1} (grid b, trace iii) can be assigned to the experimental feature at 3547.5 cm^{-1}(grid b, trace i). Thus, the closely spaced computed features of different conformers are in good agreement with the many shoulders observed in the broad spectral feature at 3545.2 cm^{-1}. It is probably due to the presence of these closely spaced features that the experimental feature observed at 3545.2 cm^{-1} was found to be unusually broad.

The experimental features observed at 3409.8 and 3405.2 cm^{-1} (grid a, trace i) are in good agreement with the computed scaled wavenumber at 3397.1 cm^{-1} (grid a, trace iii) and 3382.7 cm^{-1}(grid a, trace ii), corresponding to the υ(N-H) mode of conformer $GLU1_{II}$ and $GLU2_I$ respectively. The assignment for the experimental feature at 3470.0 cm^{-1} will be discussed later in the carbonyl stretching discussion section. The computed scaled frequencies corresponding to the C-H symmetric and asymmetric stretch of all three conformers were clustered around 2880-2900 cm^{-1} and 2930-2950 cm^{-1}, respectively. However, owing to the very low infrared absorption cross-sections these modes, the C-H features were not seen in the experiments.

In summary, the most significant difference between the backbone O-H and side-chain $O-H_{sc}$ is found in case of conformer with the backbone type I (i.e., $GLU2_I$), both experimentally and computationally. The conformers with the same backbone type (i.e., $GLU1_{II}$ and $GLU3_{II}$) are found to have closely spaced frequencies both in experiments and in computations. The observation is similar to that seen in case of L-threonine, where also a sizeable spectral shift was observed in backbone type I conformers ($THR2_I$), while the spectral features of backbone type II ($THR1_{II}$ and $THR3_{II}$) were closely spaced.[143] The conformational preference observed in this

study is the same as those reported by Alonso and co-workers, based on their microwave studies.[140]

Table 4.2 Experimental (in argon matrix) and computed scaled vibrational wavenumbers (cm^{-1}) at the MP2/6-311++G(d,p) level for the different conformers of L-glutamic acid. See text for details on scaling

Experimental	Computed scaled[a]			IR intensity[b]	Vibrational modes
	$GLU1_{II}$	$GLU2_{I}$	$GLU3_{II}$		
3553.5			3553.5	75.0	ν (O-H_{sc})
3552.1	3552.8			77.0	ν (O-H_{sc})
3547.5		3549.0		87.3	ν (O-H_{sc})
3545.2	3545.2			79.6	ν (O-H)
3536.5			3547.0	82.4	ν (O-H)
3470.0	3434.9				2ν (C=O)
3409.4		3397.1		52.3	ν (N-H)
3405.2	3382.7			8.8	ν (N-H)
3224.8		3223.8		327.0	ν (O-H)
1779.4		1766.4		299.0	ν (C=O)
1751.9			1751.9	390.0	ν (C=O_{sc})
1744.6	1746.4			271.9	ν (C=O_{sc})
1727.8	1741.7			229.0	ν (C=O)
1719.8		1739.6		217.2	ν (C=O_{sc})

[a] Scaling factors for computed vibrational wavenumbers of L-glutamic acid conformers: 0.9347 for the O−H stretch region (3800-2000 cm^{-1}) and 0.9643 for the C=O stretch and skeleton vibration (2000-400 cm^{-1}). See text for details.
[b] Computed IR intensity(km mol^{-1}) corresponding to each normal mode.

4.4.4 (b) 1850-1700 cm^{-1} region

The carbonyl stretching region (1850-1700 cm^{-1}) of L-glutamic acid shows the contribution from two C=O stretching modes, i.e., backbone C=O_b and side-chain C=O_{sc} and are shown in Fig. 4.6 (grid c) and Table 4.2. The two carbonyl groups, i.e. (C=O_b and C=O_{sc}) of conformer $GLU1_{II}$ and $GLU3_{II,}$ demonstrate coupled vibrational motions in the form of symmetric and antisymmetric stretching. However, in the case of backbone type I conformer ($GLU2_I$), the two carbonyl groups manifest uncoupled and independent C=O vibrations. The computed wavenumber for symmetric and antisymmetric C=O vibrations of $GLU1_{II}$ is found at 1746.4 and 1741.7 cm^{-1} (grid c, trace ii) respectively and are in good agreement with the experimental

features observed at 1744.6 and 1727.8 cm^{-1} (grid c, trace i) respectively. The C=O vibration (antisymmetric) of the $GLU3_{II}$ conformer is computed at 1751.9 cm^{-1} (grid c, trace iv) and matches well with the experimental feature at 1751.9 cm^{-1}. The symmetric C=O stretch of the $GLU3_{II}$ is found at 1748.1 cm^{-1}, very closely spaced to the antisymmetric mode. The carbonyl stretching's of $GLU2_{I}$ were computed to be 1766.4 and 1739.6 cm^{-1} (grid c, trace iii) and are in agreement with the experimentally observed features at 1779.4 and 1719.8 cm^{-1} respectively (grid c, trace i). The broad unassigned feature at 3470.0 cm^{-1} in the *3600-3100 cm^{-1} region* can be possibly assigned to the C=O overtone modes of the glutamic acid conformer. The anharmonic vibrational frequencies computation yielded the overtone wavenumber for conformer $GLU1_{II}$ at 3434.9 and 3427.6 cm^{-1}.

4.4.4 (c) 1450-1000 cm^{-1} region

The experimental and computed features in the spectral region of 1450-1000 cm^{-1} are shown in Fig. 4.7 and Table 4.3. The normal modes vibrations in this spectral region involve skeleton vibrations, deformation, atomic excursions involving bend and torsional motions. The prominent features observed in this region are seen at 1430.3, 1387.5, 1380.8, 1318.6, 1262.0, 1164.1, 1144.3, 1133.4, 1111.6 and 1057.8 cm^{-1} (trace i). The computed scaled wavenumber of conformer $GLU1_{II}$ occurring at 1415.5, 1391.4, 1376.5, 1313.1, 1293.0, 1158.1, 1146.5, 1111.0, 1101.8 and 1051.3 cm^{-1} (trace ii) corroborate well with the experimental features mentioned above. The computed wavenumbers of conformer $GLU2_{I}$ are seen at 1439.7, 1418.1, 1398.9, 1386.2, 1352.3, 1336.8, 1312.3, 1195.0, 1184.5, 1140.8, 1111.6 and 1001.2 cm^{-1} (trace iii) and are in good agreement with the experimental features seen at 1435.9, 1418.8, 1402.8, 1399.9, 1363.2, 1341.2, 1314.5, 1247.6, 1233.4, 1136.0, 1130.2 and 1000.1 cm^{-1} (trace i) respectively.

Table 4.3 Experimental (in argon matrix) and computed scaled vibrational wavenumbers (cm^{-1}) at the

Experimental	Computed scaled[a]			IR intensity[b]
	$GLU1_{II}$	$GLU2_{I}$	$GLU3_{II}$	
1435.9		1439.7		17.0
1430.3	1415.5			19.7
1418.8		1418.1		19.2
1414.8			1421.6	13.8
1402.8		1398.9		306.3
1399.9		1386.2		73.3
1387.5	1391.4			15.4
1382.4			1382.9	64.6
1380.8	1376.5			56.0
1376.0			1378.7	17.3
1363.2		1352.3		17.3
1341.2		1336.8		12.8
1318.6	1313.1			53.8
1314.5		1312.3		24.3
1288.1			1310.4	38.7
1262.0	1293.0			17.4
1247.6		1195.0		19.6
1233.4		1184.5		40.9
1205.6			1167.6	24.6
1164.1	1158.1			93.1
1144.3	1146.5			74.2
1141.3			1150.4	173.6
1136.0		1140.8		164.1
1133.4	1111.0			156.4
1130.2		1111.6		94.5
1123.6			1104.1	72.0
1111.6	1101.8			303.8
1109.7			1100.6	319.1
1057.8	1051.3			19.5
1043.1			1043.1	29.4
1000.1		1001.2		25.1

[a] Scaling factors for computed vibrational wavenumbers of L-glutamic acid conformers: 0.9347 for the O−H stretch region (3800-2000 cm^{-1}) and 0.9643 for the C=O stretch and skeleton vibration (2000-400 cm^{-1}). See text for details.
[b] Computed IR intensity(km mol^{-1}) corresponding to each normal mode.

The other features in experiments are observed at 1414.8, 1382.4, 1376.0, 1288.1, 1205.6, 1141.3, 1123.6, 1100.6 and 1043.1 cm^{-1} (trace i) and are assigned to the conformer $GLU3_{II}$. The computed wavenumber corresponding to the above experimental features is found at 1421.6, 1382.9, 1378.7, 1310.4, 1167.6, 1150.4, 1104.1, 1100.6 and 1043.1 cm^{-1} (trace iv) respectively. Even though the experimental features in this region are in good agreement with the computed wavenumber of three low energy conformers, the spectral complexity of this region makes the

vibrational assignment to be tentative. However, clearly, by spectral assignment in the *3600-3100 cm^{-1} region* and *1850-1700 cm^{-1} region*, we observe three lowest energy conformer of L-glutamic acid, i.e., $GLU1_{II}$, $GLU2_{I,}$ and $GLU3_{II}$ present in our experiments.

4.4.5 Atoms in molecule analysis of the conformers of L-glutamic acid

The conformational preferences of amino acids are suggested to be influenced by the intramolecular hydrogen bonding. Therefore, atoms in molecule (AIM) analysis was performed to investigate the presence of such interactions in conformers of L-glutamic acid, as was done in case of L-threonine discussed in the last chapter. Fig. 4.9a-h shows the AIM analysis of various conformers of L-glutamic acid, done at MP2/6-311++G(d,p) level of theory. The topological parameters such as electron density (**(**$\rho(r_c)$), Laplacian of electron density($\nabla^2\rho(r_c)$) and hydrogen bonding interaction energy (E_{HB}) corresponding to the bond critical points are shown in Table 4.4.

The conformer $GLU2_I$ shows bond critical points along O-H···N bond path (Fig. 4.9b) and thus indicating the presence of an intramolecular hydrogen bond between the O-H moiety of the backbone carboxyl (COOH) and the NH2 group. The presence of hydrogen-bonded interaction in $GLU2_I$ is in agreement with our experimental results, where O-H stretching feature of conformer was seen at significantly red-shifted (3224.8 cm^{-1}) compared to the spectral features of backbone type II conformers ($GLU1_{II}$ and $GLU3_{II}$). Similarly, the conformer $GLU8_I$ also shows the intramolecular hydrogen bond along (O-H···N) bond path, as seen in Fig. 4.9h. The O-H···N hydrogen bonding interactions are unique for backbone type I conformers and lend stability to such conformers. The hydrogen bonding energy of O-H···N interaction in $GLU2_I$ was estimated to be 11.8 kcal mol^{-1}. A weak interaction (E_{HB} 4.0 kcal mol^{-1}) between side-chain carbonyl ($C=O_{sc}$) and NH_2 moiety is also located in conformer $GLU2_I$.

The stability of backbone type II conformers, in earlier research work, was attributed to the presence of a bifurcated ($C=O^{\cdots}H_2N$), involving the carbonyl group of backbone and NH2 group.[64,82,88,87,140] However, the AIM analysis of backbone type II conformer does not locate such bond critical points, as shown in Fig. 4.9 (all, except 4.9b and 4.9h). However, in few conformers, such as $GLU1_{II}$, $GLU4_{II,}$ and $GLU6_{II}$, hydrogen bond due to side-chain group ($COOH_{sc}$) was observed. The AIM analysis results are consistent with our experimental observation, where backbone type I conformers were found to have $O\text{-}H^{\cdots}N$ hydrogen bonding, but, the bifurcated hydrogen bonding in backbone type II was not observed.

The backbone type I conformers are stabilized due to the backbone hydrogen bonding, while the backbone type II conformers do not show backbone hydrogen bonding, but in few conformers interaction due to the acidic side-chain group were observed. Therefore, the role of the acidic side-chain group in the stability of backbone type II conformers of L-glutamic acid was studied. The role of the side-chain was understood by a comparative study with glycine, as was done in case of L-threonine. The conformer energy ordering in case glycine was observed to be same as that in L-glutamic acid, i.e., backbone type II conformer was the most stable form in both amino acid, and 2^{nd} lowest energy conformer showed backbone type I geometry. Based on the analysis, equivalent to that was done in L-threonine, as discussed in the last chapter, the acidic side-chain group of L-glutamic acid was observed to play a supportive role in determining the conformation preference, and not the decisive.[143] Thus, like the polar side-chain, in L-threonine, the acidic side-chain also play only a supportive role in conformational preferences of α-amino acids. The backbone type II structures are stabilized due to electronic delocalization interaction, as discussed in the last chapter, whereas the backbone type I conformers are stabilized by intramolecular hydrogen bonding.

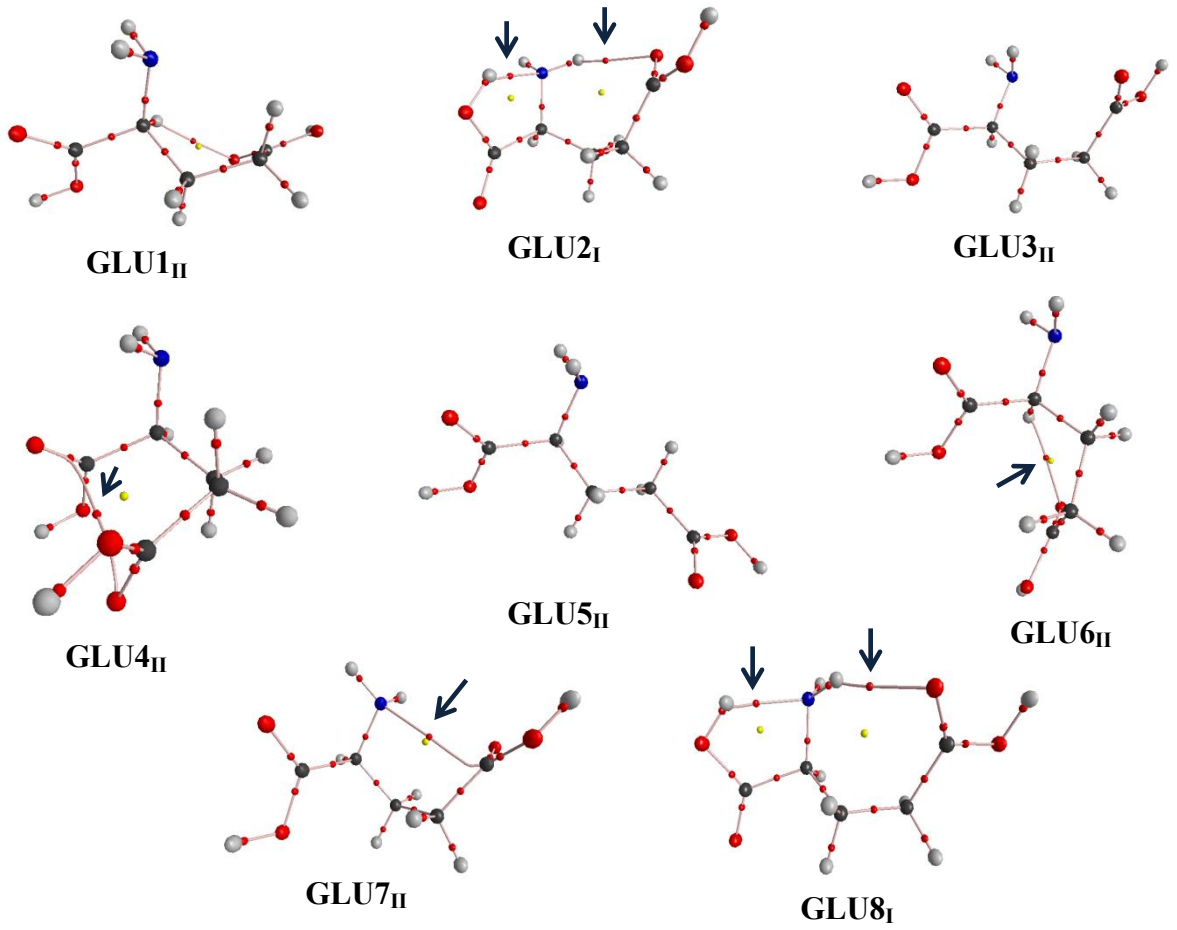


Fig. 4.9 AIM analysis of the various conformers of L-glutamic acid at MP2/6-311++G(d,p) level of theory. The bond critical points corresponding to hydrogen bond are indicated by the black arrow

Table 4.4 Electron density [$\rho(r_c)$], Laplacian of electron density [$\nabla^2\rho(r_c)$], local electronic kinetic energy density [$G(r_c)$], local electronic potential energy density [$V(r_c)$] and hydrogen bond energy (E_{HB}) values for hydrogen bonded interactions in the different conformers of L-glutamic acid, at the MP2/6-311++G(d,p) level. The values for $\rho(r_c)$, $\nabla^2\rho(r_c)$, $G(r_c)$ and $V(r_c)$ are given in atomic units and for E_{HB}, in kcal mol^{-1}

Conformer	Interactions	$\rho(r_c)$	$\nabla^2\rho(r_c)$	$G(r_c)$	$V(r_c)$	E_{HB}
$GLU1_{II}$	O-H_{SC}···H-Cα	0.0109	0.0385	0.0079	-0.0063	-1.96
$GLU2_I$	O-H···N	0.0411	0.1132	0.0329	-0.0376	-11.79
	N-H···O=Csc	0.0178	0.0675	0.0147	-0.0126	-3.96
$GLU3_{II}$						
$GLU4_{II}$	C=O···O=Csc	0.0097	0.0372	0.0075	-0.0056	-1.76
$GLU5_{II}$						
$GLU6_{II}$	O-H_{SC}···H-Cα					
$GLU7_{II}$	N-H···O=Csc	0.0102	0.0365	0.0075	-0.0058	-1.82
$GLU8_I$	O-H···N					
	N-H···O=Csc	0.0131	0.0497	0.0104	-0.0084	-2.62

4.5 Discussions

The conformational study of L-glutamic acid using semiempirical methods and *ab initio* computations at the HF/4-31G level of theory has been reported in the literature by Navarrete *et al.*[144] The optimized geometry, obtained in their study showed backbone type III structure. However, the low energy conformer observed in our computation and experiment shows backbone type II and I geometries, rather than type III. Alonso and co-workers have reported the conformation of L-glutamic acid, using rotational spectroscopy and *ab initio* computation.[140] Using Molecular Beam Fourier Transform Microwave (MB-FTMW) spectroscopy, Alonso and co-worker observed five conformers of L-glutamic acid. The solid sample was introduced into the microwave cavity employing laser ablation technique, and rotational spectra of ablated species were recorded using FTMW spectrometer. The L-glutamic acid conformers seen in their study were Iagc1, IIggc1, Igac1, IIaGc1, and IIagc1, mentioned in the decreasing order of conformational stability. These conformers correspond to and match with our conformers $GLU1_{II}$, $GLU2_{I}$, $GLU7_{II'}$, $GLU8_{I}$, and $GLU14_{I}$ respectively, both concerning structure and energy. Furthermore, the rotational constant of conformers obtained in this study is also in exact agreement with those reported by Alonso and co-workers, at MP2/6-311++G(d,p) level of theory.

The stability of conformer Iagc1 ($GLU1_{II}$ in our work) and Igac1 ($GLU7_{II}$ in our work) was suggested to be due to presence bifurcated ($C{=}O\text{-}NH_2$) hydrogen bond. However, the AIM analysis of conformer $GLU1_{II}$ and $GLU7_{II}$, does not locate any bond critical point along ($C{=}O^{\cdots}H_2N$) bond path. However, a hydrogen bond involving the NH_2 group and side-chain carbonyl group ($NH^{\cdots}O{=}C_{sc}$) was observed in conformer $GLU1_{II}$ and $GLU7_{II}$ (Fig. 4.9). The stability of conformer IIggc1 ($GLU2_{I}$ in our study), IIaGc1 ($GLU8_{I}$ in our research) and IIagc1

($GLU14_I$ in our study) was attributed to the presence of (O-H···N) hydrogen bond between the hydroxyl group of backbone COOH and NH_2 moiety. The (O-H···N) hydrogen bond of conformer $GLU2_I$ was identified both in our experiments (massive spectral shift of ~320 cm^{-1}) and AIM analysis, as discussed earlier. Likewise, the AIM analysis of other two conformers also located a bond critical point along (O-H···N) bond path. The high transient temperature attained in the laser ablation method of sample deposition results in a larger population to pool in the higher energy conformers, relative to our resistive heating method. Therefore, Alonso *et al.* observed five conformers of L-glutamic acid in their microwave study, while we saw only three.

The structure and energetics of different conformers were represented in the form of the conformational dartboard. The location of any conformer in dartboard depends upon its backbone geometry and energy, as discussed in the last chapter. The conformational dartboard of L-glutamic acid, based on the structure and energetics of conformers at MP2/6-311++G(d,p) level of theory, is shown in Fig. 4.10. The first octant region of the dartboard, which represents conformers lying within the energy of 1.25 kcal mol^{-1} relative to most stable conformer, is populated only by backbone type I and II conformers. The observation suggests that low energy conformers of L-glutamic acid prefer just backbone type I and II geometries. The conformers adopting backbone type III, IV, I', II' III' and IV' are observed to fill the dartboard only after the first octant and thus are high-energy structures. The preference of low energy conformers to adopt specific backbone geometries was also seen in the case of L-threonine, as discussed in the last chapter.[143]

Furthermore, the side-chain groups with varying chemical nature, i.e., L-glutamic acid acidic in case of L-glutamic acid and polar in case of L-threonine were observed to play a supportive role in conformational stability. Thus, irrespective of the chemical nature of the side-

chain group, the low energy conformers of α-amino acids seem to adopt only specific geometries. Therefore, the role of the non-polar side-chain in conformational preferences of α-amino acids will be exciting to explore. Moreover, it will be interesting to understand that if the above trend is pertinent over the entire naturally occurring amino acid.

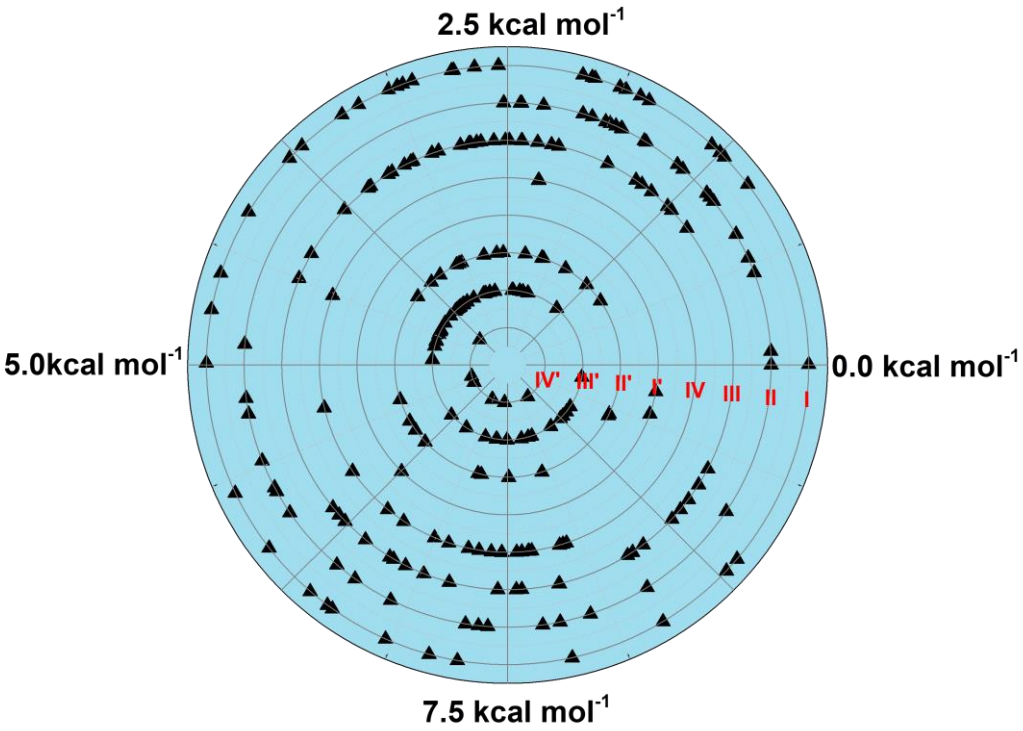


Fig. 4.10 Conformational dartboard of L-glutamic acid at MP2/6-311++G(d,p) level of theory

4.6 Conclusions

The conformational study of L-glutamic acid was performed employing matrix isolation infrared spectroscopy and *ab initio* computation methods. The solid sample was deposited on to the cold substrate KBr window, using a heated nozzle maintained at ~388 K and ~378 K. By scanning the potential energy surface of L-glutamic acid, 267 and 299 conformers were obtained at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theory, respectively. The conformers obtained

were classified into eight classes, based on their backbone structure. The spectral features seen in experiments were found to be in good agreement with the computed spectra of the three lowest energy conformers of L-glutamic acid viz. $GLU1_{II}$, $GLU2_{I}$, and $GLU3_{II}$. The backbone type I conformer ($GLU2_{I}$) showed an intramolecular hydrogen bonding, between backbone carboxylic and amino groups, both in experiments and computations (AIM analysis). However, backbone type II conformers ($GLU1_{II}$, and $GLU3_{II}$), whose stability was suggested to the presence of a bifurcated ($NH_2\cdots O=C$) hydrogen bond in earlier works, does not show such interactions. The stability of backbone type II conformers is governed by vicinal orbital delocalization interactions, as discussed in the last chapter.

The acidic side-chain group, present in L-glutamic acid, does not play a decisive role in conformational preferences, and at best, they may be supportive. Similar behavior was also seen in the case of L-threonine, where even the polar side-chain group was found to plays a supportive role in conformational preferences, as discussed in the last chapter. Therefore, after rationalizing the role of the polar and acidic side-chain group, the role of the non-polar and basic side-chain group will be interesting to explore. The structure and energetics of various conformers were presented in the form of the conformational dartboard, and it was observed that low energy conformers prefer only backbone type I and II geometries, while the high energy conformers adopt backbone type III, IV, I', II', III' and IV'. A similar trend was also found in the case of L-threonine, where the low energy conformers also adopted only backbone type I and II and instead rarely type III, as discussed in the last chapter. A generalization of this trend over entire α-amino acids, which vary with each other by the chemical nature and bulkiness of side-chain groups, can be useful in rationalizing the conformations of peptides.

Chapter 5

Conformations of L-Methionine

5.1 Introduction

L-methionine is an essential and sulfur-containing amino acid, which plays a vital role in nephrotoxicity of the cis-platin drug, used for cancer treatment.[149-151] The methionine-containing proteins competitively bind with cis-plantin, thus prohibiting it to interacts with DNA binding site and, hence, leads to inactivation of cis-platin and tumor resistance to the treatment with this drug. The bulky non-polar aliphatic side-chain of L-methionine deters the growth of good quality crystal which is required for high precision crystal analysis, and two crystallographic forms of L-methionine have been reported for methionine, in independent studies. [49,50] An infrared spectrum of the canonical and zwitterionic form of L-methionine was studied using KBr matrix and *ab initio* methods.[152]

The gas phase conformational studies, which exhibit the basic molecular form in the absence of condense phase intermolecular interactions, of L-methionine, have been explored, in the past, using *ab initio* computational method.[153,154] The low energy conformers of L-methionine, obtained in computation study, were suggested to be stabilized by hydrogen-bonded interactions. No report on the conformations of L-methionine in the isolated phase, using experimental spectroscopic techniques, has been reported in the literature. As discussed in earlier chapters, matrix isolation infrared spectroscopy has been widely used in the past, for studying the conformations of various α-amino acids. In this chapter, we explore the conformational landscape of L-methionine using matrix isolation infrared spectroscopy and *ab initio* computations. The solid L-methionine was deposited on to cold KBr window, maintained at 12 K, employing a heated nozzle. The sample vapor, produced by heating the sample, was swept away by an inert gas onto the cold substrate. The infrared spectra of matrix isolated L-

methionine were recorded, and the experimental results obtained was corroborated with the *ab initio* computed structures of L-methionine.

Apart from establishing the conformational landscape, factors determining the conformational stability in L-methionine were also examined. Presence of non-covalent interaction such as hydrogen bonding, in various conformers of L-methionine, was investigated with atoms in molecules (AIM) analysis. As discussed in earlier chapters, the polar (L-threonine) and acidic (L-glutamic acid) side-chain groups were found to play a supportive role, and not decisive, in defining the conformational choices of α-amino acids. Therefore, the role of the non-polar side-chain in case of conformation preferences of L-methionine was also investigated. The structure and energetics of conformers were represented in the form of the conformational dartboard to examine the intrinsic choice of low energy conformers to adopt only specific backbone geometries, as seen in the case of L-threonine and L-glutamic acid.

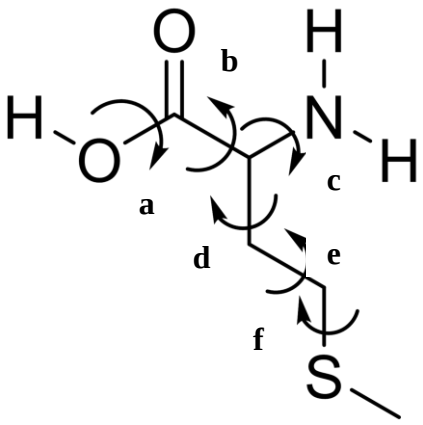
5.2 Experimental Details

The detail of the experimental setup is described elsewhere in chapter 2. A closed helium compressor-cooled cryostat (HC-4E1) was used to achieve 12 K temperature in the matrix isolation experiments. L-methionine was deposited employing the hot effusive beam nozzle, which was maintained at two different temperatures, i.e., 408 K and 423 K. The vapors of L-methionine were then carried away by flowing argon gas (Grade-I Sigma Gases and Service, 99.999%) over the hot sample, for the deposition onto the cold substrate window. The matrix gas flow rate during the deposition in all the experiments was usually kept around three mmoles/hour. The infrared spectra of L-methionine, trapped in an inert matrix, were recorded using Bruker Tensor-27 spectrometer, operating at a resolution of 0.5 cm^{-1}. The matrix was then annealed by raising the temperature of the matrix to 30 K, maintained at this elevated

temperature for an hour and then cooled back to 12 K; the spectra of the annealed matrix were then recorded.

5.3 Computational Details

The Gaussian 09 suite of the program was used for the computational study of L-methionine.[107] A large number of conformers of L-methionine were obtained by scanning the potential energy surface (PES). Initial geometries for optimization were generated by order wise rotation of six dihedral angles, as shown in Fig. 5.1. The 1944 initial geometries generated by dihedral rotation, as shown in Fig. 5.1, were optimized at the HF/6-31G level of theory. The optimization process resulted in 225 distinct stationary points. That the structures corresponding to these stationary points were indeed distinct, were verified by comparing the computed rotational constants for each of the structures. The 225 structures were then used as starting geometries for optimization at B3LYP/6-311G level of theory, and the exercise resulted in 186 stationary points. These 186 stationary points of L-methionine were then further optimized at M06-2X and MP2 level of theory, using 6-311++G(d,p) basis set.



a = 2 fold rotation ; b = 6 fold rotation
c = 3 fold rotation ; d = 3 fold rotation
e = 3 fold rotation ; f = 6 fold rotation
Total 2x6x3x3x3x6 = 1944 initial geometries

Fig 5.1 Systematic dihedral rotation of L-methionine for initial geometry generation

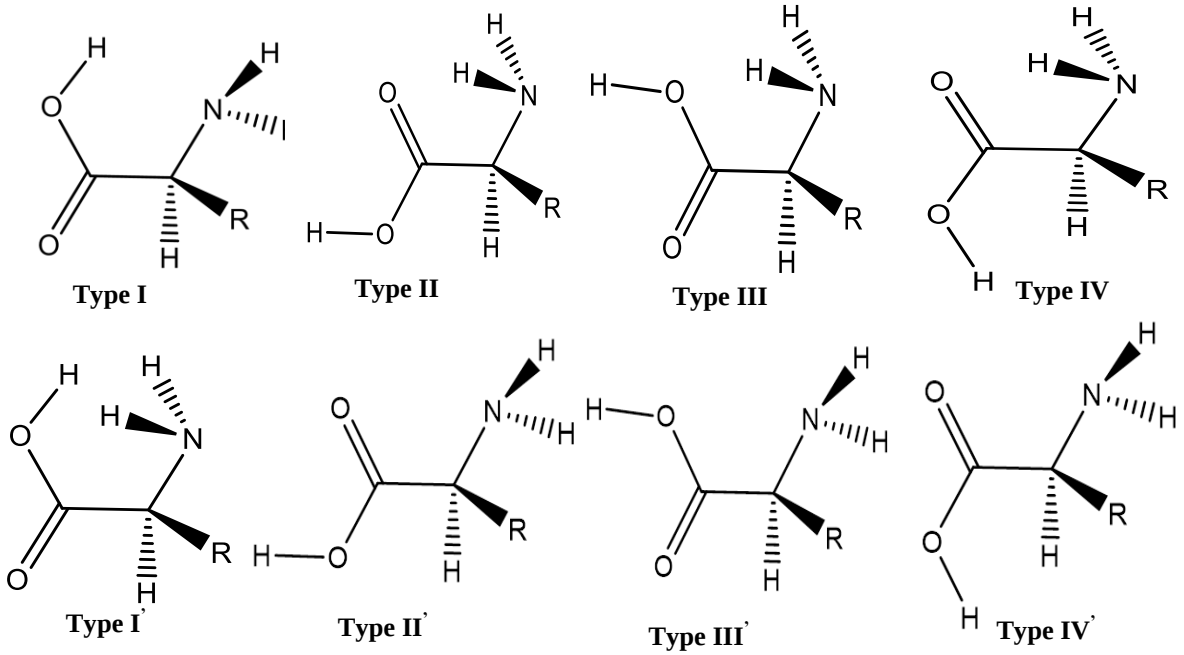


Fig 5.2 Eight backbone structures of α-amino acid

The frequency calculations of optimized geometry, using the analytical gradient method, at M06-2X and MP2 methods yielded 176 and 167 different conformers of L-methionine, and each shows real frequency for all the normal modes. The rotational constants of the optimized geometries were compared to determine that these were unique structures. The conformers of L-methionine obtained were then classified based on the backbone structure, as discussed earlier in the case of L-threonine and L-glutamic acid. The eight backbone structures of amino acids are type I, II, III, IV, I', II', III' and IV' and are shown in Fig. 5.2, and all the conformers of L-methionine could be assigned to one of the above backbone types.

The Roman numeral subscript included in the nomenclature of each conformer indicates the backbone type conformer that it belongs to. The relative energy of each conformer relative to that of the conformer, $MET1_{II}$, at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theory are shown in Fig. 5.3 and Table 5.1. In addition to the relative energies, i.e., ΔE_{MP2} and ΔE_{M06}, the relative free energy (ΔG_{MP2}) of conformers at 408 K (which is the nozzle temperature during sample deposition) were also computed at MP2/6-311++G(d,p) level of theory and are

listed in Table 5.1. The relative abundances of each conformer calculated using ΔE_{MP2} and ΔG_{MP2} are also listed as $\%P_{(\Delta EMP2)}$ and $\%P_{(\Delta GMP2)}$ respectively, in Table 5.1. The dipole moments of 20 lowest energy conformers computed at MP2/6-311++G(d,p) level of theory and are shown in the last column of Table 5.1.

The transition state geometry and the energy barrier separating various low energy conformers were computed at MP2/6-311++G(d,p) level of theory. The transition state geometries connecting a conformer pair were further verified by intrinsic reaction coordinates (IRC) calculations.

The spectral assignment between experimental and computed features was done by scaling the computed wavenumbers of low energy conformers. The computed features in the O-H and N-H stretching region, i.e., 3600-3100 cm^{-1} were scaled using a scaling factor of 0.9345. The scaling factor was obtained after assigning the most intense experimental feature at 3546.2 cm^{-1} to the intense computed wavenumber observed at 3794.8 cm^{-1}. Likewise, the calculated wavenumbers corresponding to the C-H stretch region ranging between 3100-2000 cm^{-1} was scaled with a factor of 0.9500. The vibrational assignment in the carbonyl stretching region and skeleton vibration region, i.e., 2000-4000 cm^{-1} were done by scaling the computed wavenumbers with a scaling factor of 0.9774. The scale factor was obtained by assigning the strongest computed wavenumber at 1808.9 cm^{-1} to the intense features at 1768.0 cm^{-1} observed in our experiment. Since the magnitude of perturbations due to matrix varies from one vibrational mode to differ from mode to mode;[105] therefore mode-by-mode scaling factor was used, as has been done previously in many of our work and other published work.[128-135,143]

The presence of the intramolecular hydrogen bond in the conformers of L-methionine was identified computationally with quantum topological atoms in molecules (QTAIM)

analysis.[110] The wave functions computed at MP2/6-311G++G(d,p) level of theory were used to generate the electron density topology with the help of AIM 2000 program suite. The strength of the intramolecular hydrogen bond, wherever present, was calculated using the method suggested by Espinosa *et al.*, as discussed in earlier chapters.[113]

5.4 Results

5.4.1 Conformer's energy and populations

Fig. 5.3 shows the 12 lowest energy conformers of L-methionine, computed at the MP2/6-311G++G(d,p) level of theory. The relative energy of each conformer relative to that of the conformer, $MET1_{II}$, at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theory are shown in Fig. 5.3 and Table 5.1. The backbone type II conformer, i.e., $MET1_{II}$ was found to the most stable conformer both at the MP2 and M06-2X methods, with 6-311G++G(d,p) basis set. The conformer $MET2_{I}$ was observed to be the next higher energy conformer with the energy of 0.67 kcal mol^{-1} relative to $MET1_{II}$. As seen in Fig. 5.3, 10 out of the 12 lowest energy conformer adopts either backbone type II or I. Likewise, out of the 20 lowest energy conformers listed in Table 5.1, all the conformers, except $THR3_{III}$ and $THR12_{III}$, adopt backbone types I and II. The observation that the low energy conformers of L-methionine preferentially adopt certain backbone geometry, namely type I and II, is consistent with the earlier observations on L-threonine and L-glutamic acid.

The free energy of the conformers (ΔG_{MP2}) was computed at the temperatures adopted during deposition (408 K), using MP2/6-311G++G(d,p) level of theory and are shown in Table 5.1. The conformers $MET2_{I}$ and $MET3_{III}$, which showed relative energy of 0.67 and 0.90 kcal mol^{-1}, after thermal corrections were now found to exhibit ΔG_{MP2} of 1.61 and 2.99 kcal mol^{-1}.

Table 5.1 Relative zero-point corrected energies of the conformers of L-methionine computed at M06-2X/6-311++G(d,p) ($\Delta E_{M06\text{-}2X}$) and MP2/6-311++G(d,p)(ΔE_{MP2}) levels. The relative free energies at 415 K (sample deposition temperature) at the MP2/6-311++G(d,p)(ΔE_{MP2}) levels are also shown. All energies are in kcal mol^{-1} and are relative to the energies of the conformer $MET1_{II}$. Dipole moments (D) of all the conformers at MP2/6-311++G(d,p) level of theory shown in the last column.[f]

Conformer	$\Delta E_{M06\text{-}2X}$	ΔE_{MP2}	ΔG_{MP2}	%P $_{(\Delta EMP2)}$	%P $_{(\Delta GMP2)}$	Dipole moment[f]
$MET1_{II}$	0.07	0.00	0.00	24.54	15.70	2.1
$MET2_{I}$	0.33	0.67	1.61	7.94	2.34	4.8
$MET3_{III}$	0.00	0.90	2.99	5.38	0.46	0.7
$MET4_{I}$	0.10	0.93	3.21	5.13	0.36	5.4
$MET5_{II}$	0.72	1.00	1.20	4.56	3.80	0.7
$MET6_{II}$	1.10	1.12	0.86	3.73	5.69	2.9
$MET7_{II}$	0.57	1.14	2.83	3.58	0.56	3.0
$MET8_{II}$	2.67	1.17	0.44	3.39	9.36	0.8
$MET9_{II}$	1.25	1.19	0.76	3.28	6.42	1.4
$MET10_{I}$	1.12	1.33	2.32	2.61	1.01	6.6
$MET11_{II}$	1.16	1.33	2.67	2.61	0.67	2.7
$MET12_{III}$	1.53	1.40	1.16	2.32	3.97	2.6
$MET13_{I}$	1.86	1.50	1.86	1.95	1.74	4.5
$MET14_{II}$	1.35	1.55	1.62	1.81	2.31	0.8
$MET15_{II}$	1.58	1.55	1.36	1.79	3.16	2.8
$MET16_{II}$	1.17	1.57	3.03	1.75	0.44	3.7
$MET17_{II}$	2.27	1.58	2.93	1.72	0.49	0.6
$MET18_{II}$	1.68	1.63	1.17	1.57	3.95	1.9
$MET19_{I}$	1.67	1.66	2.45	1.51	0.86	5.1
$MET20_{I}$	2.05	1.74	2.26	1.32	1.09	4.7

All energies in kcal mol^{-1}

The relative conformer stability of backbone type I conformer, which have also been found to show intramolecular hydrogen bonding, were found to be higher. The relative abundance of conformers based on the ΔE_{ZPC} and ΔG_{MP2} were computed and are also listed in Table 5.1. The vibrational assignment between experimental and calculated wavenumbers was done based on the population distribution obtained with relative ZPC energy (%P $_{(\Delta EMP2)}$) and relative free energy (%P $_{(\Delta GMP2)}$), and will be discussed later in the section. The dipole moments of 20 lowest energy conformers computed at MP2/6-311++G(d,p) level of theory and are shown in the last column of Table 5.1.

5.4.2 Conformational interconversion

As discussed earlier, the matrix isolation study of L-methionine was done using hot effusive beam nozzle. It can, therefore, be expected that conformer population distribution observed in the frozen matrix should be the same as the population distribution gas phase before deposition. However, the possibility of conformation relaxation to the lowest energy conformers during the annealing of the matrix at 30 K cannot be ignored. Based on earlier matrix isolation studies, it was established that conformational relaxation to the lowest energy conformer while annealing the matrix at 30 K is possible, provided the energy barrier of interconversion is nearly or less than ~5 kJ mol^{-1} (<1.2 kcal mol^{-1}).[80,143] Therefore, to assess the possibility of conformational interconversion during the annealing process, the energy barrier separating the conformer pairs were computed.

Conformational relaxations involving rotation of more than one bond are unlikely to have barrier lower than ~5 kJ mol^{-1}, and hence unlikely to occur during the annealing of the matrix. The energy barriers for conformal interconversion occurring involving rotation around one bond only were computed. Therefore, the interconversion between conformers corresponding to same backbone structure was considered, as the relaxation of one backbone type to another backbone type conformers are accompanied with rotations around more than one bonds and thus unlikely to occur.

The interconversion between conformers of same backbone type would, therefore, involve the only rearrangement of the side-chain orientation. The side-chain orientation of L-methionine conformers was defined based on the orientation of three dihedral angles, i.e. (N-C_α-C_β-C_Υ), (C_α-C_β-C_Υ-S) and (C_β-C_Υ-S-C). The orientations of three dihedral angles were identified using gauche and trans notation, as discussed in the previous chapter. For example, the side-chain

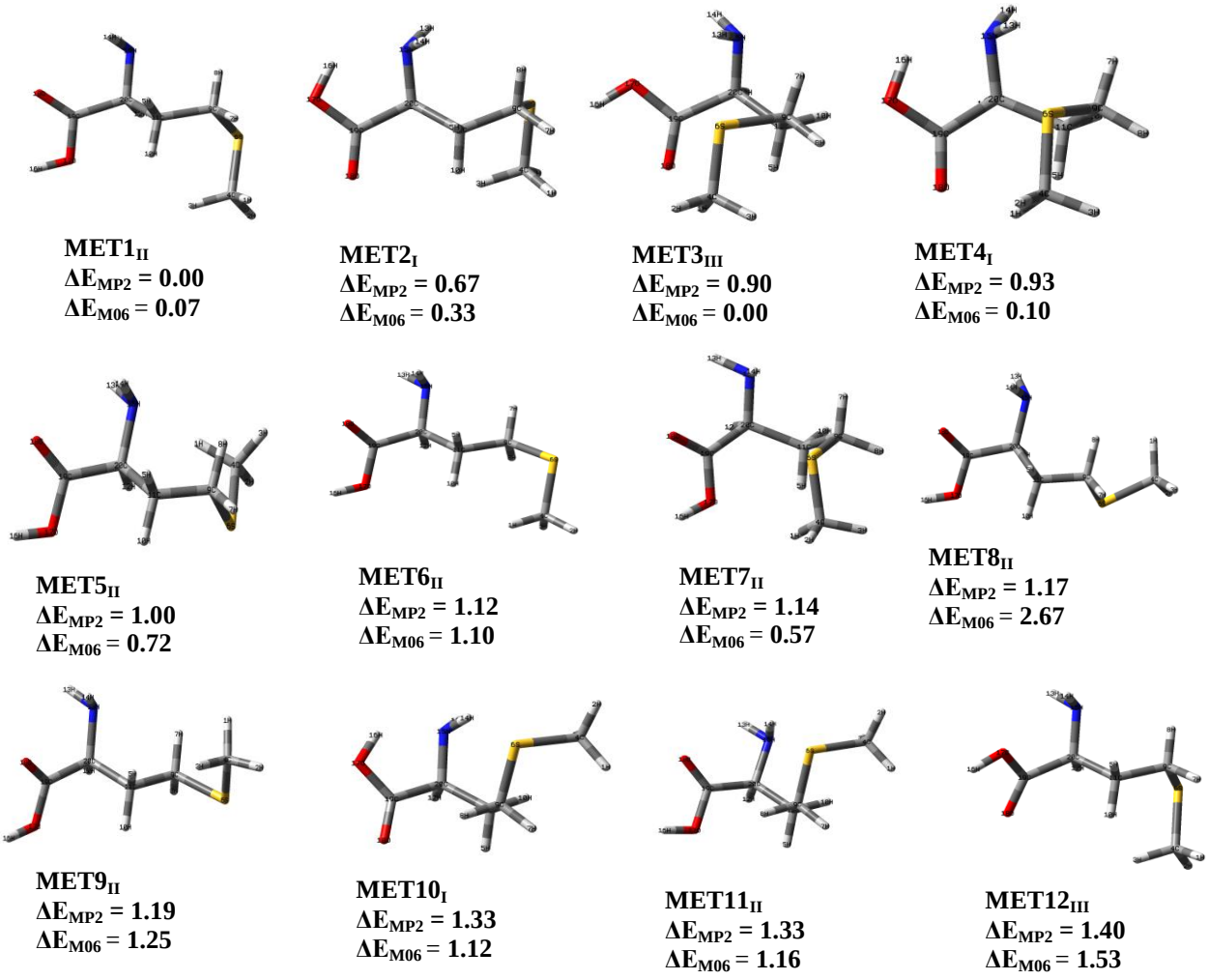


Fig 5.3 Optimized geometry of 12 lowest energy conformers of L-methionine at MP2/6-311++G(d,p) level of theory. The relative energy of the conformers at MP2 and M06 level with 6-311++G(d,p) basis set, are also shown. The energies are relative to conformer $MET1_{II}$ and are in kcal mol^{-1}.

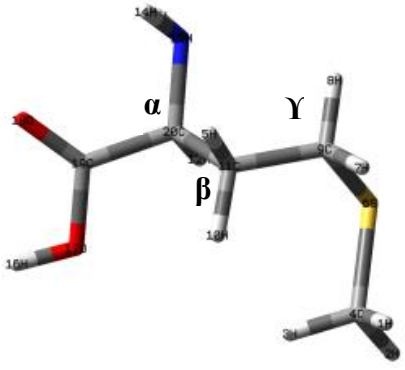


Fig. 5.4 Side-chain orientation of $MET1_{II}$ in terms of N-C_α-C_β-C_γ = G^- , C_α-C_β-C_γ-S = G^- and C_β-C_γ-S-C = G^- dihedral angles.

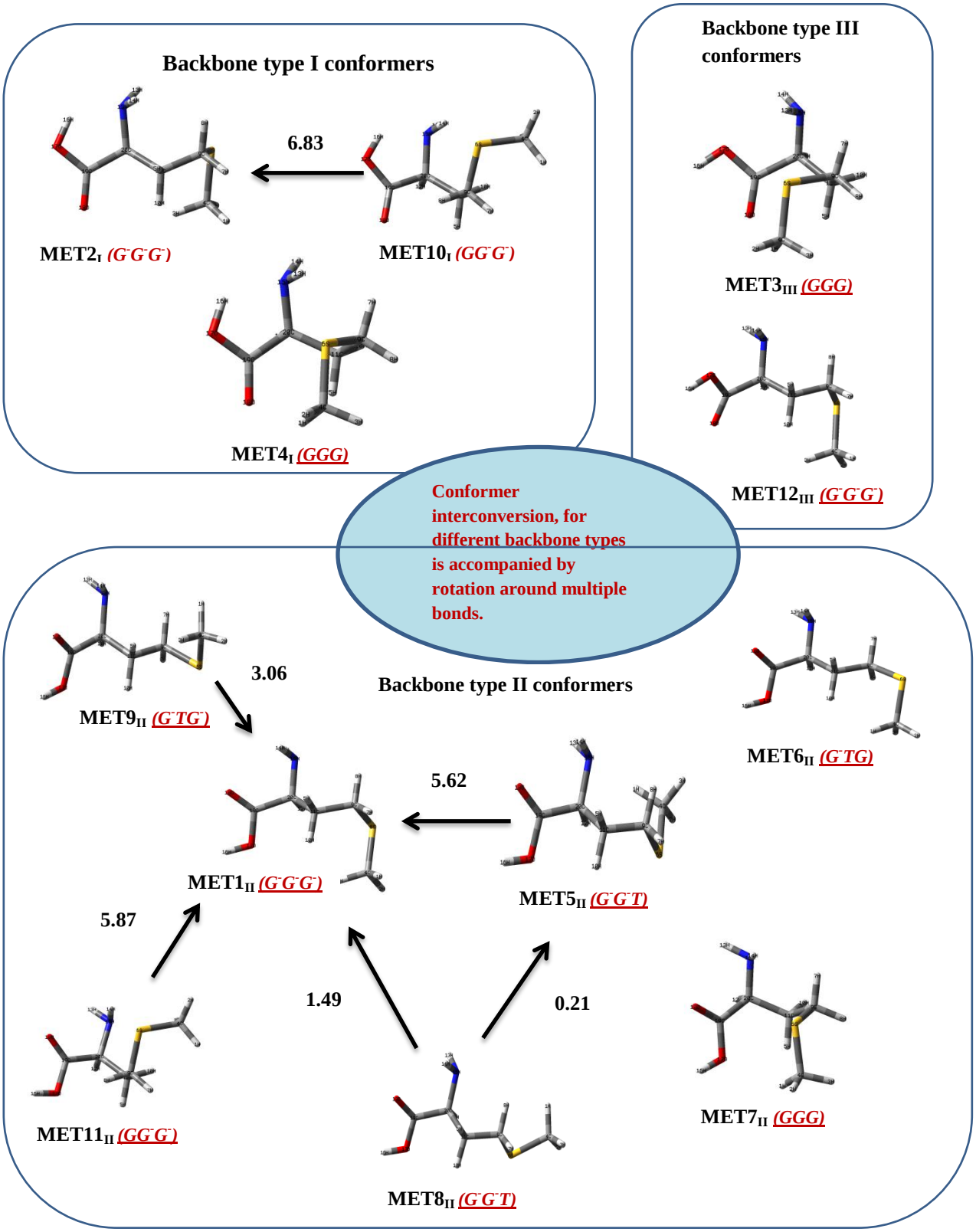


Fig. 5.5 The interconversion barrier for conformers of L- methionine; The three *italic* letters in each conformer (for ex. *(G-G-G-)* in the conformer MET1$_{II}$) represent the orientation side-chain dihedral angles i.e. (N-C_α-C_β-C_γ), (C_α-C_β-C_γ-S) and (C_β-C_γ-S-C). The energy barrier separating conformers are (shown in black) calculated at MP2/6-311++G(d,p) level of theory.

orientation of conformer $MET1_{II}$ is denoted by $G^-G^-G^-$, as shown in Fig. 5.4, and thus suggestion all three dihedral angle adopts gauche (-) orientation.

The energy barrier between conformer of the same backbone type and differ in the orientation of only one side-chain dihedral angle is shown in Fig. 5.5. For example, the conformers $MET1_{II}$ ($G^-G^-G^-$) and $MET5_{II}$ (G^-G^-T), varies only with the relative to the orientation of (C_β-C_γ-S-C_θ) dihedral angle, and the energy barrier separating the two conformers is computed to be 5.6 kcal mol^{-1}, as shown in Fig. 5.5. Thus the relaxation of $MET5_{II}$ to $MET1_{II}$ during the annealing of the matrix at 30 K is highly improbable. Likewise, the energy barriers for conformer pairs, which differ only with the relative orientation of a single dihedral angle, are shown in Fig. 5.5.

The energy barrier for all the conformer pairs was found to higher than the threshold limit of 1.2 kcal mol^{-1} (5 kcal mol^{-1}), except one, which is shown in Fig. 5.5. The lone exception is a barrier separating conformer $MET8_{II}$ and $MET5_{II}$ and is computed to be 0.2 kcal mol^{-1}. However, given the relative abundance of $MET5_{II}$ of only ~4 %, as shown in Table 5.1, it is highly unlike that conformer $MET5_{II}$ will even be observed in the first place, after the deposition. Hence, it can be assumed that the conformational population distribution present in the gas phase and which is frozen soon after deposition, will continue to be present nearly unaffected during the annealing of the matrix at 30 K.

5.4.3 IR spectra of matrix isolated L-methionine

The infrared spectra of L-methionine trapped in argon matrix are shown in Fig. 5.6, and 5.7. The IR spectra corresponding to different spectral mode have been shown separately. The O-H and N-H stretch region ranging between 3600-3100 cm^{-1} are shown in the grid (a) of Fig. 5.6, while the C-H stretching region ranging from 3100-2800 cm^{-1} is presented in the grid (b) of Fig. 5.6.

The spectral region corresponding to the carbonyl stretching mode is represented in Fig. 5.6 (grid c). Features in the O-H stretching region were seen at 3559.5, 3546.2, 3399.7 and 3183.5 cm^{-1}. The C-H stretching region shows a broad feature centered at 2927.8 cm^{-1}. However, a closer analysis reveals the broad feature to consist of some closely spaced features at 2931.4, 2925.4 and 2924.3 cm^{-1}. In the carbonyl stretching region (grid c), a strong feature is observed at 1768.0 cm^{-1} along with less intense features at 1787.2 and 1739.9 cm^{-1}. The spectral regions corresponding to the skeleton region, which range between 1450-1050 cm^{-1} are shown in Fig. 5.7.

Apart from the spectral features of amino acid, the infrared wavenumbers of H_2O and CO_2 were also seen. H_2O and CO_2 being a pervasive impurity are always observed in the matrix isolated experiments; however, the intensity of features due to H_2O and CO_2 were unchanged when after amino acid was deposited as compared to that observed with matrix gas deposited alone. The L-methionine has been reported to start decomposing upon heating at a temperature ranging from 190-240 ºC (463 K – 523K), which is higher than the temperatures used in our deposition.[157] The volatile products of thermal decomposition are found to CH_3SH and CO_2, while the non-volatile products are analyzed to be 3-methylthio-1-propanamine and CH_3SH. The matrix isolated spectra of L-methionine does not show any spectral feature of CH_3SH.

The presences of possible non-volatile impurities were examined by performing the ^{1}H and ^{13}C NMR of L-methionine before and after the deposition (after heating at 408 K), as shown in Fig. 5.8. The ^{1}H and ^{13}C NMR spectra of the sample before and after heating are found to be identical and agree with the reported spectra of L-methionine. Thus, the absence of volatile impurity features and identical NMR spectra of the sample before and after experiments rules out the possibility of thermal decomposition of L-methionine in our experiments.

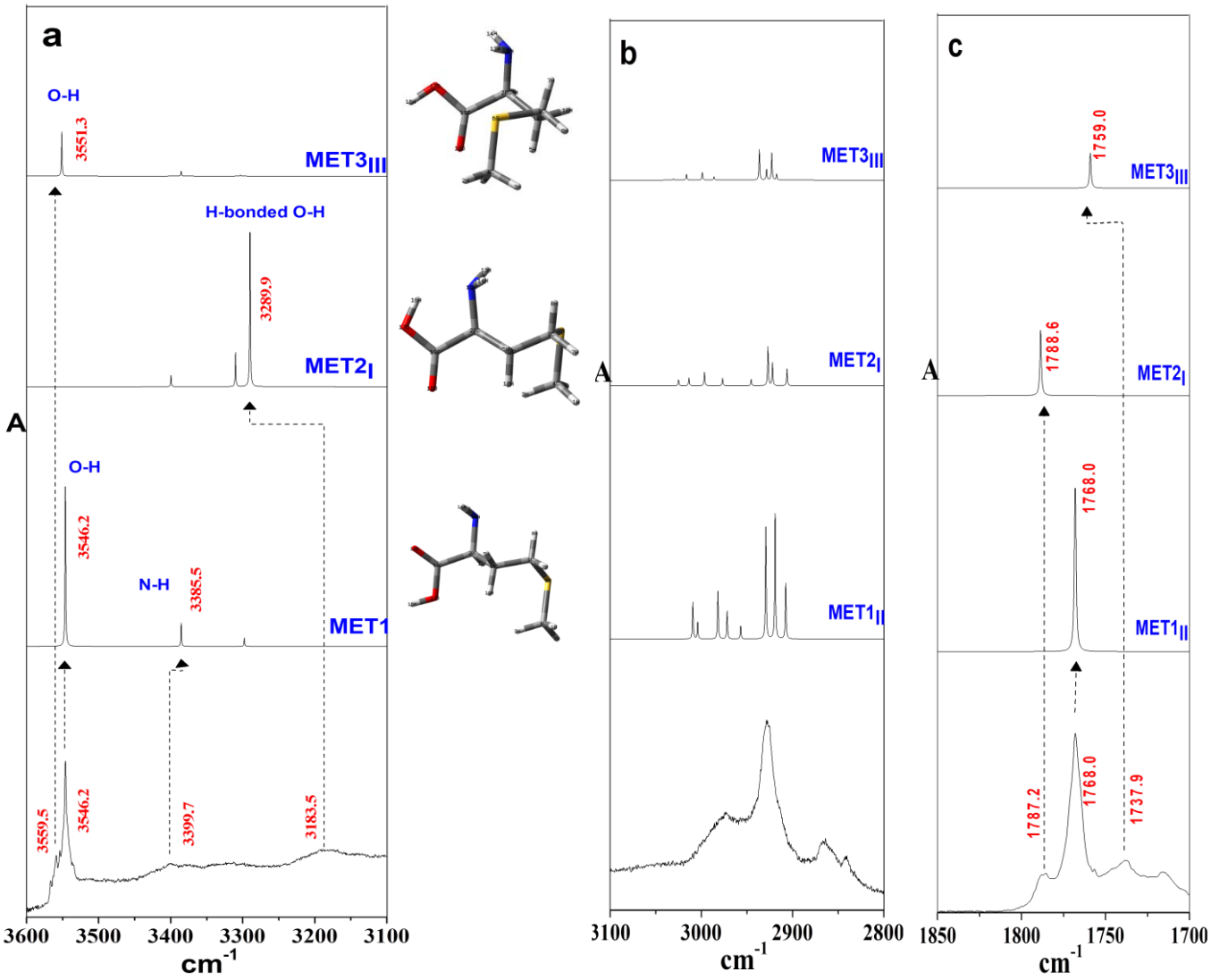


Fig 5.6 Comparison of experimental and computed spectra of L-methionine, the spectra span the region 3600-3100 cm^{-1} (grid a), 3100-2800 cm^{-1} (grid b) and 1850-1700 cm^{-1} (grid c). i) IR spectra of matrix isolated L-methionine deposited at 415 K; (ii-iv) Computed spectra showing the spectral features of the three lowest energy conformers, calculated at the MP2/6-311++G(d,p) level of theory. The computed wavenumbers in the grid 'a', 'b' and 'c' are scaled using a scaling factor of 0.9345, 0.95 and 0.9774 respectively

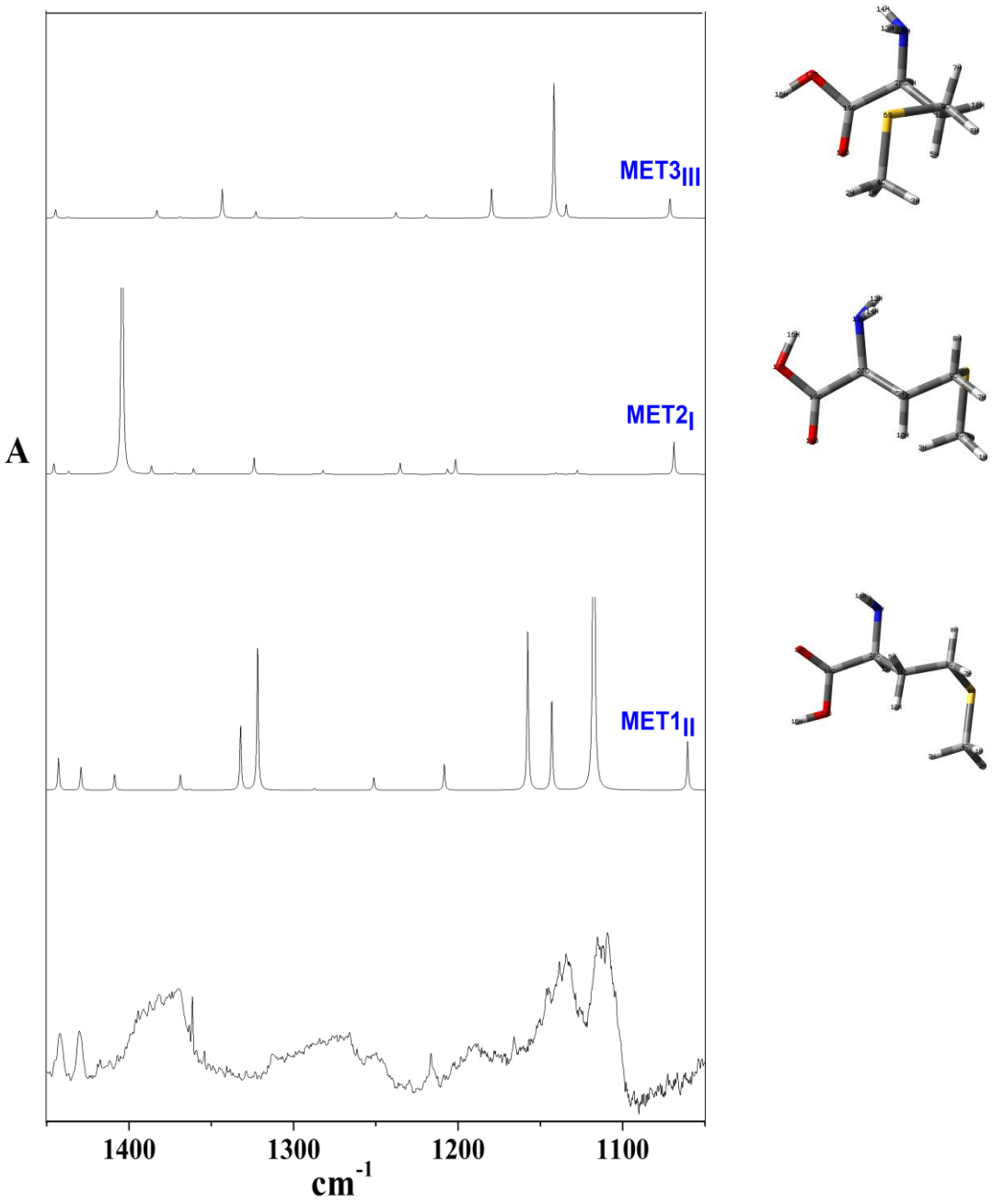
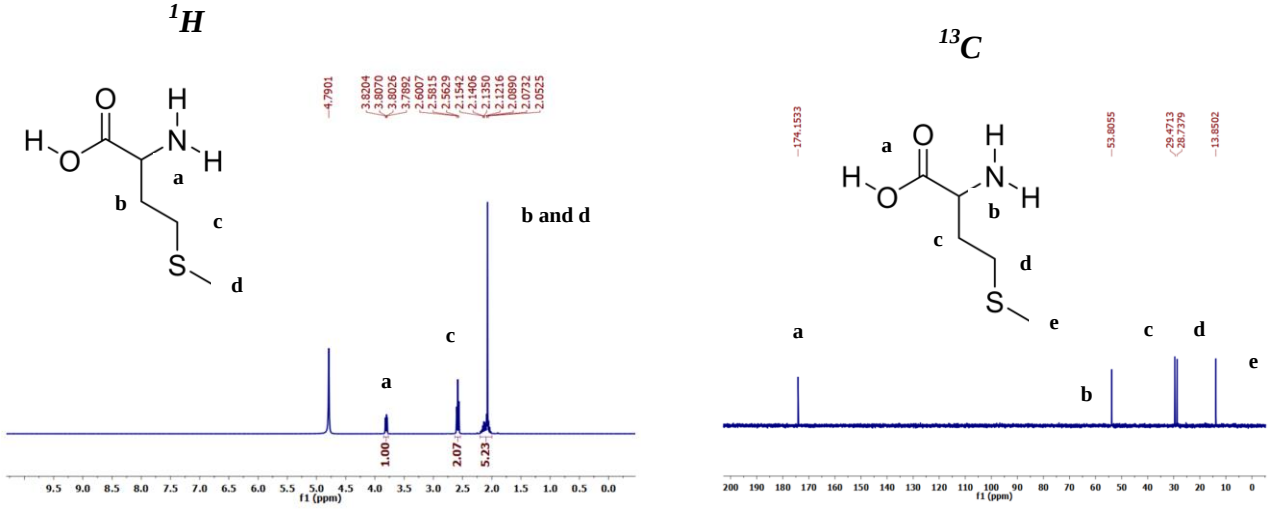


Fig 5.7 Comparison of experimental and computed spectra of L-methionine, the spectra span the region 1450-1000 cm^{-1}. i) IR spectra of matrix isolated L-methionine deposited at 415 K; (ii-iv) Computed spectra showing the spectral features of the three lowest energy conformers, calculated at the MP2/6-311++G(d,p) level of theory. The computed wavenumbers are scaled using a scaling factor of 0.9774.

L-methionine (before heating)



L-methionine (after heating at 408 K)

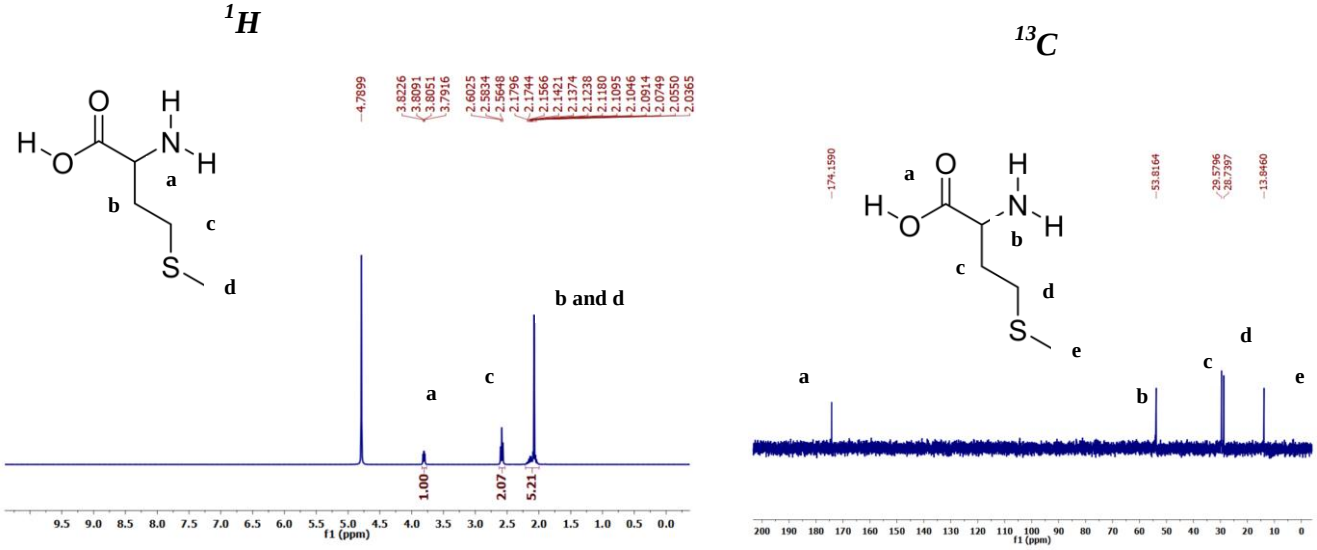


Fig 5.8 H^1 NMR and C^{13} NMR spectra of L-methionine, before (top) and after (bottom) sample heating

Table 5.2 Experimental (in argon matrix) and computed scaled vibrational wavenumbers (cm^{-1}) at the MP2/6-311++G(d,p) level for the different conformers of L-methionine. See text for details on scaling

Experimental	Computed scaled[a]			IR intensity[b]	Vibrational modes
	$MET1_{II}$	$MET2_{I}$	$MET3_{III}$		
3559.5			3551.3	81.6	ν (O-H)
3546.2	3546.2			78.7	ν (O-H)
3399.7	3385.5			9.0	ν (N-H)
3183.5		3289.9		246.6	ν (O-H)
2973.4	2981.8			11.3	ν (C_{β}-H)
2931.4		2927.0		28.6	ν (C_{γ}-H)
2927.8	2929.3			26.6	ν (C_{γ}-H) + ν(C-H;CH_3
2925.4	2919.1			29.3	ν (C-H;CH_3 group)
2924.3			2922.7	29.7	ν (C_{γ}-H)+ ν (C-H;CH_3
1787.2		1788.6		265.7	ν (C=O)
1768.0	1768.0			277.6	ν (C=O)
1737.9			1759.0	211.1	ν (C=O)

[a] Scaling factors for computed vibrational wavenumbers of L-methionine conformers: 0.9345 for the O−H stretch region (3800-3100 cm^{-1}), 0.95 for C-H stretch region (3100-2800 cm^{-1}) and 0.9774 for the C=O stretch and skeleton vibration (2000-400 cm^{-1}). See text for details.
[b] Computed IR intensity(km mol^{-1}) corresponding to each normal mode. The computed intensity was scaled based on the relative abundance of the respective conformer.

Table 5.3 Experimental (in argon matrix) and computed scaled vibrational wavenumbers (cm^{-1}) at the MP2/6-311++G(d,p) level for the different conformers of L-methionine. See text for details on scaling

Experimental	Computed scaled[a]			IR intensity[b]
	$MET1_{II}$	$MET2_{I}$	$MET3_{III}$	
1450.0	1460.6			8.1
1442.4	1442.9			9.3
1430.3	1429.2			6.8
1390.5		1404.2		441.7
1386.3			1383.1	10.3
1377.8		1386.4		7.4
1369.9	1368.8			4.4
1361.6			1343.4	39.3
1327.3	1332.2			18.4
1312.3	1321.8			41.6
1307.5		1323.9		15.6
1252.4	1251.2			3.7
1216.7	1208.4			7.4
1166.2			1179.7	40.8
1145.0	1157.7			46.1
1134.9	1143.1			25.9
1133.6			1141.8	180.7
1116.8			1134.4	17.9
1112.1		1127.6		3.7
1105.2	1117.6			193.9

[a] Scaling factors for computed vibrational wavenumbers of L-methionine conformers: 0.9774 for the C=O stretch and skeleton vibration (2000-400 cm^{-1}). See text for details.
[b] Computed IR intensity(km mol^{-1}) corresponding to each normal mode. The computed intensity was scaled based on the relative abundance of the respective conformer.

5.4.4 Vibrational assignments

The three lowest energy conformers of L-methionine, i.e., $MET1_{II}$, $MET2_{I}$, and $MET3_{III,}$ show a relative abundance of 24.54, 7.94 and 5.38 % respectively. These three conformers belong to different backbone type and thus shows varying O-H and C=O stretching vibrations. These O-H and C=O vibrational features are important signatures for identifying the various conformers present in experiments. The assignment of experimental features with computed wavenumbers of low energy conformer was aided by generating computed spectra. The synthetic spectra were generated with SYNSPEC program, and a Lorentzian line shape with a line width of 1.0 cm^{-1} was used. The computed spectra corresponding to the three lowest energy conformers were synthesized separately employing the computed scaled wavenumbers of individual conformers, as shown in Fig. 5.6 and 5.7 (trace ii-iv). The intensity of spectral features in the synthetic spectra was derived from the computed intensity of respective vibrational modes, which were then scaled for the relative population of individual conformers. The experimental features, along with the computed wavenumber of the three lowest energy conformers, are listed in Table 5.2 and 5.3.

5.4.4 (a) 3600-3100 cm^{-1} region

The IR spectra of matrix isolated L-methionine (trace i) and the computed spectra of the three lowest energy conformers (trace ii-iv) in the O-H and N-H stretching region are shown in Fig. 5.6 (grid a). The strongest feature in this spectral region is observed at 3546.2 cm^{-1} and is a potential feature to be assigned to the υ (O-H) of the lowest energy conformer. The computational study at MP2/6-311G++G(d,p) shows conformer $MET1_{II}$ to be the most stable conformer and $MET2_{I}$ to be next in energy with a ΔE_{ZPC} of 0.67 kcal mol^{-1}. The relative abundance of the conformer $MET1_{II}$ (24.5%) was, therefore, higher than that of $MET2_{I}$ (7.9%).

Thus, based on the relative stability and population distribution, the assignment of strongest feature observed in the experiment at 3546.2 cm^{-1} is made to the υ (O-H) mode of conformer $MET1_{II}$.

A scaling factor based on the above assignment of the experimental feature 3546.2 cm^{-1} to the computed wavenumber at 3794.8 cm^{-1} (O-H stretching of $MET1_{II}$) was calculated and employed to scale the computed wavenumbers of the other conformers. The scaled wavenumbers of other conformers provide a first assignment with the features observed in the experiment. However, when the assignment of the strongest feature at 3546.2 cm^{-1} is made to the intense computed feature of conformer $MET2_I$ (observed at 3520.5 cm^{-1}) and a scale factor based on this assignment is calculated (1.0007), the scaled computed features for the other conformers show a poor assignment for the different experimental wavenumbers.

Thus, the above observations clear favor the first assignment and which also provides evidence that the conformer $MET1_{II}$ is the most stable. The observation of backbone type II conformer ($MET1_{II}$) as the most stable conformer and backbone type I conformer ($MET2_I$) to be the next higher energy conformer (both in the experiments and computations) are consistent with our previous study of L-threonine [143] and L-glutamic acid. A similar conformational preference of backbone type II and type I has also been reported in the conformational study of several non-polar and polar amino acids.[64,79,80,89,87,117,158]

The vibrational assignment of the experimental features with the computed wavenumbers based on the conformational stability derived from relative free energy (ΔG_{MP2}) was also performed. While, the relative abundance of backbone type II conformers ($MET6_{II}$, $MET8_{II,}$ and $MET9_{II}$) are found to increase, the population distribution of conformer $MET2_I$ and $MET3_{III}$ were seen to decrease significantly to 2.3 % and 0.5 % respectively, as shown in Table 5.1. Thus,

based on ΔG_{MP2}, only backbone type II conformers, i.e., $MET1_{II}$, $MET6_{II}$, $MET8_{II,}$ and $MET9_{II}$ are found to be the predominant conformers. However, the vibrational assignment of the experimental feature with the computed features of only backbone type II conformer's, which are mentioned above, were found unsatisfactory. The closely spaced O-H and C=O vibrations of backbone type II conformer's, i.e., $MET1_{II}$, $MET6_{II}$, $MET8_{II,}$ and $MET9_{II}$ are found to cluster within few wavenumbers region and hence, shows a poor agreement with the several large shifted features observed in experiments. This largely shifted wavenumber are seen to be in excellent agreement with the backbone type I conformers ($MET2_I$).

The conformer stability computed at MP2/6-311++G(d,p) level of theory shows $MET1_{II}$ to be the lowest energy conformer, which is in good agreement with the experiments. However, the computations at M06-2X/6-311++G(d,p) level of the theory suggest conformers $MET3_{III}$ (0.00 kcal mol^{-1}), $MET1_{II}$ (0.07 kcal mol^{-1}) and $MET4_I$ (0.10 kcal mol^{-1}) be nearly isoergic. The almost similar stability and relative abundance of above three conformers with different backbone type, i.e., III, I, and II respectively, implies equally strong computed features. However, as shown in Fig. 5.6 (grid a, b and c), the experimental spectra in each region contain a strong feature along with other weaker features. Furthermore, the mean square deviation of the experimental and scaled computed features at M06-2X level of was found to be higher than that obtained with computed features at the MP2/6-311++G(d,p) level of theory. Thus, the computed wavenumbers at MP2/6-311++G(d,p) level yields better assignment of the experimental features and therefore have been used.

Based on the above-discussed assignment of experimental features at 3546.2 cm^{-1} to the computed scaled wavenumber corresponding to the υ (O-H) of $MET1_{II}$, the designation for other experimental features was made, as shown in Fig. 5.6 (grid a) and Table 5.2. The broad feature at

3183.5 cm^{-1} (trace i, grid a) can be assigned to the computed scaled wavenumber of backbone type I conformer, i.e., $MET2_I$ (trace iii, grid a), whose calculated scaled υ (O-H) was found at 3289.9 cm^{-1} (trace iii, grid a). The large redshift of ~350 cm^{-1} relative to the same mode of $MET1_{II}$ is due to an intramolecular hydrogen bonding interaction, and the observation is also corroborated with our computations. The observation of intramolecular hydrogen bonding in backbone type I conformer is also consistent with our previous study of L-threonine and L-glutamic acid.

The feature at 3559.5 cm^{-1} (trace i) can be assigned to the υ (O-H) of $MET3_{III}$, computed to occur at 3551.3 cm^{-1} (trace iv). Thus, in summary, the conformer with backbone type I geometry (i.e., $MET2_I$) shows the sizeable red-shifted O-H stretching features both in our computation and experiments. However, the backbone type II and III conformers are found to have comparatively closely spaced O-H stretching wavenumbers, again both in our experiments and computation. The observation is also supported by the presence of hydrogen bonding interaction in backbone type I conformers, which is absent is backbone type II and III conformers.

The experimental feature at 3399.7 (trace i, grid a) can be assigned to the υ (N-H) of $MET1_{II}$, which is computed to occur at 3385.5 cm^{-1} (trace i, grid a). The computed feature due to υ (N-H) of $MET2_I$ and $MET3_{III}$ were found to occur at 3310.3 and 3385.9 cm^{-1} respectively. However, due to poor infrared absorption cross-section and the poor abundance of these two conformers, their υ (N-H) vibrational features could not be observed in our experiments.

5.4.4 (b) 3100-2800 cm^{-1} region

The C-H stretching region which occurs between 3100-2800 cm^{-1} is shown in Fig. 5.6 (grid b) and Table 5.2. The computed wavenumbers corresponding to the C-H stretching mode were

scaled by a factor of 0.9500, which was obtained by assigning the strongest experimental feature at 2927.8 cm^{-1} (trace i) to the intense computed feature of $MET1_{II}$ (trace ii). The feature observed at 2973.4 cm^{-1} (trace i) is found in good agreement with the $\upsilon(C_{\beta}\text{-}H)$ of $MET1_{II}$, computed to occur at 2981.8 cm^{-1} (trace ii). A comparison with the computed wavenumber indicated that the features at 2925.4 cm^{-1} (trace i) could be assigned to the combination mode involving $\upsilon(C_{\Upsilon}\text{-}H)+ \upsilon(C\text{-}H_3$; methyl group) (trace ii). Likewise, a comparison of the experimental and computed feature suggests that the feature at 2931.4 and 2924.3 cm^{-1} (trace i) could be assigned to the υ(C-H) modes of conformer $MET2_I$ and $MET3_{III}$ respectively, computed to occur at 2927.0 cm^{-1} (trace iii) and 2922.7 cm^{-1} (trace iv) respectively. The Infrared absorption cross section of other υ(C-H) modes of $MET2_I$ and $MET3_{III}$ were negligible and hence were not observed in our experiments.

5.4.4 (c) 1800-1750 cm^{-1} region

The carbonyls stretching absorption of L-methionine occurs in the spectral region of 1800-1750 cm-1 and are shown in Fig. 5.6 (grid c) and Table 5.2. The computed scaled wavenumbers of υ(C=O) vibrations of $MET1_{II}$, $MET2_{I,}$ and $MET3_{III}$ are computed to occur at 1768.0 (trace ii), 1788.6 (trace iii) and 1759.0 cm^{-1} (trace iv) respectively. The strong experimental feature observed at 1768.0 cm^{-1} (trace i) is in excellent agreement with the computed wavenumber of $MET1_{II}$ (trace ii). The C=O stretching wavenumber of $MET2_I$ (trace iii) is found to be blue shifted by ~20 cm^{-1} relative to the same of the mode of $MET1_{II}$. The blue shift observed is consistent with our experiments, and the feature observed at 1787.2 cm^{-1} (trace i) can be assigned to the υ(C=O) vibration of $MET2_I$. The comparison with computed feature suggests that the experimental feature at 1737.9 cm^{-1} (trace i) can be assigned to the (C=O) vibration of $MET3_{III}$ (trace iv).

5.4.4 (d) 1450-1050 cm^{-1} region

Fig. 5.7 and Table 5.3 show the computed and experimental features in the vibrational region of 1450-1050 cm^{-1}. The vibrational motions of this spectral region occur in the form of skeleton vibration, bending, deformation, and torsional motions. The experimental spectra in this region shows dominant features at 1450.0 1442.4, 1430.3, 1369.9, 1327.3, 1312.3, 1252.4, 1216.7, 1145.0, 1134.9 and 1105.2 cm^{-1}, as shown in Fig. 5.7 (trace i) and Table 5.3. These experimental features are assigned to the computed wavenumber of conformer $MET1_{II}$, which are computed to occur at 1460.6, 1442.9, 1429.2, 1368.8, 1332.2, 1321.8, 1251.2, 1208.4, 1157.7, 1143.1 and 1117.6 cm^{-1} respectively (trace ii).

The other experimental wavenumbers are assigned to the computed wavenumber of conformer $MET2_I$ and $MET3_{III}$. The calculated scaled wavenumber of conformer $MET2_I$ are moderate, or high intensity is observed at 1404.2, 1386.4, 1323.9 and 1127.6 cm^{-1} (trace iii). These wavenumbers are in good agreement with the experimental features seen at 1390.5, 1377.8, 1307.5 and 1112.1 cm^{-1} (trace i). Likewise, the experimental features at 1386.3, 1361.6, 1145.0, 1133.6 and 1116.8 cm^{-1} (trace i) are assigned to the conformer $MET3_{III}$, which was computed to show features at 1381.1, 1343.4, 1179.7, 1141.8 and 1134.4 cm^{-1} (trace iv).

Even though the experimental features well corroborated by the computed wavenumbers of three lowest energy conformer; the vibrational assignments in this spectral region are admitted to be tentative, owing to the complexity of spectral feature in the skeleton vibration region. Therefore, based on the spectral assignment in the O-H, N-H, C-H, and C=O stretching region, we have observed three lowest energy conformer of L-methionine, i.e., $MET1_{II}$, $MET2_{I,}$ and $MET3_{III.}$

5.4.5 Atoms in molecule (AIM) analysis

As discussed earlier in the studies of L-threonine and L-glutamic acid, the backbone type I conformers are stabilized by hydrogen bonding interaction, but, backbone type II do not show such interactions. The presence of such non-covalent interactions in various conformers of L-methionine was examined by AIM analysis.[110] The criteria for defining the nature of non-covalent interaction were based on the topological parameter proposed by Koch and Popelier, as discussed in the earlier chapters.[111,112] The stabilization energy due to hydrogen bonding interaction, wherever present, was calculated employing the method suggested by Espinosa et. al.[113]

The bond critical points obtained after AIM analysis at MP2/6-311++G(d,p) level of theory, for lower energy conformer is shown in Fig. 5.9. Table 5.4 represents the topological parameters such as electron density ($\rho(r_c)$), Laplacian of electron density ($\nabla^2\rho(r_c)$) and hydrogen bond interaction energy (E_{HB}) at the bond critical points, corresponding to the non-covalent interactions site in various conformers of L-methionine.

The spectral shift of υ(O-H) mode observed in backbone type I conformer, i.e., $MET2_I$, as discussed earlier, indicated the presence of hydrogen bonding interactions, as discussed previously. The AIM analysis of conformer $MET2_I$ shows bond critical points along (O-H···N) bond path. The observation of hydrogen bonding interaction in backbone type I conformer is consistent with experimental results. The hydrogen bond energy (E_{HB}) corresponding to the (O-H···N) bond critical points was computed to be 10.7 kcal mol^{-1}. Likewise, the conformer $MET4_I$, another backbone type I also showed hydrogen bond along (O-H···N) bond path, as shown in Fig. 5.9, and the E_{HB} was calculated to be ~10.0 kcal mol^{-1}. Hence, the (O-H···N) hydrogen bond interactions, involving COOH and NH_2 group, is unique to backbone type I conformers of L-

methionine as well as other amino acids such as L-threonine, L-glutamic acid, as discussed in earlier chapters.

The AIM analysis of backbone type II conformer, i.e., $MET1_{II}$ does not show bond critical point along bifurcated ($C{=}O^{\cdots}H_2N$) bond path, involving, involving the carbonyl group of backbone carboxyl (COOH) and NH_2 group, as shown in Fig 5.9. As mentioned in the earlier chapter, the study in the past had suggested that a bifurcated hydrogen bond stabilizes backbone type II conformers of amino acids. However, our AIM analysis does not locate any hydrogen bond interaction in the backbone type II conformer of L-methionine. Furthermore, as shown in Fig. 5.9, no backbone type II conformers, i.e., $MET5_{II,}$ $MET6_{II,}$ $MET7_{II,}$ and $MET8_{II}$ shows bond critical point along bifurcated ($C{=}O^{\cdots}H_2N$) bond path. However, in some cases such as $MET5_{II}$ and $MET7_{II,}$ a weak interaction involving side-chain was observed. The observation is consistent with our previous study of L-threonine and L-glutamic acid, where also the backbone type II conformers did not show any bifurcated ($C{=}O^{\cdots}H_2N$) hydrogen bond. The stability of backbone type II structure, by NBO deletion study, was attributed to the vicinal orbital delocalization interaction, as discussed in earlier chapters.

While the AIM analysis of L-methionine is consistent with our experimental observations and results of previous studies, the issue that needs to be investigated is the role of the hydrophobic side-chain in determining the conformational preference of amino acid. The hydrophobic side-chain, unlike L-threonine and L-glutamic acid, is incapable of forming hydrogen bonding interaction with the backbone COOH and NH_2 moiety. As discussed above, the AIM analysis of conformer $MET1_{II}$ does not show any hydrogen bond either involving backbone moiety or side-chain group, as shown in Fig. 5.9. Whereas, the conformer $MET2_{I}$ exhibits a hydrogen bonding interaction, as suggested by AIM analysis and also by experiments.

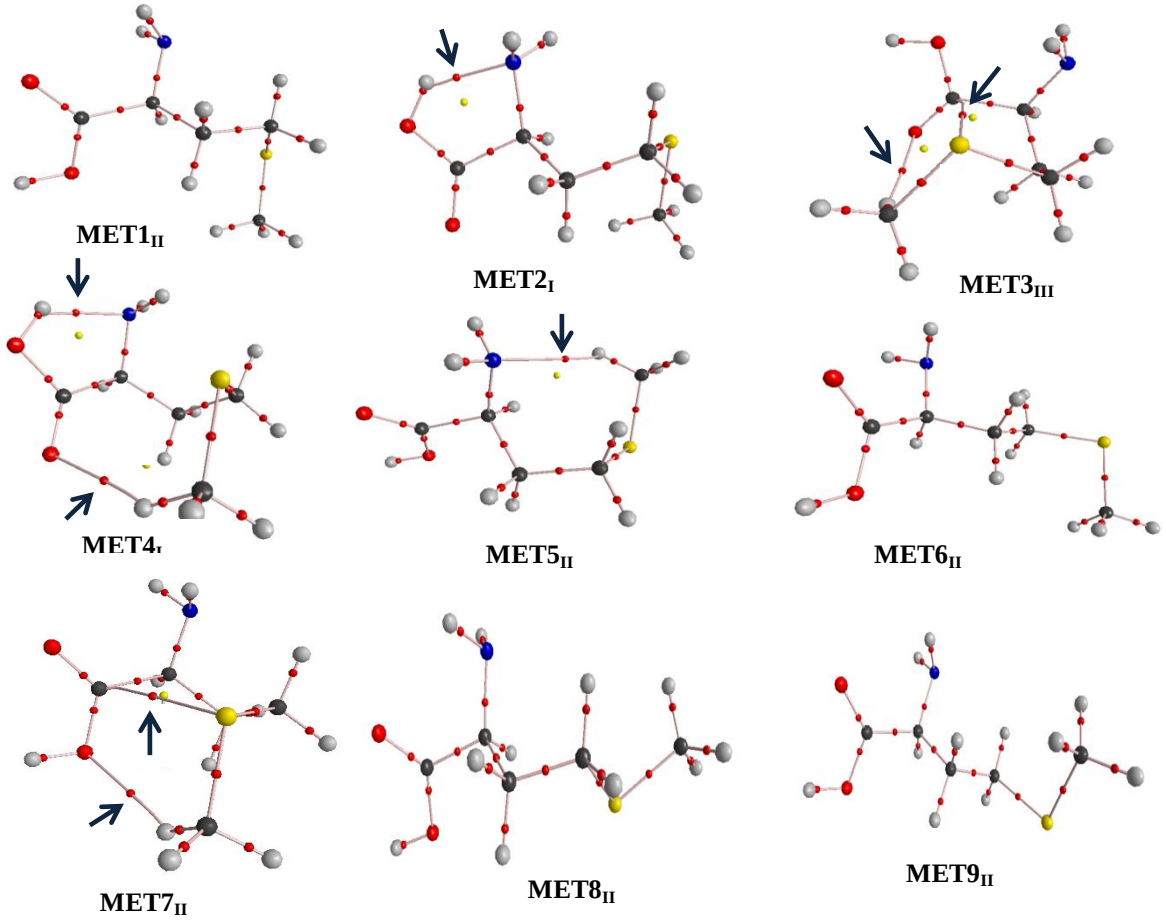


Fig 5.9 AIM analysis of the various conformers of L-methionine at MP2/6-311++G(d,p) level of theory. The bond critical points corresponding to hydrogen bond are indicated by the black arrow

Table 5.4 Electron density [$\rho(r_c)$], Laplacian of electron density [$\nabla^2\rho(r_c)$], local electronic kinetic energy density [$G(r_c)$], local electronic potential energy density [$V(r_c)$] and hydrogen bond energy (E_{HB}) values for hydrogen bonded interactions in the different conformers of L-methionine, at the MP2/6-311++G(d,p) level. The values for $\rho(r_c)$, $\nabla^2\rho(r_c)$, $G(r_c)$ and $V(r_c)$ are given in atomic units and for E_{HB}, in kcal mol^{-1}

Conformer	Interactions	$\rho(r_c)$	$\nabla^2\rho(r_c)$	$G(r_c)$	$V(r_c)$	E_{HB}
$MET2_I$	O-H···N	0.0381	0.1116	0.0310	-0.0341	-10.70
$MET3_{III}$	C=O···H-C(methyl)	0.0111	0.0380	0.0079	-0.0063	-1.99
	O=C···S	0.0113	0.0356	0.0076	-0.0062	-1.95
$MET4_I$	O-H···N	0.0360	0.1094	0.0295	-0.0317	-9.95
	C=O···H-C(methyl)	0.0124	0.0431	0.0091	-0.0074	-2.33
$MET5_{II}$	H-N···HC(methyl)	0.0070	0.0217	0.0044	-0.0033	-1.03
$MET7_{II}$	C=O···H-C(methyl)	0.0089	0.0285	0.0058	-0.0046	-1.43

Interestingly the conformer $MET1_{II}$ even without any hydrogen bonding interaction is relative more stables that conformer $MET2_{I}$ by 0.67 kcal mol^{-1} at MP2/6-311++G(d,p) level of theory, as shown in Fig. 5.3. Therefore, the role of the side-chain was examined by a comparative study with glycine, which was also used in previously studied amino acid. The comparative study between glycine and L-methionine suggest that the hydrophobic side-chain group does not play a defining role in the conformational preferences of L-methionine, and at best, they may be only supportive.

Based on the conformational studies of glycine (proton), L-threonine (polar side-chain), L-glutamic acid (acidic side-chain) and L-methionine (hydrophobic side-chain), a picture emerges that chemical nature of the side-chain group in amino acid has only little role in defining the conformational preference. The polar and acidic side-chain group, which can participate in hydrogen bonding interactions with the backbone moieties, either as hydrogen bond acceptor or donor, does not exclusively dictate the conformational preference on the concerned amino acid. Likewise, the hydrophobic side-chain group or no side-chain (glycine) plays only a supportive role and not the decisive factor for the conformational preference of respective amino acid.

5.5 Discussions

The energetics and structural preferences of various conformers of L-methionine were represented in the form of a conformational dartboard, as shown in Fig. 5.10. As discussed in earlier chapters, the eight tracks in dartboard plot represent the eight backbone structures of amino acids. Conformers with same backbone type fall on same respective track based on their relative energy. The angular position (θ) for any conformers is obtained from $\theta = \Delta E*(360/10)$, where ΔE is the energy of conformer relative to the $MET1_{II}$ at MP2/6-311++G(d,p) level of theory. The conformational dartboard of L-methionine employing structures and energy of all the

conformers obtained at MP2/6-311++G(d,p) level of theory, is shown in Fig. 5.10. The first octant of L-methionine conformational dartboard, which represents relative energy up to 1.25 kcal $mol^{-1,}$ is predominantly populated by conformers corresponding to backbone type I and II and rarely type III. The conformers corresponding to other backbone types, i.e., IV, I', II' III' and IV' begin to appear on the dartboard only beyond the first octant. This tendency of L-methionine low energy conformers to prefer backbone type I and II are also observed in Fig. 5.3 and Table 5.1. Of the ten lowest energy conformers shown in Fig. 5.3 adopts either backbone type I or II, except $MET3_{III}$. Likewise, all the 20 lowest energy conformers that are listed in Table 5.1 takes backbone type I and II, except two conformers.

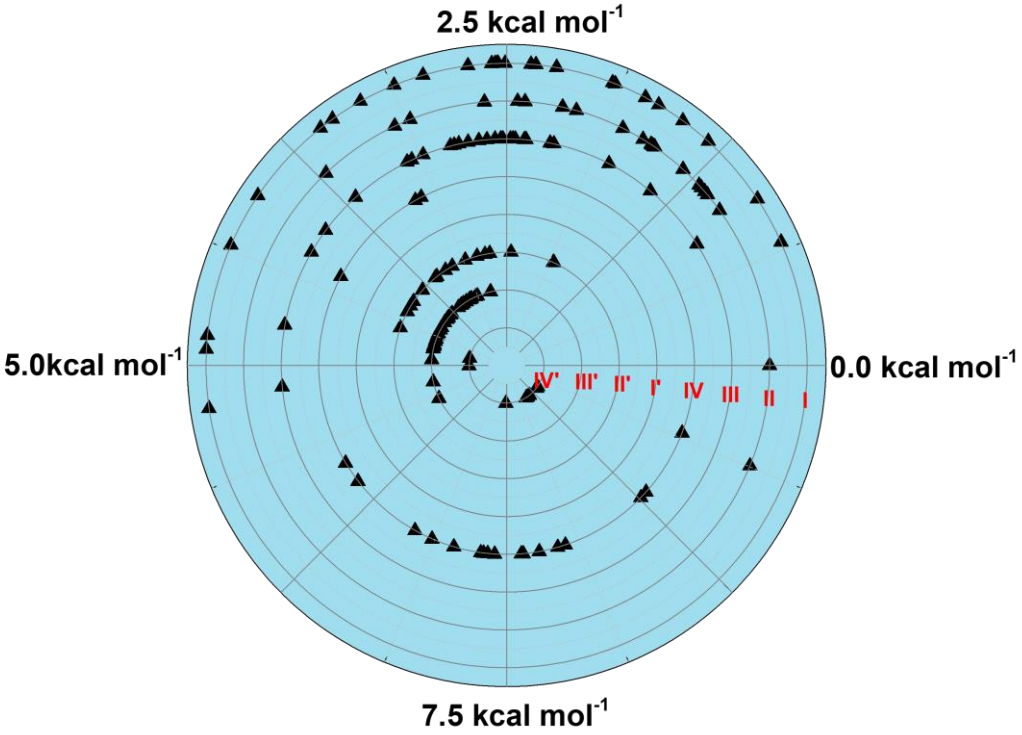


Fig. 5.10 Conformational dartboard of L-methionine at MP2/6-311++G(d,p) level of theory

The observation of low energy structure of L-methionine adopting only specific backbone geometries is consistent with our studies of other amino acids, i.e., L-threonine and L-glutamic acid, where also, of the eight backbone structures, the low energy conformers preferred mainly type I and II, as discussed in earlier chapters.[143] Furthermore, the observation also indicates the generality in the above trend, in the entire α-amino acids, irrespective of the nature of their side-chain moiety.

5.6 Conclusions

The conformations of L-methionine were studied using matrix isolation infrared spectroscopy and *ab initio* computations. The solid sample was deposited on the cold KBr window, kept at 12 K, with the help of a heated nozzle, maintained at 408 K and 423 K. The computational study was done by scanning the potential energy surface of molecule to obtain possible conformers, and a total of 167 and 176 conformers were found at MP2/6-311++G(d,p) and M06-2X/6-311++G(d,p) level of theories, respectively. The conformers of L-methionine were classified into eight classes, based on the orientation of backbone moieties. The vibrational features observed in the different regions of the experimental spectra were corroborated by the computed wavenumbers of three lower energy conformers viz. $MET1_{II}$, $MET2_{I}$, $MET3_{III}$. The backbone type I conformer ($MET2_{I}$), was found to be stabilized by an intramolecular (O-H$\cdots$N) hydrogen bond between backbone COOH and NH_2 groups. A large red-shift in the O-H stretching vibration of backbone type I conformer was observed and AIM analysis also located bond critical points along hydrogen bonding interaction. However, in case of backbone type II structure, whose stability was alluded to the presence of bifurcated ($NH_2\cdots$O=C) hydrogen bonding in earlier studies, the AIM analysis does not find a bond critical point along such interactions.

The role of the non-polar side-chain group in deciding the conformational preferences of L-methionine was analyzed by a comparative study with glycine, which is devoid of the side-chain. Like L-threonine (polar side-chain) and L-glutamic acid (acidic side-chain), the hydrophobic side-chain group of L-methionine does not play a decisive role in conformational preferences. Moreover, it was also observed that the lower energy conformers of L-methionine adopted mostly backbone type I and II structures, similar to that seen in L-threonine and L-glutamic acid. Thus, in summary, low energy conformers of α-amino acids prefer only specific backbone structures, irrespective of the chemical nature and bulkiness of the side-chain group present.

Chapter 6

α-Amino Acids, Tale of Two Structures

6.1 Introduction

Naturally occurring α-amino acids are an essential subset of biologically relevant molecules and are building blocks of peptides and proteins. The twenty naturally occurring α-amino acids are composed of common backbone structure, which consists of a carboxyl group (COOH), an amino group (NH_2), Cα and H, as shown in Fig. 6.1. The nature and composition of side-chain moiety vary for each amino acid. A broad classification of α-amino acids is based on the chemical nature of side-chain group *viz.* non-polar, polar, acidic and basic amino acids.

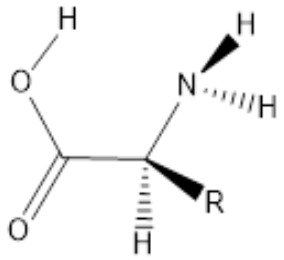


Fig. 6.1 Structure of α-amino acids

The diverse chemical nature and bulkiness of side-chain moieties of α-amino acids results in a complex but richer conformational landscape of α-amino acids. For example, glycine which has proton as a 'side-chain group' and is a three-rotor molecule, shows a simpler conformational landscape as compared with L-lysine, which has butylamine as a 'side-chain group'. Moreover, the amine side-chain group in L-lysine can also participate in an intramolecular hydrogen bonded interactions with backbone moieties. Thus, the varying nature of the side-chain group facilitates or prohibits the possibilities of non-covalent hydrogen-bonded interactions between the side-chain and backbone moiety of amino acids. It is therefore conceivable that the nature of the side-

chains groups in the various amino acids may play a vital role in deciding the conformational stability.

The structures of amino acids have been widely studied in the solid and solution phase, where the Zwitterionic form or charged species plays an important role in conformational stabilization. The gas phase study of amino acids conformations has its advantages of revealing the structure of neutral forms of amino acids (HOOC-CH(R)-NH_2, the canonical forms that are present in the peptide side-chains). Furthermore, the gas phase conformational study also presents the inherent molecular properties, free from the interactions occurring in the condensed phase. In the past, the conformations of various α-amino acids, individually, have been extensively studied using various experimental techniques and computational methods. The experimental techniques such as microwave spectroscopy, matrix-isolation FTIR spectroscopy, photoelectron spectroscopy, and others have been employed to explore the conformational landscape of individual α-amino acids. Likewise, in the computational regime, the conformations of α-amino acids have been understood using quantum mechanical calculations, as discussed in earlier chapters. Few reports, discussing the conformational landscape of the entire α-amino acids family using computational methods are present in the literature.[159-162] However, to the best of our knowledge, no report describing any general trends in the conformational preference of all the α-amino acids has been reported. In this chapter, we present a generalization in the structural preferences displayed by the amino acids.

As discussed in the earlier chapters, the side-chains with the different chemical nature, i.e., polar (L-threonine), acidic (L-glutamic acid) and non-polar (L-methionine) do not manifest a decisive role in the conformational preferences, and at best, they were found to be only supportive. Our experimental and computational studies these amino acids, i.e., L-threonine, L-

glutamic acids, and L-methionine, also revealed that the low energy conformers of all three amino acids show the preference for only specific backbone structures viz. type I and II. This observation is clearly presented using the "conformational dartboard" for these three amino acids, where the first octant (representing conformational energies of up to 1.25 kcal mol^{-1} above the lowest energy form) was found to be populated predominantly by the conformers adopting backbone type I and II structures, and to lesser extent type III. Therefore, the above observation of the low energy conformational preference for only certain backbone types and the side-chains group playing a supportive role provided an impetus for us to explore any possible generality that may be present over the entire α-amino acids. An understanding of the conformational landscape of all naturally occurring α-amino acids and rationalization of current preference for specific backbone types is useful in understanding their role in the structure of proteins.

The conformational landscape of α-amino acids other than glycine, L-threonine, L-glutamic acid, and L-methionine (which have been discussed in earlier chapters) was studied by *ab initio* computations. The potential energy surface (PES) scan was conducted for all these α-amino acids, except L-arginine, to obtained possible conformers. The experimental studies were also done for L-aspartic acid, L-glutamine, and L-histidine. But, owing to sample decomposition in the process of deposition by heating, the vibrational assignments were not done. Apart from exploring the backbone preference by low energy conformers of α-amino acids, we also attempted to quantitatively estimate the conformational preference, in terms of backbone propensity, which is a quantitative measure of the population of the backbone structure in question for any given amino acid. Furthermore, any prevailing preferred orientation of the side-chain groups of α-amino acids was also investigated.

6.2 Computational Methods

As already described, potential energy surface (PES) scan was carried out to obtain possible conformers of α-amino acids, following the schemes described in the earlier chapters. The final numbers of conformers obtained for different α-amino acids are listed in the last column of Table1, where for glycine only seven conformers were located, but in the case of L-lysine 1607 local minima were calculated. The PES scan of L-arginine was not carried out owing to the limitation of computational resources. L-arginine exists in two isomeric forms and the large side-chain results in ~28500 initial geometries for each isomeric forms, thus working up a total ~57000 initial geometries for PES scans.

All the conformers of α-amino acids were classified into eight classes, as shown in Fig. 6.2, based on their backbone geometry, a scheme which has been discussed in the earlier chapters. The same type of nomenclature described earlier was used for all the amino acids, which were based on their backbone structure. The conformational dartboard based on conformers backbone structure and energy for various α-amino acids were plotted. The population distribution using the Boltzmann equation was calculated for the different conformers of a given α-amino acid using their energy relative to the lowest energy conformer of that particular α-amino acid. Likewise, the population of various backbone types or backbone propensity for a corresponding α-amino acid were also computed.

To comprehend any general preference of side-chain orientations in α-amino acid, the side-chain dihedral angle χ_1 (N-C_α-C_β-X), where X is the heaviest atom or group, of all the conformers of various α-amino acids, were extracted. The preference in the side-chain orientations was analyzed by plotting the side-chain dihedral angle χ_1 and the relative abundance of conformers of various α-amino acids.

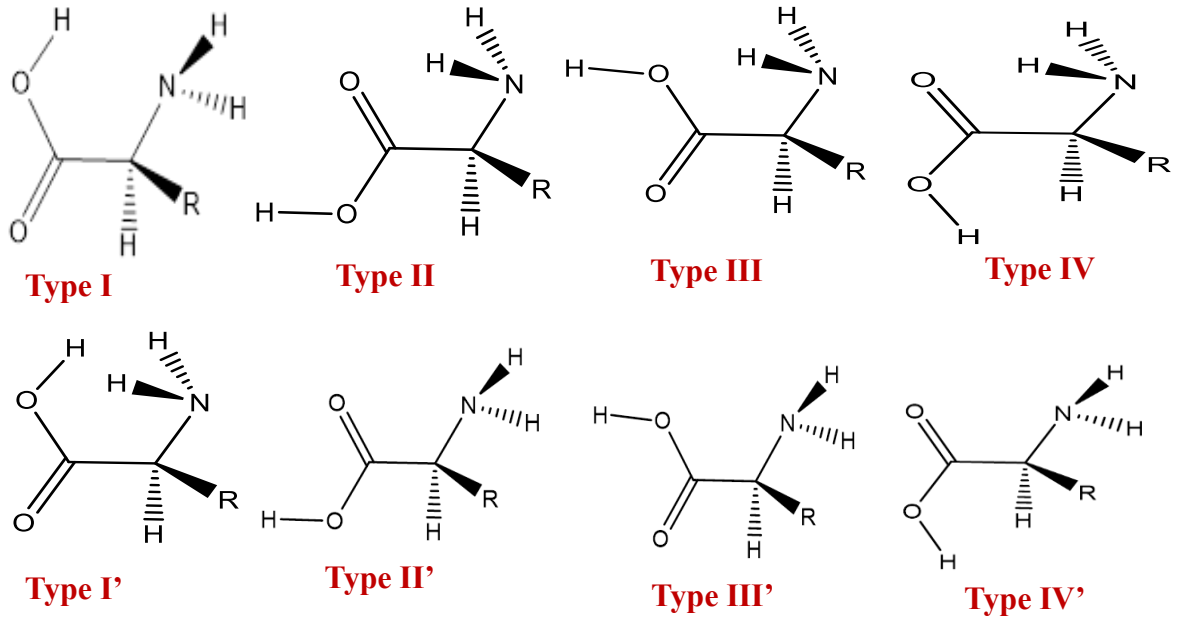


Fig. 6.2 Eight backbone structures of α-amino acid

Table 6.1 List of the total number of geometries optimized and final number of conformers obtained for each amino acid, except arginine, in the potential energy surface scan method

Amino acid	a	b	c	d
Gly	36	11	8	7
Ala	36	10	8	8
Val	108	25	21	21
Leu	648	99	83	80
Ile	648	71	65	64
Ser	648	58	45	45
Thr	648	51	44	43
Cys	648	40	31	30
Met	1944	225	186	176
Asp	864	107	86	83
Asn	1296	61	48	43
Glu	2592	394	317	299
Gln	3888	189	126	114
Phy	648	33	31	30
Trp	648	71	69	67
Tyr	2160	70	59	56
His	648	64	52	50
Lys	8748	2968	1832	1609
Pro	24	10	9	9
Arg	~57000	-	-	-

a; Number of initial geometries optimized at HF/6-31G level of theory
b; Number of stationary points *(obtained after optimization in previous step)* optimized at B3LYP/6-311G level of theory
c; Number of stationary points *(obtained after optimization in previous step)* optimized at M06/6-311++G(d,p) level of theory
d; Final number of conformers *(each conformer shows real frequency for all the normal modes.)*

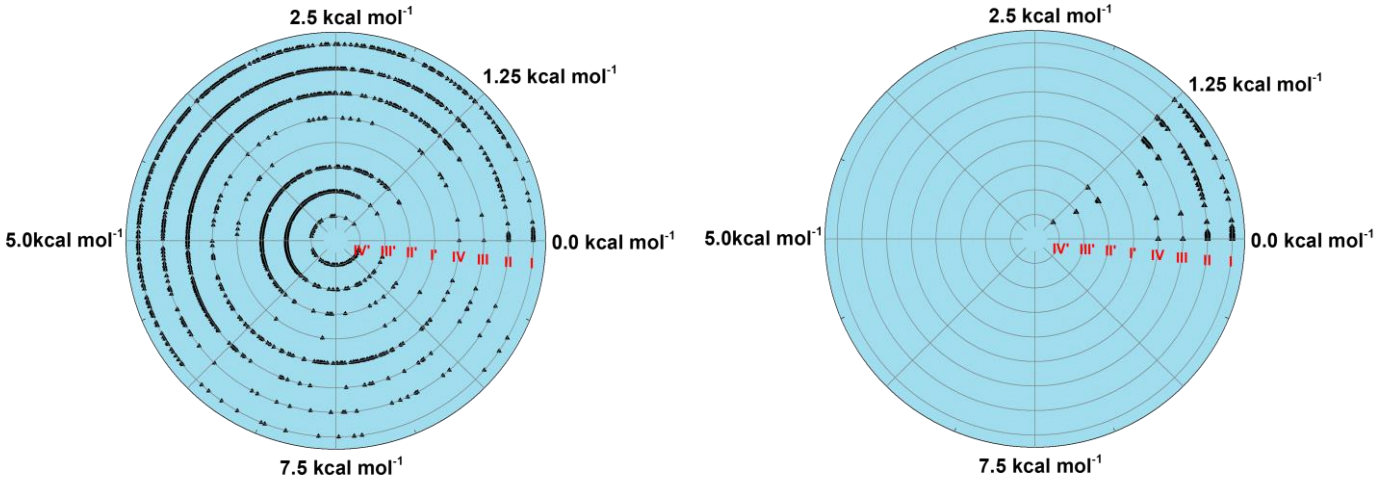


Fig. 6.3 a) Conformational dartboard of α-amino acid representing the conformers up to relative energy of 10.0 kcal mol^{-1}; b) The conformational dartboard of α-amino acids constituting the conformers up to relative energy of 1.25 kcal mol^{-1}

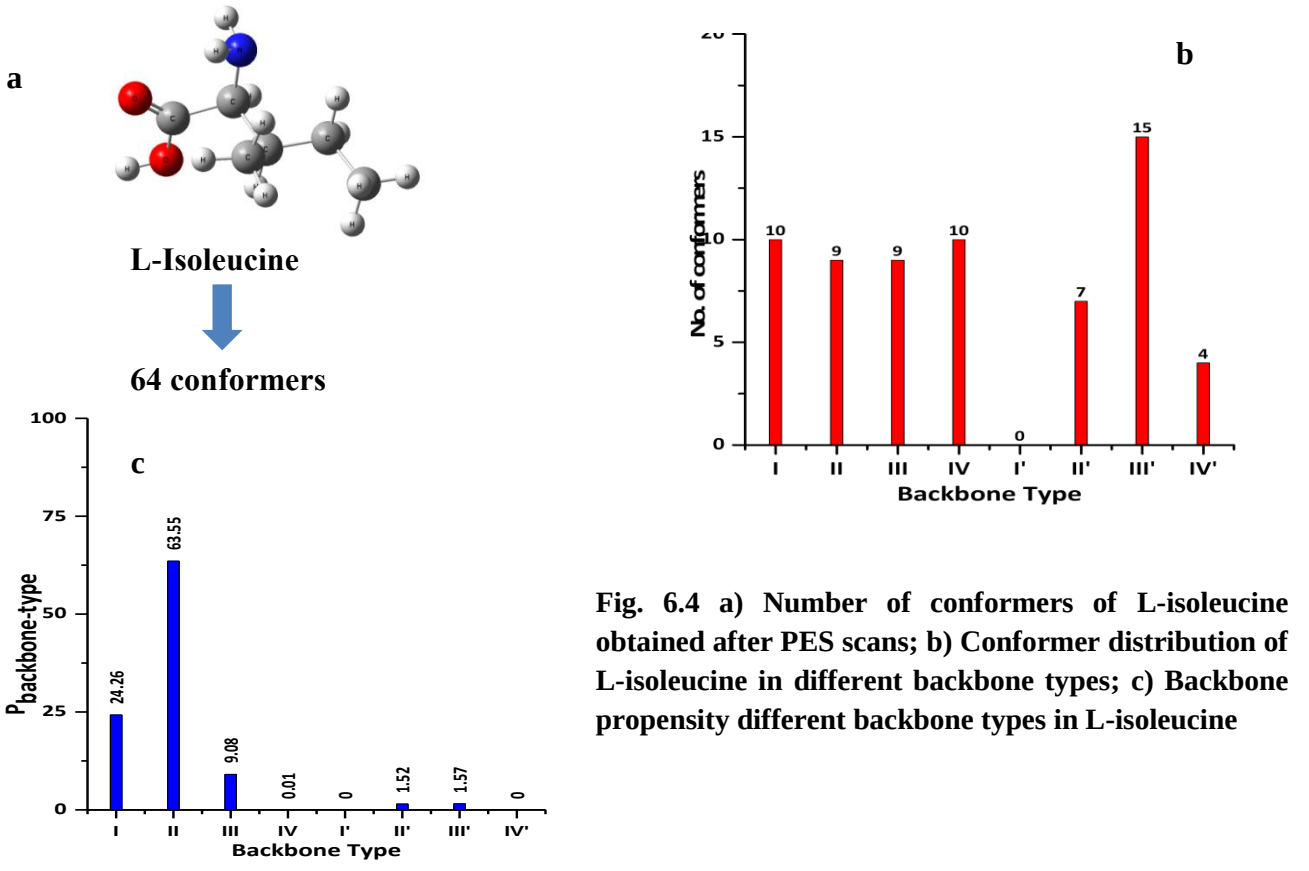


Fig. 6.4 a) Number of conformers of L-isoleucine obtained after PES scans; b) Conformer distribution of L-isoleucine in different backbone types; c) Backbone propensity different backbone types in L-isoleucine

6.3 Results and Discussions

6.3.1 The conformational dartboard of α-amino acids

The conformational dartboard comprising of different conformers of various α-amino acids, except L-arginine, plotted together is shown in Fig. 6.3a. The conformational dartboards of various α-amino acids were also plotted separately, but have not been shown here to avoid repetition. All the eight tracks of the conformational dartboard, which corresponds to the eight different backbone structures viz. type I, II, III, IV, I', II', III' and IV' are populated by a large number of conformers corresponding to each backbone types, but with varying relative energy. A closer look at the first octant of the α-amino acid conformational dartboard, which represents conformers with relative energy up to 1.25 kcal mol^{-1}, indicates a preferential distribution, of conformers belonging to type I and II backbone structures. To better observe this trend, the conformational dartboard for all the conformers of the different α-amino acids with relative energy up to 1.25 kcal mol^{-1}, were plotted separately and is shown in Fig. 6.3b. As shown in Fig. 6.3b, in the first octant, the number density on track I and II are distinctly higher when compared with the other tracks, which are comparatively populated only rarely. The large number density on the track I and II of the conformational dartboard in the first octant region, clearly indicates the intrinsic tendency of low energy conformers of α-amino acids, to prefer only backbone type I and II structures. This inherent tendency of low energy conformers of α-amino acids to adopt only specific backbone structures, irrespective of the chemical nature and bulkiness of side-chain group, further support similar observations we made earlier with the polar (L-threonine), acidic (L-glutamic acid) and non-polar (L-methionine) amino acids.

It is noteworthy to mention that the conformations of α-amino acids have been extensively studied for the past two decades, using various experimental and computations

methods. However, to the best of our knowledge, there are no reports discussing the intrinsic backbone preference of the low energy conformers in α-amino acids, irrespective of nature of side-chain moiety. This report, therefore, constitutes the first attempt to rationalize the conformational landscape of all the α-amino acids, except L-arginine, based on the intrinsic backbone preference of the low energy conformers of α-amino acids.

6.3.2 Backbone propensity, a quantitative estimate of intrinsic backbone preference in the α-amino acids

In the above section, we qualitatively observed that the low energy conformers of various α-amino acids exhibit an intrinsic tendency to adopt backbone type I and II geometries. To quantitatively describe this inherent behavior of α-amino acids, we propose a term "backbone propensity," which is a quantitative measure of the population of a given backbone structure of amino acid. To explain this concept, let us take a case study of L-isoleucine, as an example. The PES scan for this amino acid resulted in 64 conformers, which were then classified based on the backbone structure of conformers into eight classes, which are shown in Fig. 6.2. Fig. 6.4b is a bar graph showing the number of conformers in each backbone type for this amino acid. From this bar graph, it can be that there are ten conformers of L-isoleucine in type I geometry, and nine conformers in type II and III and interestingly, 15 conformers with backbone type III'. It must be noted that these are the total number of conformations in each backbone type; the dartboard clearly shows that the *lower energy forms* adopt predominantly type I and II only. The normalized Boltzmann distribution for each conformer at room temperature (298 K) can be obtained using the formula shown below. For all the α-amino acids the conformational degeneracy was taken to be unity.

$$P_{conformer} = \frac{e^{-\frac{\Delta E}{kt}}}{\sum_1^n e^{-\Delta E/kt}}$$

ΔE= Relative energy; T = 298K; n= no of conformer

The backbone propensity, which is the quantitative estimate of the population of a given backbone type can thus be obtained by summing up the normalized Boltzmann distribution of all the conformers belonging to the backbone type under consideration, using the equation shown below.

$$P_{backbone\ type} = \sum_{1}^{m} P_{conformer}$$

m= no of conformer in a given backbone type

The backbone propensity of all eight backbone structures in L-isoleucine, which were computed using the above equations, were then plotted as a bar graph for all the eight backbone geometries, as shown in Fig. 6.4c. The population of backbone type II, which represents nine conformers of L-isoleucine, as shown in Fig. 6.4a was calculated to be ~64 % (P_{II}), whereas the ten conformers of backbone type I geometry sums up to yield backbone propensity of ~24 % (P_{I}). Likewise, the nine conformers of type III backbone yielded backbone propensity of ~9 % (P_{III}). Interestingly, the 15 conformers, which adopted backbone type III' structure contribute only ~1.5 % (P_{III}) to the total population of L-isoleucine at room temperature, hence implying that all those conformers are high energy form. Thus, the significant fraction of L-isoleucine population at room temperature, i.e., ~88 %, adopts only backbone type II and I (P_{II+I}).

The backbone propensities of different backbone types for all the α-amino acids were estimated, employing the above procedure. The sum of the backbone propensity of type I and II, (which are the most preferred geometries) was then plotted as a bar graph for each amino acid, as shown in Fig. 6.5. Amino acids such as L-histidine and L-proline shows 98% and 99 %, of the total population, respectively, in the backbone type I and II(P_{II+I}), whereas the sum of backbone propensity of type I and II(P_{II+I}) in case of L-aspartic acid and L-cysteine were calculated to be 69% and 73 % respectively, as shown in Fig. 6.5. Of all the α-amino acids, only L-lysine has the

least propensity of backbone type I and II, which was estimated to be 40 %. The average backbone propensity of type I and II (*avg.* P_{II+I}) over entire α-amino acids was found to be ~81% with a standard deviation of ~13%. Hence, the quantitative analysis of the backbone population of various α-amino acids establishes that at room temperature, a significant population of α-amino acids exist in two backbone structures. Moreover, this backbone propensity analysis is consistent with our conformational dartboard results, where the first octant had shown exceedingly large number density on track I and II relative to other tracks, and, hence qualitatively demonstrated the existing intrinsic preference of low energy conformers of α-amino acids to adopt backbone type I and II geometries.

6.3.3 Relative preference between backbone type I and II geometries

Based on our study discussed in previous sections we have demonstrated, both qualitatively and quantitatively, the existence of an intrinsic backbone preference in α-amino acids, and significant population pooling into conformers with backbone type I and II structures. We then examined the relative choice between these two backbones. We had already indicated that, on an average, as much as ~81% of the population existed in these two forms only. The selectivity between these two backbone types was explored using the Boltzmann distribution equation. As shown in Fig. 6.6a, the population in backbone type II (P_{II}) and type I (P_I) in case of L-alanine was estimated to be 61 % and 16 % respectively. Substituting the P_{II} and P_I population values in the Boltzmann equation, as shown in Fig. 6.6b, one can compute the average relative stability of backbone type II relative to I $(E_{II}\text{-}E_I \text{ or } \Delta E_{II\text{-}I})$, which in the case of L-alanine, was found to be -0.82 kcal mol^{-1} (Fig. 6.6a). The negative sign implies that backbone type II is relatively more stable than type I or conversely the preference for backbone type II is higher than type I, which is

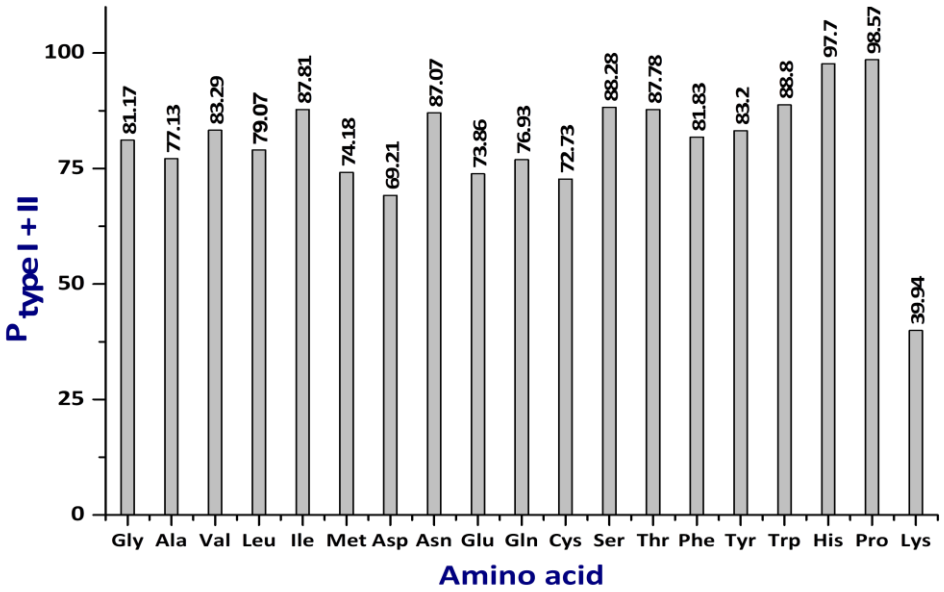
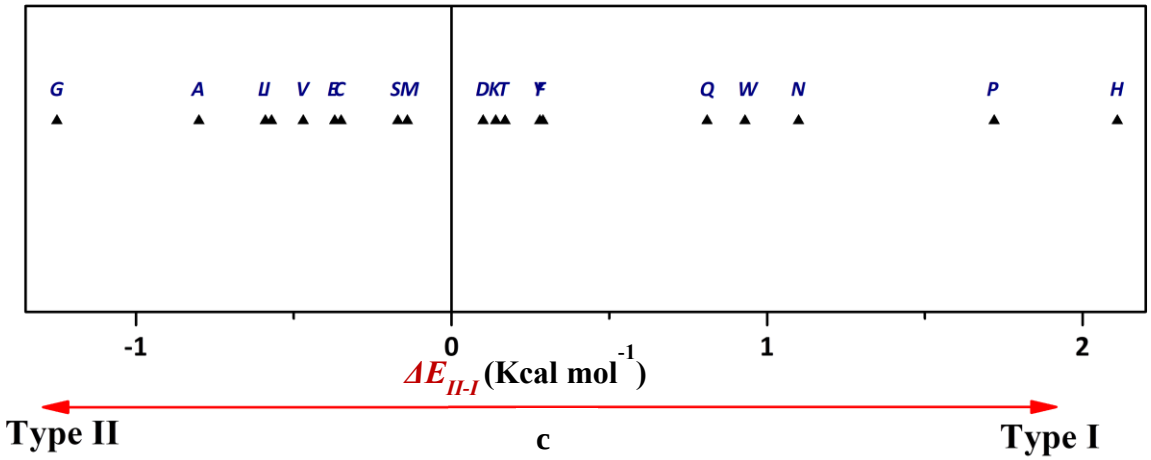


Fig. 6.5 Sum of the backbone propensity of type I and II of naturally occurring α-amino acid (except arginine)

a

Amino Acid	P_{II} (%)	P_I (%)	ΔE_{II-I}
Ala	61.17	15.96	-0.80
Asn	11.78	75.29	1.10

b $\frac{P_{II}}{P_I} = g\, e^{-\frac{(EII - EI)}{kT}}$ ΔE_{II-I} Relative stability of ***type II*** w.r.t to ***type I***



c

Fig. 6.6 a) Backbone type I and type II population of alanine and asparagine; b) Boltzmann equation for backbone stability of type II relative to type I (ΔE_{II-I}); c) Relative backbone stability plot for α-amino acid

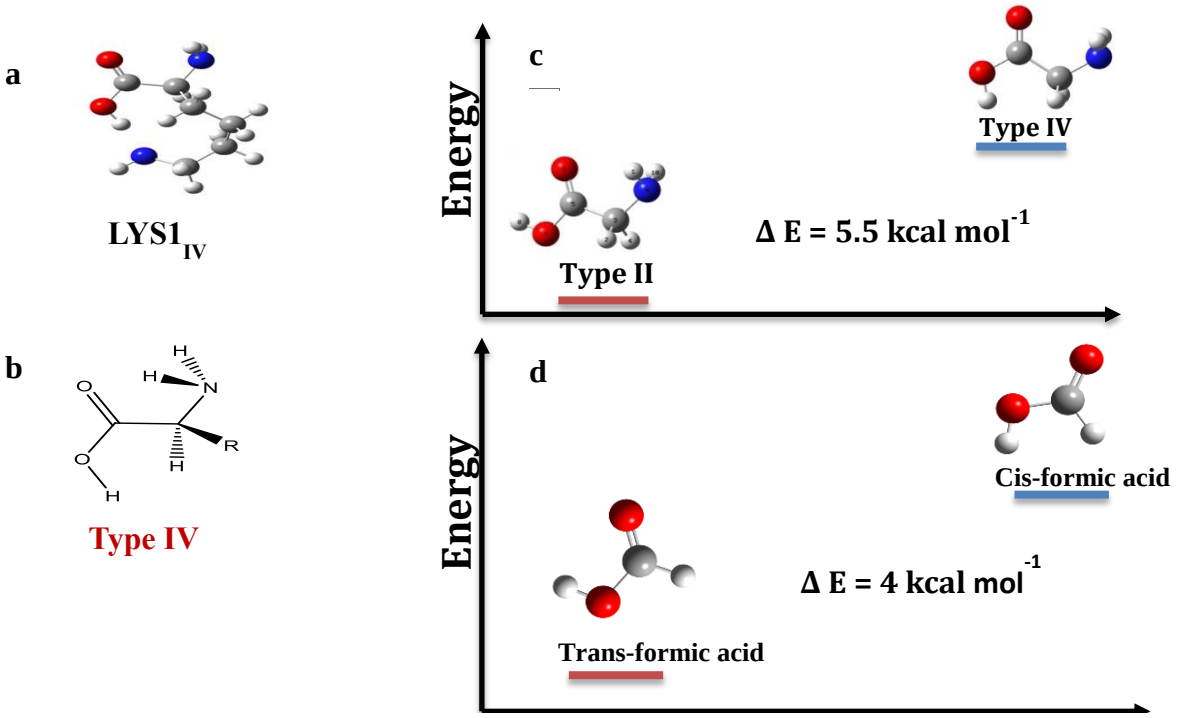


Fig. 6.7 a) Lowest energy conformer of Lysine; b) Backbone type IV structure; c) Relative stability of backbone type II and IV conformers of glycine; d) Relative stability of trans and cis conformers of formic acid

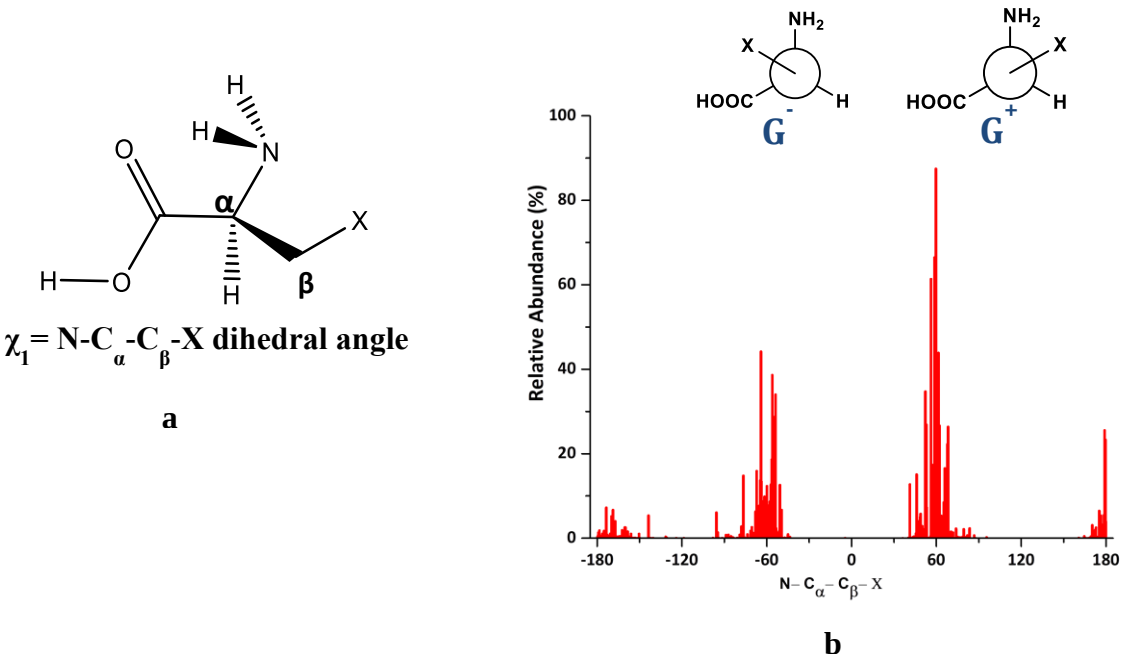


Fig. 6.8 a) Side-chain dihedral angle (χ_1) of α-amino acids; b) Distribution of side-chain dihedral angles (χ_1) in α-amino acids

also evident from the population distribution of these two backbone geometries in L-alanine. Likewise, in the case of L-asparagine, the stability of type II relative to type I *(ΔE_{II-I})* was computed to be 2.38 kcal mol^{-1}. The positive value of ΔE_{II-I} suggests that in L-asparagine, the backbone type II is less stable than type I, or conversely, the backbone type II is less preferred than type I, which is again evident from the population of these two backbone structures in L-asparagine, as shown in Fig. 6.6a.

Using the above method, we computed the average relative stability of backbone type II relative to type I *(ΔE_{II-I})* for the rest of the α-amino acids. The ΔE_{II-I} values obtained for various amino acids were plotted together on a relative backbone stability plot, as shown in Fig. 6.6c. The group of α-amino acids lying in the positive region (right side) manifest a lower preference for backbone type II structure and a higher choice for type I geometry. The bunch of amino acids falling in the negative region of the plot shows a higher preference for backbone type II structures as compared to type I.

The amino acids such as L-histidine (shown as H), L-proline (P) and L-asparagine (N), shows large positive ΔE_{II-I}, as shown in Fig. 6.6c, has a significant fraction of their population in the form of backbone type I. But, in the case of Glycine (G), L-alanine (A) and L-leucine (L) a significant negative value of ΔE_{II-I} were observed, thus suggesting that the large fractions of population in these amino acids are present in backbone type II geometry. The amino acids present in the center region exhibit a smaller value of ΔE_{II-I} that indicates a nearly equal preference for both backbone types I and II structures. Hence, based on relative stability plot, we have rationalized the relative choice for backbone type I and II in various α-amino acids.

6.3.4 Curious case of L-lysine

As discussed in the earlier section, a significant fraction of the population of different α-amino acids, at room temperature, exist in the backbone type I and II. L-lysine, however, appears to be an exception, where the sum of backbone propensity of type I and II was estimated to be only ~40%. In L-lysine, the most stable conformer adopts type IV backbone geometry, as shown in Fig. 6.7a. The factors driving this unusual behavior of L-lysine were therefore explored. Interestingly, in glycine, the conformer with a type IV backbone structure was found to be a high energy form, being 5.5 kcal mol^{-1} higher relative to the most stable conformer, which adopts backbone type II structure (Fig. 6.7c).

The lower stability of the type IV structure of glycine relative to type II structure appears to be due to the orientation of carboxyl (COOH) group, very much on the lines of the trans geometry of formic acid being more stable than the cis. The conformer of formic acids have been extensively studied in the literature, and the geometry of trans formic acid has been observed to be more stable than the cis by ~ 4 kcal mol^{-1} (Fig. 6.7d).[163] The COOH group orientation in trans formic acid is similar to that present in type II conformer of glycine, whereas the orientation in cis formic acids parallels that observed in type IV conformer of glycine. Hence, it is not surprising that the type IV structure in glycine has a higher form and is not preferred. In principle, this unfavorable orientation of the carboxylic acid group as seen in type IV geometry can be sometimes be offset in amino acids, where the side-chain groups can non-covalently interact with backbone moiety and provide stability. However, no other amino acid, except lysine, shows a preference for backbone type IV structures.

In the case of L-lysine, the butylamine side-chain group was found to be long enough to turn around and establish a stable hydrogen bonded interaction between the backbone COOH

group and side-chain NH_2 moiety. The long, flexible, basic side-chain group enables this strong hydrogen bonding with the COOH group, thus making type IV structure as the most stable form, notwithstanding the unfavorable orientation of the carboxylic group.

One can also expect, that there exists the possibility of L-glutamine and L-asparagine to adopt the type IV structure, as they possess long side-chains, such as propionamide (CH_2-CH_2-CO-NH_2) and ethanamide (CH_2-CO-NH_2) respectively. However, unlike L-lysine, in the case of L-glutamine and L-arginine, the side-chain is not long enough to turn around and adopt the energetically favorable form and thus assist in stabilizing the type IV form. Furthermore, the NH_2 group in L-glutamine and L-asparagine are in the form of an amide where the nitrogen, compared with that in the amine, is a weak hydrogen bond acceptor. In L-lysine, the NH_2 group exists in the amine form where the nitrogen can act as strong hydrogen bond acceptor. Therefore, the backbone type IV structures are not preferred in L-glutamine and L-asparagine.

To summarize, all the α-amino acids, except lysine shows an intrinsic tendency for backbone type I and II structures and the side-chain group plays at best a supporting role and not the decisive. However, in L-lysine, the low energy form adopts backbone type IV structure, due to strong hydrogen bonding interaction between the side-chain and backbone group. Thus, unlike other amino acids, only in L-lysine, the side-chain plays a vital role in determining conformational preference and stability.

6.3.5 Side-chain orientations

In the previous sections, we observed an inherent tendency of low energy conformers of α-amino acids, except L-lysine, to adopt only backbone type I and II structures. We next examined to understand if the orientation of side-chain groups in the various conformers of different α-amino acids displayed any preference in its orientation. The alignment of the side-chain group in

various conformers of different α-amino acids was rationalized in terms of side-chain dihedral angle $\chi_{1,}$ defined as the dihedral angle between N-C_α-C_β-X, where, X is the heaviest atom or group present in the side-chain (Fig. 6.8a). The nature of the X atom varies for different α-amino acids; being for example O in the case of L-serine and L-threonine, S in the case of L-cysteine, and C in the case of L-aspartic acid.

The χ_1 dihedral values of the different conformers of various α-amino acids were then plotted against the relative abundance of such conformers, as shown in Fig. 6.8b. The χ_1 dihedral angle shows a preference for +60º and -60º and is only rarely observed to adopt ±180º (Fig. 6.8b). The preferred orientation in the area of +60º and -60º corresponds to the (+) *gauche* (G^+) and (-) *gauche* (G^-) respectively, whereas the dihedral region of ±180º represents the *trans*. Hence, the side-chain dihedral angle (χ_1) distribution analysis establishes that the conformer's α-amino acids prefer only specific side-chain orientation viz. (±) *gauche* forms. This behavior is not surprising, though, and it parallels the behavior of butane, which also prefers a *trans (T)*, (+) *gauche* (G^+) and (-) *gauche* (G^-) orientation. However, the *trans* form which happens to be the most stable in case of butane is found to be less preferred in α-amino acids relative to the (±) *gauche* forms, as shown in Fig. 6.8b. This behavior in the dihedral preferences of the side-chains in α-amino acids could be attributed to the presence of COOH and NH_2 moieties in the α-amino acids. In the (±) *gauche* forms, the side-chain of the α-amino acids can approach in relatively close proximity the COOH and NH_2 groups, thus participating in non-covalent interaction such as hydrogen bonding. However, the *trans* orientation likely precludes such a possibility, thus making the trans, less preferred relative to (±) *gauche* forms in the α-amino acids. Thus, the conformational landscape of the α-amino acids adopts not only specific backbone structures as low energy forms but also prefer specific side-chain orientations.

6.4 Conclusions

The conformational landscapes of α-amino acids were explored using the *ab initio* quantum mechanical computational studies. The potential energy surfaces (PES) of various α-amino acids were scanned to obtain possible conformers. All the conformers of α-amino acids were classified into eight classes based on the backbone structures. The energetics and backbone structures of different conformers of various α-amino acids were represented in the form of the conformational dartboard. The conformational dartboard comprising of the conformers of various α-amino acids together depicted an inherent tendency of the low energy conformers of α-amino acids to prefer only backbone structure I and II and rarely other backbone geometries. The intrinsic preference for backbone type I and II was independent of chemical nature and steric bulkiness of side-chain moieties of various α-amino acids, except lysine. Thus, the side-chain group of most of the α-amino acids, except lysine, were observed to play a supportive role in determining the conformational stability and not the decisive.

The backbone preferences in various α-amino acids were also quantitatively rationalized based on backbone propensity, which is the quantitative measure of the population of a given backbone structure of corresponding α-amino acid. At the room temperature ~81% of the total population of all α-amino acids, except L-lysine, were found to exist in the form of backbone type I and II. The relative preferences between backbone type I and II in various α-amino acids were understood in terms of the backbone type II stability relative to type I (ΔE_{II-I}), which was deduced using the Boltzmann distribution equation and the population of backbone type I and II, as discussed earlier. The set of α-amino acids having a negative value for ΔE_{II-I} shows a higher preference for backbone type II relative to type I and vice-versa.

Of all the α-amino acids, only L-lysine showed an unusual behavior, in which the lowest energy conformer adopted backbone type IV structure. The sum of backbone propensity of type I and II was estimated to be ~41%, much lesser to that calculated for other α-amino acids. The backbone type IV structure in L-lysine is stabilized due to a strong intramolecular hydrogen bonded interaction between the side-chain and backbone COOH group. Unlike other α-amino acids where side-chain groups were found to play an only supportive role, in the case of L-lysine, the butylamine side-chain group plays a decisive role in defining the conformational preference and stability.

The preference of the side-chain orientations in the conformers of various α-amino acids was understood by analyzing the relative distribution of side-chain dihedral angle χ_1 (N-C_α-C_β-X). Like the selectivity observed in the backbone structures, the side-chain groups majorly adopted (±) gauche orientations.

Thus, based on our systematic and extensive study of the conformational landscape of α-amino acids, we observed the existence of an inherent tendency of low energy conformers of various α-amino acids to prefer only certain backbone structures and specific side-chain orientations. The work presented in this chapter and earlier chapters of the thesis reveals a pattern in an otherwise seemingly complex conformational landscape of α-amino acids. A general rationalization of the conformational landscape of α-amino acids is necessary for understanding its role of this building block in the structure of peptides and proteins, which will be discussed in the next chapter.

Chapter 7

What is the Role of the Amino Acids Conformational Dichotomy on the Conformational Landscape of the Dipeptides?

7.1 Introduction

The twenty naturally occurring α-amino acids are the building blocks of proteins and peptides, whose functions intimately depend on the structure that the biomolecules adopt. Protein folding, a very complex phenomenon, is a process by which a chain of polypeptides acquires its native or functional form.[5] Based on the reversible nature of the protein folding process, it was theorized by Anfinsen, that the native state of a protein is intrinsic to its amino acids sequence.[164] Furthermore, it is likely that there exists some inherent relationship between the structure of the proteins or peptides and the molecular conformations of the individual amino acids residues that make up the proteins. This intrinsic relationship between the native form of protein and the amino acid sequence has been rationalized based on *conformational propensity*, which is the quantitative estimate of the occurrence of a given amino acid residue in the different secondary structure region, i.e., helix, strand, turn and others.[9]

In the literature, there exist extensive studies on the *conformational propensities* of amino acid residues in various proteins and peptides.[10,11,12] However, the conformational propensity of amino acid residues reflects both the intrinsic propensity of amino acid residue and the influence of neighboring residues (environmental effects). Hence, the conformational propensity of any given amino acids residues varies from one protein to the other, and thus, conformational propensity is 'context dependent.' Moreover, these conformational propensities do not replicate the unfolded structure of the protein or the initial steps of secondary structure formation.

The understanding of the unfolded structure of proteins is based on *intrinsic conformational propensities* of the amino acid residues, which is a quantitative estimate of the

tendency of amino acid residues to adopt helix, strand or other secondary structure, in the absence of environmental influence.[17-25] The *intrinsic conformational propensities* of amino acid residues have been extensively studied in the literature, employing various small peptides as a model system and with the use of different experimental and computational methods.[26-36] However, the physical origin of *intrinsic conformational propensities* is not well established and is still a matter of debate. According to Vymetal *et al.*[36], *"The physical origin of this distinct propensities is still unclear, and it is questionable whether it can be investigated by computer simulation using the current method and models."*

Exploring the conformational landscape of dipeptides, the smallest unit of a peptide, is, therefore, necessary for understanding the conformational preference or *intrinsic conformational propensities* of different amino acid residues. It will also be useful to explore if the preferred backbone structures (Type I and II) of α-amino acids play any role in determining the structural preference of dipeptides. The relationship between conformational preferences of monomeric amino acids with that of dipeptides could possibly result in a vital link for rationalizing the physical origin of the *intrinsic conformational preferences* of dipeptides, which is discussed in this chapter.

The relationship between conformational preferences of the α-amino acids and intrinsic propensities of dipeptides was justified based on the geometrical similarity of preferred structures in both cases. Moreover, the association between the structural preferences of α-amino acids and dipeptides were also understood based on the Ramachandran region or (φ, ψ) space adopted by the stable conformers.

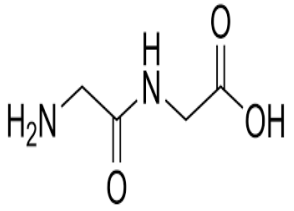
The conformations of dipeptides have been widely studied using various experimental and computational methods.[159-184] The conformational study in gas phase provides an

opportunity to explore the intrinsic propensities of these molecules, under a solvent-free environment and free from intermolecular interactions occurring in the condensed phase. The conformational stability of natural dipeptides such as Gly-Gly (Fig 7.1a), and others, are governed by the interactions due to terminal COOH (C- terminal) and NH_2 (N- terminal).[183,184] However, in a regular polypeptide chain (Fig. 7.1b), the amino acid residue lying in the inner region are far from the terminal moieties. Therefore, the conformational preference or the *intrinsic conformational preferences* of these amino acid residues, in polypeptides, remain unaffected by these terminal groups. Thus, to rationalize the *intrinsic conformational propensities* or conformational preferences of such residues, in small peptide structures, the effect of the terminal group has to be avoided. For this purpose, a molecule mimicking the central segment of such a polypeptide chain is considered, i.e., capped peptides, chemically protected on both ends. Capped peptides of varying length have been used in the past (Fig. 7.1c and 7.1d). The smallest capped peptide is a 'model dipeptide,' (Fig. 7.1c), where X represents different amino acid residues. The longer capped peptides for exploring the *intrinsic conformational propensities* are tripeptide and pentapeptides. The tri and pentapeptides, apart from being chemically protected at both the ends, also contain an equal number of glycine residues at both sides of X. In the current study, we explored the potential energy surface (PES) of model dipeptides, employing *ab initio* QM computation method to investigate the conformational preference or *intrinsic conformational propensity*. The conformational preferences of longer capped peptides; using *ab initio* computational study is highly expensive and has not been addressed in this work.

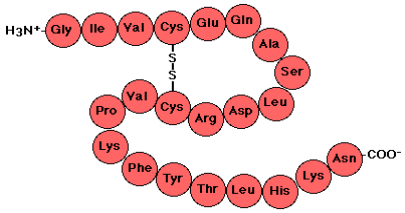
7.2 Computational Methods

The conformational landscapes of all model dipeptides were explored by *ab initio* QM study, using Gaussian 09 package.[107] Potential energy surface (PES) scan was done to obtain possible minima of various model dipeptides. The initial geometries were generated by the systematic rotation of backbone and side-chain dihedral angles. For example, in the case of the glycine model dipeptide, the six-fold rotation around C-Cα bond and Cα-N bonds resulted in 36 starting geometries. The two peptide bonds present in glycine model dipeptide were retained in the anti-periplanar arrangement, as the orientation is known to be energetically favorable, and the peptide or amide planes, in proteins and polypeptides exits only in anti-periplanar geometry.

The generated initial structures were optimized at the HF/6-31G level of theory. The optimized geometries obtained were then used as starting geometry for optimization at B3LYP/6-311G level of theory. The stationary point obtained were then refined at M06-2X/6-311++G (d,p) level of theory, using opt = tight and int = ultrafine condition. The non-redundant sets of stationary points from a large number of optimized geometries, at every stage of calculation, were extracted by screening the structures, which had distinct rotational constants. Frequency calculations were carried out at M06-2X method, using analytical gradients to identify if the stationary points corresponded to minima or maxima. The number of initial geometries, stationary points obtained at the various level of optimization and final numbers of conformers obtained for all the dipeptides, except arginine, are listed in Table 7.1. In the case of arginine dipeptide, the number of initial geometries obtained after regular rotation of backbone and the side-chain dihedral angles was ~57000, and therefore the PES scan for arginine dipeptide was not performed.

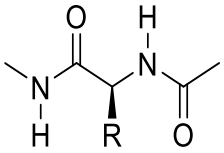


(a)



(b)

Me-NH-X-O-Me



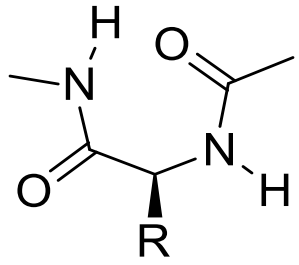
(c)

(d)

Me-NH-GXG-O-Me (Model tetrapeptide)

Me-NH-GGXGG-O-Me (Model hexapeptide)

Fig. 7.1 a) Chemical structure of glycine dipeptide; b) A polypeptide; c) Chemical structure of model dipeptide; d) Chemical forms of model tetra and hexapeptides, where X represents the different amino acid residue

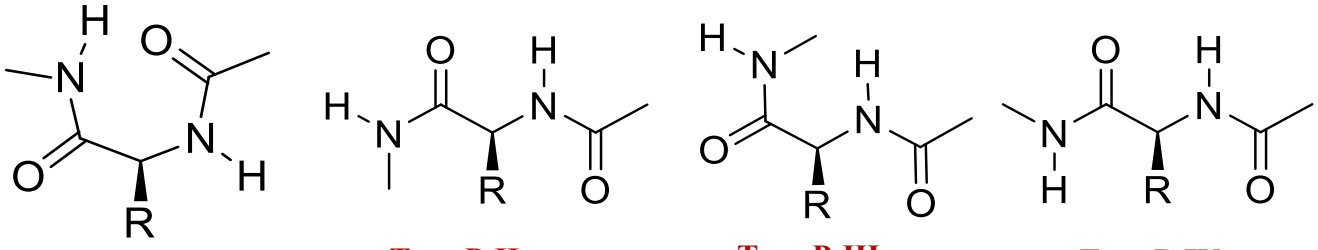


CCC or Type P-I

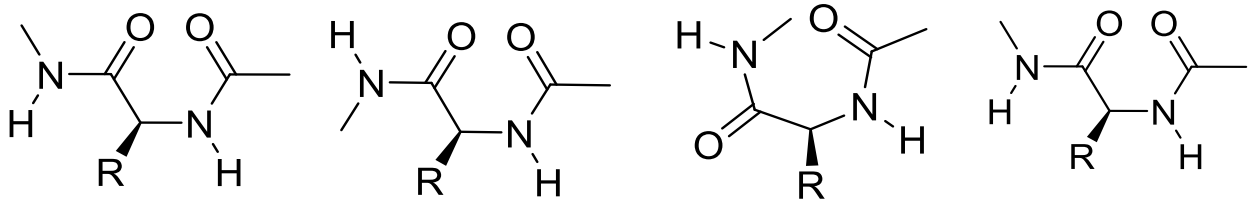
Three key Dihedral angles in model dipeptides

1. ***H-N-C-Cα ; (cis configuration) (ω)***
2. ***N-C-Cα-N ; (cis configuration) (ψ)***
3. ***C-Cα-N-C; (cis configuration) (φ)***

Fig.7.2 Backbone definition of model dipeptides



Type P-I Type P-II Type P-III Type P-IV



Type P-I' Type P-II' Type P-III' Type P-IV'

Fig. 7.3 Eight backbone structures of model dipeptide

Table 7.1 List of the total number of geometries optimized and final number of conformers obtained for each model dipeptides, except arginine; in the potential energy surface scan method.

Model dipeptides	a	b	c	d
Gly	36	6	5	3
Ala	36	10	6	5
Val	108	23	19	18
Leu	648	85	55	48
Ile	648	84	57	54
Ser	648	50	30	28
Thr	648	42	32	28
Cys	648	65	41	38
Met	1944	209	141	129
Asp	864	104	73	69
Asn	1296	49	28	23
Glu	2592	348	249	218
Gln	3888	133	83	72
Phy	648	32	19	17
Trp	648	63	38	35
Tyr	2160	63	37	36
His	648	49	27	27
Lys	8748	1943	1191	1015
Pro	12	4	4	3
Arg	~57000	-	-	-

a; Number of initial geometries optimized at HF/6-31G level of theory
b; Number of stationary points *(obtained after optimization in previous step)* optimized at B3LYP/6-311G level of theory
c; Number of stationary points *(obtained after optimization in previous step)* optimized at M06/6-311++G(d,p) level of theory
d; Final number of conformers *(each conformer shows real frequency for all the normal modes.)*

A classification method based on the backbone structure of model dipeptide was proposed. The orientation of the backbone structure of model dipeptide could be defined by three dihedral angle viz. H-N-C-Cα, N-C-Cα-N, and C-Cα-N-C (Fig. 7.2). The three dihedral angles used for determining backbone orientation of model dipeptide are equivalent to torsional angle ω, ψ, and φ respectively, of the peptide. For the backbone structure shown in Fig. 7.2, the above three dihedral angles adopts cis configuration, and hence, the orientation of backbone is CCC or Type P-I. Since, the backbone geometry is defined in term of three dihedral angles and in the situation of each of this dihedral angles adopting two configurations, i.e., cis and trans, a total 8

(2^3) different backbone structures are possible, including CCC or type P-I structure. The eight backbone types of model dipeptides are P-I, P-II, P-III, P-IV, P-I', P-II', P-III' and P-IV' (Fig. 7.3). The prefix 'P' used in the nomenclature of model dipeptide backbone types stand for 'peptide' and have been used to avoid confusion with the backbone types of α-amino acids. The three dihedral angles used for backbone classification of dipeptides is roughly equivalent to the three dihedral angles used in the characterization of monomeric α-amino acids. The basis of the backbone structure classification of model dipeptide is equivalent to that used for α-amino acids.

All the conformers of various model dipeptides (numbers listed in Table 7.1) were classified into these eight classes, based on their backbone geometries. Conformer's nomenclature based on their backbone type and energy was also proposed, similar to that done in α-amino acids. The structural preferences in model dipeptides were qualitatively and quantitatively analyzed by methods identical to that used for α-amino acids. The structures and energetics of conformers of various model dipeptides were represented in the form of the conformational dartboard, and the propensities of different backbone types were also estimated using the method discussed in the last chapter.

7.3 Results and Discussions

7.3.1 Conformational dartboard of model dipeptides

The conformational dartboard representation, similar to that used for α-amino acids, was used to analyze the structural preferences of various model dipeptides. The different conformers of various model dipeptides were presented in the dartboard fashion, based on their backbone structure and energetics (Fig. 7.4a). The eight tracks in the dartboard, represent the eight backbone types of model dipeptides viz. P-I, P-II, P-III, P-IV, P-I', P-II', P-III' and P-IV'. The conformer distribution in the first octant region of the dartboard, which also represents

conformers with energy up to 1.25 kcal mol^{-1} relative to the most stable conformer (Fig. 7.4b). In the first octant region of the dartboard, track P-I and P-IV are densely populated, whereas the other tracks are empty, thus implying that low energy conformers of model dipeptides adopt only backbone type P-I and P-IV structures. This observation also suggests that the low energy conformers of model dipeptides, irrespective of the chemical nature and bulkiness of the side-chain group, prefer only backbone type P-I and P-IV.

The track corresponding to backbone type P-II, P-III, P-II' and P-III' are found empty in the entire range of dartboard, i.e., 10.0 kcal mol^{-1} (Fig. 7.4b). The empty tracks in dartboard indicate that conformers of model dipeptides do not adopt these four backbone geometries. Interestingly, the amide or peptide groups in these four backbone structures are in syn-periplanar orientation, i.e., between the carbonyl and N-H group, as shown by boxes in Fig. 7.4c. As discussed earlier, the amide or peptide plane (carbonyl and N-H group) in proteins and peptides adopt energetically favorable anti-periplanar geometry, due to the partial double bond character and not syn-periplanar form. Thus, the conformer of model dipeptides does not choose these four backbone types and hence results in the tracks corresponding to these structures in the dartboard being empty.

The conformational dartboard analysis of model dipeptides conformers qualitatively present the intrinsic tendencies of low energy conformers of model dipeptides to prefer only specific backbone types, i.e., type P-I and P-IV. Furthermore, as observed for the amino acids, the side-chain groups with varying chemical nature in dipeptides does not play a decisive role in determining the conformational preferences of dipeptides, and at best, they may only be supportive. The above observed intrinsic conformational preferences in model dipeptides are

similar to that seen in the case of α-amino acids, wherein the backbone types I and II were preferred.

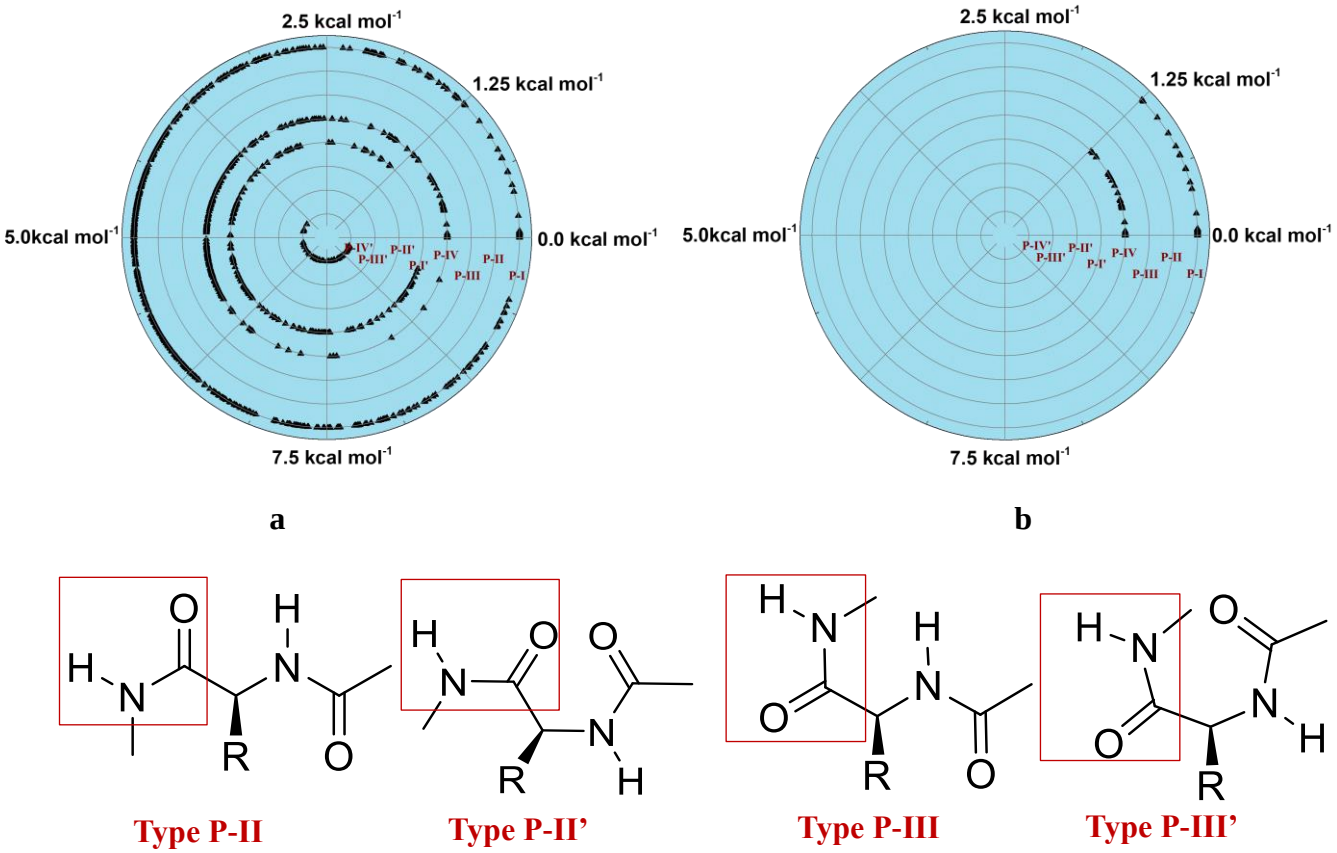


Fig. 7.4 a) Conformational dartboard of model dipeptides; b) The first octant of conformational dartboard of model dipeptides; c) Backbone types with syn-periplanar amine plane (shown in the red box)

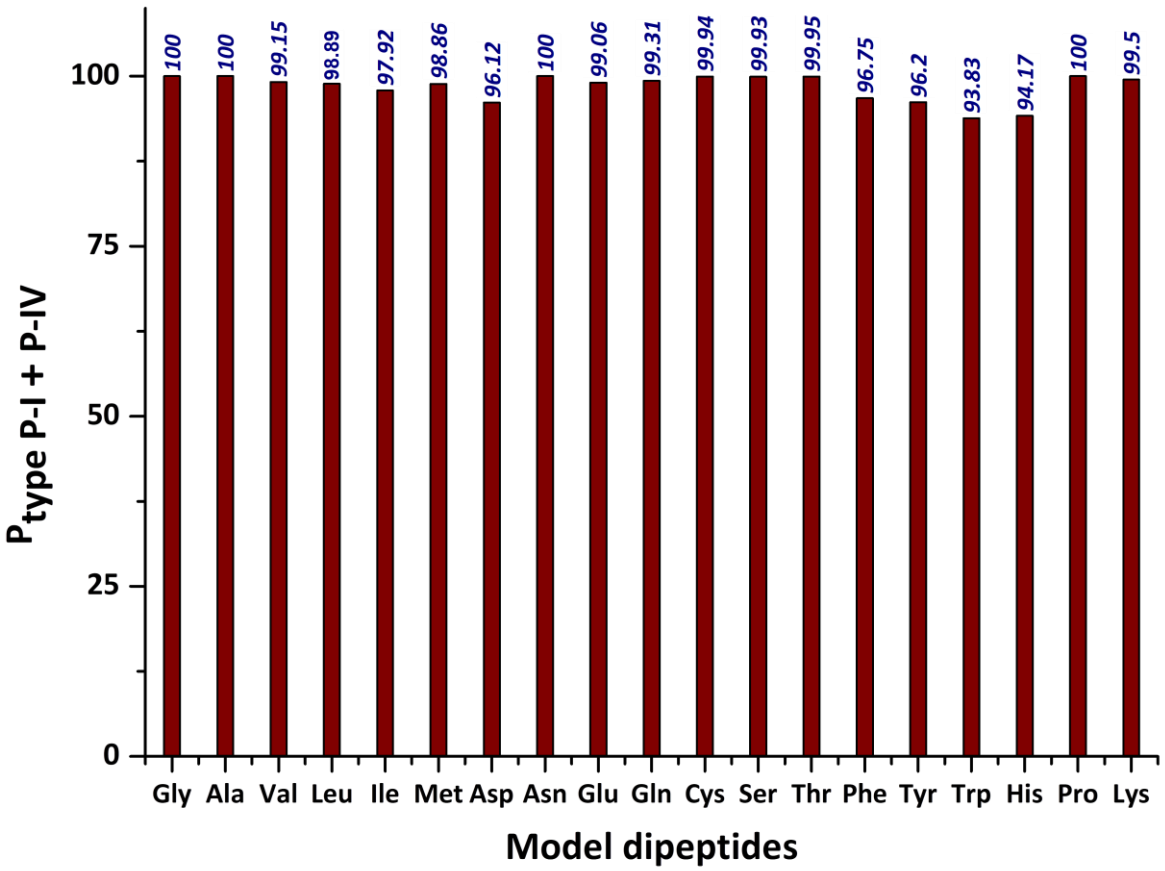


Fig. 7.5 Sum of the backbone propensity of type P-I and P-IV of model dipeptides (except arginine)

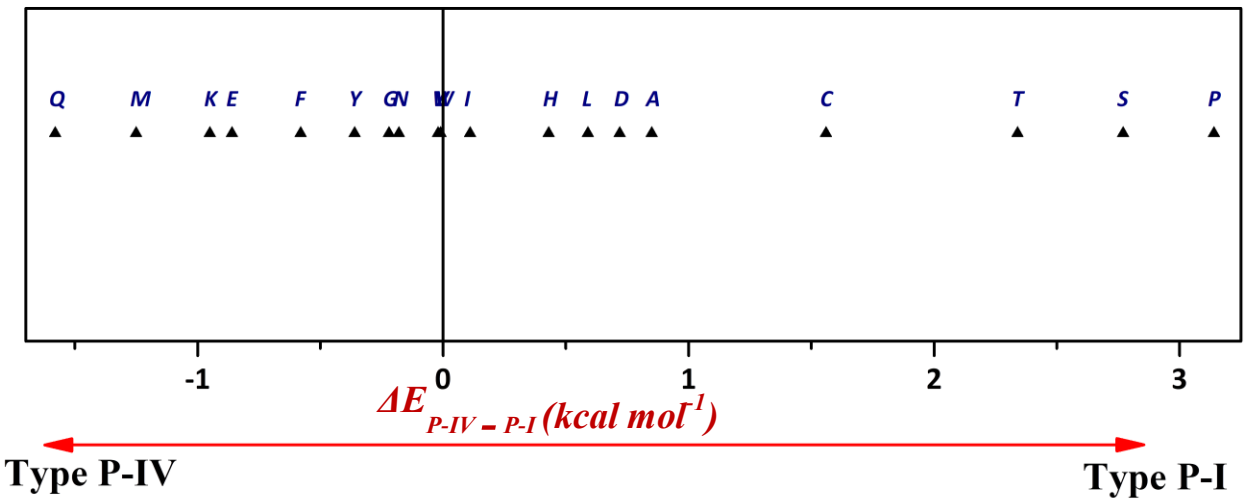


Fig. 7.6 Relative backbone stability plot for model dipeptides

7.3.2 Backbone propensity, a quantitative estimate of population distribution

The intrinsic backbone preferences of model dipeptides were quantitatively rationalized based on backbone propensity, which is the quantitative measure of the population of a particular backbone structure of any given model dipeptide. As has been discussed in the last chapter, the backbone propensity is estimated by summing up the relative abundance of all the conformers belonging to a given backbone type. The sum of backbone propensity of two preferred backbone types, i.e., type P-I and P-IV ($P_{\textbf{type P-I + P-IV}}$) of various dipeptides were plotted against respective model dipeptides (Fig. 7.5). The $P_{\textbf{type P-I + P-IV}}$ values in case of glycine, alanine, asparagine, and proline dipeptides were observed to be 100%, whereas many others dipeptide show the population over 98%, for these two backbone structures. The average $P_{\textbf{type P-I + P-IV}}$ over all the dipeptides was calculated to be 98.3%, with a standard deviation of 1.9%. Thus, in model dipeptide, ~98% of the total population at room temperature, are present in backbone type P-I and P-IV forms. This observation is similar to that seen in α-amino acids, where ~81 % of the total population existed in backbone type I and II, as discussed in the last chapter. Furthermore, like α-amino acids, in dipeptides too, the side-chain group was found to play only a supportive role in determining the structural preferences. Therefore, the relationship between preferred backbone structures of dipeptides and α-amino acids is worth exploring.

The relative preference between backbone type P-I and P-IV in various dipeptides was rationalized in term of the relative backbone stability of two backbone structures ($\Delta E_{P\text{-}IV\ -\ P\text{-}I}$). The stability of backbone type P-IV relative to P-I ($\Delta E_{P\text{-}IV\ -\ P\text{-}I}$) for the various dipeptides was estimated using the population distribution for type P-IV relative to type P-I, computed as shown below;

$$\frac{N_{P\,IV}}{N_{P\,I}} = e^{-\frac{(EP_{\,IV} - EP_{\,I})}{kT}} \quad \Delta E_{P\text{-}IV\,-\,P\text{-}I}$$

Stability of type P-IV relative to type P-I

The $\Delta E_{P\text{-}IV\ -\ P\text{-}I}$ values estimated for various dipeptides were then plotted on an energy scale (Fig. 7.6). The dipeptides lying in the positive region of the plot (right side) had a higher population in backbone type P-I relative to that in type P-IV. For example, dipeptides such as proline, serine, threonine, and others, show a sizeable positive value of $\Delta E_{P\text{-}IV\ -\ P\text{-}I}$, and, hence, have a significant fraction of population in backbone type P-I and a minor population in the type P-IV. Likewise, the dipeptides in the negative region (left side) of the plot, show a higher preference for backbone type P-IV structures as compared to backbone type P-I structure. The set of dipeptides lying in the central region have nearly equal propensities for both the backbone structures.

7.3.3 The relationship between preferred backbone geometries of α-amino acids and dipeptides

The low energy conformers of monomeric amino acids were found to prefer backbone type I and II, as discussed in the last chapter, whereas in the case of dipeptides backbone type P-I, P-IV were preferred. Therefore, exploring if there is any correlation between preferred backbone structures of monomeric amino acids and dipeptides is essential to understand any possible influence the conformers of α-amino acids may have on the structural preferences of dipeptides. Closer structural analysis of backbone types for the monomeric amino acids and the dipeptides is therefore called for to see if there are any interesting relationships. One can ask the question if type I of the amino acid structure leads preferentially to P-I or P-IV structures on peptide formation? Likewise, does the type II structure of the amino acid lead to a preferred dipeptide backbone? It may be noted from Fig. 7.7a that the P-I backbone of the dipeptide leads to a turn-like structure while the type P-IV can give rise to a linear form. From the figure, it can also be

seen that the turn like dipeptide structure can arise from a type I backbone of the amino acid, while the linear form can result from the type II amino acid backbone. The question, therefore, amounts to asking if the amino acid backbone influences the formation of a turn or a linear form in the dipeptide.

We, therefore, set out to see if there existed a quantitative relationship between the backbone geometries, type I and II of the α-amino acids and the P-I and P-IV backbones of the dipeptides.

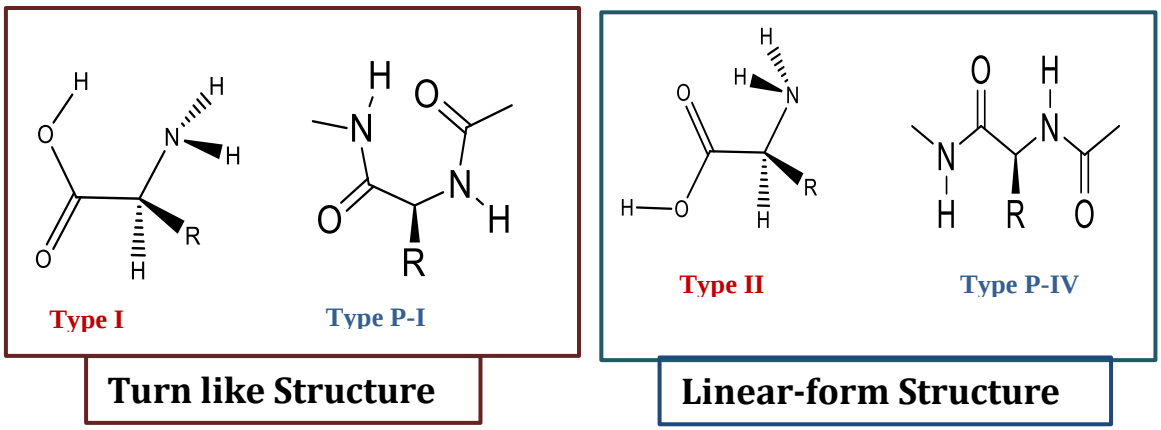
A plot was made between propensities of the type I structure of monomer and dipeptides (type P-I) (Fig. 7.7b), for various amino acids. This plot has at its origin a propensity of 50 for type I and 50 for type P-I. It will be recalled that propensity is defined as the population of the amino acid or the dipeptide existing in the given backbone form. In the plot shown (Fig. 7.7b), points in quadrant 1, indicate a high propensity for type I and type P-I, implying these amino acid adopt preferentially type I backbone in monomer and result in a type P-I structure for the dipeptide, which is a turn-like structure. Clearly, proline (P) belongs to this category, having a type I propensity of 90 and a propensity of almost 100 for type P-I. Another amino acid in quadrant one is "H" (histidine) which has a type I propensity of 90, and shows a P-I propensity of 60, implying that it is more likely to adopt a turn-like structure.

Interestingly a few other amino acids, such as T (threonine), S (serine), C (cysteine), A (alanine), L (leucine), and D (aspartic acid) show a preference for type P-I backbone in their dipeptides, even though these amino acids show somewhat equal preference for type I and II in their monomeric forms. The amino acids, Q, M, K, E, F, and Y, display a tendency for type P-IV dipeptide backbone even though their monomeric units show an approximately equal preference for type I and II structures. The amino acid "G" (glycine) adopts a type II backbone in its

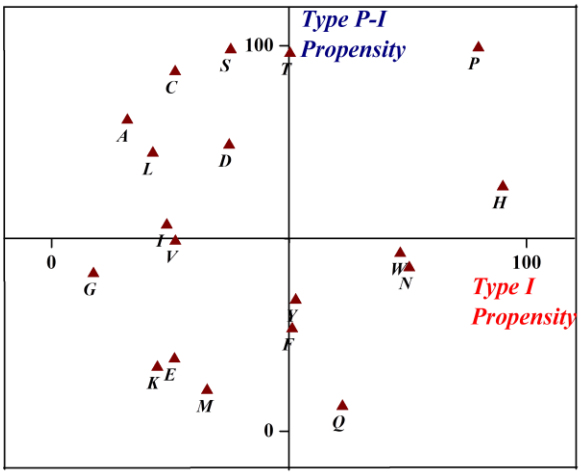
monomeric form, but interestingly displays flexibility in adopting both type P-I and P-IV backbones for the dipeptides. All the other amino acids in this study lie in the third quadrant. This analysis indicates that while some amino acids prefer to adopt a turn like structure irrespective of their backbone preference for the monomers, certain other amino acids prefer a linear form. Some amino acids which occur near 50% propensity for P-I show an equal preference for both linear and turn-like structure and are therefore flexible in their behavior and probably adopt a linear or turn-like form based on the context in the peptide structure.

A similar picture indicating the above discussed emerges if one re-plot the same data with the origin as (0,0), where it is clear that one set of amino acids prefer type P-I while the others prefer the type P-IV (Fig. 7.7c). The amino acids placed closer to the top of the Y-axis are those who have a high propensity for turn-like structure in dipeptides (high population of type P-I) relative to that in monomeric form (less population of type I). Similarly, another set of amino acids lying closer to the X-axis shows lower preferences for turn-like structure in dipeptides (minor population of type P-I) as compared to that in the monomer (higher population of type I).

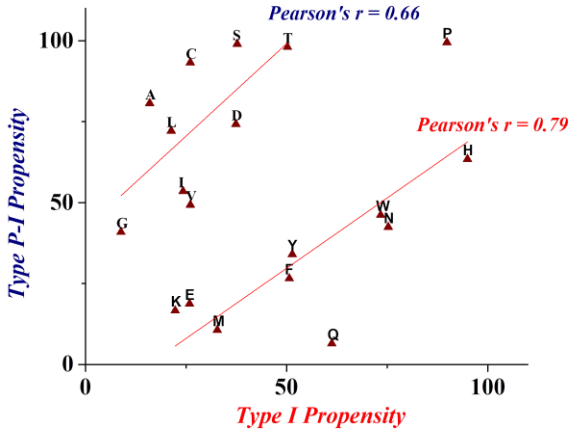
Thus, in summary, this analysis reveals the set of amino acids preferring a turn-like structure and those that prefer a linear form and many others that independent of their monomeric backbone structure, demonstrate flexibility in adopting either turn-like or linear backbone in the dipeptides. Amino acid such as proline adopt only one structure preferentially in both the monomeric amino acid and in the dipeptide, indicating that such amino acids when present in a chain, may determine the structure at the point of their occurrence on the peptide chain; i.e., make the chain go strand-like or turn-like. Amino acids that show flexibility in their peptide structures, (i.e., those have a near 50% propensity for the P-I and P-IV forms) may adopt



a



b



c

Fig. 7.7 Geometrically equivalent backbone structures of amino acids and model dipeptide i.e. turn-like structures (type I and P-I) and strand like structures (type II and P-IV); b) Correlation between propensities of turn-like structures of amino acids and model dipeptides (type I and P-I) with origin at (50,50); c) Correlation between propensities of turn-like structures of amino acids and model dipeptides (type I and P-I) with origin at (0,0)

a

ΔE_{II-I} Stability of *type II* relative to *type I*

=

$\Delta E_{linear(a.a.)}$ Stability of linear-form relative to turn-like in amino acids

	P_{II} (%)	P_I (%)	ΔE_{II-I} or ΔE_{linear} (a.a.)
Ala	61.17	15.96	-0.80
Asn	11.78	75.29	1.10

b

$\Delta E_{P\text{-}IV\ -\ P\text{-}I}$ Stability of type P-IV relative to type P-I

=

$\Delta E_{linear(d.p.)}$ Stability of linear-form relative to turn-like in dipeptides

	$P_{P\text{-}IV}$ (%)	$P_{P\text{-}I}$ (%)	$\Delta E_{P\text{-}IV\ -\ P\text{-}I}$ or ΔE_{linear} (d.p.)
Ala	19.29	80.71	0.85
Asn	57.54	42.46	-0.18

c

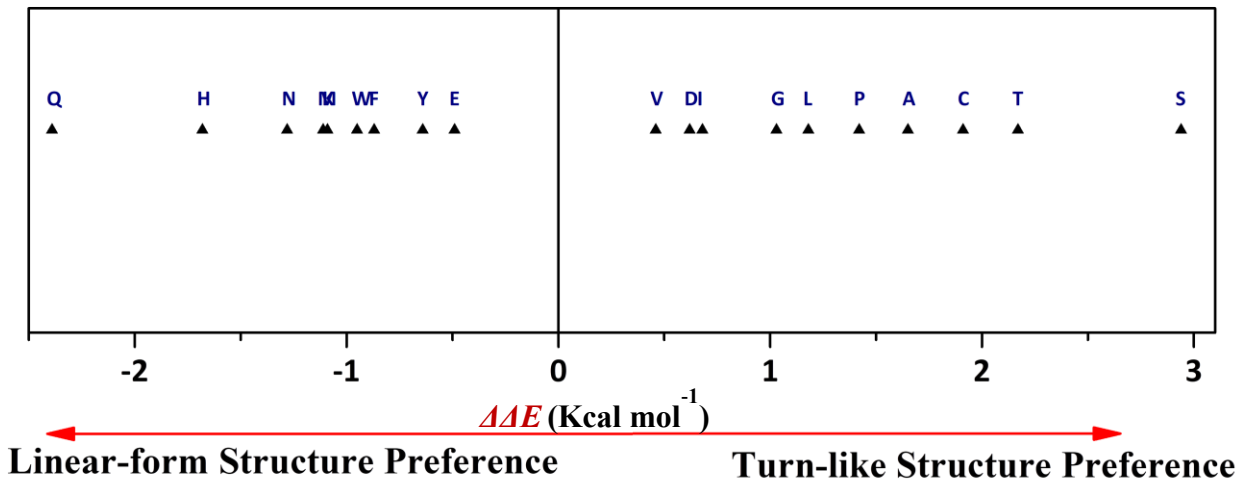


Fig. 7.8 a) The stability of linear-form backbone structure (type II) relative to turn-like backbone structure (type I) in α-amino acid, with examples of L-alanine and L-asparagine; b) The stability of linear-form backbone structure (type P-IV) relative to turn-like backbone structure (type P-I) in model dipeptides, with examples of alanine and asparagine dipeptides; c) A ΔΔE plot is showing the change in the relative preferences of linear-form and turn-like backbone structures going from monomeric amino acids to dipeptides

structures depending on the context. However, the combined preference for geometrically equivalent backbone structures, i.e., turn-like and linear form backbone structures together, in monomeric amino acids (~81 %) and dipeptides (~98 %) are nearly conserved.

7.3.4 Quantifying the variation in structural preference from α-amino acids to dipeptides

The change in the preferences of turn-like and linear-form backbone structures of monomeric amino acids and dipeptides were rationalized based on the relative stability of the geometrically similar backbone types. As discussed in the earlier section, the relative stability between backbone type P-IV and P-I in dipeptides was estimated by $\Delta E_{P\text{-}IV\ -\ P\text{-}I}$ values. It may be recalled that the backbone type P-IV has linear-form geometry and backbone type P-I has turn-like backbone geometry (Fig. 7.7a). Therefore, the stability of linear-form geometry (or type P-IV) relative to turn-like backbone geometry (or type P-I) in dipeptides, is indicated by ΔE_{linear} (dipeptide).

Likewise, the preference between backbone type I and II in α-amino acids were quantified based on $\Delta E_{II\text{-}I}$ values, which is the estimate of the stability of backbone type II relative to type I for any given α-amino acids, as discussed in the last chapter. The backbone type I of α-amino acids shows turn-like geometry, whereas the backbone type II adopts linear-form backbone geometry (Fig 7.7a). Therefore, the stability of the linear-form backbone relative to turn-like backbone structure in α-amino acids is represented by ΔE_{linear} (amino acid) and is the same as the $\Delta E_{II\text{-}I}$ value. For example, the $\Delta E_{II\text{-}I}$ or ΔE_{linear} (amino acid) estimated for L-alanine and L-asparagine were found to be -0.80 and 1.10 kcal mol^{-1}, respectively (Fig. 7.8a).

The variation in the preferences of linear-form structures between alanine dipeptide and monomeric alanine can be obtained by subtracting the ΔE_{linear} (dipeptide) by ΔE_{linear} (amino

acid). The ΔΔE value (ΔE_{linear} (dipeptide) - ΔE_{linear} (amino acid)) was computed to be 1.65 kcal mol^{-1} for alanine. The positive ΔΔE value for alanine indicates that the propensity of linear-form structures decreases on going from monomeric alanine to the alanine dipeptide, which is also evident from the population of linear-form structures of monomeric alanine(type II) and alanine dipeptide (type P-IV) (Fig. 7.8a and 7.8b). The monomeric-alanine has ~61 % of the total population in linear-form structure (type II), whereas in alanine dipeptide the population of linear-form backbone structure (type P-IV) was only ~19 %. Conversely, the positive ΔΔE value also implies that the preference for turn-like structure increases from monomeric alanine (type I) to alanine dipeptide (type P-I). In the case of asparagine, the ΔΔE value was computed to be -1.28 kcal mol^{-1}, thus indicating that linear-form structures preference increases from monomeric asparagine to the asparagine dipeptides. The populations of linear-form backbone structure (type II) in monomeric asparagine was estimated to be ~12 % whereas the population in asparagine dipeptide (type P-IV) was obtained to be ~58 %.

Based on the above method, the variation in the structural preference of geometrically similar backbone structures between each pair of monomeric amino acids and dipeptides were estimated. The ΔΔE values obtained for each pair of α-amino acids and dipeptides were plotted together on a ΔΔE scale (Fig. 7.8c). For the amino acids in the negative region, the preference for linear-form backbone increases going from monomeric form (minor population of type II) to dipeptide (high population of type P-IV). Conversely, the same group of amino acid also exhibit a lower preference for turn-like structure in dipeptides (smaller population of P-I) relative to that in monomeric form (higher population of type I). The amino acids in the positive region of ΔΔE plot show a higher preference for a turn-like structure in dipeptides (higher population of type P-I) as compared to that of monomer (minor population of type I). Conversely, these amino acids

also exhibit a lower preference for linear-form structures in dipeptides (smaller population of type P-IV) relative to that of the monomeric form (higher population of type II).

7.4 Preferred Backbone Structure of α-Amino Acid and Dipeptides in the Ramachandran Plot

The relationship between preferred backbone types of α-amino acids and dipeptides were also explored based on the (φ,ψ) region occupied by these backbones in the Ramachandran plot. First developed in 1963 by G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, the Ramachandran Plot is widely employed in proteins and peptide structural studies, with the purpose to visualize the energetically allowed regions for backbone dihedral angles φ and ψ of amino acid residues of in proteins.[13] The dihedral angle φ and ψ represents the torsional motions of Cα-N and Cα-C bonds. The Ramachandran plot of dipeptides, except glycine and proline, shows two major energetically favored region, i.e., helix and strand regions.

As discussed in the earlier section, the backbone of dipeptides was classified based on the relative orientation of three backbone dihedral angles, i.e., H-N-C-Cα, N-C-Cα-N, and C-Cα-N-C (Fig. 7.9a). These three backbone dihedral angles are also equivalent to ω, ψ, and φ torsional angles of a peptide, respectively. The Ramachandran regions (φ, ψ space) adopted by dipeptides conformers were investigated based on the orientation of backbone dihedral angles N-C-Cα-N (ψ), and C-Cα-N-C (φ).

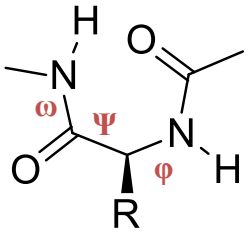
The backbone geometry of α-amino acids was defined by the orientation of three backbone dihedral angles, i.e., H-O-C-Cα, O-C-Cα-N, and C-Cα-N-lp(N), as discussed in earlier chapters (Fig. 7.9b). The backbone dihedral angles O-C-Cα-N, and C-Cα-N-lp(N) in α-amino acids, represent the torsional motion along Cα-C and Cα-N bonds, respectively. These two dihedral angles of monomeric amino acids are equivalent to the ψ and φ dihedral angles of

dipeptides. Since the dihedral angles, φ and ψ were initially used representing for the torsional angle of dipeptides and not α-amino acids. Therefore, the equivalent dihedral angles in monomeric amino acids were termed as φ' and ψ', and the Ramachandran region occupied by the α-amino acids were rationalized, based on these dihedral angles.

The Ramachandran region adopted by the preferred backbone structures of monomeric amino acids and dipeptides, i.e., type I, II and type P-I, P-IV, are shown in Fig. 9.9c. The backbone type I and type P-I of α-amino acids and dipeptides, respectively, attain the central rectangular region of the Ramachandran plot (Fig. 9.9c). As discussed in the earlier section, the backbone type I and type P-I also shows turn-like structures, and hence adopt the same region of the Ramachandran plot. The other pair of geometrically similar backbone structures of α-amino acids and dipeptides, i.e., type II and type P-IV, falls in the corner regions of the Ramachandran space.

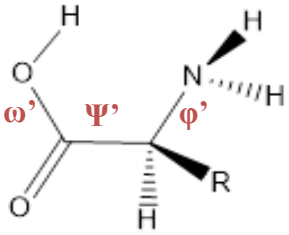
7.4.1 Helix and extended strand propensities of α-amino acids and dipeptides

The two major secondary structure regions observed in the Ramachandran plot are helix (*H*) and extended-strand (*E*), as discussed above. The φ, ψ space corresponding to different secondary structures of dipeptides has been reported in the literature by Vymetal *et al.*[36] Based on the free energy profiles of dipeptides in terms of φ and ψ torsion, the φ, ψ space of different secondary structures were defined. The Ramachandran space of 5-secondary state regions, i.e., *ext, ppII,* α_R*,* α'_R*, and* α_L were defined in terms of the regular rectangular area, (Fig. 9.9d). Since the boundary between *ext and ppII* regions (as well as α_R and α'_R regions) is not well differentiated and ambiguous in case of dipeptides, therefore, a 3-secondary state region was also adapted by Vymetal et. al. The 3-secondary state regions are *E* (*ext + ppII*), *H* ($\alpha_R + \alpha'_R$), and α_L, where '*E'*



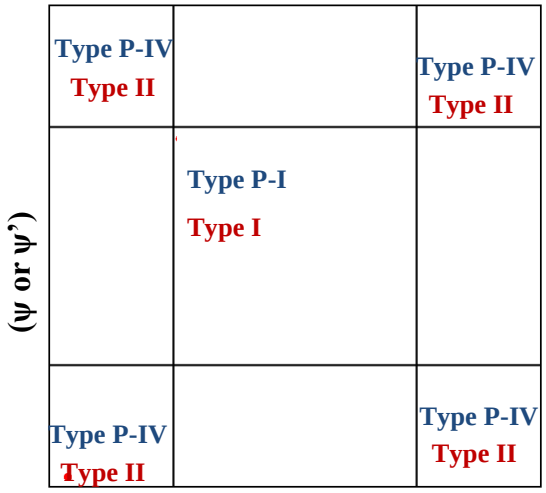
H-N-C-Cα (ω)
N-C-Cα-N (Ψ)
C-Cα-N-C (ϕ)

a

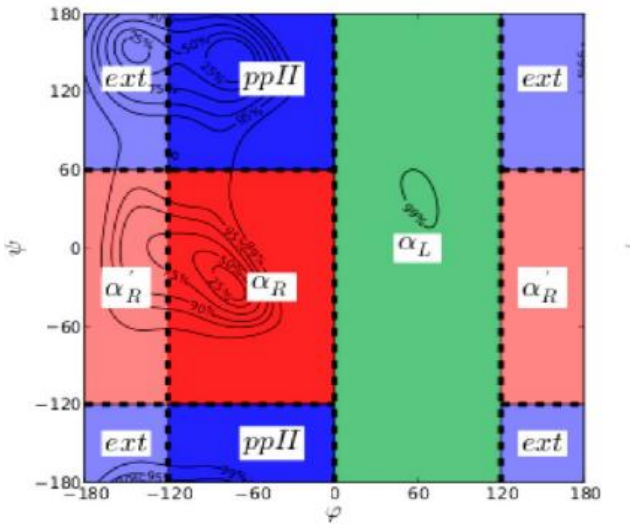


H-O-C-Cα (ω')
O-C-Cα -N (Ψ')
C-Cα-N-lp(N) (φ')

b



c



J. Chem. Theory Comput. 2013, *9*, 441−451

d

5-state boundary regions

- ***ext*** = extended strand region
- ***ppII*** = polyproline type II region
- α_R = Right handed helix region
- α'_R = near right handed helix region
- α_L = Left handed helix region

e

3-state boundary regions

- E = *ext* + *ppII*
- H = α_R + α'_R
- α_L

f

Fig. 7.9 a) The Ramachandran dihedral angles (φ,ψ) for dipeptides; b) The Ramachandran dihedral angles (φ',ψ') for α-amino acids; c) The Ramachandran region occupied by backbone type I and II (α-amino acids) and backbone type P-I and P-IV (dipeptides); d) The (φ,ψ) regions corresponding to different secondary structures (as suggested in literature); e) 5-state boundary regions (as proposed in literature); f) 3- state boundary condition, as the border between *ext* and *ppII* (as well as α_R and α'_R) are ambiguous among dipeptides and sometime not differentiated (as suggested in literature)

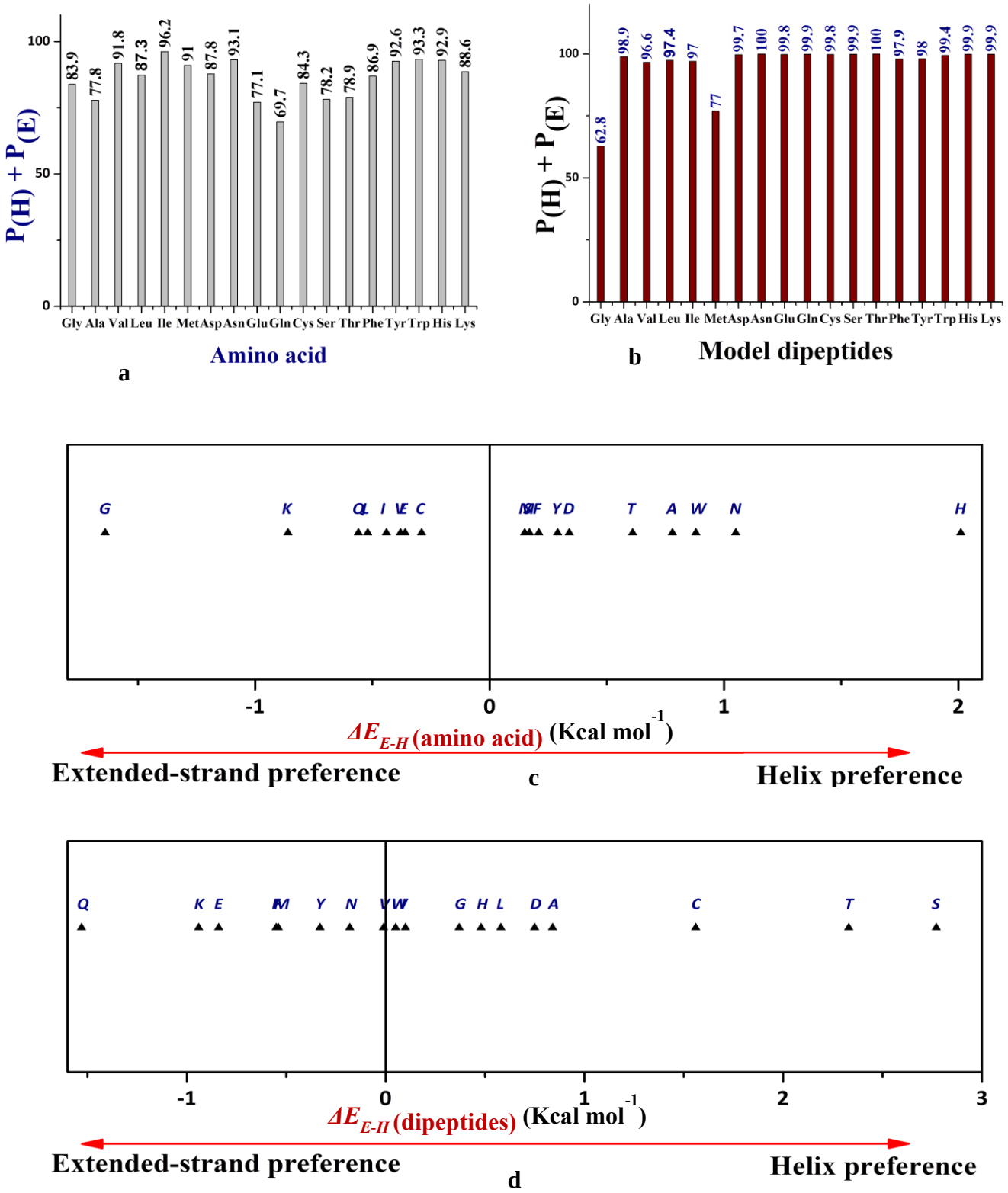


Fig. 7.10 a) Sum of the helix and extended-strand propensities of α-amino acids; b) Sum of the helix and extended-strand propensities of dipeptides; c) Relative preference between the extended-strand and helix structures in α-amino acids; d) Relative preference between the extended-strand and helix structures in dipeptides

stands for the extended-strand region and '*H'* is the helix region. A closer analysis of Fig. 9.9c and 9.9d, suggest that the φ, ψ space of backbone type I and P-I of α-amino acids and dipeptides are partially similar to that of helix region ('H'), whereas, the φ, ψ space of backbone type II and P-IV are almost similar to that of the extended-strand region ('E').

The preference of helix '*H'* and extended-strand '*E'* form in monomeric amino acids and dipeptides was understood in term of helix propensity and extended-strand propensities. The helix 'P_H (%)' or strand propensity 'P_E (%)', like backbone propensity, is the quantitative estimate of the population of the helix or strand structure in concerned α-amino acids or dipeptides. The helix or strand propensity for a concerned α-amino acid or dipeptide is estimated by summing up the relative abundance of all the conformers, present in the respective region of the Ramachandran Plot.

The sum of the helix and extended-strand propensities (P_H (%) + P_E (%)) of various monomeric amino acids plotted against corresponding α-amino acid (Fig 7.10a). The average P_H (%) + P_E (%) was estimated to be ~ 86 %, with a standard deviation of ~7%. The observation is similar to our earlier backbone propensity analysis, where ~ 81 % of the total population existed in the backbone type I and II. The preference of extended-strand 'E' form relative to helix form 'H', in α-amino acids was analyzed using the previously discussed method of relative stability $\Delta E_{E\text{-}H}$ (amino acid). The $\Delta E_{E\text{-}H}$ (amino acid) of monomeric amino acids was plotted on the energy scale, as shown in (Fig 7.10c). The α-amino acids lying in the positive region show a high preference for helix, whereas the α-amino acids present in the negative region exhibit high extended-strand propensity.

Similarly, in the case of dipeptides, the sum of helix and extended-strand propensities (P_H (%) + P_E (%)), plotted against respective dipeptide (Fig 7.10b). The average P_H (%) + P_E (%) in

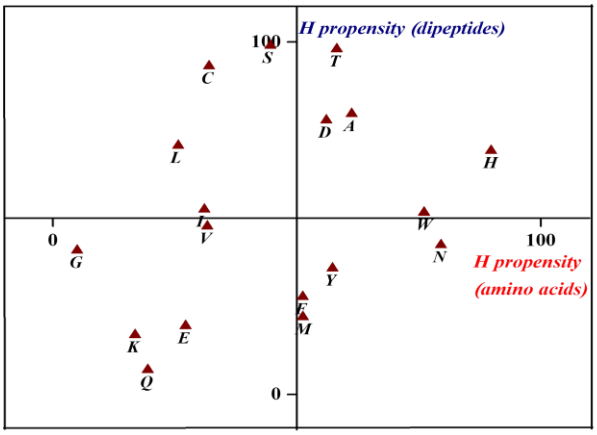
dipeptides was found to be ~96 %, with a standard deviation of ~9 %. The observation is again similar to our previous analysis, were ~98 % of the total population of dipeptides were present in backbone type P-I and P-IV. The preference between the extended-strand structures and helix form was analyzed using the above-discussed method of ΔE_{E-H} (dipeptides) (Fig. 7.10d). The dipeptides found in the positive region shows high helix propensity (P_H (%)) and those present on the negative side of the plot shows high extended-strand propensity (P_E (%)).

7.4.2 The relationship between the helix and extended strand propensity of the α-amino acids and dipeptides

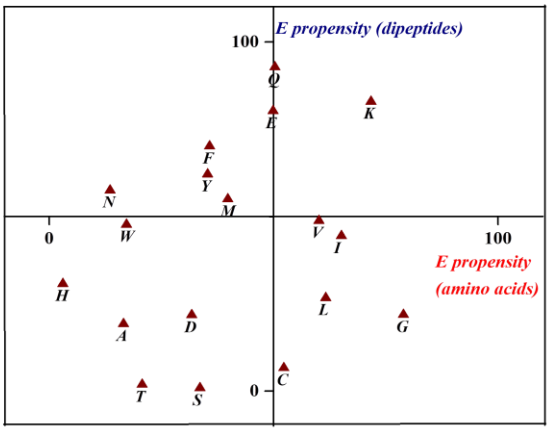
The relationship between the conformational preferences of monomeric amino acids and dipeptides was analyzed by correlating the helix and extended-strand propensities of monomer and dipeptides. The helix propensities for α-amino acids were plotted against that of dipeptides (Fig. 7.11a and 7.11b). The correlation plots are similar to that observed previously in the correlation of turn-like backbone structures (backbone type I and P-I). The amino acids in the bifurcated distribution are linearly related, which is again similar to that seen in the earlier analysis. Likewise, the correlation between the extended-strand propensities of α-amino acids and dipeptides also shows a bisected linear relationship (Fig. 7.11c and 7.11d). The observation is again, similar to the previously observed correlation between linear-form backbone structures (backbone type-II and P-IV).

7.4.3 Rationalizing the observed variation in individual structural preferences between α-amino acids and dipeptides

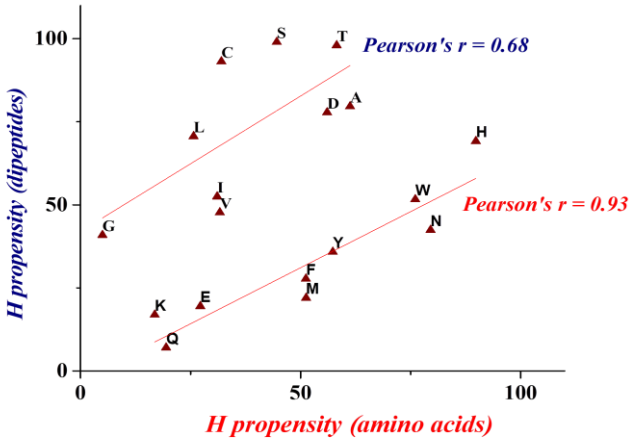
The variation observed in the individual structural choices between α-amino acids and dipeptides were quantified based on the relative stability scale. The stability of extended-strand form



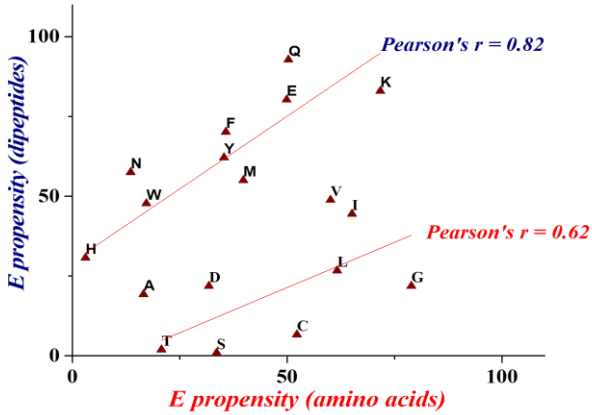
a



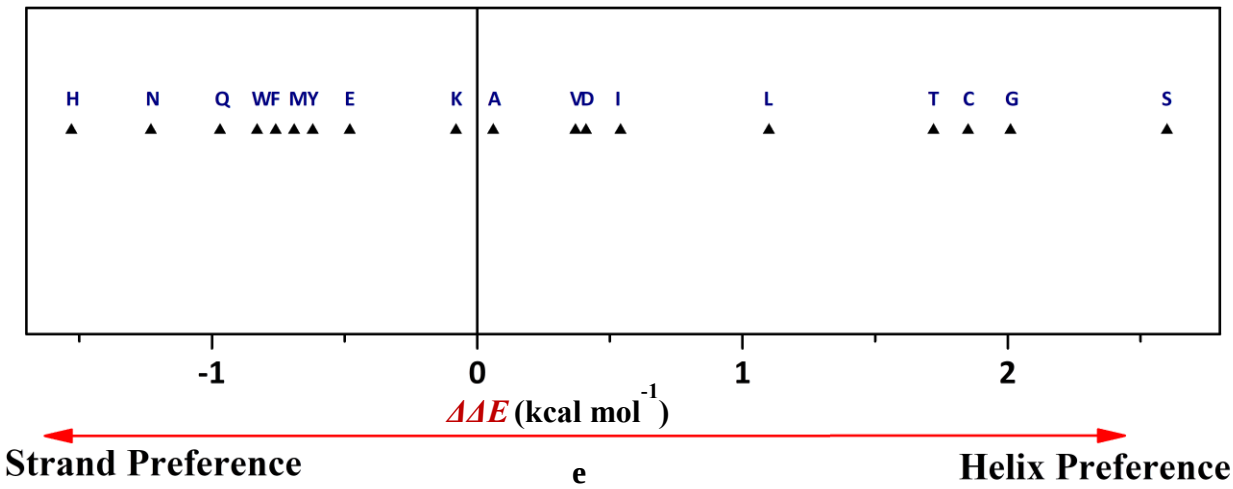
c



b



d



e

Fig. 7.11 a-b) Correlation between the helix propensities of α-amino acids and dipeptides, with origin at (50,50) and (0,0); c-d) Correlation between the strand propensities of α-amino acids and dipeptides, with origin at (50,50) and (0,0); e) An empirical ΔΔE plot showing the variation in the helix (H) and strand (E) propensities going amino acids to dipeptides

relative to helix structure for a concerned α-amino acid or dipeptide was quantified by ΔE_{E-H} (amino acid) or ΔE_{E-H} (dipeptides), respectively, as discussed in the earlier section. Therefore, the variation in the preference of extended-strand between dipeptides and monomeric amino acid can be arrived by subtracting ΔE_{E-H} (dipeptides) by ΔE_{E-H} (amino acids), as shown below

$$\mathbf{\Delta\Delta E = \Delta E_{E-H}\ (dipeptides) - \Delta E_{E-H}\ (amino\ acids)}$$

The ΔΔE values estimated for each pair of monomeric amino acids and dipeptides were presented together on a ΔΔE scale (Fig. 7.11e). The amino acids in the negative region show a high preference for the extended-strand structure in dipeptides as compared to that of monomer. The amino acids in the positive region exhibit low preference for the extended-strand in dipeptide, relative to that in monomeric form. The variation observed in the structural preferences between α-amino acids and dipeptide as nearly the same as that seen previously in Fig 7.8c.

Thus, in summary, the preferences of individual structure, i.e., helix or extended-strand, vary between monomeric amino acids and dipeptides. But the variation observed is not random, and shows a linear relationship in a bifurcated manner, similar to that observed in our earlier analysis. However, the combined preference for helix and extended-strand structures in monomeric amino acids (~86 %) and dipeptides (~96 %) are nearly conserved. Therefore, the intrinsic structural choices of monomeric amino acids appear to have a certain influence in determining the structural preferences of dipeptides. Our Study shows that monomeric amino acids adopt prefer certain backbone conformers, which also likely to influence the conformations of dipeptides. This may suggest that allowed Ramachandran angles in dipeptides could be inherently restricted due to the angles preferences of monomeric amino acids. Essentially, monomeric amino acids have restricted conformational space, described in terms of φ',ψ'

dihedrals angles or backbone geometries, which in dipeptides is are further influenced by the steric hindrance from atoms of neighboring amino acids. While the observations made here are at best strongly suggestive, they do provide directions for rationalizing the very complicated process of protein structures. More investigations clearly need to be done to understand this problem better.

7.5 Conclusions

The conformational landscape of model dipeptides were explored employing *ab initio* quantum mechanical studies. The different conformers of dipeptides were obtained by scanning the potential energy surface, which was then classified into eight classes based on their backbone geometry. The conformational dartboard shows that low energy conformers of dipeptides adopted only backbone type P-I and P-IV structures, and these two backbone structures constituted ~98 % of the total populations of dipeptides. The relative preferences between backbone type P-I and P-IV in various dipeptides were understood based on the average stability of backbone structures.

The relationship between preferred conformers of dipeptides (type P-I and P-IV) and monomeric amino acids (type I and II) was rationalized based on the geometrical equivalence of these backbone structures. The two preferred backbone structures of monomeric amino acids and dipeptides shows turn-like (type I and P-I) and linear-form geometries (type II and type P-IV). The correlation analysis between the propensity of turn-like and linear-form backbone structures of monomeric amino acids and dipeptides shows a segregated distribution. The amino acids in segregated form also exhibit linear relationship, thus suggesting that a group of amino acids shows a high preference for a turn like backbone structures in dipeptides (type P-I) relative to that in monomeric amino acids (type I). Likewise, other groups of amino acids exhibit a high

preference for linear-form backbone structures in dipeptides (type P-IV) as compared to that in monomeric form (type II). However, the combined preferences of these two structures, i.e., turn-like and linear form in monomeric form (type I and II, ~81%), as well as dipeptides (type P-I and P-IV, ~98%), remains nearly same. Thus, our study shows that the conformational preferences of monomeric amino acids are likely to influence the conformations of dipeptides.

The relationship between the conformations of monomeric amino acids and dipeptides was also rationalized based on the Ramachandran Plot. Preference of amino acids and dipeptides to adopt two major regions of the Ramachandran plot, i.e., helix, and extended strand region was understood. The turn-like backbone structures of monomeric form and dipeptides (type I and P-I) adopted the helix region, whereas the linear-form backbones (type II and P-IV) adopted extended-strand region of the Ramachandran plot. The correlation analysis based on the helix and extended strand propensities of amino acids and dipeptides were observed to be nearly the same as those seen in the correlation analysis of backbone structures. A set of amino acids exhibited a preference for helix form in dipeptides, relative to monomeric amino acids, whereas another group of amino acids shows a high preference for extended strand structures in dipeptide. Once again, the combined preferences for helix and extended strand structures were observed to be nearly the same in monomeric amino acids (~86%) and dipeptides (~96%).

Thus, in summary, our study suggests that the allowed Ramachandran regions (φ, ψ space) of dipeptides, which is also known as intrinsic conformation preferences, could possibly be inherently restricted due to the dihedral angle preferences (φ', ψ') of monomeric amino acids. However, unlike in monomeric amino acids, in case of dipeptides, the steric interactions arising from the neighboring residue atoms further influence the energetically favorable regions in the Ramachandran plot.

Chapter 8

Summary

The conformational landscape of naturally occurring α-amino acids was explored employing matrix isolation infrared spectroscopy and *ab initio* computations. The conformational study of amino acids in an isolated environment provides the information about the structural preference of neutral or canonical form and is also useful for comprehending the inherent molecular properties free from the intermolecular interactions present in the condensed phase. The vapors of solid amino acid samples were generated employing a heated nozzle placed near a cold substrate window (KBr) and were then trapped in a solid matrix of inert gas. The experimental results observed were corroborated with the computational studies. The potential energy surface was scanned to obtained possible conformers of α-amino acids, and the conformers were then classified into eight classes based on their backbone structures, labeled as type I, II, III, IV, I', II', III' and IV'.

The vibrational assignments of the experimental spectra were done using the computed wavenumbers of the amino acids. In the case of L-threonine, this exercise suggested the presence of four low energy conformers viz. $THR1_{II}$, $THR2_{I}$, $THR3_{II,}$ and $THR4_{III}$, where the subscript Roman numeral indicates the backbone type that the conformer belongs to. Likewise, in the case of L-glutamic acid and L-methionine, three conformers were observed, i.e., $GLU1_{II}$, $GLU2_{I}$, $GLU3_{II}$, and $MET1_{II}$, $MET2_{I}$, $MET3_{III}$, respectively. The lowest energy conformers in above three amino acids adopt backbone type II structure, whereas 2^{nd} lowest energy conformer has backbone type I structure. Conformers with a backbone type I were stabilized by an intramolecular hydrogen bonding interaction, involving backbone O-H and lone pair of nitrogen (O-H$^{\cdots}$N). The stability of the backbone type II structure in earlier works was attributed to the

presence of bifurcated intramolecular hydrogen bonding ($C=O^{\cdots}{}_2HN$), involving backbone carboxylic and NH_2 groups. However, AIM analysis in our study does not locate any bond critical points along such interactions. The NBO deletion study suggested that backbone type II structures were stabilized by vicinal orbital delocalization interaction. The role of the side-chain group in determining conformational preferences was also explored, and it was found that the side-chains at best played a supportive and not a decisive role.

The structure and energetics of conformers of α-amino acids were presented in the form of the conformational dartboard. The lower energy conformers of α-amino acids, irrespective of side-chain group, exhibited a tendency to adopt backbone types I, II and rather rarely III. The quantitative analysis of α-amino acids trend to adopt specific backbone structures suggested that on an average more than 80% of the total population at room temperature existed in conformers which assume backbone type I and II, except L-lysine. The lowest energy conformers of L-lysine shows backbone type IV structure, which was stabilized due to intramolecular hydrogen bonding between the backbone and butylamine side-chain group. The relative preference between backbone type I and II were rationalized based on the average stability of backbone type I and II and a group of amino acid preferred backbone type I while other set shows the significant population in backbone type II. The side-chain moieties of α-amino acids preferentially adopted $gauche^-$ (-60.0) and $gauche^+$ (60.0) orientations.

The conformations of α-amino acids have been widely studied in the past, employing different experimental techniques and computational methods, as mentioned in the earlier chapters. However, to our best knowledge, no report describing the presence of any general trends in the conformational preferences of all α-amino acids have been reported. Based on our systematic and extensive study of conformations of α-amino acids, the presence of an inherent

tendency of low energy conformers of various α-amino acids, to prefer only certain backbone structures and specific side-chain orientations is reported. The thesis work also reveals a pattern in an otherwise seemingly complex conformational landscape of α-amino acids.

The role of preferred backbone structures of α-amino acids in determining the conformational preferences of dipeptides was also explored. The conformers of model dipeptides were obtained after scanning the potential energy surface, and it was found that the conformers of dipeptides too could be classified into eight backbone classes. The low energy conformers of dipeptides adopted backbone type P-I and P-IV and more than 98% of the total populations at room temperature existed in these two backbone forms only, irrespective of side-chain group. The conformational preferences in dipeptides there can be seen to be similar to that seen in monomeric amino acids, i.e., there is a clear preference for certain backbone geometries. The relationship between conformational preferences of monomeric amino acid and dipeptides was therefore sought to be explored based on the geometrical equivalence of preferred backbone structures.

The backbone type I and P-I of monomeric amino acid and dipeptides, respectively, shows turn-like geometry, whereas backbone type II and P-IV depicts linear-form geometry. The correlation plot between the propensity of turn-like backbone structures (type I and P-I) of monomeric amino acids and dipeptides exhibit a bifurcated linear relationship. The bifurcated distribution suggested that for one group of amino acids the preference for turn-like backbone structures increases going from monomeric form to dipeptide, whereas for another set of amino acids the linear-form backbone structure is more preferred in dipeptides as compared to monomeric amino acids. The observation was further understood by fixing the origin of the plot at (50, 50). There are a few amino acids such as proline which prefer a linear structure in both

the free amino acid and the dipeptide, indicated by the location of this amino acid into the top right-hand corner of this plot. Alternatively, some of the amino acids prefer to adopt a turn-like backbone structure in dipeptide, irrespective of their backbone preferences in monomeric form. Yet again there is another group of amino acids which prefers a linear backbone for the dipeptide. A group of amino acids show an equal preference for both turn-like and linear-form backbone structure in dipeptides (~50% propensity for type P-I or P-IV) and display flexibility in their behavior, and probably adopt linear or turn-like structure based on the context in the peptide structure.

The variation in the conformational preferences between monomeric amino acids and dipeptides was quantitatively rationalized based on an empirical ΔΔE scale, which is a measure of the change in the relative preference of linear-form backbone structure going from the monomeric amino acid to its dipeptide. The amino acids preferring turn-like backbone in dipeptides were observed at the extreme positive end of the plot, whereas those preferring linear-form were located in the far negative region. The amino acids having equal preference were seen lying in the near central area of the plot. While the observations made here are at best strongly suggestive, they do provide directions for rationalizing the very complicated process of protein structures. More investigations clearly need to be done to understand this problem better.

The relationship between amino acid and dipeptide conformer was also understood based on the Ramachandran plot. The φ, ψ space corresponding to different secondary structure region of dipeptides was adopted from an earlier reported literature. The φ, ψ space taken by turn-like backbone structure (type I and P-I) and linear-form (type II and P-IV) were nearly the same as a helix and extended-strand region respectively. Therefore, the correlation analysis between helix propensity of monomeric amino acids and dipeptides were the same as those observed for turn-

like structure. Likewise, the correlations between extended-strand propensities were also similar to that found in the case of linear-form backbone structure.

The conformational choices in peptides are also known as intrinsic conformational propensity, which is the quantitative estimate of the occurrence of a concerned amino acid residue in the helical, sheet, or turn region of Ramachandran plot. The physical origin of the intrinsic conformational preferences of peptides or the tendency to adopt only certain φ,ψ space, is attributed to the steric factor[37,38] and is still a matter of debate.[36] Our study shows that monomeric amino acids adopt specific backbone structures, which also likely to influence the conformations of the dipeptides. This may suggest that allowed Ramachandran angles in dipeptides could be inherently restricted due to the angles preferences of monomeric amino acids. Essentially, monomeric amino acids display a restricted conformational space, described in terms of φ',ψ' dihedrals angles or backbone geometries, which in the dipeptides is further influenced by the steric environment offered by the atoms of the neighboring amino acids. However, as mentioned earlier, the observations made here should be considered as being strongly suggestive and providing a direction for the understanding of the conformation of peptides. More work clearly needs to be done for a more thorough understanding of this problem.

Bibliography

(1) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*; Fourth edition. New York : W.H. Freeman, 2005., 2005.

(2) Walsh, G. *Proteins: Biochemistry and Biotechnology*; Wiley, 2014.

(3) Whitford, D. *Proteins: Structure and Function*; Wiley, 2005.

(4) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry, Fifth Edition*; W.H. Freeman, 2002.

(5) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884–890.

(6) Road, H. Protein Structure and Function. *Annual review of physical chemistry* **1972**, No. 14, 165–191.

(7) Pauling Linus, Corey B Robert, B. H. R. Protein Content, While. *Proceedings of the National Academy of Sciences* **1951**, *37*, 205–211.

(8) Anfinsen, C. B., Haber, E, Sela, M.; White, F. H. The Kinetics of Formation of Native Ribonuclease during Oxidation of the Reduced Polypeptide Chain. *Proceedings of the National Academy of Sciences* **1961**, *47* (9), 1309–1314.

(9) Chou, P. Y.; Fasman, G. D. Conformational Parameters for Amino Acids in Helical, β-Sheet, and Random Coil Regions Calculated from Proteins. *Biochemistry* **1974**, *13* (2), 211–222. https://doi.org/10.1021/bi00699a001.

(10) Jr, D. L. M.; Kimt, P. S. Context-Dependent Secondary Structure Formation of a Designed Protein Sequence. **1996**, *380*, 730–734.

(11) Dill, K. A.; Shortie, D. Denatured States of Proteins. *Annual Review of Biochemistry* **1991**, *60*, 795–825.

(12) Kihara, D. The Effect of Long-Range Interactions on the Secondary Structure Formation of Proteins. *Protein science : a publication of the Protein Society* **2005**, *14* (8), 1955–1963. https://doi.org/10.1110/ps.051479505.

(13) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of Polypeptide Chain Configurations. *Journal of molecular biology* **1963**, *7*, 95–99.

(14) Brant, D. A.; Flory, P. J. The Configuration of Random Polypeptide Chains. I. Experimental Results. *Journal of the American Chemical Society* **1965**, *87* (13), 2788–2791. https://doi.org/10.1021/ja01091a002.

(15) Flory, P. J.; Volkenstein, M. Statistical Mechanics of Chain Molecules. *Biopolymers* **1969**, *8* (5), 699–700. https://doi.org/10.1002/bip.1969.360080514.

(16) Tanford, C. Protein Denaturation. In *Advances in protein chemistry*; Anfinsen, C. B., Anson, M. L., Edsall, J. T., Richards, F. M. B. T.-A. in P. C., Eds.; Academic Press, 1968; Vol. 23, pp 121–282. https://doi.org/https://doi.org/10.1016/S0065-3233(08)60401-5.

(17) Tiffany, M. L.; Krimm, S. New Chain Conformations of Poly(Glutamic Acid) and Polylysine. *Biopolymers* **1968**, *6* (9), 1379–1382. https://doi.org/10.1002/bip.1968.360060911.

(18) Tiffany, M. L.; Krimm, S. Circular Dichroism of the "Random" Polypeptide Chain. *Biopolymers* **1969**, *8* (3), 347–359. https://doi.org/10.1002/bip.1969.360080306.

(19) Shortle, D.; Ackerman, M. S. Persistence of Native-Like Topology in a Denatured Protein in 8 M Urea. *Science* **2001**, *293* (5529), 487–489. https://doi.org/10.1126/science.1060438.

(20) Prestegard, J. H.; al-Hashimi, H. M.; Tolman, J. R. NMR Structures of Biomolecules Using Field Oriented Media and Residual Dipolar Couplings. *Quarterly reviews of biophysics* **2000**, *33* (4), 371–424.

(21) Bax, A. Weak Alignment Offers New NMR Opportunities to Study Protein Structure and Dynamics. *Protein science : a publication of the Protein Society* **2003**, *12* (1), 1–16. https://doi.org/10.1110/ps.0233303.

(22) Tjandra, N.; Grzesiek, S.; Bax, A. Magnetic Field Dependence of Nitrogen−Proton J Splittings in 15N-Enriched Human Ubiquitin Resulting from Relaxation Interference and Residual Dipolar Coupling. *Journal*

of the American Chemical Society **1996**, *118* (26), 6264–6272. https://doi.org/10.1021/ja960106n.

(23) Meiler, J.; Blomberg, N.; Nilges, M.; Griesinger, C. A New Approach for Applying Residual Dipolar Couplings as Restraints in Structure Elucidation. *Journal of Biomolecular NMR* **2000**, *16* (3), 245–252. https://doi.org/10.1023/A:1008378624590.

(24) Skrynnikov, N. R.; Kay, L. E. Assessment of Molecular Structure Using Frame-Independent Orientational Restraints Derived from Residual Dipolar Couplings. *Journal of Biomolecular NMR* **2000**, *18* (3), 239–252. https://doi.org/10.1023/A:1026501101716.

(25) Delaglio, F.; Kontaxis, G.; Bax, A. Protein Structure Determination Using Molecular Fragment Replacement and NMR Dipolar Couplings. *Journal of the American Chemical Society* **2000**, *122* (9), 2142–2143. https://doi.org/10.1021/ja993603n.

(26) Eker, F.; Griebenow, K.; Cao, X.; Nafie, L. A.; Schweitzer-Stenner, R. Preferred Peptide Backbone Conformations in the Unfolded State Revealed by the Structure Analysis of Alanine-Based (AXA) Tripeptides in Aqueous Solution. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101* (27), 10054–10059. https://doi.org/10.1073/pnas.0402623101.

(27) Woutersen, S.; Hamm, P. Structure Determination of Trialanine in Water Using Polarization Sensitive Two-Dimensional Vibrational Spectroscopy. *The Journal of Physical Chemistry B* **2000**, *104* (47), 11316–11320. https://doi.org/10.1021/jp001546a.

(28) Woutersen, S.; Hamm, P. Isotope-Edited Two-Dimensional Vibrational Spectroscopy of Trialanine in Aqueous Solution. *The Journal of Chemical Physics* **2001**, *114* (6), 2727–2737. https://doi.org/10.1063/1.1336807.

(29) Eker, F.; Cao, X.; Nafie, L.; Huang, Q.; Schweitzer-Stenner, R. The Structure of Alanine Based Tripeptides in Water and Dimethyl Sulfoxide Probed by Vibrational Spectroscopy. *The Journal of Physical Chemistry B* **2003**, *107* (1), 358–365. https://doi.org/10.1021/jp026958t.

(30) Chellgren, B. W.; Creamer, T. P. Short Sequences of Non-Proline Residues Can Adopt the Polyproline II Helical Conformation. *Biochemistry* **2004**, *43* (19), 5864–5869. https://doi.org/10.1021/bi049922v.

(31) Grdadolnik, J.; Golič Grdadolnik, S.; Avbelj, F. Determination of Conformational Preferences of Dipeptides Using Vibrational Spectroscopy. *The Journal of Physical Chemistry B* **2008**, *112* (9), 2712–2718. https://doi.org/10.1021/jp7096313.

(32) Mikhonin, A. V; Ahmed, Z.; Ianoul, A.; Asher, S. A. Assignments and Conformational Dependencies of the Amide III Peptide Backbone UV Resonance Raman Bands. *The Journal of Physical Chemistry B* **2004**, *108* (49), 19020–19028. https://doi.org/10.1021/jp045959d.

(33) Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. Tripeptides Adopt Stable Structures in Water. A Combined Polarized Visible Raman, FTIR, and VCD Spectroscopy Study. *Journal of the American Chemical Society* **2002**, *124* (48), 14330–14341. https://doi.org/10.1021/ja027381w.

(34) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. Stable Conformations of Tripeptides in Aqueous Solution Studied by UV Circular Dichroism Spectroscopy. *Journal of the American Chemical Society* **2003**, *125* (27), 8178–8185. https://doi.org/10.1021/ja034625j.

(35) Hagarman, A.; Measey, T. J.; Mathieu, D.; Schwalbe, H.; Schweitzer-Stenner, R. Intrinsic Propensities of Amino Acid Residues in GxG Peptides Inferred from Amide I′ Band Profiles and NMR Scalar Coupling Constants. *Journal of the American Chemical Society* **2010**, *132* (2), 540–551. https://doi.org/10.1021/ja9058052.

(36) Vymětal, J.; Vondrášek, J. Critical Assessment of Current Force Fields. Short Peptide Test Case. *Journal of Chemical Theory and Computation* **2013**, *9* (1), 441–451. https://doi.org/10.1021/ct300794a.

(37) Lovell, S. C.; Davis, I. W.; Iii, W. B. A.; Bakker, P. I. W. De; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure Validation by Cα Geometry: φ,ψ and Cβ Deviation. *Proteins: Struct,. Funct.,*

Bioinf. **2003**, *450*, 437–450.

(38) Ho, B. K.; Thomas, A.; Brasseur, R. Revisiting the Ramachandran Plot : Hard-Sphere Repulsion , Electrostatics , and H-Bonding in the α -Helix. *Protein science : a publication of the Protein Society* **2003**, *12*, 2508–2522. https://doi.org/10.1110/ps.03235203.of.

(39) Heijenoort, J. van. Formation of the Glycan Chains in the Synthesis of Bacterial Peptidoglycan. *Glycobiology* **2001**, *11* (3), 25R–36R. https://doi.org/10.1093/glycob/11.3.25R.

(40) Pisarewicz, K.; Mora, D.; Pflueger, F. C.; Fields, G. B.; Marí, F. Polypeptide Chains Containing D-γ-Hydroxyvaline. *Journal of the American Chemical Society* **2005**, *127* (17), 6207–6215. https://doi.org/10.1021/ja050088m.

(41) Albrecht, G.; Corey, R. B. The Crystal Structure of Glycine. *Journal of the American Chemical Society* **1939**, *61* (5), 1087–1103. https://doi.org/10.1021/ja01874a028.

(42) Marsh, R. E. A Refinement of the Crystal Structure of Glycine. *Acta Crystallographica* **1958**, *11* (9), 654–663. https://doi.org/10.1107/S0365110X58001717.

(43) Levy, H. A.; Corey, R. B. The Crystal Structure of Dl-Alanine. *Journal of the American Chemical Society* **1941**, *63* (8), 2095–2108. https://doi.org/10.1021/ja01853a020.

(44) Donohue, J. The Crystal Structure of DL-Alanine. II. Revision of Parameters by Three-Dimensional Fourier Analysis1. *Journal of the American Chemical Society* **1950**, *72* (2), 949–953. https://doi.org/10.1021/ja01158a079.

(45) Hirokawa, S. A New Modification of L-Glutamic Acid and Its Crystal Structure. *Acta Crystallographica* **1955**, *8* (10), 637–641. https://doi.org/10.1107/S0365110X55001990.

(46) Gaffney, J. S.; Pierce, R. C.; Friedman, L. Mass Spectrometer Study of Evaporation of .Alpha.-Amino Acids. *Journal of the American Chemical Society* **1977**, *99* (13), 4293–4298. https://doi.org/10.1021/ja00455a015.

(47) Edsall, J. T.; Blanchard, M. H. The Activity Ratio of Zwitterions and Uncharged Molecules in Ampholyte Solutions. The Dissociation Constants of Amino Acid Esters. *Journal of the American Chemical Society* **1933**, *55* (6), 2337–2353. https://doi.org/10.1021/ja01333a019.

(48) Ellzy, M. W.; Jensen, J. O.; Hameka, H. F.; Kay, J. G. Correlation of Structure and Vibrational Spectra of the Zwitterion L-Alanine in the Presence of Water: An Experimental and Density Functional Analysis. *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy* **2003**, *59* (11), 2619–2633.

(49) Torii, B. Y. K. Crystal Structures and Molecular Conformations of L-Methionine and L-Norleucine. *Acta Crystallographica* **1973**, *B29*, 2799–2807.

(50) Dalhus, B, Gorbitz, C. H. Crystal Structures of Hydrophobic Amino Acids. I. Redeterminations of L-Methionine and L-Valine at 120 K. *Acta chemica Scandinavia* **1996**, *50*, 544–548.

(51) Lesarri, A.; Mata, S.; Cocinero, E. J.; Blanco, S.; Lûpez, J. C.; Alonso, J. L. The Structure of Neutral Proline **. *Angew. Chem. Int. Ed.* **2002**, *41* (24), 4673–4676.

(52) Sanz, M. E.; Cabezas, C.; Mata, S.; Alonso, J. L. Rotational Spectrum of Tryptophan. *The Journal of Chemical Physics* **2014**, *140* (20), 204308. https://doi.org/10.1063/1.4876001.

(53) Cocinero, E. J.; Lesarri, A.; Grabow, J.; López, J. C.; Alonso, J. L. The Shape of Leucine in the Gas Phase. *ChemPhysChem* **2007**, *8*, 599–604. https://doi.org/10.1002/cphc.200600730.

(54) Alonso, J. L.; Cocinero, E. J.; Lesarri, A.; Sanz, M. E.; López, J. C. The Glycine–Water Complex**. *Angew. Chem. Int. Ed.* **2006**, *45*, 3471–3474. https://doi.org/10.1002/anie.200600342.

(55) Sanz, M. E.; Lesarri, A.; Pen, M. I.; Vaquero, V.; Cortijo, V.; Lo, J. C.; Alonso, L. The Shape of -Alanine. *J. Am. Chem. Soc.* **2006**, *128*, 3812–3817. https://doi.org/10.1021/ja058194b.

(56) Blanco, S.; Lesarri, A.; Lo, J. C.; Alonso, L. The Gas-Phase Structure of Alanine. *J. Am. Chem. Soc.* **2004**, *126*, 11675–11683. https://doi.org/10.1021/ja048317c.

(57) Sanz, M. E.; Blanco, S.; López, J. C.; Alonso, J. L. Rotational Probes of Six Conformers of Neutral Cysteine **. *Angew. Chem. Int. Ed.* **2008**, *47*, 6216–6220. https://doi.org/10.1002/anie.200801337.

(58) Lesarri, A.; Sa, R.; Cocinero, E. J.; Lo, J. C.; Alonso, L. Coded Amino Acids in Gas Phase : The Shape of Isoleucine. *J. Am. Chem. Soc.* **2005**, *127* (8), 12952–12956. https://doi.org/10.1021/ja0528073.

(59) Sanz, M. E.; Cortijo, V.; Caminati, W.; Lopez, J. C.; Alonso, J. L. The Conformers of Phenylglycine. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2006**, *12* (9), 2564–2570. https://doi.org/10.1002/chem.200501013.

(60) Alonso, J. L.; Vaquero, V.; Peça, I.; López, J. C.; Mata, S. All Five Forms of Cytosine Revealed in the Gas Phase ** Angewandte. *Angew. Chem. Int. Ed.* **2013**, *52*, 2331–2334. https://doi.org/10.1002/anie.201207744.

(61) Blanco, S.; Sanz, M. E.; Lopez, J. C.; Alonso, J. L. Revealing the Multiple Structures of Serine. *Proceedings of the National Academy of Sciences* **2007**, *104* (51), 20183–20188. https://doi.org/10.1073/pnas.0705676104.

(62) Mata, S.; Blanco, S.; Juan, C. L.; Alonso, L.; Química-física, D.; Tecnol, P. Jet-Cooled Rotational Spectrum of Laser-Ablated Phenylalanine. *J. Phys. Chem. B* **2011**, *115*, 9653–9657. https://doi.org/10.1021/jp200800a.

(63) Mata, S.; Vaquero, V.; Cabezas, C.; Pen, I.; Alonso, L.; Lo, J. C. Observation of Two New Conformers of Neutral Proline W. *Physical Chemistry Chemical Physics* **2009**, *11*, 4141–4144. https://doi.org/10.1039/b904633j.

(64) Alonso, L.; Perez, C.; Sanz, M. E.; Lo, J. C.; Blanco, S. Seven Conformers of L -Threonine in the Gas Phase : A LA-MB-FTMW Study. *Physical Chemistry Chemical Physics* **2009**, *11*, 617–627. https://doi.org/10.1039/b810940k.

(65) Pen, I.; Sanz, M. E.; Lo, J. C.; Alonso, J. L. Preferred Conformers of Proteinogenic Glutamic Acid. *J. Am. Chem. Soc.* **2011**, *134*, 2305–2312. https://doi.org/10.1021/ja2101449.

(66) Császár, A. G. On the Structures of Free Glycine and α-Alanine. *Journal of Molecular Structure* **1995**, *346*, 141–152. https://doi.org/https://doi.org/10.1016/0022-2860(94)09017-J.

(67) Csaszar, A. G. Conformers of Gaseous Glycine. *Journal of the American Chemical Society* **1992**, *114* (24), 9568–9575. https://doi.org/10.1021/ja00050a041.

(68) Gronert, S.; O'Hair, R. A. J. Ab Initio Studies of Amino Acid Conformations. 1. The Conformers of Alanine, Serine, and Cysteine. *Journal of the American Chemical Society* **1995**, *117* (7), 2071–2081. https://doi.org/10.1021/ja00112a022.

(69) Császár, A. G. Conformers of Gaseous α-Alanine. *The Journal of Physical Chemistry* **1996**, *100* (9), 3541–3551. https://doi.org/10.1021/jp9533640.

(70) Zhang, M.; Huang, Z.; Lin, Z. Systematic Ab Initio Studies of the Conformers and Conformational Distribution of Gas-Phase Tyrosine. *The Journal of Chemical Physics* **2005**, *122* (13), 134313. https://doi.org/10.1063/1.1869471.

(71) Zhang, M.; Lin, Z. Ab Initio Studies of the Conformers and Conformational Distribution of the Gaseous Hydroxyamino Acid Threonine. *Journal of Molecular Structure: THEOCHEM* **2006**, *760* (1), 159–166. https://doi.org/https://doi.org/10.1016/j.theochem.2005.12.008.

(72) Huang, Z.; Yu, W.; Lin, Z. Exploration of the Full Conformational Landscapes of Gaseous Aromatic Amino Acid Phenylalanine: An Ab Initio Study. *Journal of Molecular Structure: THEOCHEM* **2006**, *758* (2), 195–202. https://doi.org/https://doi.org/10.1016/j.theochem.2005.10.043.

(73) Czinki, E.; Császár, A. G. Conformers of Gaseous Proline. *Chemistry – A European Journal* **2003**, *9* (4), 1008–1019. https://doi.org/10.1002/chem.200390103.

(74) Allen, W. D.; Czinki, E.; Császár, A. G. Molecular Structure of Proline. *Chemistry – A European Journal*

2004, *10* (18), 4512–4517. https://doi.org/10.1002/chem.200400112.

(75) Kasalová, V.; Allen, W. D.; Schaefer III, H. F.; Czinki, E.; Császár, A. G. Molecular Structures of the Two Most Stable Conformers of Free Glycine. *Journal of Computational Chemistry* **2007**, *28* (8), 1373–1383. https://doi.org/10.1002/jcc.20680.

(76) Wilke, J. J.; Lind, M. C.; Schaefer, H. F.; Császár, A. G.; Allen, W. D. Conformers of Gaseous Cysteine. *Journal of Chemical Theory and Computation* **2009**, *5* (6), 1511–1523. https://doi.org/10.1021/ct900005c.

(77) Kesharwani, M. K.; Karton, A.; Martin, J. M. L. Benchmark Ab Initio Conformational Energies for the Proteinogenic Amino Acids through Explicitly Correlated Methods. Assessment of Density Functional Methods. *Journal of Chemical Theory and Computation* **2016**, *12* (1), 444–454. https://doi.org/10.1021/acs.jctc.5b01066.

(78) Ivanov, A. Y.; Sheina, G.; Blagoi, Y. P. FTIR Spectroscopic Study of the UV-Induced Rotamerization of Glycine in the Low Temperature Matrices (Kr , Ar , Ne). *Spectrochim. Acta, A* **1999**, *55*, 219–228.

(79) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M. L. T. S.; Fausto, R.; Adamowicz, L. Matrix-Isolation Infrared and Theoretical Studies of the Glycine Conformers. *J. Phys. Chem. A* **1998**, *102* (97), 1041–1054.

(80) Nunes, C. M.; Lapinski, L.; Fausto, R.; Reva, I. Near-IR Laser Generation of a High-Energy Conformer of L-Alanine and the Mechanism of Its Decay in a Low-Temperature Nitrogen Matrix. *Journal of Chemical Physics* **2013**, *138* (12), 0–12. https://doi.org/10.1063/1.4795823.

(81) Lambie, B.; Ramaekers, R.; Maes, G. On the Contribution of Intramolecular H-Bonding Entropy to the Conformational Stability of Alanine Conformations. *Spectrochimica Acta PartA* **2003**, *59*, 1387–1397.

(82) Kaczor, A.; Reva, I. D.; Proniewicz, L. M.; Fausto, R. Importance of Entropy in the Conformational Equilibrium of Phenylalanine : A Matrix-Isolation Infrared Spectroscopy and Density Functional Theory Study. *J. Phys. Chem. A* **2006**, *110*, 2360–2370.

(83) Jarmelo, S.; Lapinski, L.; Nowak, M. J.; Carey, P. R.; Fausto, R. Preferred Conformers and Photochemical ($\lambda > 200$ Nm) Reactivity of Serine and 3 , 3-Dideutero-Serine In the Neutral Form. *J. Phys. Chem. A* **2005**, *109*, 5689–5707.

(84) Lambie, B.; Ramaekers, R.; Maes, G. Conformational Behavior of Serine : An Experimental Matrix-Isolation FT-IR and Theoretical DFT (B3LYP)/ 6-31 ++ G ** Study. *J. Phys. Chem. A* **2004**, *108*, 10426–10433.

(85) Najbauer, E. E.; Apo, R. Identi Fi Cation of Serine Conformers by Matrix-Isolation IR Spectroscopy Aided by Near-Infrared Laser-Induced Conformational Change , 2D Correlation Analysis , and Quantum Mechanical Anharmonic Computations. *J. Phys. Chem. B* **2015**, *119*, 10496–10510. https://doi.org/10.1021/acs.jpcb.5b05768.

(86) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. Combined Matrix-Isolation Infrared and Theoretical DFT and Ab Initio Study of the Nonionized Valine Conformers. *The Journal of Physical Chemistry A* **1999**, *103* (22), 4404–4412. https://doi.org/10.1021/jp984457v.

(87) Boeckx, B.; Nelissen, W.; Maes, G. Potential Energy Surface and Matrix Isolation FT-IR Study of Isoleucine. *J. Phys. Chem. A* **2012**, *116*, 3247–3258.

(88) Boeckx, B.; Maes, G. Experimental and Theoretical Observation of Di Ff Erent Intramolecular H - Bonds in Lysine Conformations. *J. Phys. Chem. B* **2012**, *116*, 12441–12449.

(89) Stepanian, S. G.; Yu, A.; Adamowicz, L. Conformational Composition of Neutral Leucine . Matrix Isolation Infrared and Ab Initio Study. *Chemical Physics* **2013**, *423*, 20–29. https://doi.org/10.1016/j.chemphys.2013.06.018.

(90) Dobrowolski, J. C.; Jamróz, M. H.; Kołos, R.; Rode, J. E.; Sadlej, J. Theoretical Prediction and the First IR Matrix Observation of Several l -Cysteine Molecule Conformers. *ChemPhysChem* **2007**, *8*, 1085–1094.

https://doi.org/10.1002/cphc.200600784.

(91) Najbauer, E. E.; Bazsó, G.; Apóstolo, R.; Fausto, R.; Biczysko, M.; Barone, V.; Tarczay, G. Identification of Serine Conformers by Matrix-Isolation IR Spectroscopy Aided by Near-Infrared Laser-Induced Conformational Change, 2D Correlation Analysis, and Quantum Mechanical Anharmonic Computations. *The Journal of Physical Chemistry B* **2015**, *119* (33), 10496–10510. https://doi.org/10.1021/acs.jpcb.5b05768.

(92) Ramaekers, R.; Pajak, J.; Rospenk, M.; Maes, G. Matrix-Isolation FT-IR Spectroscopic Study and Theoretical DFT (B3LYP)/ 6 – 31 ++ G ** Calculations of the Vibrational and Conformational Properties of Tyrosine. *Spectrochimica Acta PartA* **2005**, *61*, 1347–1356. https://doi.org/10.1016/j.saa.2004.10.003.

(93) Whittle, E.; Dows, D. A.; Pimentel, G. C. Matrix Isolation Method for the Experimental Study of Unstable Species. *The Journal of Chemical Physics* **1954**, *22* (11), 1943. https://doi.org/10.1063/1.1739957.

(94) Becker, E. D.; Pimentel, G. C. Spectroscopic Studies of Reactive Molecules by the Matrix Isolation Method. *The Journal of Chemical Physics* **1956**, *25* (2), 224–228. https://doi.org/10.1063/1.1742860.

(95) Van Thiel, M.; Becker, E. D.; Pimentel, G. C. Infrared Studies of Hydrogen Bonding of Methanol by the Matrix Isolation Technique. *The Journal of Chemical Physics* **1957**, *27* (1), 95–99. https://doi.org/10.1063/1.1743725.

(96) Milligan, D. E.; Pimentel, G. C. Matrix Isolation Studies: Possible Infrared Spectra of Isomeric Forms of Diazomethane and of Methylene, CH2. *The Journal of Chemical Physics* **1958**, *29* (6), 1405–1412. https://doi.org/10.1063/1.1744730.

(97) Van Thiel, M.; Becker, E. D.; Pimentel, G. C. Infrared Studies of Hydrogen Bonding of Water by the Matrix Isolation Technique. *The Journal of Chemical Physics* **1957**, *27* (2), 486–490. https://doi.org/10.1063/1.1743753.

(98) NORMAN, I.; PORTER, G. Trapped Atoms and Radicals in a Glass 'Cage.' *Nature* **1954**, *174* (4428), 508–509. https://doi.org/10.1038/174508a0.

(99) Sankaran, K.; Vidya, V.; Viswanathan, K. S.; George, L.; Singh, S. Trimethyl Phosphate−Water Interaction: A Matrix-Isolation Infrared and Ab Initio Study. *The Journal of Physical Chemistry A* **1998**, *102* (17), 2944–2953. https://doi.org/10.1021/jp9733330.

(100) Sundararajan, K.; Viswanathan, K. S. Conformations of Triethylphosphate : A Reanalysis of the Matrix Isolation Spectra Conformations of Triethylphosphate : A Reanalysis of the Matrix Isolation Spectra. *J. Indian Inst. Sci.* **2005**, *85*, 403–418.

(101) Hobe, M. Von; Stroh, F.; Beckers, H. The UV / Vis Absorption Spectrum of Matrix-Isolated Dichlorine Peroxide , ClOOCl. *Physical Chemistry Chemical Physics* **2009**, *11*, 1571–1580. https://doi.org/10.1039/b814373k.

(102) Costa, P.; Sander, W. Hydrogen Bonding Switches the Spin State of Diphenylcarbene from Triplet to Singlet. *Angewandte Chemie International Edition* **2014**, *53* (20), 5122–5125. https://doi.org/10.1002/anie.201400176.

(103) Bondybey, V. E.; English, J. H. Spectroscopy and Relaxation of Pb2 in Rare Gas Solids. *The Journal of Chemical Physics* **1977**, *67* (8), 3405–3411. https://doi.org/10.1063/1.435335.

(104) Olbert-Majkut, A.; Ahokas, J.; Lundell, J.; Pettersson, M. Raman Spectroscopy of Formic Acid and Its Dimers Isolated in Low Temperature Argon Matrices. *Chemical Physics Letters* **2009**, *468* (4), 176–183. https://doi.org/https://doi.org/10.1016/j.cplett.2008.12.011.

(105) G. C. Pimental, S. W. C. INFRARED SPECTRAL PERTURBATIONS IN MATRIX EXPERIMENTS. *Pure Appl. Chem* **1963**, *7*, 111–123.

(106) Cradock, S.; Hinchcliffe, A. J. *Matrix Isolation: A Technique for the Study of Reactive Inorganic Species*; Cambridge University Press, 1975.

(107) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian 09, Revision B.01. *Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT*. Wallingford CT 2009.

(108) Vincent, M. A.; Hillier, I. H. The Structure and Interaction Energies of the Weak Complexes of CHClF 2 and CHF 3 with HCCH : A Test of Density Functional Theory Methods. *Physical Chemistry Chemical Physics* **2011**, *13*, 4388–4392. https://doi.org/10.1039/c0cp02626c.

(109) Bader, R. F. W.; Bader, R. F. *Atoms in Molecules: A Quantum Theory*; International series of monographs on chemistry; Clarendon Press, 1990.

(110) Biegler-König, F.; Schönbohm, J. Update of the AIM2000-Program for Atoms in Molecules. *Journal of Computational Chemistry* **2002**, *23* (15), 1489–1494. https://doi.org/10.1002/jcc.10085.

(111) Koch, U.; Popelier, P. L. A. Characterization of C-H-O Hydrogen Bonds on the Basis of the Charge Density. *Journal of Physical Chemistry* **1995**, *99* (24), 9747–9754. https://doi.org/10.1021/j100024a016.

(112) Popelier, P. L. A. *Atoms in Molecules. An Introduction*; Pearson Education, 2000.

(113) Espinosa, E.; Molins, E.; Lecomte, C. Hydrogen Bond Strengths Revealed by Topological Analyses of Experimentally Observed Electron Densities. *Chemical Physics Letters* **1998**, *285* (3), 170–173. https://doi.org/https://doi.org/10.1016/S0009-2614(98)00036-0.

(114) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO Version 3.1. *NBO Version 3.1*.

(115) Yip, G. M. S.; Chen, Z.; Edge, C. J.; Smith, E. H.; Dickinson, R.; Hohenester, E.; Townsend, R. R.; Fuchs, K.; Sieghart, W.; Evers, A. S.; et al. Receptors Identified by Photolabeling. *Nature Chemical Biology* **2013**, *9* (11), 715–720. https://doi.org/10.1038/nchembio.1340.

(116) Petrella, R. J.; Karplus, M. The Role of Carbon-Donor Hydrogen Bonds in Stabilizing Tryptophan Conformations. *Proteins: Struct,. Funct., Bioinf.* **2004**, *726* (June 2003), 716–726.

(117) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. Combined Matrix-Isolation Infrared and Theoretical DFT and Ab Initio Study of the Nonionized Valine Conformers. *J. Phys. Chem. A* **1999**, *103*, 4404–4412.

(118) Reva, I. D.; Stepanian, S. G.; Plokhotnichenko, A. M.; Radchenko, E. D.; Sheina, G. G. Infrared Matrix Isolation Studies of Amino Acids. Molecular Structure of Proline. *J. Mol. Struct.* **1994**, *18*, 1–13.

(119) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. Conformers of Nonionized Proline . Matrix-Isolation Infrared and Post-Hartree - Fock Ab Initio Study. *J. Phys. Chem. A* **2001**, *105*, 10664–10672.

(120) Schäfer, L.; Kulp-Newton, S. Q.; Siam, K.; Klimkowski, V. J.; Van Alsenoy, C. Ab Initio Studies of Structural Features Not Easily Amenable to Experiment: Part 71. Conformational Analysis and Structural Study of Valine and Threonine. *Journal of Molecular Structure: THEOCHEM* **1990**, *209* (3), 373–385. https://doi.org/https://doi.org/10.1016/0166-1280(90)80089-7.

(121) Lakard, B. Ab Initio Study of Amino Acids Containing Hydroxy Groups (Serine, Threonine and Tyrosine). *Journal of Molecular Structure: THEOCHEM* **2004**, *681* (1–3), 183–189. https://doi.org/10.1016/j.theochem.2004.04.067.

(122) Xu, X.; Lin, Z. Comprehensive Ab Initio Study on the Conformations of L-Threonine and L-Allo-Threonine and Related Species in Gas Phase. *Journal of Molecular Structure: THEOCHEM* **2010**, *962* (1–3), 23–32. https://doi.org/10.1016/j.theochem.2010.09.010.

(123) Zhang, M.; Lin, Z. Ab Initio Studies of the Conformers and Conformational Distribution of the Gaseous Hydroxyamino Acid Threonine. *Journal of Molecular Structure: THEOCHEM* **2006**, *760*, 159–166. https://doi.org/10.1016/j.theochem.2005.12.008.

(124) Szidarovszky, T.; Czakó, G.; Császár, A. G. Conformers of Gaseous Threonine. *Molecular Physics* **2009**, *107*, 761–775. https://doi.org/10.1080/00268970802616350.

(125) Quesada-Moreno, M. M.; Márquez-García, A. Á.; Avilés-Moreno, J. R.; López-González, J. J. Conformational Landscape of L-Threonine in Neutral, Acid and Basic Solutions from Vibrational Circular Dichroism Spectroscopy and Quantum Chemical Calculations. *Tetrahedron Asymmetry* **2013**, *24* (24), 1537–1547. https://doi.org/10.1016/j.tetasy.2013.09.025.

(126) Najbauer, E. E. Exploring the Conformational Space of Cysteine by Matrix Isolation Spectroscopy Combined with Near-Infrared Laser Induced Conformational Change. *J. Phys. Chem. B* **2014**, *118*, 2093–2103.

(127) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *The Journal of Chemical Physics* **1997**, *106* (23), 9639–9646. https://doi.org/10.1063/1.473863.

(128) Karir, G.; Viswanathan, K. S. Phenylacetylene-Water Complex: Is It N···σ or H···π in the Matrix? *Journal of Molecular Structure* **2016**, *1107*, 145–156. https://doi.org/10.1016/j.molstruc.2015.11.030.

(129) Raut, A. H.; Karir, G.; Viswanathan, K. S. Matrix Isolation Infrared and Ab Initio Study of the Interaction of N-Heterocyclic Carbene with Water and Methanol: A Case Study of a Strong Hydrogen Bond. *The Journal of Physical Chemistry A* **2016**, *120* (47), 9390–9400. https://doi.org/10.1021/acs.jpca.6b08148.

(130) Karir, G.; Viswanathan, K. S. H−π Landscape of the Phenylacetylene–HCl System: Does This Provide the Gateway to the Markovnikov Addition? *The Journal of Physical Chemistry A* **2017**, *121* (31), 5797–5808. https://doi.org/10.1021/acs.jpca.7b04853.

(131) Karir, G.; Kumar, G.; Kar, B. P.; Viswanathan, K. S. Multiple Hydrogen Bond Tethers for Grazing Formic Acid in Its Complexes with Phenylacetylene. *The journal of physical chemistry. A* **2018**, *122* (8), 2046–2059. https://doi.org/10.1021/acs.jpca.7b11428.

(132) Saini, J.; Viswanathan, K. S. Does a Hydrogen Bonded Complex with Dual Contacts Show Synergism? A Matrix Isolation Infrared and Ab-Initio Study of Propargyl Alcohol-Water Complex. *Journal of Molecular Structure* **2016**, *1118*, 147–156. https://doi.org/10.1016/j.molstruc.2016.04.005.

(133) Saini, J.; Viswanathan, K. S. Discerning Near-Isoergic Isomers. A Matrix Isolation Infrared and Ab Initio Study of the Propargyl Alcohol Dimers. *Journal of Physical Chemistry A* **2017**, *121* (7). https://doi.org/10.1021/acs.jpca.6b12702.

(134) Mishra, P.; Verma, K.; Bawari, D.; Viswanathan, K. S. Does Borazine–water Behave like Benzene-Water? A Matrix Isolation Infrared and Ab Initio Study. *The Journal of Chemical Physics* **2016**, *144* (23), 234307. https://doi.org/10.1063/1.4953793.

(135) Verma, K.; Viswanathan, K. S. The Borazine Dimer: The Case of a Dihydrogen Bond Competing with a Classical Hydrogen Bond. *Physical Chemistry Chemical Physics* **2017**, *19* (29), 19067–19074. https://doi.org/10.1039/C7CP04056C.

(136) Dubey, P.; Saini, J.; Verma, K.; Karir, G.; Mukhopadhyay, A.; Viswanathan, K. S. Chapter 14 - Matrix Isolation Spectroscopy—A Window to Molecular Processes A2 - Gupta, V.P. BT - Molecular and Laser Spectroscopy. In *Molecular and Laser Spectrosocopy*; Elsevier, 2018; pp 317–340. https://doi.org/https://doi.org/10.1016/B978-0-12-849883-5.00014-0.

(137) Verma, K.; Dave, K.; Viswanathan, K. S. Hydrogen-Bonded Complexes of Phenylacetylene-Acetylene: Who Is the Proton Donor? *Journal of Physical Chemistry A* **2015**, *119* (51), 12656–12664. https://doi.org/10.1021/acs.jpca.5b08559.

(138) Verma, K.; Viswanathan, K. S. "A Tale of Two Structures": The Stacks and Ts of Borazine and Benzene Hetero and Homo Dimers. *ChemistrySelect* **2018**, *3* (3), 864–873. https://doi.org/10.1002/slct.201703005.

(139) Contineanu, I.; Neacşu, A.; Gheorghe, D.; Tănăsescu, S.; Perişanu, Ş. The Thermochemistry of Threonine Stereoisomers. *Thermochimica Acta* **2013**, *563*, 1–5. https://doi.org/https://doi.org/10.1016/j.tca.2013.04.001.

(140) Peña, I.; Sanz, M. E.; López, J. C.; Alonso, J. L. Preferred Conformers of Proteinogenic Glutamic Acid.

Journal of the American Chemical Society **2012**, *134* (4), 2305–2312. https://doi.org/10.1021/ja2101449.

(141) Tapiero, H.; Mathe, G.; Couvreur, P.; Tew, K. D. II. Glutamine and Glutamate. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* **2002**, *56* (9), 446–457.

(142) Najbauer, E. E.; Bazsó, G.; Góbi, S.; Magyarfalvi, G.; Tarczay, G. Exploring the Conformational Space of Cysteine by Matrix Isolation Spectroscopy Combined with Near-Infrared Laser Induced Conformational Change. *The Journal of Physical Chemistry B* **2014**, *118* (8), 2093–2103. https://doi.org/10.1021/jp412550q.

(143) Dubey, P.; Mukhopadhyay, A.; Viswanathan, K. S. Do Amino Acids Prefer Only Certain Backbone Structures? Steering through the Conformational Maze of l-Threonine Using Matrix Isolation Infrared Spectroscopy and Ab Initio Studies. *Journal of Molecular Structure* **2019**, *1175* (5), 117–129. https://doi.org/10.1016/j.molstruc.2018.07.066.

(144) Navarretea, J. T. L.; Bencivennib, L.; Ramondob, F.; Hernhndeza, V.; Ramirez, F. J. Theo Chem. *J. Mol. Struct. (threochem)* **1995**, *330* (94), 261–266.

(145) Sun, W.; Kinsel, G. R.; Marynick, D. S. Computational Estimates of the Gas-Phase Basicity and Proton Affinity of Glutamic Acid. *The Journal of Physical Chemistry A* **1999**, *103* (20), 4113–4117. https://doi.org/10.1021/jp9908101.

(146) Bouchoux, G.; Bimbong, R. N. B.; Nacer, F. Gas-Phase Protonation Thermochemistry of Glutamic Acid. *The Journal of Physical Chemistry A* **2009**, *113* (24), 6666–6676. https://doi.org/10.1021/jp902438a.

(147) Weiss, I. M.; Muth, C.; Drumm, R.; Kirchner, H. O. K. Thermal Decomposition of the Amino Acids Glycine , Cysteine , Aspartic Acid , Asparagine , Glutamic Acid , Glutamine , Arginine and Histidine. *BMC Biophysics* **2018**, *11* (2), 1–15.

(148) Barnes, A. J.; Kuzniarski, J. N. S.; Mielke, Z. Strongly Hydrogen-Bonded Molecular Complexes Studied by Matrix-Isolation Vibrational Spectroscopy. Part 1.-Amine-Hydrogen Chloride Complexes. *Journal of the Chemical Society, Faraday Transactions 2* **1984**, *80* (4), 465–463. https://doi.org/10.1039/f29848000465.

(149) Wheate, N. J. The Side Effects of Platinum-Based Chemotherapy Drugs: A Review for Chemists. **2018**, *47* (19). https://doi.org/10.1039/c8dt00838h.

(150) Reedijk, J. Why Does Cisplatin Reach Guanine-N7 with Competing S-Donor Ligands Available in the Cell? *Chemical Reviews* **1999**, *99* (9), 2499–2510. https://doi.org/10.1021/cr980422f.

(151) Silva, D. A.; Anto, L.; Costa, N. I. O. S.; Nu, N. Ab Initio Reaction Path for Cisplatin Interaction With L -Cysteine And. **2007**, *108* (2008), 401–414. https://doi.org/10.1002/qua.

(152) Cao, X.; Fischer, G. Conformational and Infrared Spectral Studies of L -Methionine and Its N-Deuterated Isotopomer as Isolated Zwitterions. **2002**, 41–50.

(153) Gronert, S.; Hair, R. A. J. O. Ab Initio Studies of Amino Acid Conformations . 1 . The Conformers of Alanine , Serine , and Cysteine. **1995**, No. 12, 2071–2081.

(154) Ramabhadran, R. O.; Sengupta, A.; Raghavachari, K. Application of the Generalized Connectivity-Based Hierarchy to Biomonomers : Enthalpies of Formation of Cysteine and Methionine. **2013**. https://doi.org/10.1021/jp403123c.

(155) Raut, A. H.; Karir, G.; Viswanathan, K. S. Matrix Isolation Infrared and *Ab Initio* Study of the Interaction of N-Heterocyclic Carbene with Water and Methanol: A Case Study of a Strong Hydrogen Bond. *The Journal of Physical Chemistry A* **2016**, *120* (47), 9390–9400. https://doi.org/10.1021/acs.jpca.6b08148.

(156) Verma, K.; Viswanathan, K. S. The Borazine Dimer: The Case of a Dihydrogen Bond Competing with a Classical Hydrogen Bond. *Physical Chemistry Chemical Physics* **2017**, *19* (29), 19067–19074. https://doi.org/10.1039/C7CP04056C.

(157) Yablokov, V. A.; Vasina, Y. A.; Zelyaev, I. A.; Mitrofanova, S. V. Kinetics of Thermal Decomposition of Sulfur-Containing Amino Acids. **2009**, *79* (6), 1141–1145. https://doi.org/10.1134/S1070363209060188.

(158) Blanco, S.; Sanz, M. E.; López, J. C.; Alonso, J. L. Revealing the Multiple Structures of Serine. *Proceedings*

of the National Academy of Sciences **2007**, *104* (51), 20183 LP-20188.

(159) Ropo, M.; Schneider, M.; Baldauf, C.; Blum, V. First-Principles Data Set of 45,892 Isolated and Cation-Coordinated Conformers of 20 Proteinogenic Amino Acids. *Scientific Data* **2016**, *3*, 160009.

(160) Ropo, M.; Blum, V.; Baldauf, C. Trends for Isolated Amino Acids and Dipeptides: Conformation, Divalent Ion Binding, and Remarkable Similarity of Binding to Calcium and Lead. *Scientific reports* **2016**, *6*, 35772. https://doi.org/10.1038/srep35772.

(161) Kishor, S.; Dhayal, S.; Mathur, M.; Ramaniah, L. M. Structural and Energetic Properties of α-Amino Acids: A First Principles Density Functional Study. *Molecular Physics* **2008**, *106* (19), 2289–2300. https://doi.org/10.1080/00268970802422577.

(162) Matta, C. F.; Bader, R. F. W. Atoms-in-Molecules Study of the Genetically Encoded Amino Acids. II. Computational Study of Molecular Geometries. *Proteins: Structure, Function, and Bioinformatics* **2002**, *48* (3), 519–538. https://doi.org/10.1002/prot.10170.

(163) Pettersson, M.; Maçôas, E. M. S.; Khriachtchev, L.; Fausto, R.; Räsänen, M. Conformational Isomerization of Formic Acid by Vibrational Excitation at Energies below the Torsional Barrier. *Journal of the American Chemical Society* **2003**, *125* (14), 4058–4059. https://doi.org/10.1021/ja0295016.

(164) ANFINSEN, C. B.; HABER, E.; SELA, M.; WHITE Jr, F. H. The Kinetics of Formation of Native Ribonuclease during Oxidation of the Reduced Polypeptide Chain. *Proceedings of the National Academy of Sciences of the United States of America* **1961**, *47* (9), 1309–1314.

(165) Yu, W.; Xu, X.; Li, H.; Pang, R. U. I.; Fang, K. U. N.; Lin, Z. Extensive Conformational Searches of 13 Representative Dipeptides and an Efficient Method for Dipeptide Structure Determinations Based on Amino Acid Conformers. *Journal of Computational Chemistry* **2009**, *30* (13), 2105–2120. https://doi.org/10.1002/jcc.

(166) Császár, A. G.; Perczel, A. Ab Initio Characterization of Building Units in Peptides and Proteins. *Progress in Biophysics and Molecular Biology* **1999**, *71* (2), 243–309. https://doi.org/https://doi.org/10.1016/S0079-6107(98)00031-5.

(167) Bouchoux, G. Gas Phase Basicities of Polyfunctional Molecules. Part 3: Amino Acids. *Mass spectrometry reviews* **2012**, *31* (3), 391–435.

(168) Hermans, J. The Amino Acid Dipeptide: Small but Still in Fl Uential after 50 Years. *Proceedings of the National Academy of Sciences* **2011**, *108*, 3095–3096. https://doi.org/10.1073/pnas.1019470108.

(169) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. Accurate Ab Initio Quantum Chemical Determination of the Relative Energetics of Peptide Conformations and Assessment of Empirical Force Fields. *Journal of the American Chemical Society* **1997**, *119* (25), 5908–5920. https://doi.org/10.1021/ja962310g.

(170) Yuan, Y.; Mills, M. J. L.; Popelier, P. L. A.; Jensen, F. Comprehensive Analysis of Energy Minima of the 20 Natural Amino Acids. *The Journal of Physical Chemistry A* **2014**, *118* (36), 7876–7891. https://doi.org/10.1021/jp503460m.

(171) Heaton, A. L.; Moision, R. M.; Armentrout, P. B. Experimental and Theoretical Studies of Sodium Cation Interactions with the Acidic Amino Acids and Their Amide Derivatives. *The Journal of Physical Chemistry A* **2008**, *112* (15), 3319–3327. https://doi.org/10.1021/jp711649g.

(172) Kesharwani, M. K.; Karton, A.; Martin, J. M. L. Benchmark Ab Initio Conformational Energies for the Proteinogenic Amino Acids through Explicitly Correlated Methods. Assessment of Density Functional Methods. *Journal of Computational Chemistry* **2016**, *12* (1), 444–454. https://doi.org/10.1021/acs.jctc.5b01066.

(173) Perczel, A.; Jákli, I.; McAllister, M. A.; Csizmadia, I. G. Relative Stability of Major Types of β-Turns as a Function of Amino Acid Composition: A Study Based on Ab Initio Energetic and Natural Abundance Data.

Chemistry – A European Journal **2003**, *9* (11), 2551–2566. https://doi.org/10.1002/chem.200204393.

(174) Bakker, J. M.; Plützer, C.; Hünig, I.; Häber, T.; Compagnon, I.; von Helden, G.; Meijer, G.; Kleinermanns, K. Folding Structures of Isolated Peptides as Revealed by Gas-Phase Mid-Infrared Spectroscopy. *ChemPhysChem* **2005**, *6* (1), 120–128. https://doi.org/10.1002/cphc.200400345.

(175) Gord, J. R.; Hewett, D. M.; Hernandez-castillo, A. O.; Blodgett, K. N.; Rotondaro, M. C.; Varuolo, A.; Kubasik, A.; Zwier, T. S. As a 3 10 -Helix Former †. *Physical Chemistry Chemical Physics* **2016**, *18*, 25512–25527. https://doi.org/10.1039/c6cp04909e.

(176) Blodgett, K. N.; Zhu, X.; Walsh, P. S.; Sun, D.; Lee, J.; Choi, S. H.; Zwier, T. S. Conformer-Speci Fi c and Diastereomer-Speci Fi c Spectroscopy of Aβα Synthetic Foldamers: Ac − Ala − β. *Journal of Physical Chemistry A* **2018**, *122*, 3697–3710. https://doi.org/10.1021/acs.jpca.8b01273.

(177) Blodgett, K. N.; Fischer, J. L.; Lee, J.; Choi, S. H.; Zwier, T. S. Conformation-Speci Fi c Spectroscopy of Asparagine-Containing Peptides : In Fl Uence of Single and Adjacent Asn Residues on Inherent Conformational Preferences. *The Journal of Physical Chemistry A* **2018**, *122*, 8762–8775. https://doi.org/10.1021/acs.jpca.8b08418.

(178) Deblase, A. F.; Harrilal, C. P.; Lawler, J. T.; Burke, N. L.; Mcluckey, S. A.; Zwier, T. S. Conformation-Speci Fi c Infrared and Ultraviolet Spectroscopy of Cold [YAPAA + H] + and [YGPAA + H] + Ions : A Stereochemical " Twist " on the β - Hairpin Turn. *J. Am. Chem. Soc.* **2017**, *139*, 5481–5493. https://doi.org/10.1021/jacs.7b01315.

(179) Wieczorek, R.; Dannenberg, J. J. Hydrogen-Bond Cooperativity, Vibrational Coupling, and Dependence of Helix Stability on Changes in Amino Acid Sequence in Small 310-Helical Peptides. A Density Functional Theory Study. *Journal of the American Chemical Society* **2003**, *125* (46), 14065–14071. https://doi.org/10.1021/ja034034t.

(180) Cable, J. R.; Tubergen, M. J.; Levy, D. H. Laser Desorption Molecular Beam Spectroscopy: The Electronic Spectra of Tryptophan Peptides in the Gas Phase. *Journal of the American Chemical Society* **1987**, *109* (20), 6198–6199. https://doi.org/10.1021/ja00254a057.

(181) Evans, D. A.; Wales, D. J.; Dian, B. C.; Zwier, T. S. The Dynamics of Conformational Isomerization in Flexible Biomolecules. II. Simulating Isomerizations in a Supersonic Free Jet with Master Equation Dynamics. *The Journal of Chemical Physics* **2003**, *120* (1), 148–157. https://doi.org/10.1063/1.1626541.

(182) Weinkauf, R.; Schermann, J.-P.; de Vries, M. S.; Kleinermanns, K. Molecular Physics of Building Blocks of Life under Isolated or Defined Conditions. *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics* **2002**, *20* (3), 309–316. https://doi.org/10.1140/epjd/e2002-00185-0.

(183) Hünig, I.; Seefeld, K. A.; Kleinermanns, K. REMPI and UV–UV Double Resonance Spectroscopy of Tryptophan Ethylester and the Dipeptides Tryptophan–serine, Glycine–tryptophan and Proline–tryptophan. *Chemical Physics Letters* **2003**, *369* (1), 173–179. https://doi.org/https://doi.org/10.1016/S0009-2614(02)01966-8.

(184) Abo-Riziq, A. G.; Bushnell, J. E.; Crews, B.; Callahan, M. P.; Grace, L.; De Vries, M. S. Discrimination between Diastereoisomeric Dipeptides by IR–UV Double Resonance Spectroscopy and Ab Initio Calculations. *International Journal of Quantum Chemistry* **2005**, *105* (4), 437–445. https://doi.org/10.1002/qua.20719.