

# BIOPHYSICAL STUDIES ON STRUCTURAL TRANSITIONS OF T7 BACTERIOPHAGE ENDOLYSIN

Ph.D THESIS

by

MEENAKSHI SHARMA



DEPARTMENT OF BIOTECHNOLOGY  
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE  
ROORKEE-247667, INDIA  
NOVEMBER, 2017



# **BIOPHYSICAL STUDIES ON STRUCTURAL TRANSITIONS OF T7 BACTERIOPHAGE ENDOLYSIN**

**A THESIS**

*Submitted in partial fulfilment of the  
requirements for the award of the degree*

*of*

**DOCTOR OF PHILOSOPHY**

*in*

**BIOTECHNOLOGY**

*by*

**MEENAKSHI SHARMA**



**DEPARTMENT OF BIOTECHNOLOGY  
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE  
ROORKEE-247667, INDIA  
NOVEMBER, 2017**





**©INDIAN INSTITUTE OF TECHNOLOGY ROORKEE, ROORKEE- 2017  
ALL RIGHTS RESERVED**





# INDIAN INSTITUTE OF TECHNOLOGY ROORKEE ROORKEE

## CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled “ **BIOPHYSICAL STUDIES ON STRUCTURAL TRANSITIONS OF T7 BACTERIOPHAGE ENDOLYSIN**”, in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from July, 2013 to November, 2017, under the supervision of Dr. Krishna Mohan Poluri, Assistant Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institute.

(**MEENAKSHI SHARMA**)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(**KRISHNA MOHAN POLURI**)  
Supervisor

The Ph. D. Viva-Voce examination of **Ms. Meenakshi Sharma**, Research Scholar has been held on .....

**Chairman, SRC**

**Signature of External Examiner**

This is to certify that student has made all the corrections in the thesis.

**Signature of Supervisor**

**Head of the Department**

**Dated:** \_\_\_\_\_





## Table of Contents

Acknowledgments	V
Abstract	VII
List of publications	XI
List of presentations in conference	XII

### Chapter 1

<b>Introduction to bacteriophage endolysin and biophysical techniques</b>	1
<b>1.1 Bacterial cell wall</b>	3
1.1.1 Gram-positive cell wall	4
1.1.2 Gram-negative cell wall	5
<b>1.2 Bacteriophages</b>	5
1.2.1 Bacteriophage classification	6
1.2.2 Structural organization of bacteriophage	7
1.2.3 Phage multiplication cycles	8
1.2.4 Phage therapy	9
<b>1.3 Bacteriophage derived peptidoglycan degrading enzymes</b>	10
1.3.1 Virion-associated peptidoglycan hydrolases (VAPGHs)	11
1.3.2 Holins, Pinholins and Spanins	12
<b>1.4 Bacteriophage endolysins</b>	12
1.4.1 Structural architecture of endolysins	14
1.4.1.1 Endolysins infecting Gram-positive bacteria	14
1.4.1.2 Endolysins infecting Gram-negative bacteria	15
1.4.2 Bacterial resistance to endolysin	17
1.4.3 Endolysin engineering	18
1.4.3.1 Domain shuffling/swapping/recombination: rational design	18
1.4.3.2 Mutagenesis approaches	19
1.4.3.3 Endolysin translocation	20
1.4.4 Artilynsins	21
1.4.5 Endolysin applications	22
1.4.5.1 Food preservation	22
1.4.5.2 Therapeutics	22
1.4.5.3 Biofilm elimination and disinfestation	23
1.4.5.4 Agriculture	23
1.4.5.5 Endolysin CBD mediated detection and quantification	24
<b>1.5 Protein structure and folding</b>	24
1.5.1 Protein folding	25
1.5.2 Folding energy landscape or folding funnel	26
1.5.3 Protein misfolding and aggregation	27
1.5.4 Intrinsically disordered proteins (IDPs)	28

1.5.5 Factors affecting protein stability	29
<b>1.6 Biophysical techniques used to study protein structure and stability</b>	29
1.6.1 Circular dichroism (CD) spectroscopy	29
1.6.2 Fluorescence spectroscopy	31
1.6.3 Fluorescence lifetime spectroscopy (FLS)	32
1.6.4 Nuclear magnetic resonance (NMR) spectroscopy	33
1.6.4.1 Protein structure determination	33
1.6.4.2 Protein dynamic studies using NMR spectroscopy	36
1.6.4.3 Hydrogen exchange (HX)	37
1.6.4.4 Translational diffusion measurements	37
1.6.4.5 Temperature dependence of amide proton chemical shifts	38
<b>1.7 Thesis outline and objectives</b>	39
1.7.1 Thesis objectives	40
1.7.1 Scope of the thesis	40
<b>1.8 References</b>	41

## **Chapter 2**

### **Mechanistic insights into pH-dependent structure-function relationship of T7 endolysin**

<b>2.1 Introduction</b>	67
<b>2.2 Materials and Methods</b>	68
2.2.1 Construct design	68
2.2.2 Overexpression and purification	69
2.2.3 Cell lysis turbidimetric assay	72
2.2.4 Circular dichroism (CD) spectroscopy	73
2.2.5 Fluorescence spectroscopy	73
2.2.6 Fluorescence lifetime spectroscopy (FLS)	73
2.2.7 Nuclear magnetic resonance (NMR) spectroscopy	74
2.2.8 Measurement of translational diffusion characteristics of T7L using 2D-DOSY	74
2.2.9 Calculation of buried index and pKa Values	75
<b>2.3 Results</b>	75
2.3.1 pH-dependent enzymatic efficiency of T7L endolysin	75
2.3.2 pH-dependent secondary structural modulations of T7 endolysin	76
2.3.3 pH-dependent tertiary structural changes in T7L	77
2.3.4 Reversibility of the pH-dependent T7L structural transition	83
2.3.5 Effect of ZnSO <sub>4</sub> on pH-induced structural transitions	84
<b>2.4 Discussion</b>	85
2.4.1 Mechanistic insights into the pH-dependent structural transition of T7 endolysin	85

<b>2.5 Concluding remarks</b>	88
<b>2.6 References</b>	88

### **Chapter 3**

## **Conformational heterogeneity, dynamics and stability features of T7 endolysin native conformations**

<b>3.1 Introduction</b>	93
<b>3.2 Materials and methods</b>	94
3.2.1 Protein expression and purification	94
3.2.2 Circular dichroism (CD) spectroscopy	94
3.2.3 Steady state fluorescence measurements	95
3.2.4 Size exclusion chromatography (SEC)	95
3.2.5 Nuclear magnetic resonance (NMR) spectroscopy	95
3.2.5.1 Backbone resonance assignment	96
3.2.5.2 Hydrogen-exchange (HX) experiment	96
3.2.5.3 NMR relaxation experiments	96
<b>3.3 Results</b>	97
3.3.1 Backbone resonance assignment of T7L	97
3.3.2 Existence of multiple conformations in native state ensemble	98
3.3.3 Residue level insights into the T7L native state stability	100
3.3.4 Assessing the pH dependent conformational heterogeneity in native state ensemble	102
3.3.5 Structural dynamics of T7L native conformations	105
3.3.6 Assessing the structural stability features of T7L native state conformations	107
3.3.6.1 Chemical denaturation of T7L native conformations	108
3.3.6.2 Thermal stability and reversibility of low pH conformations	110
<b>3.4 Discussion</b>	111
3.4.1 Heterogeneity and differential structural stability features of T7L native state ensemble	111
<b>3.5 Concluding remarks</b>	113
<b>3.6 References</b>	113

### **Chapter 4**

## **Deciphering the differential structural stability and dynamics of partially folded conformations of T7 endolysin**

<b>4.1 Introduction</b>	117
<b>4.2 Materials and methods</b>	118
4.2.1 Protein expression and purification	118
4.2.2 Circular dichroism (CD) spectroscopy	118

4.2.3 Steady state fluorescence measurements	119
4.2.4 Cell lysis turbidimetric assay	119
4.2.5 Nuclear magnetic resonance (NMR) spectroscopy	120
<b>4.3 Results</b>	121
4.3.1 Structural heterogeneity of T7L partially folded conformations	121
4.3.2 Residue level insights into structure and dynamics of T7L partially folded states	123
4.3.3 Differential stability characteristics of T7L partially folded states	125
4.3.3.1 Chemical denaturation profiles of low pH conformations	125
4.3.3.2 Thermal stability and reversibility of low pH conformations	130
4.3.4 pH and temperature dependent amidase activity	132
4.3.5 Effect of temperature on amide proton chemical shifts	133
4.3.6 Comparative studies on structural- dynamics of PF conformations	134
<b>4.4 Discussion</b>	135
4.4.1 Mechanistic insights into differential biophysical characteristics of T7L PF conformations	135
<b>4.5 Concluding remarks</b>	137
<b>4.6 References</b>	137

## **Chapter 5**

### **Concluding remarks and future perspectives**

5.1 Concluding remarks	141
5.2 Future perspectives	142

## **Appendices**

**Appendix-I** NMR chemical shift values of T7L backbone resonance assignment at pH 7 i

**Appendix-V** NMR chemical shift values of T7L backbone resonance assignment at pH 3 iv

## ACKNOWLEDGEMENTS

*I would like to express my deep love and affection to my family members specially 'Mummy', for their infinite love, trust and immense support throughout my journey. I dedicate this thesis to my mother for her great engorgement.*

*I would like to express my deepest gratitude to my supervisor, Dr. Krishna Mohan Poluri for his guidance and encouragement during my Ph. D work. I have been inspired by his meticulousness and deep research thoughts.*

*I am grateful to Dr. Dinesh Kumar (CBMR, Lucknow) for his support in my Ph. D research work and for sharing valuable information.*

*I am thankful to my SRC committee members: Dr. Bijan Choudhury, Dr. A. K. Sharma, HOD and Dr. Kaushik Ghosh; DRC chairman: Prof. Vikas Pruthi, and Prof. Partha Roy for their constructive input. I also sincerely thank Prof. Ritu Barthwal, Dr. Sulakshana P. Mukherjee, Dr. Pravindra Kumar, Dr. P. Gopinath and other faculty members of Department of Biotechnology for their help and support during my Ph.D.*

*My gratitude goes to all my labmates Khushboo Gulati, Nidhi Joshi, Krishnakant Gangele, Jaya Lakshmi, Nipanshu Agarwal, Mukesh Kumar Meher, Nancy Jaiswal (CBMR, Lucknow), Dr. Shikha Sen, Apporva, Minal, Gaurav, Gopal, Rohan, Mohd. Yousuf and Neelanchal for helping me and providing friendly environment in the lab. I had really great time with you all.*

*I also thank the office staff, Department of Biotechnology, Institute Instrumentation Centre (IIC), Centre of Nanotechnology, campus administration and staff of my hostels, Kasturba & Sarojini bhawans, who rendered their help and made a comfortable stay in the campus.*

*Thanks to all my colleagues Deepa, Shashank, Amol, Poonam, Nishu, Anjali, Benazir and seniors, Pradeep sir, Padma mam and Sweta mam for their kind support and help.*

*I am grateful to University Grant Commission UGC for providing me financial assistance (JRF/SRF fellowship).*

*I humbly acknowledge several of my friends and well-wishers, whom I have not mentioned above and whose best wishes have always encouraged me.*

*Last but not the least my heartiest thanks to my friends Zia Tariq, Arun, Nidhi, Anita, Nupur and Suparna for making life so ecstatic and euphoric inside and outside the campus and for all the crazy moments.*



## Abstract

Bacteriophages are extremely diversified prokaryotic viruses that act on bacteria. Certain classes of bacteriophage known as lytic phages cause complete lysis of susceptible bacterial culture with the aid of endolysins. Endolysins hydrolyze specific bonds in the peptidoglycan, which triggers bursting and release of mature phage particles from bacterial host. Endolysins have been categorized into four classes that include: glycosidase (muramidase), endopeptidase, amidohydrolase (amidase) and lytic transglycosylase. Endolysins infecting Gram-positive bacteria contain an enzymatic catalytic domain (ECD) and a cell binding domain (CBD), whereas most endolysins from a Gram-negative background possess only an enzymatic catalytic domain. Endolysins from T7L, T4L,  $\lambda$  phage, K11, and KP32 etc., are the major representatives of a single-domain family.

Among these endolysins, amidases (also called as an N-acetylmuramoyl-L-alanine amidases) are known for their rapid lysis activity. They destabilize the peptidoglycan by separating the glycan polymer from the stem peptide by cleaving a critical amide bond between the glycan moiety (MurNAc) and the peptide moiety (L-alanine). Bacteriophages such as T7, T3, K11, Thermus scotoductus phage vB\_Tsc2631, and Bacillus anthracis prophage etc., comprise the single domain endolysin amidase family. These proteins have a molecular weight around 16-20 kDa and share the conserved  $\alpha$ - $\beta$  fold.

T7 endolysin (T7 lysozyme-T7L) is produced by T7 bacteriophage from *Podoviridae* family. It is a 151 amino acid long (17 kDa) globular bi-functional protein containing  $Zn^{+2}$  as a co-factor. T7L facilitates the rapid lysis of Gram-negative bacteria and inhibit T7 RNA polymerase at the end of their lytic multiplication cycle. Zinc is required for its amidase activity but not for RNAP inhibition. T7 endolysin contains five  $\beta$ -strands and three  $\alpha$ -helices, and has a prominent cleft. Zinc atom is located in the catalytic cleft and anchored to the protein by three side-chain ligands, His17, His122, and Cys130, and connected to the hydroxyl group of Tyr46 through a water molecule.

The current thesis unravels the molecular insights on pH dependent structure-folding-dynamics-function relationship of T7 endolysin by employing various biophysical techniques.

**Chapter 1** includes an overview of the origin of endolysins, classification, domain organization, structural diversity, mechanism of action, and comparative structural analysis of different phage endolysins with mammalian peptidoglycan recognition proteins (PGRPs) /bacterial

endolysins with a special emphasis on the structural-enzymatic relationship of phage endolysins. The chapter also presents the concepts and applications of engineered endolysins, which includes optimization and development of endolysin into potential therapeutic compounds by means of rational design, recent applications and patents in the field of medicine, food safety, agriculture and biotechnology. Further, a brief overview of biophysical techniques used to achieve the proposed research work was presented.

**Chapter 2** unravels the characterization of T7 endolysin structural features, and influence of pH on its structural stability and enzymatic activity. T7L was successfully expressed and purified by transforming *E. coli* BL 21 (DE3) cells with plasmid (pAR4593) containing T7L gene. Purification of T7L was done by using combination of chromatographic techniques and confirmed its monomeric state. The mechanism of pH-dependent differential activity profiles and structural features were investigated by employing cell lysis assay and various biophysical techniques including CD, fluorescence, size-exclusion chromatography and NMR techniques etc. Studies established that pH plays a major role in maintaining the structural integrity of the T7L structure and lowering the pH results in the loss of lysis activity. T7L exhibited a pH dependent reversible structural transition from native state to partially folded (PF) conformation below pH 6. Structural studies indicated a partial loss of secondary/tertiary structures in PF states (pH range 5-3) compared to its native state. Moreover, T7L PF states displayed differential binding characteristics of 8-anilino-1-naphthalenesulfonic acid (ANS) dye. Although zinc is essential for T7L amidase activity, CD experiments indicated that zinc does not influence this pH-induced structural transition. Structural analysis of T7L suggested that a network of buried His residues are the source for the observed pH induced structural transition

**Chapter 3** provides a detailed study of structure, stability and dynamic features of T7L native conformation at pH 8, 7 and 6. Temperature dependent far-UV CD measurements indicate thermal melting of secondary structure above 50 °C and formation of insoluble aggregates or precipitates at ~ 70 °C. Urea dependent unfolding experiments based on CD and fluorescence evidenced for differential structural stabilities and unfolding characteristics of pH 6 conformation compared to pH 8 and 7. Backbone resonance assignment of T7L at pH 7 was obtained using a set of conventional triple-resonance NMR experiments. Heterogeneous intensities of the NH cross peaks and lack of several sequential connections in the 3D experiments indicating an inherent dynamic nature of T7L



native state. The structural heterogeneity of the native state is further confirmed by assigning the multiple conformations of T7L segments. The dynamic behavior and the heterogeneity are regulated by the pH, and the low pH state (pH 6) is comparatively more dynamic than the higher pH (8, 7) native state conformations. Dynamics measurements using NMR relaxation experiments provided detailed sequence specific information along the polypeptide chain at various time scales. It is observed that the T7L shows significant chain dependent conformational dynamics in its  $R_1$  and  $R_2$  values. Steady state NOEs show constant values ( $\sim 0.82$ ) over the polypeptide chain, evidencing for a rigid T7L backbone across fast (nano to picosecond) time scale motions. Residues at pH 6 experienced more fluctuations in  $R_1$ ,  $R_2$  and NOE values as compared to pH 8 and 7. Further, hydrogen exchange (HX) experiment reveals that majority of protected residues belong to the structural elements and account for the stability of the native state.

**Chapter 4** delineates the structure-stability-dynamics relationship of T7L low pH partially folded conformations. Detailed biophysical and NMR studies revealed that these PF states differ in their structural and stability characteristics and exhibit differential dynamics. Urea dependent unfolding features of PF state at pH 5 and 4 evidenced for a collapsed conformation at intermediate urea concentrations (i.e. 2 M and 4 M), although pH 3 conformation resulted in uniform melting of structural elements. The pH 3 PF conformation is less stable as compared to pH 5 and 4 conformations as evidenced by free energy ( $\Delta G$ ) calculations. Thermal denaturation experiments delineated that the pH 5 conformation is thermally irreversible in contrast to pH 3 due to formation of insoluble aggregates. Temperature dependent amidase activity measurements show that thermally challenged pH 3 conformation achieve 100 % of its activity when the pH is adjusted back to 7 although the PF conformations do not possess lysis activity. In order to unravel the structural characteristics of T7L low pH PF states, backbone resonance assignments of pH 3 conformation was obtained by using triple resonance NMR experiments. A total of 37 residues were assigned out of  $\sim 42$  observable peaks. Amide proton temperature coefficient analysis indicated that these assigned residues were devoid of H-bonding and located in the unstructured elements. Few residues such as S68, H69 (loop 34) and S148 (C-terminal) are notable as their temperature coefficients are  $< -4$  ppb/K at pH 5 and are  $> -5$  ppb/K at pH 4 and 3 suggesting that these residues might be involved in some local H-bonding interactions at pH 5 that melted upon lowering the pH. Further,  $^{15}\text{N}$  transverse relaxation rate measurements ( $R_2$ ) were used to probe their dynamics and conformational exchange behavior in  $\mu\text{s}$  to  $\text{ms}$  time scale. It was observed that residues having enhanced  $R_2$  values are mainly

concentrated in the loop 34 and C-terminal region. Residue level analysis of the NMR based urea and temperature dependent HSQC spectra established that the structural elements  $\alpha$ 1-helix and  $\beta$ 3- $\beta$ 4 segment of the N-terminal half of T7L is the major source for differential unfolding characteristics of PF states. All these results established that the low pH PF states of T7L are heterogeneous and exhibit differential structural, unfolding, thermal reversibility, and dynamic features.

In summary, the current thesis investigated the structure-function relationship of T7L at residue level. Detailed insights were obtained on the stability and dynamic features of various T7L conformations such as native state and partially folded. Unraveling these biophysical features of a single domain bacteriophage endolysins serve as guiding attributes in assessing the structural and functional constraints of a single domain endolysins that are specifically designed as next generation enzybiotics by engineering them into chimeolysins and artilysins.



### *List of publications*

1. **Sharma M**, Kumar D, Poluri KM. Elucidating the pH-dependent structural transition of T7 bacteriophage endolysin. *Biochemistry*. 2016; 55(33):4614-25.
2. **Sharma M**, Kumar D, Poluri KM. Unraveling the differential structural stability and dynamics features of T7 endolysin partially folded conformations. *BBA-General Subjects*. In Press-Manuscript Accepted, 2018.
3. **Sharma M**, Jaiswal N, Kumar D, Poluri KM. Deciphering the structural heterogeneity and stability-dynamic features of T7 endolysin native conformations (Manuscript under preparation).
4. Gupta P<sup>#</sup>, **Sharma M**<sup>#</sup>, Arora N, Pruthi V, Poluri KM . Chemistry and biology of farnesol: a quorum sensing molecule with immense therapeutic potential. Submitted 2017. #: equal contribution.
5. Arora N, Patel A, **Sharma M**, Mehtani J, Pruthi PA, Pruthi V, Poluri KM. Insights into the enhanced lipid production characteristics of a fresh water microalga under high salinity conditions. *Industrial & Engineering Chemistry Research*. 2017; 56(25):7413-7421.
6. Arora N, Dubey D, **Sharma M**, Patel A, Guleria A, Pruthi PA, Kumar D, Pruthi V, Poluri KM. Elucidating the differential metabolic responses of a green microalga during arsenic (III, V) mitigation. Submitted 2017.

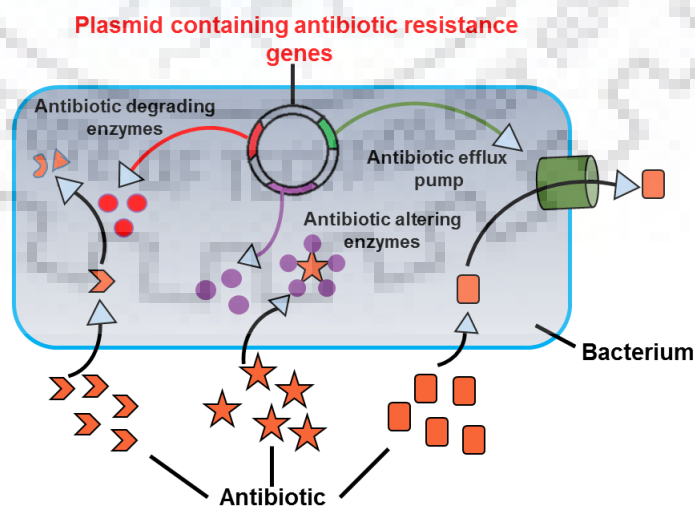
## *List of conference presentations*

### **Poster presentations**

1. **Sharma M** and Poluri KM “Biophysical analysis on structure and stability of T7 bacteriophage zinc amidase”, Indo-US Conference and Workshop on Recent Advance in Structural Biology & Drug Discovery (RASBDD) at Indian Institute of Technology (IIT) Roorkee, Uttarakhand, India, 9-11 October, 2014.
2. **Sharma M** and Poluri KM “Structural and stability features of a novel zinc amidase from T7 bacteriophage”, National Symposium on Biophysics & Golden Jubilee Meeting of the Indian Biophysical Society at Jamia Milia Islamia, New Delhi, India, 14-17 February, 2015.
3. **Sharma M** and Poluri KM “Structure and stability features of T7 endolysin molten globule: a third phase of protein”, the 22<sup>nd</sup> Conference of the National Magnetic Resonance Society (NMRS) at IIT Kharagpur, West Bengal, India, 18-21 February 2016.
4. **Sharma M**, Poluri KM and Kumar D “Biophysical investigations on the ph dependent structural fragility of T7 endolysin”, the 22<sup>nd</sup> Conference of the Asia Pacific NMRS (APNMR) at IISc Bangalore, Karnataka, India, 16-19 February 2017.
5. **Sharma M**, Poluri KM and Kumar D “Deciphering the pH-dependent structure-function relationship of T7 endolysin as a model protein of single domain amidases”, during the Biotech day at Department of Biotechnology, IIT Roorkee, Uttarakhand, India, 4 March 2017.
6. **Sharma M**, Poluri KM and Kumar D “Biophysical Characterization of Partially Folded Conformation of T7 Endolysin”, Annual Symposium of the Indian Biophysical Society (IBS) at IISER, Mohali, Punjab, India, 22-25 March, 2017.

## Introduction

Bacteria are unicellular prokaryotes and most abundant living entities on earth. They have adapted to nearly all ecological habitats due to their diverse metabolic competency. Pathogenic bacteria cause serious illnesses and life threatening diseases by secreting toxins (endo- or exotoxins). Although awareness of proper sanitization, vaccination and antibiotics decrease the bacterial infections, antibiotic-resistant bacterial strains have caused revival in some cases [1]. The bacterial antibiotic resistance emerges through several ways, such as genetic mutation, vertical or horizontal gene transfer, drug's efflux and biofilm formation [2-6] (**Figure 1.1**). Genetic mutations are the important sources that enable the bacteria to release antibiotic inactivating enzymes and remove or alter the cell targets susceptible to antibiotic intrusion. In order to survive against antibiotics attack, bacteria contain specific genes that code for resistance. Due to natural selection mechanism, they survive to reproduce and pass down (vertical gene transfer) their resistant gene to their offspring. Bacteria can transfer (horizontal gene transfer) their genetic material along with resistance encoding genes (resistomes) by means of plasmids, transposons or viruses (bacteriophages) to the susceptible bacteria. These mechanisms are responsible for antibiotic inactivation and target modifications. Multidrug efflux pumps of bacteria allow extrusion of antibiotics. Antibiotics induce and/or regulate the overexpression of efflux pumps, thus leading to broad spectrum resistance to bacteria. Bacteria that lack protective mutation and resistance genes become less susceptible when grown in biofilm forms.



**Figure 1.1:** Schematic representation of bacterial resistance mechanism towards traditional antibiotics.

In 2013, the center for disease control and prevention (CDC) announced that now humans are residing in “post-antibiotic era” [7]. Similarly, the world health organization (WHO) also warned about the development of extreme antibiotic resistance. Some of such common bacterial species that have developed resistance towards conventional antibiotics are summarized in **Table 1.1**. A latest WHO report on September 2016 reveals that “antimicrobial resistance is everywhere and has the potential to affect anyone, of any age, in any country, and antibiotic multidrug resistance (AMR) is a severe threat to global public health and demands serious action across all government sectors and society”.

**Table 1.1:** Some of the most common bacterial pathogens with antibiotic resistance [1,7].

<b>Bacterial species</b>	<b>Antibiotics</b>
<b>Gram-positive bacteria</b>	
<i>Staphylococcus aureus</i>	Methicillin, Linezolid, Daptomycin, Fluoroquinolones, Vancomycin, Tetracycline, Gentamicin, Streptomycin, Penicillin
<i>Enterococcus faecium</i>	Vancomycin, Linezolid, Daptomycin, Fluoroquinolones, Tetracycline, Gentamicin, Streptomycin, Penicillin
<i>Enterococcus faecalis</i>	Vancomycin, Linezolid, Daptomycin, Tetracycline, Gentamicin, Streptomycin, Penicillin
<i>Streptococcus pneumonia</i>	Penicillin
<b>Gram-negative bacteria</b>	
<i>Neisseria gonorrhoeae</i>	Penicillin
<i>Acinetobacter baumannii</i>	Fluoroquinolones, Tetracycline, Gentamicin, Streptomycin, Penicillin
<i>Pseudomonas aeruginosa</i>	Fluoroquinolones, Tetracycline, Gentamicin, Streptomycin, Penicillin
<i>Salmonella typhimurium</i>	Erythromycin, Gentamicin
<i>Klebsiella pneumonia</i>	Cephalosporin, Carbapenems, Gentamicin, Streptomycin, Penicillin
<i>Escherichia coli</i>	Cephalosporin, Fluoroquinolones, Tetracycline, Gentamicin, Streptomycin, Penicillin
<i>Shigella</i>	Tetracyclin

Rapid evolution of antibiotic resistant bacteria, prevalent use of antibiotics, over prescription, and inadequacy of contemporary drug development by pharmaceutical industries allow bacterial diseases to be widespread, despite decades of fighting with antibiotics and vaccines [8-10]. Lack of efficient antimicrobial techniques are also making various medical procedures (viz. organ transplantation, chemotherapy, diabetes management, surgery etc.) highly compromised [11,12].

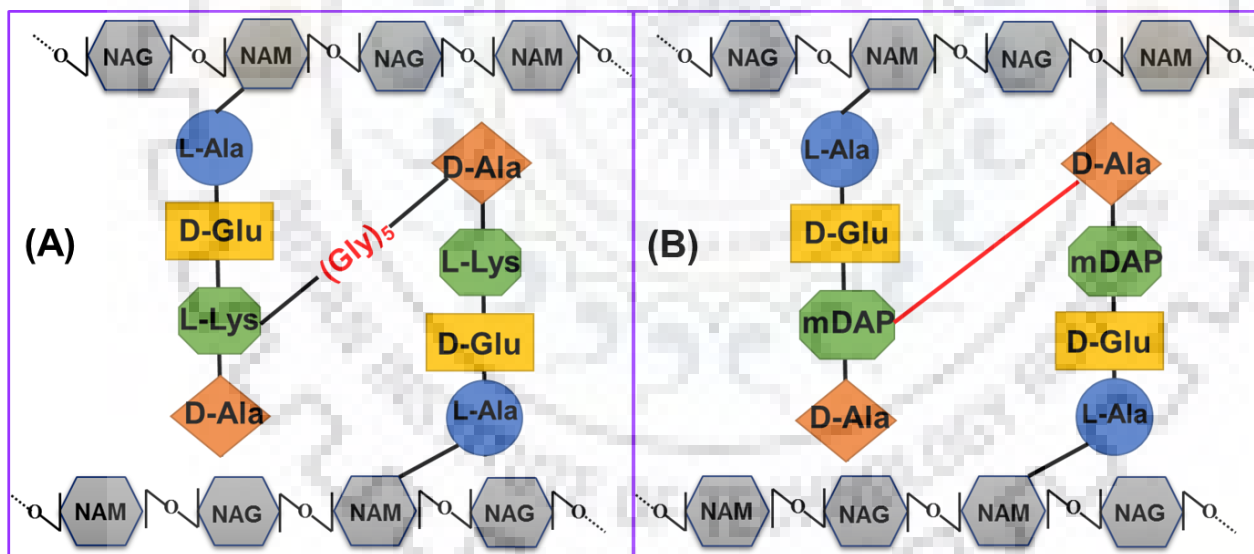
Various approaches have been employed to deal with the issue of antibiotic resistance. Antimicrobial peptides (AMPs) and/or proteins are widely considered to be the alternative of antibiotics [13,14]. One of the possible approaches for antibiotic replacement is use of bacteriophages and/or their peptidoglycan degrading enzymes (enzybiotics) as antimicrobial agents [15]. The present thesis aims at understanding the structure-function relationship of a bacteriophage encoded peptidoglycan degrading enzyme. Chapter 1 deals with introduction and literature survey on bacteriophage and phage encoded peptidoglycan degrading enzymes with a special emphasis on endolysins as potential antimicrobial agents. The later part of the chapter presents the concepts of protein structure and folding, applications of various biophysical techniques that are crucial in decoding the protein structure-function relationships.

### **1.1 Bacterial cell wall: The protecting layer**

Cell wall is the outermost covering of bacteria and responsible for rigidity, shape, replication and protection against mechanical stress or damage from osmotic lysis. Peptidoglycan (PG) or murein is the major constituent of the cell wall. It is composed of glycan strands and stem peptides. The glycan strand consists of alternating sugar moieties N-acetylglucosamine (GlcNAc or NAG) and N-acetylmuramic acid (MurNAc or NAM) connected by  $\beta$  (1-4) glycosidic bonds. These glycan strands are cross-linked to one another by stem peptide or tetrapeptide that is attached to a NAM sugar unit. Stem peptides are made up of four or five amino acids [16]. The cell wall is also decorated with important ligands that provide adherence and receptor binding sites for viruses or antibiotics. Indeed, peptidoglycan is the target of numerous antibacterial antibiotics [17]. For example, penicillin suppresses the activity of transpeptidase and carboxypeptidase enzymes that are involved in peptidoglycan production [18-20]. Lysozyme is able to lyse the peptidoglycan by cleaving the bond between the NAM and NAG [21]. However, peptidoglycan is integral part of all bacterial cell walls. These are divided into two major types: Gram-positive and Gram-negative cell wall based on the cell wall composition [22].

### 1.1.1 Gram-positive cell wall

Gram-positive bacteria have thick cell wall (15-80 nm) and contain several layers of peptidoglycan which make about 50 % of total cell wall dry weight [23]. Cell wall contains unique molecules such as teichoic acid, amino acids and sugars. Teichoic acid molecules are responsible for the overall negative charge to Gram-positive cell wall due to the presence of phosphodiester bonds between teichoic acid monomers [24]. They are imperative for the survival of Gram-positive bacteria by providing chemical and physical protection. There are various types of peptide arrangements reported in Gram-positive bacterial peptidoglycans. One of the extensively studied Gram-positive cell wall is the composition of *Staphylococcus aureus* peptidoglycan (**Figure 1.2 A**) [25]. The stem peptide composed of four amino acids: L-alanine linked to D-glutamic acid, linked to L-lysine, which is finally linked to D-alanine via pentaglycine bridge. Gram-positive bacteria are more susceptible for penicillin mediated degradation than Gram-negative bacteria due to absence of extra protecting layer called outer membrane (OM) [26].



**Figure 1.2:** Peptidoglycan composition of (A) Gram-positive and (B) Gram-negative bacteria.



### **1.1.2 Gram-negative cell wall**

In the Gram-negative bacteria, the cell wall is composed of a single layer of peptidoglycan (PG) protected by an outer membrane (OM) [27]. Lipopolysaccharide (LPS) is the major constituent of OM. Proteins such as the porins, OmpF, and OmpC, serve as anchor between OM and PG [28,29]. These membrane proteins allow the passive diffusion of small molecules such as mono- and disaccharides and amino acids across the OM. LPS serves as a protecting barrier, and can be subdivided into three regions: outer O-antigen or O-polysaccharide, middle core polysaccharide and inner part lipid A [30]. The hydrophobic head of Braun's lipoproteins are covalently attached to PG layer [31,32]. A gel-like matrix, periplasm is present between the OM and the plasma membrane. It is an integral compartment of the Gram-negative cell wall, and accommodates sugars, vitamins, metal ion, and enzymes that are essential for bacterial nourishment [33]. Gram-negative stem peptides are also composed of four amino acids: L-alanine, D-glutamic acid, diaminopimelic acid (DAP) and D-alanine. The D-alanine on one chain is connected to the diaminopimelic acid on adjacent chain via peptide bridge (**Figure 1.2 B**).

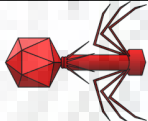



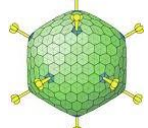
### **1.2 Bacteriophages: Viruses that infect bacteria**

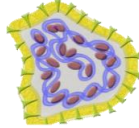
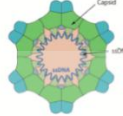
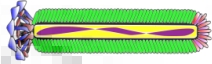
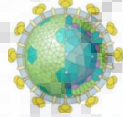
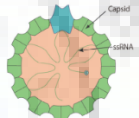
Bacteriophages are obligate intracellular parasites that lack an independent metabolism and require a bacterial host in order to replicate themselves. They make use of some or all the host biosynthetic machinery to propagate and cause the complete lysis of susceptible host bacteria for their maturation. Bacteriophages are also known to infect single-celled prokaryotic organisms of archaea and eubacteria [34,35]. Like bacteria, bacteriophages are ubiquitously present in all natural niches. The total population of phage is estimated to be about  $10^{32}$  phage particles with  $10^8$  species making them most abundant and diverse 'life' system biological entities on the planet earth [36-39]. They can be frequently found everywhere from water, plants, sewage to mosquitos, fish, humans [40-44]. Bacteriophages are the important source of genetic and ecological versatility of bacteria [45]. They can potentially influence the pathogenicity of bacteria by converting a non-pathogenic strain to pathogenic and vice-versa. The relationship of phages and bacteria varies enormously from simple predator-prey to symbiotic relationship promoting the survival and evolutionary success of both. Phages also significantly affect the global carbon cycle by facilitating the bioavailability of organic carbon in the form of lysis products [46,47].

### 1.2.1 Bacteriophage classification

In early studies Felix d'Herelle (1918) proposed only one species of bacteriophage i.e. *Bacteriophagum intestinale* [48,49]. In 1943, Ruska proposed classification of viruses by electron microscopy (EM) and proved the morphological diversity of phages [50]. Holmes in 1948 divided viruses into three families and constituted phages into *Phaginea* family. Further, in 1967, Bradley proposed six basic morphological forms: tailed phages (contractile, non-contractile and short tail), filamentous phages, small isometric phages with ssDNA and ssRNA phages [51]. Bradley established a milestone in the history of bacteriophage classification and his scheme was adopted by international committee on taxonomy of viruses (ICTV), the basis of present classification. ICTV classified bacteriophages according to their morphology and genetic material. Bacteriophages have been classified under the order of Caudovirales with 13 families and 40 genera [52]. The broad classification of phages with detailed shape and structural features are presented in **Table 1.2**.

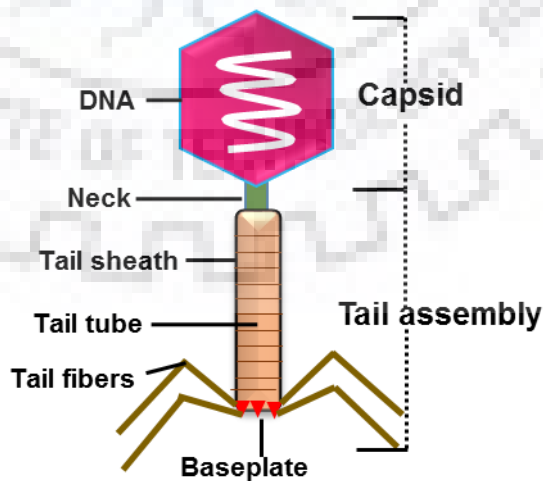
**Table 1.2:** Bacteriophage classification and virion morphotypes.

<b>Double stranded DNA: Non-enveloped</b>			
<b>Family</b>		<b>Morphology</b>	<b>Members</b>
<i>Myoviridae</i>	Icosahedral capsid, helical rigid tail surrounded by contractile sheath and have tail plate, fibers and spikes.		T2, T4, T6
<i>Styloviridae</i>	Icosahedral capsid, long flexible tail, no contractile sheath		KP36, F20, T1, T5
<i>Podoviridae</i>	Icosahedral capsid, short tail with contractile sheath		T3, T7
<b>Double stranded DNA: Enveloped</b>			
<i>Corticoviridae</i>	Icosahedral protein-lipid containing capsid, tailless		<i>Pseudoalteromonas</i> phage MP2
<i>Tectiviridae</i>	Icosahedral capsid-doubly coated, tailless		phage PRD1, AP50, Bam35

<i>Plasmaviridae</i>	Pleomorphic Virion		MV-L1, MV-L2, MV-L3,
<b>Single stranded DNA: Non-enveloped</b>			
<i>Microviridae</i>	Icosahedral capsid, tailless		Φ X 174
<i>Inoviridae</i>	Helical or filamentous shape		MV-L51, M13
<b>Double stranded RNA: Enveloped</b>			
<i>Cystoviridae</i>	Icosahedral capsid, spherical virion		Phage Φ 6
<b>Single stranded RNA: Non-enveloped</b>			
<i>Leviviridae</i>	Icosahedral capsid		Phage MS2 and Q beta

### 1.2.2 Structural organization of bacteriophage

The morphology of bacteriophages has been studied by electron microscopy (EM). Bacteriophages comprise of two main components; the protein and the nucleic acid collectively called nucleoprotein. The size of phages varies from 24 to 200 nm. T4 bacteriophage (**Figure 1.3**) is the largest group with 200 nm long and 80-100 nm wide [53].



**Figure 1.3:** Typical structure of a bacteriophage.

**Head or capsid:** Phages contain head with different size and shapes such as icosahedral, spherical, filamentous etc. The capsid is made up of several subunits called capsomeres, which consists of many protein subunits called promoters. Head or capsid is composed of different proteins. Nucleic acid of phage is coated inside the capsid and, therefore, head provides protection to genetic material from the external environment and nucleic acid degrading enzymes.

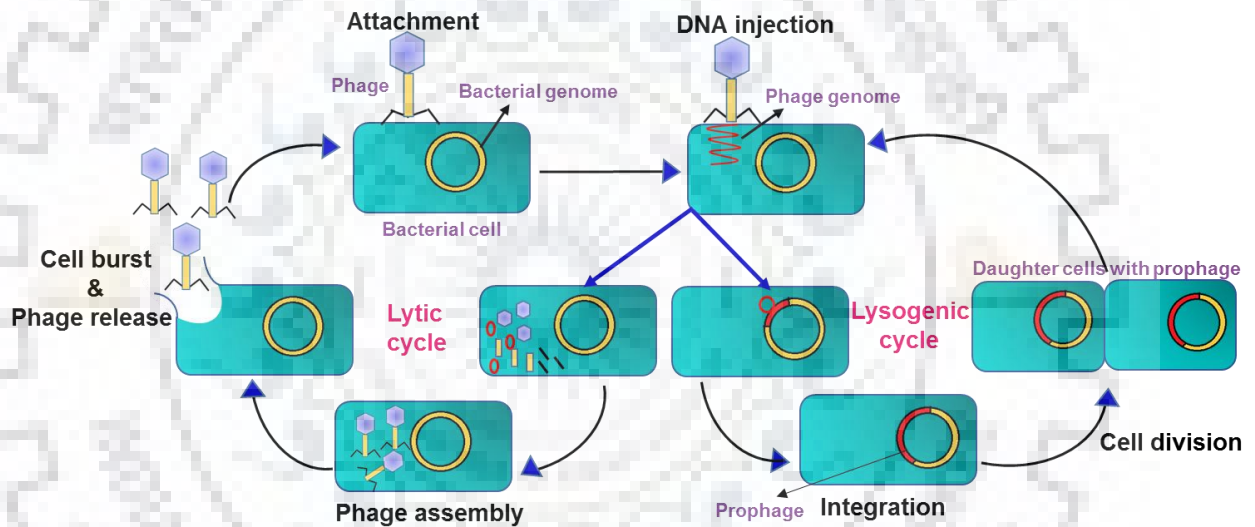
**Tail Assembly:** Many but not all phages have tails attached to the capsid. It is a hollow tube like structure through which the nucleic acid crosses during bacterial infection and present in varying sizes. More complex systems (such as T4 phage) have a contractile sheath outside the tail which contract during the host infection. They also have base plate and tail fibers at the end of tail that facilitate the attachment and binding of virion to the bacterial cell wall. Phages that lack such complex tail assembly have other structures for the attachment and binding mechanisms.

### **1.2.3 Phage multiplication cycles**

Bacteriophages require bacterial host to reproduce. Adsorption or attachment of phage to host is initiated by recognition of bacterial cell surface receptors by phage tail fibers or other equivalent structures [54,55]. Phages irreversibly bind with these receptors, and small change in receptors (due to mutation or evolutionary events) will prevent the adsorption. Once the attachment is done, phages release their genetic material into the cytoplasm of bacteria via formation of holes on cell wall. The empty capsid is called a ghost or doughnut. After injection, phages establish two different types of pathways on the basis of their type: virulent or lytic cycle; and temperate or lysogenic cycle (**Figure 1.4**) [56].

**Virulent or lytic cycle:** Virulent phages reproduce themselves inside bacteria by lytic cycle and cause cell lysis at the end of the cycle to release phage progeny. Example: T-series phages. After penetration event, viral genomes are transcribed by host transcription machinery and thereafter early phage proteins are produced by translation. These early proteins overtake the bacterial cell synthetic machinery and redirect them to synthesize proteins required for viral capsid and genome packaging [56]. Once the virion is completed by capsid proteins assembly, genome packing inside the head and tail joining, it produces late proteins to rupture the bacterial cell wall. These proteins are known as lytic enzyme or endolysins. Lysis or burst open of bacterial cell wall release hundreds of phage progeny. Released phage particles again infect other bacterial host and re-initiate the cycle.

**Temperate or lysogenic cycle:** Temperate or non-virulent phages reproduce themselves inside bacteria by lysogenic cycle and do not cause cell lysis at the end of their cycle. Example:  $\lambda$ -phage. After penetration, genome of temperate phages integrates into the chromosome of bacteria and stay in dormant state, called a prophage. The prophage harboring bacterial genome called a lysogen or lysogenic bacterium. The prophage replicates host genome and keeps on transmitting to daughter bacterial cells for many generations by cell divisions. The relationship of phage and bacterium in lysogenic state is a symbiotic association where prophage blocks its own transcription and protects the cell from lysis as well from infection from other phage by producing repressor proteins. Detachment of prophage from bacterial genome is often associated with taking some neighboring bacterial gene with prophage, which is an important source of horizontal gene transfer [57-59].



**Figure 1.4:** Multiplication cycles of bacteriophage inside the host bacterium.

#### 1.2.4 Phage therapy

As discussed in the above sections, many *Staphylococcus* and *Mycobacterium* species have developed resistance to commonly used antibiotics [60]. A group of bacteria (superbugs) named as ESKAPE (*E. faecalis*, *S. aureus*, *K. pneumoniae*, *A. baumannii*, *P. aeruginosa* and *E. coli*) is one of the serious concern to the health care worldwide [61]. Bacteriophages can be used to overcome such therapeutic problems due to their bactericidal activity and ability to evolve with bacterial host [15]. Unlike antibiotics, isolation of phages is very easy and cost effective as they are usually isolated from natural sources. Due to their strain-specific activity they are very less harmful for natural flora

and being used in very low quantity due to their self-multiplication potency at bacterial infection site. They can be used in liquid or solid form or with hydrogels and deliver through oral, exogenously, parenteral or by inhalation to cure some diseases associated with antibiotic resistance [62,63].

Extensive studies have been done on the phage therapy such as ulcer patients were treated with phage containing biodegradable films [64]. *P. aeruginosa* phage was found to reduce the mice infection when injected intraperitoneally [65], and phages against *Acinetobacter baumannii*, *P. aeruginosa* and *S. aureus* also showed impressive results in mice models [66]. Similarly, oral administration of phage found to prevent diarrhea and sepsis caused by *E. coli* and *P. aeruginosa*, respectively [67,68]. PyoPhage is a commercially available bio-product and effective against infection caused by *E. coli*, *P. aeruginosa*, *S. aureus*, *Proteus* and *Streptococcus* [69].

Although bacteriophage therapy serves as a unique approach for alternatives to antibiotics in the era of bacterial multidrug resistance, there are many regulatory and scientific challenges with phage therapy. The success of phage therapy is hampered by a lack of investment support from large pharmaceutical companies, due to very large costs associated with clinical trials using variety of phages. Now days application of phage therapy is also limited due to: (a) chance of emergence of resistance against phages due to horizontal gene transfer followed with natural selection and bacterial evolution; (b) high bacterial strain specificity of phage therapy require making of different cocktails for treatment of the same infection or disease caused by different strains and; (c) stimulation of immunologic responses due to repeated exposure of same phage evoke the immune system which yield antibodies and reduce the effectiveness towards infection.

### **1.3 Bacteriophage derived peptidoglycan degrading enzymes**

Due to increasing antibiotic and bacteriophage resistance among variety of bacterial strains, peptidoglycan degrading enzymes (PDEs) or peptidoglycan degrading hydrolases (PGHs) have emerged as potential novel source of new antimicrobials. Phage encoded cell wall degrading enzymes lyse the bacterial cell wall to facilitate phage release. PDEs are considered to be effective against many prevalent pathogens such as *Staphylococcus* species, *Streptococcus pneumoniae*, *Enterococcus faecium*, *Bacillus anthracis*, *Clostridium* species, *Klebsiella pneumoniae* and *P. aeruginosa* [70]. There are majorly three types of PDEs: virion-associated peptidoglycan hydrolases

(VAPGHs); endolysins and holins. These enzymes share similar mode of action but have different structures and time of actions during phage lytic cycle. VAPGHs initiate the phage life cycle by penetrating the bacterial cell envelope. Further, in the late stage of the cycle, lysis occurs when the phage endolysin and a holin molecule are produced [70].

### 1.3.1 Virion-associated peptidoglycan hydrolases (VAPGHs) or tail-associated muralytic enzymes

They act in the first step of phage life cycle. VAPGHs locally degrade the peptidoglycan and helps virus in initial cell wall penetration and allow genome injection inside the host bacterium. They generate a small hole which facilitate the phage tail tube across the cell wall to eject its genome. The phenomenon of antimicrobial activity of VAPGHs are known as ‘lysis from without’, which initiate with the adsorption of phages onto bacterial cell surface and disruption of cell envelope before they insert their genetic material inside the bacterium.

**Table 1.3:** List of some virion-associated peptidoglycan hydrolases.

Phage	Host	Protein
φ6	<i>P. syringae</i>	P5
φKMV	<i>P. aeruginosa</i>	Gp36
φK2	<i>P. aeruginosa</i>	Gp181
T4	<i>E. coli</i>	Gp5
T5	<i>E. coli</i>	Pb2
T7	<i>E. coli</i>	Gp16
P22	<i>Salmonella typhimurium</i>	Gp4
P68	<i>S. aureus</i>	P17
φ29	<i>B. subtilis</i>	Gp3
Sp-β	<i>B. subtilis</i>	CwIP

Homology analysis of various sequenced phages or prophage genomes suggested that VAPGHs (**Table 1.3**) belong to bacteriophages infecting both Gram-positive and Gram-negative bacteria [71]. However, not all phages encode for VAPGHs. VAPGHs have modular structural

organization with one or two catalytic domains [72]. Presence of modular organization allows protein engineering (such as domain swapping) to alter as well as to improve their bacteriolytic properties [73].

### **1.3.2 Holins, Pinholins and Spanins**

Holins are phage-encoded hydrophobic membrane proteins that are responsible for the permeabilization of cytoplasmic membrane of bacteria and allow muralytic enzymes (canonical or SAR endolysin) to translocate into the periplasm and degrade the peptidoglycan. Holins represents the family of small membrane proteins that control the timing of lysis [74,75]. Holins belong to  $\alpha$ -helix type class of proteins and creates lesions called as ‘holes’ into the bacterial cell membrane [76]. They don’t have the potential to lyse the bacterial cells, instead they trigger the formation of holes and make plasma membrane permeable for endolysin to attack peptidoglycan layer [74].

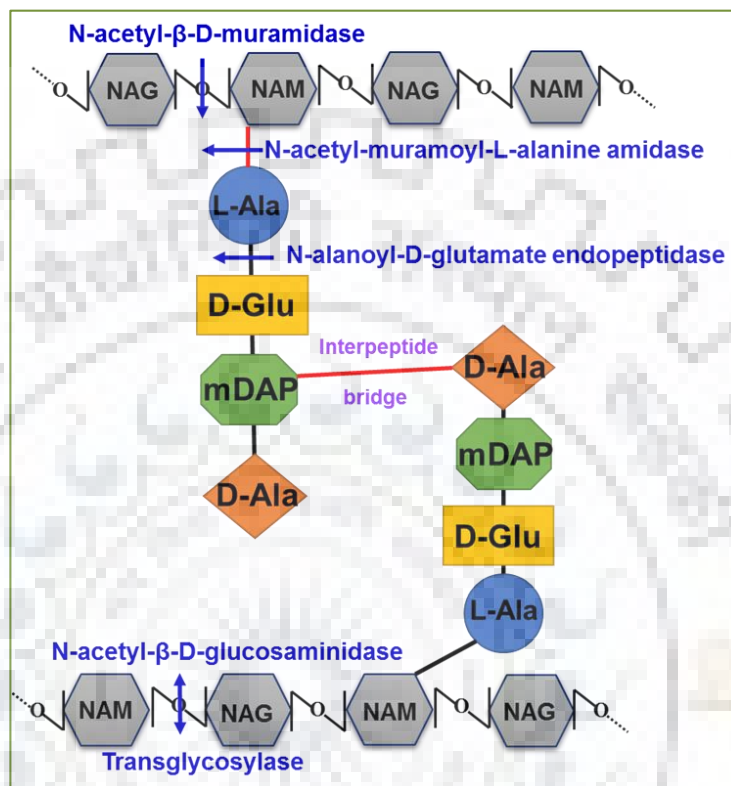
Recently different class of holins called “pinholins” is discovered. They creates small, heptameric channel that leads to depolarize the bacterial membrane [77]. They promote devastation of proton motive force (pmf), enable endolysins to acquire an active form and then invade the PG. An essential but different class of protein, known as spanins causes the disruption of bacterial outer membrane [78,79].

### **1.4 Bacteriophage endolysins**

Endolysins are encoded by dsDNA bacteriophages during the late phase of gene expression in the lytic cycle to degrade the host peptidoglycan [80,81]. First phage encoded lysin was found from *S. aureus* infecting phage [82,83]. The ‘endolysin’ term was proposed to describe a proteinaceous lytic compound that lyses the target bacterial cell wall from inside, once produced during the phage multiplication. Endolysins are also effective in rapid lysis of bacterial cell when applied externally as a purified recombinant protein. Hence, they can serve as alternative of current antibiotics [84]. Application of controlled endolysin as a single protein entity, provides an improvement for health care and regulation over the use of complex self-replicating viruses [85,86]. Unlike phages, endolysins have a broader host range makes them preferable for treatment of various bacterial infection. Phage endolysins exert their uniqueness or specificity on the basis of following three factors: a) enzyme activation by specific components present in host bacterial cell wall, b) specificity in substrate recognition and cell wall binding, c) cleave unique linkages. Endolysins can cleave glycosidic, peptide or amide bonds present in bacterial peptidoglycan. On the basis of



cleavage of specific linkages endolysins can be classified into following categories: (a) N-acetylmuramidase; (b) endo- $\beta$ -N-acetylglucosaminidase; (c) transglycosylase; (d) endopeptidase and (e) N-acetylmuramoyl-L-alanine [80,81,87,88] (**Figure 1.5**).



**Figure 1.5:** Mode of action of different endolysins in Gram-negative peptidoglycan layer.

**N-acetylmuramidase:** It hydrolyzes the  $\beta$ -1,4 glycosidic linkage between the N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM), thereby reduces the glycan chain into disaccharide units. Example: Phage T4 endolysin [89].

**Endo- $\beta$ -N-acetylglucosaminidase:** It hydrolyzes the N,N'-diacetyl chitobiose of asparagine linked sugar moieties of peptidoglycan. Example: *Streptococcus agalactiae* phage B30 [90].

**Transglycosylase:** Their action is similar to N-acetylmuramidase. Transglycosylases cleave the  $\beta$ -1,4 glycosidic linkage between the N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) and form 1,6-anhydromuramic acid [91]. In contrast to muramidase, which use water molecule from solvent to cleave the substrate, transglycosylases do not show hydrolysis activity. Their catalytic residue (generally glutamate) donates a proton to oxygen of glycosidic linkage, which leads to the formation of 1,6-anhydromuramyl residues [80,87]. Example: phage  $\Phi$  KZ endolysin gp144 [92].

**Endopeptidase:** It cleaves the peptide moiety of peptidoglycan such as tetrapeptide linkage [91].

Example: *S. aureus* phage  $\Phi$ H5 endolysin [93].

**N-acetylmuramoyl-L-alanine:** It is a specific amidohydrolase that cleaves a critical amide bond between the NAM and L-alanine of stem peptide. Example: T7 endolysin, *S. aureus* phage  $\Phi$ 11 endolysin etc. [72,80,94].

Majority of endolysins display single type of muralytic activity however, few bifunctional endolysins have also been reported. These endolysins harbour two distinct and independent muralytic activities such as endolysin encoded by *S. aureus* phage  $\Phi$ 11 and *Staphylococcus warneri* M phage  $\Phi$  WMY. They have both amidase and endopeptidase activity [95-97]. Similarly *Streptococcus agalactiae* phage NCTC 11261 possesses endopeptidase and muramidase activity [96,97] and *Streptococcus agalactiae* phage B30 harbours glucosaminidase as well as endopeptidase [98]. Interestingly, T7 endolysin, an amidase, is a unique bifunctional protein with muralytic activity, can also bind to T7 RNA polymerase (RNAP) and inhibits the transcription of T7 related genes during infection of bacterial cells [94].

#### **1.4.1 Structural architecture of endolysins**

Although endolysins share a common biological function of lysing infected bacterial cell wall, they do possess extensive structural and biochemical diversity. The structure of endolysins can be divided into two categories (a) endolysins infecting Gram-positive bacteria and (b) endolysins infecting Gram-negative bacteria. Such a classification reflects the differences between the cell wall structures of these two bacterial groups.

##### **1.4.1.1 Endolysins infecting Gram-positive bacteria**

Endolysins lysing Gram-positive bacteria possess at least two distinct polypeptide modules. One is devoted to substrate recognition termed as cell wall binding domain (CBD) and the other is enzymatically active domain (EAD) or catalytic domain (CD) [80,84,99]. Generally CBDs are located on the N-terminal and EADs are located on the C-terminal, and are connected with by a short linker region (**Figure 1.6**). This linker region sometime contains amino acid cleavage site for auto-proteolysis of CBDs. However, these orientations and proximities of CBDs and EADs are not universally conserved as both the domains can be present on either terminals as well in the middle of the endolysins, and in few case CBDs can be absent. In some cases, endolysins have been found

to composed of more than one (or some time two) different domain for catalytic activity or substrate binding. For examples, *Bacillus* phage endolysin plyG displays one EAD, two clearly separate CBDs and a spore binding domain (SBD) [100]. The sequence comparison of various endolysins shows that the EADs are more conserved, whereas the CBDs are more variable [101].



**Figure 1.6:** Modular structure of Gram-negative bacteria infective endolysin.

Other important domain of the Gram-positive bacterial endolysins are cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain has been identified in nearly 87 phage endolysins. The CHAP domain contains cysteine and histidine residues at active site [102,103]. It is strictly encoded by Gram-positive infecting phages and predominantly found in *Staphylococcal* and *Streptococcal* endolysins. The CHAP domain associated with amidase and/or endopeptidase activity and often found with SH3 domains [104,105]. It can act as D-alanyl-L-alanyl enopeptidase in *Streptococcus* phage B30, D-alanyl-glycyl enopeptidase in *Staphylococcus* phage  $\Phi$ 11 or as an N-acetylmuramoyl-L-alanine amidase in *Streptococcus* phage C1 [101]. The association of CHAP with other amidase and enopeptidase families refers to molecular mechanism of cooperative cleavage of multiple PG bonds rather than one specific site. In some *Streptococcal* endolysins, the catalytic domains need activation from their less-active (E-form) to full-active form (C-form). Such conversion takes place by means of binding of endolysin with cell wall of bacteria, which leads to irreversible conformational changes in catalytic region [106-108].

#### 1.4.1.2 Endolysins infecting Gram-negative bacteria

In contrast to endolysins lysing Gram-positive background, phage endolysins lysing Gram-negative bacteria (endolysins of *E. coli* phage T4 and T7) possess single-domain structure of an EAD or catalytic domain [94,109]. They are globular proteins and usually shorter in length with 15-20 kDa in weight. The Gram-negative bacterial PG is located between the OM and periplasmic space and devoid of any surface protein and carbohydrates. This sort of architecture prevents the collateral damage by limiting the access to the PG from outside and that might be the reason of having single-domain by Gram-negative endolysins. These endolysins have efficient catalytic function (aid

multiple lytic reaction at the time of cell lysis), in contrast to their Gram-positive counterparts, which are proposed to bind with cell wall substrates and exhibit slower off-rate [110]. Some phage endolysins have been reported with multi EADs. For example *Prochlorococcus marinus* phage PSS2 and *Synechococcus* phage endolysins consist N-terminal glycoside hydrolase (GH) activity (MURO or GLUCO) and peptidase activity (PET-C39-2, YKUD, or PETM15-3) [101,111]. EADs associated with different mode of actions have been summarized in **Table 1.4**.

**Table 1.4:** Summary of phage endolysin enzymatic active domains.

Class	Sub-class	Endolysin
Glycosidase	N-acetylglucosaminidase	Lambda Sa2
	N-acetylmuramidase (lysozyme)	T4, P1, Cp-1, PlyBa04, Bcp1
	Lytic Transglycosylase	ΦKZgp144 and Lambda endolysins
Amidase	N-acetylmuramoyl-L-alanine amidase	T7, LytA, LysK, Lyt2638A, CS74L, BPS13
Endopeptidase	D-alanyl-glycyl	LytA
	L-alanoyl-D-glutamate	Ply118, Ply500, LysB4

However, few Gram-negative endolysins are reported to have modular domain structure. Endolysin KZ144 (*P. aeruginosa* phage ΦKZ) and EL188 (*P. aeruginosa* phage EL) are composed of a N-terminal cell wall binding domain and a C-terminal enzymatic active domain [92,112]. Whereas, another *P. aeruginosa* Lu11 endolysin Lu11gp113 has an N-terminal enzymatic active domain and a C-terminal cell wall binding domain [113]. Recently, three more endolysins with modular structures namely OBP gp279, 201Φ2-1gp229 and PVP-SE1gp146 were analyzed from Gram-negative infecting *Myoviridae* phages [114]. The possible reason for such modular structure in Gram-negative endolysins could be explained by the early phage evolution phenomenon. It was suggested that different domain-coding genes or modules were exchanged between phages during

era of early phage evolution and further rearrangement of gene fragments gave birth to new proteins with diverse combinations [115,116].

#### **1.4.2 Bacterial resistance to endolysin**

According to data available to date, no resistance has been reported against endolysins, suggesting imperishable lytic activity [117-121]. Several approaches (such as exposure of endolysins in sub-lethal concentrations or mutagenesis) had been employed. But researchers have failed to promote resistance in sensitive bacterial host, although these approaches easily cause resistance against antibiotics. Endolysins target ligands that are essential for survival of various bacterial species, and these ligands cannot be modified for the sake of bacterial viability. For example, choline of *S. pneumoniae*, polyrhamnose of *S. pyogenes* and polysaccharide of *B. anthracis* [120,122].

The bacteriophages coevolved with their host in such a way that binding domain of their lytic enzymes has evolved to target unique and immutable bonds in the cell wall. This type of coevolution resulted in binding and cleaving of highly conserved ligands in cell wall by endolysin, and thereby, believed to be the reason for ultimate survival and release of phages without on shoot of endolysin resistance [88]. The extracellular (on PG) localization of endolysin targets makes resistance a very rare event. Furthermore, several strategies have been attempted to create resistance strains to native or recombinant phages [88]. In order to achieve resistance *S. pneumoniae* cell were repeatedly challenged with lethal dose of Pal endolysin [117]. Similarly *B. anthracis* cells were subjected with methane-sulphonic acid ethyl ester to produce mutation in order to survive against PlyG endolysin [120]. However, no strains with resistance phenotype were recovered, although these strategies yielded thousand fold increased antibiotic resistance for novobiocin and streptomycin. Similar results were obtained in attempt to create methicillin resistant *Staphylococcus aureus* (MRSA) strain against chimeric endolysin ClyS [121].

Despite these promising outcomes, there are reports for bacterial resistance to other peptidoglycan hydrolases (PGHs). For example, bacterial resistance against human lysozyme has been accomplished by variety of secondary modification in cell wall, like N-deacetylation or O-acetylation of PG and alanylation of teichoic acids [25,123]. Resistance to another PGH from bacterial origin such as lysostaphin has been identified [124]. The resistance could be attributed to the modification within pentaglycine bridge (lysostaphin target), such as deletion of one glycine residue or incorporation of a serine residue [125,126].

### **1.4.3 Endolysin engineering**

These enzyme-based antimicrobials efficiently lyse Gram-positive bacterial cell. However, impermeable LPS containing OM is the major obstacle for exogenous application of endolysins on Gram-negative bacteria. Therefore, engineering approaches are required for further improvement and development of recombinant endolysins to make them efficient OM invaders, and more applicable in complex environments such as food matrix, mucous membrane or blood [127]. Various protein engineering strategies like domain shuffling or random mutagenesis or other translocation modifications have been utilized on endolysins to improve their functionality. These techniques can efficiently alter the binding and lytic properties of endolysins.

#### **1.4.3.1 Domain shuffling/swapping/recombination: rational design**

The modular architecture of bacteriophage endolysins and relative independence of their EADs and CBDs endow them to recombine different domains from different origins and make it plausible to create functional chimeras. Domain swapping is therefore a robust technique for rational design and optimization of novel lytic enzymes with enhanced catalytic activity and broad host range [110,128]. The domain shuffling exists naturally [129]. It also indicates that horizontal gene transfer occurs naturally within phage population or between phage and bacterial host by recombination driven exchange of endolysin domains or modules [130]. For example, the naturally occurring chimeras are *Listeria* phage endolysin PlyPSA [131], and pneumococcal phage Dp-1 endolysin Pal [132]. Numerous studies have been reported in literature on endolysin engineering involving fusion of catalytic or binding domain to construct novel antimicrobials [129,133-135].

The earliest approach of domain swapping was initiated by Diaz and coworkers in early 1990s. They exchanged the enzymatic domains between pneumococcal autolysin LytA and the phage lysin Cpl-1, and reported the catalytic activity switched in resulting chimeras [136]. They further swap the cell wall binding specificity of LytA and Cpl-7 endolysins by exchanging choline binding domain of LytA with choline independent binding domain of Cpl-7 and *vice-versa* [137]. Croux and coworkers constructed a chimera by fusing EAD of *Clostridium acetobutylicum* autolysin Lyc and the choline binding domain of Cpl-1. The resultant construct was active against choline-containing pneumococcal cell walls but devoid of *Clostridial* cell wall binding [138]. Recently, two head-to-tail fusion proteins were created by combining lysostaphin, *Staphylococcal* bacteriocin with endolysin LysK [139] and PGH HydH5 of phage vB\_SauS-ΦIPL188 [140]. LysK possess CHAP

and amidase domains and HydH5 has a CHAP and lysozyme domain. The resulting fusion protein shows activity against three unique bonds in *Staphylococcus* PG. Synergistic effect of LysK and lysostaphin proved to be destructive for different *S. aureus* strains, including MRSA [141]. All these fusion-constructs show stronger lytic activity, than separate enzyme.

Domain shuffling approach was also exploited for increasing cell wall binding affinity among various endolysins [142]. Combination of lysostaphin C-terminal SH3 (Src homology 3) domain with full-length HydH5 enzyme yields 1.7 fold increased catalytic activity [140]. Protein engineering studies have also been utilized to improve the thermostability and solubility of endolysins. Improved engineered thermostable protein was created by replacing CBD of *Clostridium perfringens* lysin with thermophilic phage lysin [143]. The poor solubility of *Staphylococcal* phage endolysin was improved by fusing its CBD with highly soluble EAD of *Enterococcal* phage, resultant protein not only showed improved solubility but also had broad lytic activity against range of *Staphylococcal*, *Streptococcal* and *Enterococci* strains [144]. Chimeric endolysin (P16-17), fusion of N-terminal D-alanyl-glycyl endopeptidase domain of *S. aureus* phage P16 endolysin and C-terminal domain of *S. aureus* phage P17 minor coat protein, shows increased solubility over the parental hydrolases [135].

#### **1.4.3.2 Mutagenesis approaches**

Mutagenesis approach has been utilized to improve the function of endolysins. Poor solubility of *Staphylococcal* phage endolysin was improved by fusing its CBD with highly soluble EAD of *Enterococcal* phage mutation into parental nucleotide sequence [144]. The prime target of mutagenesis is to replace specific amino acid residues in order to alter the conformation and function of endolysins. Random mutagenesis approach was used for making endolysin with enhanced activity under high salt, extreme pH, or in food products. Gene coding for PlyGBS endolysin from *S. agalactiae* phage was subjected with DNA mutagenesis using error-prone PCR (polymerase chain reaction) approach or *E. coli* mutator strain. After repeated rounds of mutagenesis, a mutant endolysin with 28-fold increased activity was identified on the basis of screening of lysis zone formation [145]. Another mutant endolysin PlyGBS90-1 was found to eliminate the streptococcal colonization in vaginal tract of mice, thereby can be used as an antibiotic prophylactic to reduce neonatal infections [96,145]. The influence of deleting the CBD on the lytic activity of endolysin was studied on various phages and was associated to variable effects. Several site directed mutational

studies were also performed to examine the residues involved and their effect on lytic activity. Alteration of conserved amino acids in CHAP and amidase domains of B30 endolysin resulted in sequential loss of activity from both the domains [98]. From site-directed mutational studies, it was concluded that histidine residues are essential for M23 endopeptidase activity in *Staphylococcal* glycyL-glycine PGH ALE-1 [146].

### **1.4.3.3 Endolysin translocation**

It involves the active translocation of endolysins across the bacterial cells or eukaryotic cells by the fusion of signal peptide to the protein. This approach increases the invasion and retention of endolysins into specific cell systems. One of the proposed strategies involves the attachment of cell-penetrating peptides (CPPs) or protein transduction domains (PTDs) or Trojan horse peptides to the PGHs [147-150]. PTDs are eminently composed of basic residues or amphipathic region of proteins that facilitate their easy translocation across the cytoplasmic membrane of target cell. For penetration of eukaryotic cells, Schwarz and Dowdy (2000) made a suggestion on the use of protein transduction domain (PTD) of cell-penetrating peptides [151]. For example TAT (Trans-activator of transcription) is a cell-penetrating peptide of HIV virus contains a PTD with a nucleus localization signal (GRKKR) [152]. Fusion of such domains with lytic enzymes (such as lysostaphin) facilitates their entry into eukaryotic cell and also helps in efficient targeting of intracellular *S. aureus* infections [147]. A patent has been filed for TAT containing lysostaphin (Lyso-TAT) (<http://www.pat2pdf.org/patents/pat20110027249.pdf>), which eliminates *S. aureus* ex-vivo from infecting MAC-T mammary epithelial cells, bovine brain epithelia, human keratinocytes, and murine osteoblasts cultured cells. It has been reported that fusion of GFPs (green fluorescent proteins) with TAT-PAD construct remarkably enhance its uptake by mammalian cells [153]. Gaeng and coworkers described that fusing the *Lactobacillus brevis* S-layer signal peptide to *L. monocytogenes* phage lysin A511, which enable the active translocation of protein from *L. lactis* host cells to the surrounding environment [154]. A similar approach was used for translocation of *C. perfringens* lysin CP25L. This lysin was found to be capable of lysing *C. perfringens* cell but did not harm other members of gastrointestinal tract microflora, suggesting that endolysins have the potential to control specific pathogens residing in the gut.



#### 1.4.4 Artilysins

Artilysins are engineered endolysins or enzybiotics that are developed to attain increased antibacterial activity against Gram-negative bacteria. The ‘Artilysin’ term was coined by Briers and coworkers, to describe the group of engineered endolysins that can efficiently penetrate the OM of range of Gram-negative bacteria, like *E. coli*, *P. aeruginosa*, *S. enterica* etc. [155]. They are able to cross the OM to reach that are PG with the help of LPS-destabilizing peptides or membrane penetrating peptides (MPPs) fused to either N- or C- terminal without affecting the structure of endolysins (**Figure 1.7**) [156].



**Figure 1.7:** Schematic representation of artilysin structure.

LPS-destabilizing peptides contain amphipathic or polycationic regions that act as a wedge to rupture the LPS layer using stabilizing ionic or hydrophobic forces, and enables passage of endolysin through OM [4,157]. LPS-destabilizing peptides include; polycationic peptides (PCNPs), hydrophobic pentapeptide (HPP), parasin (Pa1) and lycotoxin [155]. They have potent ability of ‘lysis from without’ when applied exogenously, and cause rapid lysis of Gram-negative bacteria. In order to construct artilysin (Art-175), sheep myeloid antimicrobial peptide, a 29 amino acids long  $\alpha$ -helical cathelicidin (SMAP-29), was covalently fused with *P. aeruginosa* phage KZ144 endolysin [158]. The natural form of *P. aeruginosa* phage KZ144 endolysin could not kill *P. aeruginosa* cells but fusion with SMAP-29 allows it to lyse target cells from outside [112]. Although it is well known that such membrane penetrating peptides are highly cytotoxic for mammalian cells, however, fusion of MPPs with endolysin abolish their cytotoxicity [159,160]. The synergy of artilysins with OM permeabilizers, such as EDTA (ethylenediaminetetraacetic acid) can cause complete elimination of bacterial culture. For instance, LoGT-022 artilysin is lethal against *P. aeruginosa*, *A. baumannii*, *E. coli* and *S. typhimurium* and causes reduction of bacterial cultures within 30 minutes in the presence of 0.5 M EDTA [155].

### **1.4.5 Endolysin applications**

Endolysins are non-immunogenic and do not cause any non-desirable side effects or allergic reactions. Owing to unique substrate specificity and remarkable activity, endolysins have been applied to food science, biotechnology and medicine.

#### **1.4.5.1 Food preservation**

Endolysins offer a unique possibility for biological control of pathogenic bacteria in food and feed industry. Endolysins can be used as food preservatives. Zhang and coworkers described that *Listeria monocytogenes* phage lysin LysZ5 can be successfully used in control of *Listeria monocytogenes* in soya milk even at refrigeration temperature [161]. Similarly *Staphylococcal* lysin LysH5 can rapidly kill *S. aureus* present in milk [93]. It was found that in synergy with bacteriocin lysin, LysH5 could cause complete elimination of *S. aureus* within 4 hours [162]. In addition, the chimera of B30 and Ply700 endolysins been reported to be effective against *Streptococci* as well as *Staphylococci* in milk and its products [163]. *Clostridium butyricum* phage  $\Phi$ CTP1 lysin can also be used as milk preservative. Thermostable endolysins such as *Listeria* phage endolysins Ply118, Ply511 and Ply35 that can retain lytic activity even after heating at 90 °C for 30 minutes, serve as potential source of antimicrobial in food products that experienced high temperature treatment as well as also control bacterial growth in iceberg lettuce [128].

#### **1.4.5.2 Therapeutics**

Purified endolysins serve as excellent therapeutic agents for the treatment of pathogenic bacteria involved in variety of human and animal infections, either alone or in combination with antibiotics. Fishetti and coworkers reported first *in vivo* experiment and demonstrated the potential of *Streptococcal* lysin PlyC (C1 lysin) to prevent *Streptococcus pyogenes* colonization in upper respiratory tract in mouse model [164]. Nelson and coworkers used single dose of Pal endolysin for significant reduction of nasopharyngeal colonization of mice by *S. pneumoniae* within 5 hours [117], and PlyGBS for elimination of group B *Streptococci* in vaginal and oropharyngeal colonization [96]. Recently Cpl-1 was used for treatment of *S. pneumoniae* induced endocarditis in rats, resulting significant decrease of aortic vegetation and CFUs (colony-forming units) in the bloodstream [165]. In another *in vivo* study PlyG endolysin was found to be effective against *B. anthracis* infection demonstrated by 70-80 % survival of PlyG treated mice [120]. Furthermore, MV-L endolysin from

phage  $\Phi$ MR11 and chimeric lysin ClyS were reported against MRSA [118,166]. A chimeric endolysin Cly-S found to be affective for bacterial decolonization from infected skin of mice [118].

Endolysin applications in animal or human have some safety issues. However, first GLP (good laboratory practice) compliant toxicology and safety study was conducted on endolysin trial in rats by Jun and coworkers revealed no adverse effect or toxicity [167]. These encouraging results further strengthen the approach of using endolysins as potential therapeutics in the control of animal and human bacterial infections.

#### **1.4.5.3 Biofilm elimination and disinfection**

Biofilms are assembly of sessile surface-associated bacterial colonies that composed of extensive extracellular matrix. Biofilm associated bacteria establish independent relationships via active communication and local regulation of gene expression. They are attached to wide variety of surfaces, such as living tissues, mechanical or prosthetic devices, industrial or portable water piping systems or natural or artificial aquatic systems. Biofilms show high level of antibiotic resistance than their planktonic (freely suspended) counterparts [168-173]. Bacteriophage endolysins are potential source for eradicating biofilms. Endolysins treatment of biofilms destroys the matrix and eliminates the bacterial community.  $\Phi$ 11 endolysin has been successfully utilized for *Staphylococcal* biofilms elimination [174]. LysK and SAL-2 endolysins were also reported to destabilize *S. aureus* biofilms by rapid lysis of sessile cells embedded in extracellular matrix [175,176]. LysH5 was also suggested to eradicate *S. aureus* and *S. epidermidis* biofilms [177]. Furthermore, Cpl-1 in synergy with lytA eliminates the biofilm build by *S. pneumoniae*, *Streptococcus pseudopneumoniae* and *Streptococcus oralis*, similarly amidase PlyLM with protease K lyse *L. monocytogenes* monolayers [178,179]. Due to high biodegradable nature, endolysins represent a better source of decontamination over the chemical disinfectants against MRSA, *Listeria monocytogenes*, and *B. anthracis*.

#### **1.4.5.4 Agriculture**

Transgenic plants have been made by incorporating endolysin genes to gain protection against various phytopathogenic bacteria [180-182]. The T4 endolysin (muramidase) containing potato shows resistance against *Erwinia carotovora* mediated damages [180]. It is accumulated in plant tissue and is released from hair roots to hydrolyze the *Erwinia carotovora* cell wall when *Erwinia's* pectinases start destroying the plant cell wall. The effect of T4 lysozyme to other

surrounding soil bacteria was found to be negligible. Further, by using this concept, plants may also be used as bioreactor for large scale production of antimicrobial proteins. Oey and coworkers have utilized this strategy and produced *S. pneumoniae* phage endolysins Cpl-1 and Pal and the group B *Streptococcal* lysin PlyGBS in the chloroplast of tobacco plants [183].

#### **1.4.5.5 Endolysin CBD mediated detection and quantification**

The endolysin CBDs are responsible for specific activity of endolysins due to their ability binding specificity towards host cell wall. Modular architecture of endolysins make their CBD's an excellent candidates for detection, differentiation and quantification of host-bacterial cells [110,184]. CBDs can be chemically cross-linked to fluorescent dye or fused with GFPs for spectrophotometric or microscopic detection and quantification of bacterial pathogens. Eight different *Listeria* CBDs fused with fluorescent marker GFP were coupled with different cell walls in order to develop a tool for detection and differentiation of various *Listeria* strains in mixed bacterial population [110]. CBDs can also be immobilized on host cell onto solid surface using magnetic beads. For example, coating of paramagnetic beads with CBDs of *Listeria* phage endolysins enable immobilization and recovery (>90%) of *L. monocytogenes* strains from contaminated food samples [185,186]. CBD-based detection method was also utilized for the *B. anthracis*. PlyG expresses a ten amino acid synthetic peptide at C-terminal which allows stain specific detection of *B. anthracis* when combined with fluorescent quantum dot nanocrystals (Qdot<sup>®</sup>-Invitrogen) [187].

### **1.5 Protein structure and folding**

Proteins are bio-macromolecules containing four levels of structural organization. The primary structure is the sequence of amino acids themselves connected with peptide bonds. The DNA encodes the amino acid sequence of protein, and are synthesized by series of transcription and translation events followed by post-translation modifications as needed to fulfill respective biological functions. The order of amino acid sequence and their subsequent interactions define the secondary and tertiary structure of protein. The secondary structure of protein is made of  $\alpha$ -helices and  $\beta$ -sheets and is dependent on hydrogen bonding pattern. The tertiary structure of protein is the overall three-dimensional shape which takes the secondary structural elements and allows them to fold with maximum stability and minimum energy. The hydrophobic core of protein is surrounded by the external hydrophilic layer under physiological conditions. The bonding interactions between

the amino acid side chains stabilize the overall global structure and create channels and binding sites for protein interactions. Many proteins made up of multiple polypeptide chains called as subunits. These subunits interact via hydrogen bonding, salt- and disulfide-bridges to give quaternary structure, the final level of protein structural organization.

In order to execute specific functions endolysins must fold into unique three-dimensional structure. Understanding the mechanism of protein folding for these class of proteins provides a comprehensive picture of their structure-stability and functional relationship, which helps in engineering endolysins with enhanced and robust lytic activity.

### **1.5.1 Protein folding**

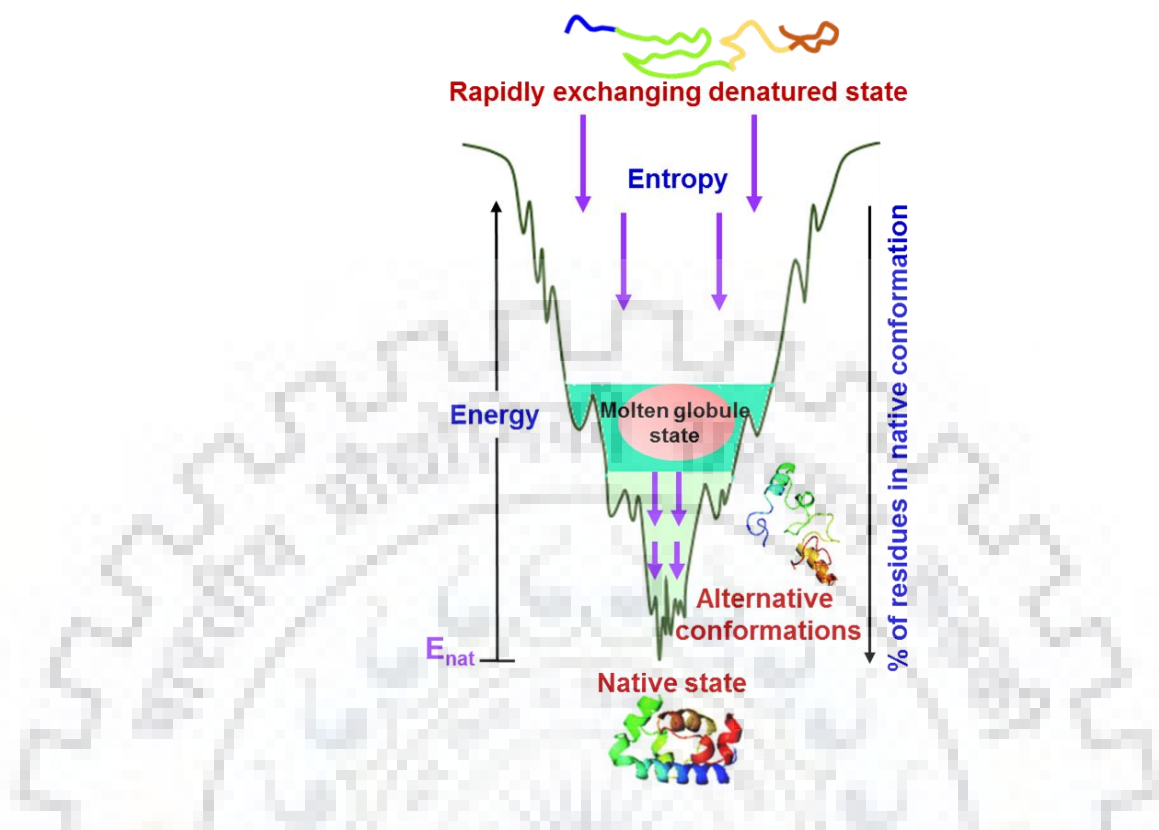
Protein folding is a physical process that helps protein to acquire a characteristic and functionally active three dimensional structure known as 3D-fold. When a protein is translated from mRNA to a linear chain of amino acids, it exists in an extended/unfolded or random coil conformation. The amino acids interact with each other and produce a well-defined 3D structure, the folded protein, known as native protein. The Anfinsen's dogma postulates that three dimensional structures of small globular proteins are determined only by the amino acid sequence [188]. The dogma is also known as 'thermodynamic hypothesis' which states that environmental conditions (such as solvent, pH, ionic strength, temperature, metal ions or prosthetic groups, etc.) in which protein folding occurs also affect the three dimensional structure of native protein [189]. Native states are thermodynamically more stable and kinetically accessible with minimum free energy under various physiological conditions [190,191]. Designing or engineering a new protein requires full understanding of protein folding mechanism to generate a stable fold with kinetic accessibility but without compromising the stability.

Levinthal paradox suggests that due to very large number of degree of freedom in an unfolded polypeptide chain, the protein has enormous number of conformational possibilities [191,192]. It would require a longer time to sequentially scan all these possible conformations in order to fold correctly. Levinthal hypothesis explains that, instead of trying each available conformation, protein folding occurs through a defined route from unfolded conformation to native state via approaching intermediates [193]. There are three models of protein folding; (1) framework model [194], (2) Hydrophobic collapse model [195,196] and, (3) nucleation condensation model

[197]. More recently protein folding funnel is emerged as most widely accepted model of protein folding pathway [198-201].

### **1.5.2 Folding energy landscape or folding funnel**

Numerous proteins fold through multiple downhill routes by approaching consecutive assembly of basic folding units, called as ‘foldons’ [202]. The energetics of protein folding mechanism can be best described by free-energy landscape model. It represents the thermodynamic and kinetic aspects of transition of unfolding ensemble to a native state. The energy landscape has a funnel topology, since the conformational space accessible to polypeptide chain decreases as the native state is reached. The depth of funnel describes the free energy of a conformational state and width represents the entropy (**Figure 1.8**). The wide end corresponds to heterogeneous unfolded state and the narrow end reflects the homogenous native state [203-205]. The surface of energy funnel consists of all conformational states accessible to the polypeptide chain such as unfolded, partially folded intermediate states, low energy excited state and completely folded native state [203,206-209]. These states are heterogeneous and highly dynamic in nature. Protein folding process takes place via a series of conformational changes by acquiring multiple routes (or intermediate conformations) on energy landscape [210,211]. Free energy landscapes of various proteins have been extensively characterized by using NMR spectroscopy [205,212-214]. Hydrogen exchange NMR spectroscopy serve as a most widely accepted tool for the study of thermodynamic and kinetic properties of protein folding intermediates [215,216].



**Figure 1.8:** Schematic representation of energy landscape of protein folding.

### 1.5.3 Protein misfolding and aggregation

The correct three dimensional architecture of protein is essential to perform biological functions. When the protein fails to achieve correct native conformation, it becomes functionally quiescent, which could be sometimes toxic and responsible for many neurodegenerative disorders [217,218]. The general pattern that emerges in all neurodegenerative diseases is an abnormal protein to aggregate in the form of amyloid fibrils [219]. Many neurodegenerative diseases are associated with increased level of metal ions and studies have been reported that metal ions (like Cu, Zn, Al etc.) act as inducer for protein aggregation [220,221]. Various biophysical and NMR techniques have been employed for the detailed studies of protein misfolding/aggregation [222-225]. The aggregation of proteins favoured by intermolecular interaction thereby, considered as an alternative pathway of protein folding. Accumulation of these misfolded amyloid fibrils causes amyloid diseases as summarized in **Table 1.5** [226]. Furthermore, some unstructured proteins or domains remain functional active under various physiological conditions, called as intrinsically disordered proteins (IDPs).

**Table 1.5:** List of diseases associated with protein aggregation.

Disease	Aggregation prone protein
Huntington's disease	Huntington
Kuru	Amyloid
Sickle cell anemia	Hemoglobin
Cystic fibrosis	Cystic fibrosis transmembrane regulator
Macular degeneration	$\beta$ -amyloid, Crystallin
Alzheimer's disease	$\beta$ -amyloid peptide, tau
Parkinson's disease	Alpha-synuclein
Polyglutamine disease	Atrophin-1, ataxins
Cancer	p53
Islet amyloid	Insulin
Cataracts	Crystallin
Scrapie Creutzfeldt-Jakob disease	Prion
Retinitis pigmentosa	Rhodopsin

#### 1.5.4 Intrinsically disordered proteins (IDPs)

Although, it is well known that a proper folded 3D structure is essential to determine the biological activity of protein, but many studies proved the existence of functionally active unstructured proteins or domains under various physiological conditions [227,228]. This raises the importance that formation of partially folded/intrinsically disordered proteins cannot be necessary related to loss of functionality. Various proteins have been observed with thermodynamically stable conformation with complete activity even in partially folded state. The structural characterization of such partially folded states elucidates that they can perform their biological function even in structural heterogeneity of the catalytic site, as rest of the conformation acts as a scaffold. Many proteins are found to be active even in their non-native partially folded state such as acylphosphatase from *Sulfolobus solfataricus* [229], VH domain of anti-human ferritin monoclonal antibody F11 [230], nuclear coactivator binding domain (NCBD) of CREB binding protein (CBP) [231]. The presence of intrinsic disorder has been reported predominantly in various classes of proteins such as regulatory, cell signaling, and cancer-associated proteins.



### **1.5.5 Factors affecting protein stability**

Environmental factors strictly affect the performance of protein by regulating its structural stability. Several environmental factors such as pH, temperature, and chemical denaturants significantly regulate the folding energy landscape of the protein thus influencing its conformation.

**pH:** It plays a crucial role in influencing the protein stability by altering the electrostatic interactions between the charged amino acids [232,233]. Alteration in pH sometimes also result in denatured state or partially unfolded state as change in pH alters the state of ionization of acidic or basic amino acids, which leads to change in 3D structure of protein [234-237].

**Chemical denaturants:** Hydrophobic and H-bonding forces are predominant interactions that govern the globular structure of protein in aqueous solution [238]. Urea and guanidine are chaotropic agents that interfere with intermolecular interactions mediated by non-covalent forces such as hydrogen bonds, hydrophobic and Van der Waals interactions, thus facilitating the protein denaturation [239].

**Temperature:** Changing the temperature perturbs the native structure of protein thus affects its stability to a sufficient extent [240-243]. Temperature mainly affects two components of folded protein: entropy and stabilizing interactions. Stabilizing interactions such as non-polar hydrophobic interactions, hydrogen and ionic bonds are weaker in nature. Therefore, increasing the temperature of protein disturb these interactions by increasing kinetic energy, which causes the molecules to vibrate more rapidly and also increases the dynamics [244]. Increase in temperature results in unfolded state of protein having high entropy (more disordered) and low stabilizing interactions.

## **1.6 Biophysical techniques used to study protein structure and stability**

### **1.6.1 Circular dichroism (CD) spectroscopy**

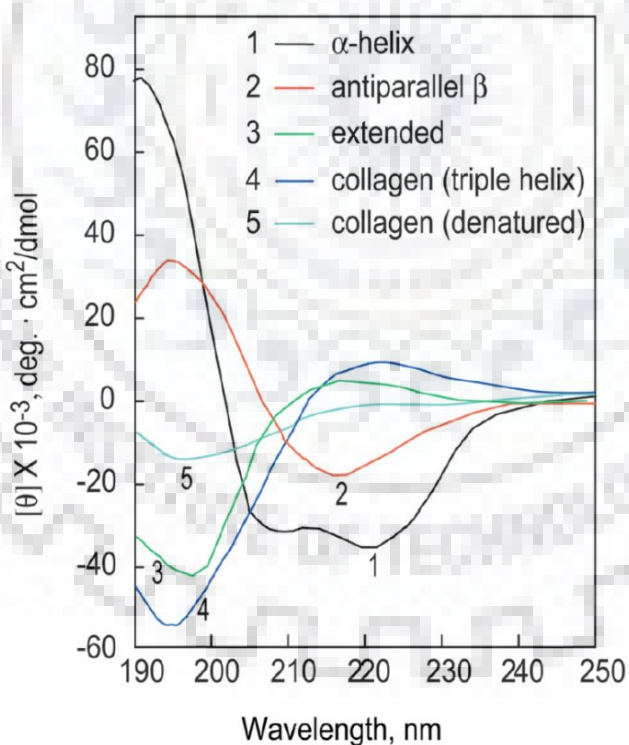
Circular dichroism spectroscopy is a type of absorption spectroscopy, which measures the difference between right- and left-handed circularly polarized light. It is a very sensitive technique for the measurement of conformational features of biomolecules such as nucleic acids, peptides, proteins etc., at very less concentration. It is immensely useful to elucidate structural characteristics of unknown biomolecules; monitor folding-unfolding transitions using temperature, pH, chemical denaturants; ligand/metal binding or mutation induced conformational changes [245,246]. The phenomenon of circular dichroism is displayed exclusively by chiral molecules. CD measures the

difference in the absorbance of left- and right-handed circularly polarized components as a function of wavelength ( $CD = A_L - A_R$ ) in terms of ellipticity ( $\theta$ ), which is calculated using differential absorption [247].

$$\theta = \frac{2.303(A_L - A_R)180}{4\pi} = 33.0 (A_L - A_R) \quad (1.1)$$

### Circular dichroism of proteins

Except glycine, all amino acids are chiral due to asymmetric conjugation of their  $\alpha$ -carbon atoms. In proteins, the peptide bonds are optically active. The secondary structure of protein can be determined by far-UV CD spectroscopy over the range of 190-250 nm. Far-UV CD measurements are due to the excitation of electronic transitions ( $n-\pi^*$  and  $\pi-\pi^*$ ) of amide groups in a polypeptide backbone. Proteins are made up of different secondary structural elements with different  $\Phi$ - $\Psi$  angles and H-bonding patterns; each of them show characteristic CD spectra as illustrated in **Figure 1.9** [247,248].

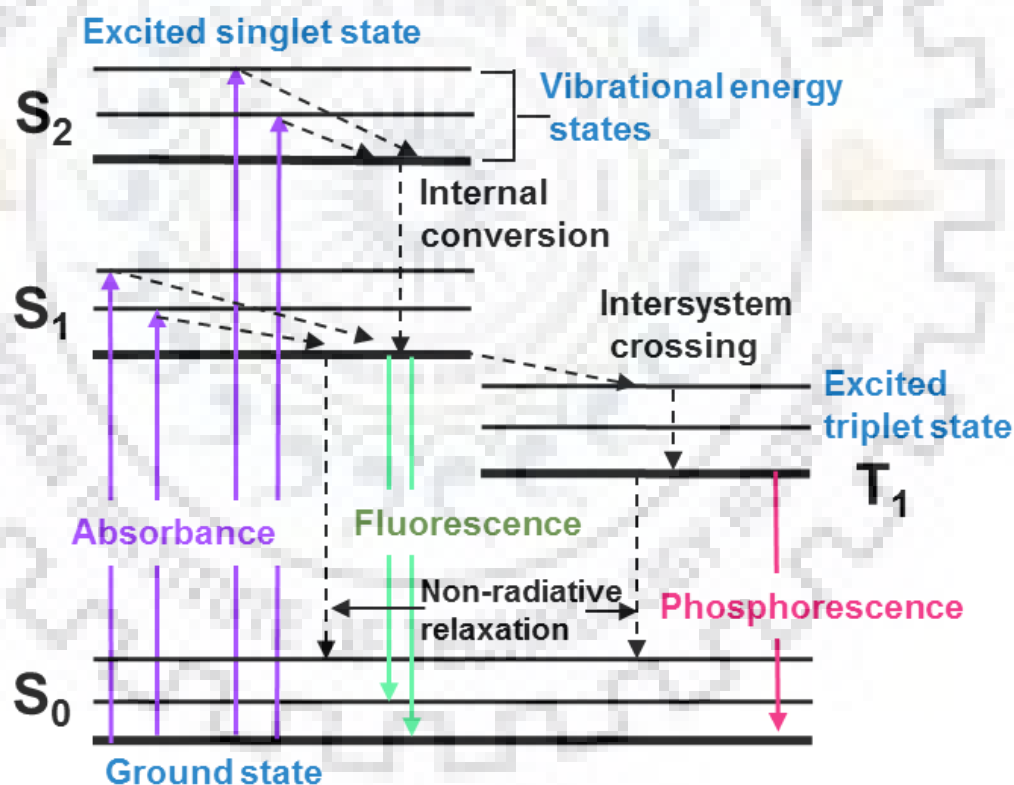


**Figure 1.9:** Characteristic far-UV CD spectra of  $\alpha$ -helix,  $\beta$ -sheet and random coil. Figure adopted from Greenfield 2006 [251]. Reprinted by permission from Nature publishing group, Copyright clearance center (nature protocols, copyright © 2006) [251].

The tertiary structure of the protein can be analyzed by using near-UV CD spectroscopy. The aromatic chromophores such as Trp, Phe and Tyr are responsible for dichroic signals due to dipole orientation and  $\pi\text{-}\pi^*$  transition band in near-UV region between 250 to 300 nm [249]. Trp shows band around at 290 nm, Tyr shows band between 275-282 nm and Phe has band between 255-270 nm. Cysteine (or disulfide bonds), and prosthetic groups also exhibit CD signals in near-UV range. Near-UV CD spectroscopic technique also used for the investigation of ‘molten globule’ state of protein which is characterized by reduced signals due to high mobility of aromatic side chains [250].

### 1.6.2 Fluorescence spectroscopy

Fluorescence is the process of emission of photons in which fluorescent capable molecules, known as fluorophores emit photons while returning from electronically excited states to the ground state as described by Jablonski diagram (**Figure 1.10**).



**Figure 1.10:** Jablonski diagram illustrating relaxation of excited state molecules via different processes.

Conformational transitions, protein folding/unfolding studies can be done by using two types of fluorophores (extrinsic and intrinsic). Intrinsic fluorophores occur naturally and involve aromatic

amino acids (Tryptophan, Tyrosine, and Phenylalanine), NADH (nicotinamide adenine dinucleotide), flavins etc. Extrinsic fluorophores such as fluorescent dyes (e.g. ANS) are applied externally to the molecules. The wavelength and intensity of photons emitted by the fluorophores are measured by fluorescence spectroscopy, and the obtained fluorescence signals provide information about structure and environment of the fluorophores/biomolecules.

For example, Trp and Tyr present in hydrophobic environment show maximum fluorescence intensity, but in contrast, the fluorescence intensity decrease in a hydrophilic environment due to the quenching of fluorescence emission by solvent molecules. Trp serves as a dominant intrinsic fluorophore and henceforth mainly used to measure the fluorescence intensity of protein upon their structural transition [252]. 8-Anilinonaphthalene-1-sulfonic acid (ANS), is an organic compound containing both a sulfonic acid and an amine group and strongly binds to cationic groups of proteins and polyamino acids through ion pair formation. Its fluorescence alters as it bind to the hydrophobic regions of proteins. It is widely used for the detection and characterization of intermediate conformations of proteins such as partially folded or molten globules due to their extensive binding with ANS molecules [253-255].

### **1.6.3 Fluorescence lifetime spectroscopy (FLS)**

It is used to measure the time spend by a fluorophore in the excited state before returning to the ground state by emission of photons, which range from picoseconds to nanoseconds. Fluorescence lifetime (FLT) measurements are based on time correlated single-photon counting (TCSPC), which is devoid of fluctuations in intensity of excitation source. FLT is an intrinsic property of fluorophore and independent of concentration, photo bleaching, sample absorption and sample thickness, but affected by temperature, quenching and polarity. It is described by the time measured for exponential decay of number of excited molecules to  $N/e$  (36.8 %) of initial population through dissipating energy in the form of fluorescence.  $N/e$  defines the number of excited molecules (N) to decay to  $1/e$  of original population. The fluorescence intensity decay can be measured by exponential function as a function of time (equation 3):

$$I(t) = I_0 \times e^{(-t/\tau)} \quad (1.2)$$

Where,  $I(t)$ =fluorescence intensity at time  $t$ ,  $I_0$ =intensity right after excitation and  $\tau$ =fluorescence lifetime.

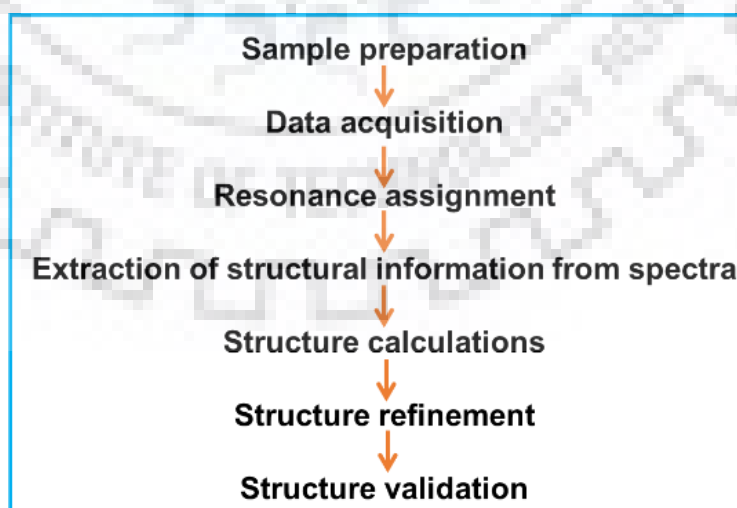
### 1.6.4 Nuclear magnetic resonance (NMR) spectroscopy

NMR spectroscopy is a powerful technique to study the structure, interactions and dynamics of various synthetic or biological molecules. NMR phenomenon relies on the nuclear spin and absorption of electromagnetic radiation from the radiofrequency region by the nuclei of atoms. NMR chemical shift depends on local neighborhood of nuclei and strongly affected by surrounding electronic configuration. Such a phenomenon allows us to distinguish different nuclei in a given NMR spectrum that appears at specific resonance frequencies.

NMR is a very sensitive technique and used in variety of industries such as food industry, agriculture, pharmaceuticals, polymer industry, research and development of various chemical compounds, drug discovery and development etc., [256]. MRI (magnetic resonance imaging) is an application of NMR and used as an analytical tool in biomedical field. NMR is also extensively used in metabolomics research. Quantitative NMR spectroscopy serves as easy and sensitive technique for the analysis and measurement of small molecules or metabolites present in crude form or in mixture [257]. Bimolecular NMR deals with the investigations on structure-dynamics-folding-stability characteristics of biomolecules such as proteins, DNA, RNA, carbohydrates and their complexes among themselves or with other synthetic/natural molecules [258-262].

#### 1.6.4.1 Protein structure determination

The process of NMR based structure determination of protein involves the following steps (Figure 1.11):



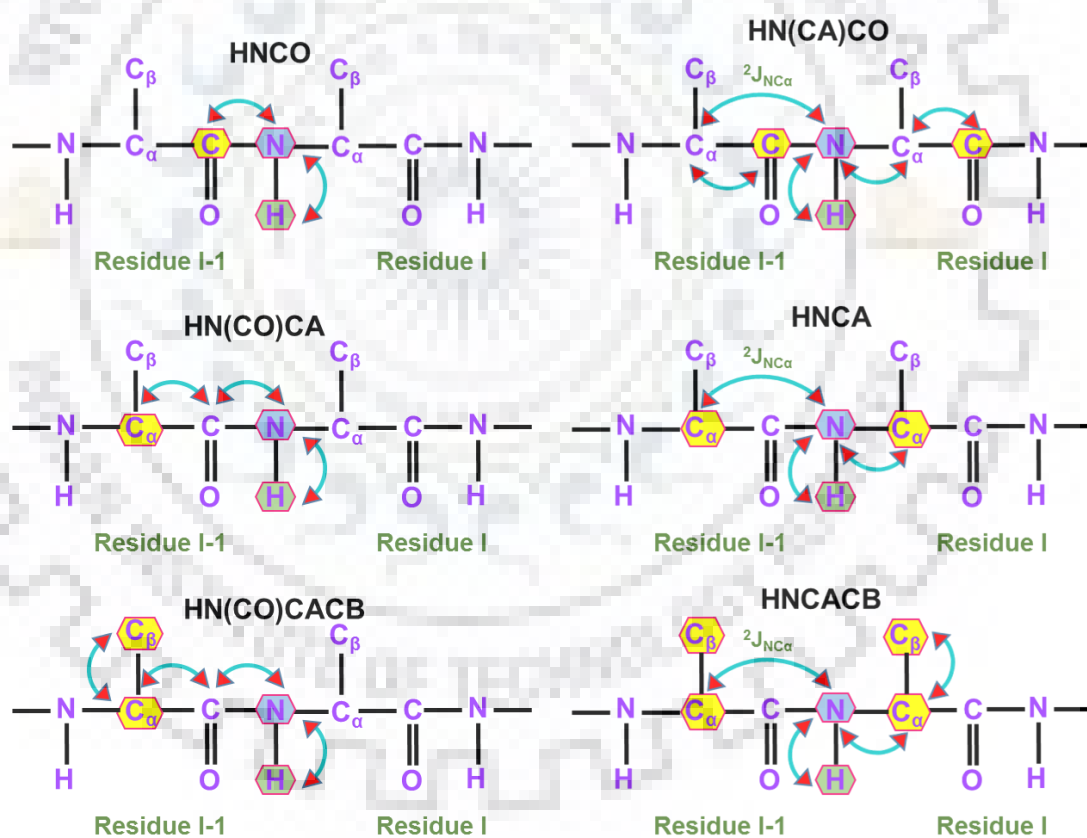
**Figure 1.11:** Typical scheme of protein structure determination using NMR spectroscopy.

**Sample preparation:** The conventional approach of NMR sample preparation is based upon the use of recombinant prokaryotic/eukaryotic expression systems. To achieve successful structure determination of protein systems it is necessary to label them with specific NMR active isotopes. In general, for bulk production of isotope labeled proteins,  $^{15}\text{N}$  labeled ammonium chloride ( $\text{NH}_4\text{Cl}$ ) and  $^{13}\text{C}$  labeled glucose are used as nitrogen and carbon sources respectively. Some other isotope labeling techniques used in protein NMR are:

- (a) **Deuterium labeling:** It is also called as negative labelling because  $^2\text{H}$  vanishes the  $^1\text{H}$  signals without disturbing  $^1\text{H}$  NMR spectra. It results in simplified  $^1\text{H}$  1D-NMR spectra in contrast to fully protonated one [263]. This strategy is very useful for the study of high molecular weight proteins [264].  $^2\text{H}$  labeling is successfully employed in 2D/3D-NMR experiments to obtain remarkably resolved spectra with intense and sharp NMR signals. Amino acid type  $^{13}\text{C}/^{15}\text{N}$  selective labelling: It involves the selective incorporation of  $^{13}\text{C}$  or  $^{15}\text{N}$  or  $^{13}\text{C}$ - $^{15}\text{N}$  labeled amino acids by growing the bacteria in minimal medium containing labelled and unlabeled amino acids. This strategy helps to identify the resonances of buried regions of NMR spectra and ease the sequential assignment of proteins [265]. Deuterium labelling also reduces relaxation by proton and increases resolution.
- (b) **Reverse labelling of selective amino acids:** In this method bacterium is grown in the medium supplemented with unlabeled amino acid against a label background ( $^{13}\text{C}/^{15}\text{N}$ ) [266,267].
- (c) **Segmental labelling:** This technique involves the labelling of either N or C-terminal of amino acids and is useful in labeling of proteins with multiple domains. It can be achieved by three methods: Native chemical ligation (NCL), Expressed protein ligation (EPL) and Protein trans-splicing (PTS) [268].
- (d) **Cell free labelling:** Protein is expressed in vitro by adding DNA or mRNA to a mixture containing transcription, translation machinery, standard amino acids (labelled and unlabeled), nucleotide triphosphates (NTPs), salt, buffer etc. It is widely used for labelling of membrane proteins [269].
- (e) **Other methods:** Stereo array isotope labelling (SAIL), IVL (Ile-Val-Leu)-side chain methyl group labelling, 10 % glucose labelling, stereospecific  $^2\text{H}$  labeling [270-273].

### Resonance assignments of proteins

Assignment of NMR resonances of protein can be achieved by sequential assignment, which involves assignment of backbone and side chain atoms by connecting individual spin systems with their sequential neighbor. The sequential assignment can be done by using triple resonance assignment using set of 3D-NMR experiments. Protein backbone contains a series of scalar coupled nuclei, therefore can be assigned by recording 3D-heteronuclear correlation spectra. The experiments that are routinely used for backbone assignment are illustrated in **Figure 1.12**. These experiments allow us to assign  $H\alpha$ ,  $H\beta$ ,  $C\alpha$ ,  $C\beta$ ,  $HN$  and  $CO$  resonance of amino acids. Chemical shift values of  $C\alpha$  and  $C\beta$  of any spin system are helpful in initial recognition of types of residue and help the identification of stretches of sequentially linked residues by comparing them with the primary sequence of the protein.



**Figure 1.12:** Schematic representation of 3D-NMR experiments used in protein backbone resonance assignments. The arrows are showing the path of magnetization transfer and observable nuclei are indicated in colored circles.

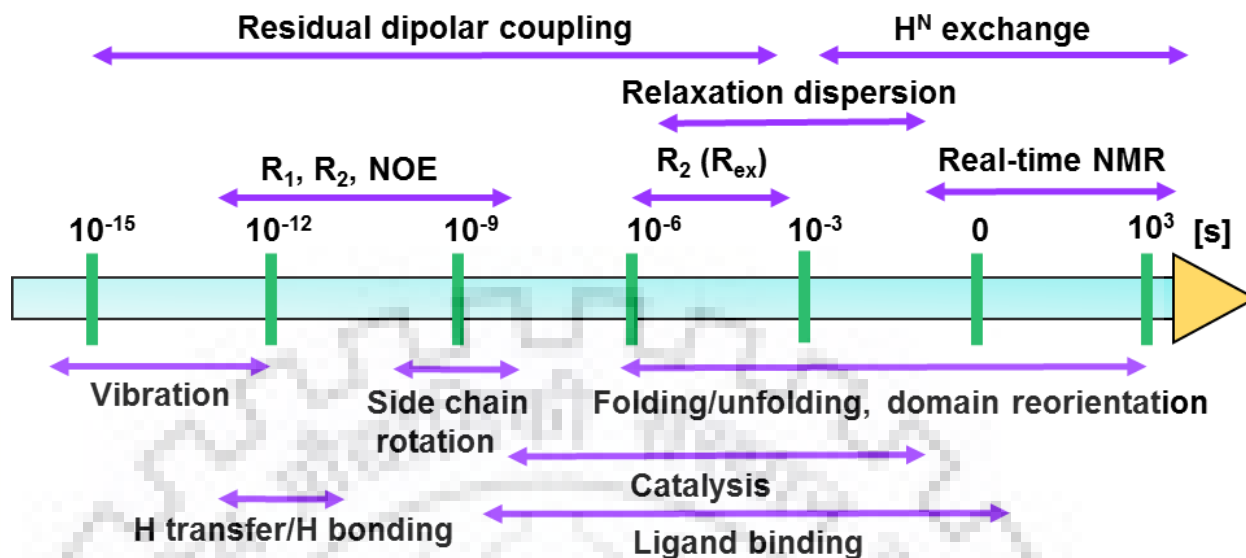
Once the backbone assignments are achieved, a set of 3D-experiments such as HCCH-TOCSY (total correlation spectroscopy) or  $^{15}\text{N}$ -TOCSY-HSQC, HCCH-COSY (correlation spectroscopy), H(CCO)NH and CC(CO)NH etc., are used to assign the side chain resonances. By using the resonance assignments of all the nuclear spins, one can measure the dipolar distances/orientations using experiments such as NOESY (nuclear Overhauser spectroscopy), RDC (residual dipolar coupling), PRE (paramagnetic relaxation enhancement) etc. These experiments provide distance restraints and thus aid us in calculating the high resolution 3D structures of biomolecules/proteins/protein complexes using NMR spectroscopy [274].

#### **1.6.4.2 Protein dynamic studies using NMR spectroscopy**

Protein dynamics is the phenomenon of time-dependent fluctuations in protein conformation. It affects variety of protein functions, thermostability, folding pathways (including misfolding and aggregation), binding interactions etc., [275-277]. NMR is a useful technique to study the protein dynamics at different time scales over the range of slow to fast motions (**Figure 1.13**).

In NMR spectroscopy, relaxation refers to the restoration of thermal equilibrium by nuclei or spins after being perturbed by radiofrequency (RF) pulse. There are three types of relaxation parameters used in NMR experiments to characterize the protein motions; (1) Longitudinal or spin-lattice relaxation ( $T_1$ ), (2) Transverse or spin-spin relaxation ( $T_2$ ), and steady state nuclear Overhauser effect (het NOE). The  $R_1$  rate (spin-lattice relaxation rate) describes the relaxation of excited state longitudinal components of spin vector to the Boltzmann equilibrium by transferring energy to the lattice (molecular framework). The  $R_2$  rate (spin-spin relaxation rate) represents the relaxation of transverse components of spin vector from coherent spin population to incoherent (i.e. dephasing) by transferring energy to nearby spins, without changing spin energy levels. Conformational exchange due to chemical exchange of nucleus in millisecond-microsecond (ms- $\mu$ s) timescale and magnetic field inhomogeneity contributes in fluctuating  $R_2$  values. Heteronuclear steady state  $^1\text{H}$ - $^{15}\text{N}$  NOE (het NOE) involves the cross relaxation of  $^{15}\text{N}$  nucleus due to attached amide  $^1\text{H}$  nucleus. Its magnitude depends upon nanosecond-picosecond (ns-ps) timescale internal motions. In addition to above relaxation experiments, more quantitative information of protein dynamics can be obtained by using model free analysis and by performing Carr-Purcell-Meiboom-Gill (CT-CPMG)  $R_2$  relaxation dispersion experiments.





**Figure 1.13:** NMR timescale of protein motions and associated experiments.

#### 1.6.4.3 Hydrogen exchange (HX)

The backbone amide hydrogen atoms have the tendency to exchange with solvent hydrogen atoms due to conformational fluctuations. The solvent exposed amide protons are more prone to exchange faster than buried or strong hydrogen bonded protons. Local conformational fluctuations or global unfolding reactions disturb H-bonds and thus responsible for hydrogen exchange (HX) process. HX experiments can be performed by adding  $D_2O$  in protein solution or incubating  $^2H$  labelled protein in water ( $H_2O$ ). The exchange rate of backbone amide protons, which reflects the local environment of each residue in protein 3D fold can be used to determine the stability constants, rate of protein folding transition, disordered regions of folded protein, interacting surfaces and solvent accessibility [278-281]. HX measurements are very sensitive to characterize the short lived folding intermediates such as partially folded conformations [215].

#### 1.6.4.4 Translational diffusion measurements

Internal thermal energy driven translational motions or Brownian motions or diffusion depends upon molecular size, shape, temperature and viscosity. Diffusion constant can be calculated by using Stokes-Einstein equation. NMR spectroscopy is a sensitive technique to measure diffusion constants. Diffusion ordered spectroscopy (DOSY) using Pulse field gradient NMR (PFG-NMR) experiments  $[(90^\circ-\tau(\text{PFG})-180^\circ-\tau(\text{PFG})-\text{acq})]$  are used to measure diffusion constants, which is used to study translation diffusion and hydrodynamic radius of protein [282,283]. It is a very sensitive

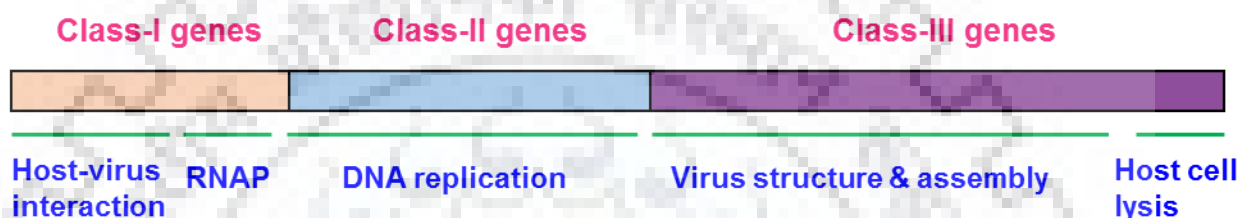
technique to determine protein size, aggregation, oligomerization, degree of ligand binding, protein folding, protein hydration etc., [284-288]. Due to their ability to distinguish the different population in a mixture, the technique named as 'NMR chromatography' [289]. In a typical DOSY experiment with Fourier transformation on one dimension and second dimension comprises of diffusion coefficient. The diffusion constant indicates the hydrodynamic properties of proteins. Aggregation or oligomerization cause extensive change in hydrodynamic radii and molecular size.

#### **1.6.4.5 Temperature dependence of amide proton chemical shifts**

NMR chemical shifts are indicative of surrounding environment of nuclei and molecular conformations of protein. Temperature dependent fluctuations can change the amide proton chemical shifts [290]. Amide proton chemical shifts depend on bond magnetic anisotropy, length of inter- or intramolecular H-bonds, ring current and electrostatic interactions. Amide protons show linear dependency toward temperature change over a small range (about below 15 °C) of their denaturation temperature. Temperature coefficient is a measure of H-bonding strength. It ranges from -16 to +2 ppb/K where, slowly exchanged H-bonded amides show more positive value than -4.5 ppb/K than rapidly exchanged and weakly H-bonded to water [291,292]. Random-coil like proteins show temperature coefficient in the range of -6 to -10 ppb/K [293]. In some cases amide protons chemical shifts show non-linear (curved) dependency toward temperature change due to presence of alternative conformations which affected by local structural and dynamic fluctuations [294,295]. The alternative conformations are also known as low-energy excited states or near-native states are within 2-3 kcal/mol of energy from the native states. These states depict the ruggedness/shallowness of the protein's native state conformation. The curvatures can be either concave or convex, which are depicted by the chemical shift position of the NH resonance, and the nature of structural element [294,295]. Non-linear temperature dependence of chemical shifts along with the native state hydrogen exchange (HX) experiments provide a detailed energy landscape analysis of the protein folding pathway.

## 1.7 Thesis outline and objectives

The main objective of this thesis is to characterize the pH dependent structure-stability-function relationship of T7 bacteriophage endolysin (T7 endolysin / T7 lysozyme / T7L) at molecular level. The natural source of T7 endolysin is the T7 bacteriophage, which belongs to *Podoviridae*. It is a non-enveloped, head-tail architecture with an icosahedral (T=7 symmetry) head (60 nm diameter). It consists of 40 kbp long linear dsDNA (**Figure 1.14**) which encodes 55 types of structural and catalytic proteins. *E. coli* is the most susceptible host for T7 Phage [296].



**Figure 1.14:** Genomic organization of T7 phage.

During the lytic cycle of T7 phage multiplication T7 RNA polymerase (RNAP) directs the transcription of class-I and class-II genes. T7 endolysin is encoded by class-II gp3.5 gene thus also known as Gp3.5 protein. It controls the phage transcription and genome replication by performing two different types of functions. When the synthesis of T7 endolysin rises above certain level, T7L bind to N-terminal and palm-finger domains of T7 RNA polymerase (T7-RNAP) and inhibits the transcription [297,298]. The endolysin binding site is located opposite to the active site cleft of RNAP with which it interacts DNA, NTPs and ssRNA [299]. All these studies suggested that T7 endolysin is a natural inhibitor of T7-RNAP initiation process.

T7 endolysin degrades the peptidoglycan due to its amidase activity with the aid of holin and spanin [300]. Its structure was first reported by Cheng and co-workers in 1994 [94]. T7 endolysin based expression systems reduce the leaky transcription of toxic bacterial genes by blocking basal expression by T7 RNAP [301,302]. Therefore, T7L is extensively used for large scale overexpression of desired recombinant genes in bacterial and eukaryotic cells containing T7-RNAP [303-305].

For the last decade, extensive studies are going on for the development of endolysins as a potent alternative of antibiotic therapy. A major challenge associated with engineering of endolysins to improve their properties is the lack of molecular information on their structure-function

relationship. Diversified enzymatic activity is associated with their structure-stability features and folding-unfolding transitions. Detailed biophysical characterization makes it possible to engineer novel endolysins with increased antimicrobial and therapeutic potential. Thus, considering the importance of structural characterization, the thesis has been designed to explore the pH dependent structural-stability-dynamics-function characteristics of native and intermediate states of T7 endolysin.

**1.7.1 Thesis objectives:** The specific objectives of the thesis are as follows

- Mechanistic insights into pH-dependent structure-function relationship of T7 endolysin.
- Conformational heterogeneity, dynamics and stability features of T7 endolysin native conformations.
- Deciphering the differential structural stability and dynamics of partially folded conformations of T7 endolysin.

**1.7.2 Scope of the thesis**

**Chapter 2** revealed the effect of pH on stability and activity of T7 endolysin using various biophysical and biochemical techniques. It also unfolded the possibility of existence of partially folded conformations of T7L at low pH. **Chapter 3** unravels the residue-level heterogeneity, conformational dynamics of T7L native state using protein NMR experiments. It also presents a comparative study of stability features associated with chemical and thermal denaturation. Finally, **Chapter 4** deciphered the molecular level characterization of differential structural stability features, diffusion characteristics of low pH partially folded conformations.

## 1.8 References

1. Sengupta S, Chattopadhyay MK, Grossart H-P. The multifaceted roles of antibiotics and antibiotic resistance in nature. *Frontiers in Microbiology*, 4 (2013).
2. Stewart PS, Costerton JW. Antibiotic resistance of bacteria in biofilms. *The Lancet*, 358(9276), 135-138 (2001).
3. Poole K. Mechanisms of bacterial biocide and antibiotic resistance. *Journal of Applied Microbiology*, 92(s1) (2002).
4. Nikaido H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology reviews*, 67(4), 593-656 (2003).
5. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. *Nature Reviews Microbiology*, 2(5), 414-425 (2004).
6. Garcia-Vallvé S, Romeu A, Palau J. Horizontal gene transfer in bacterial and archaeal complete genomes. *Genome Research*, 10(11), 1719-1725 (2000).
7. Control CfD, Prevention. Office of Infectious Disease. Antibiotic resistance threats in the United States, 2013. April 2013. (Ed.^(Eds) (2015)
8. Gould IM, Bal AM. New antibiotic agents in the pipeline and how they can help overcome microbial resistance. *Virulence*, 4(2), 185-191 (2013).
9. Viswanathan V. Off-label abuse of antibiotics by bacteria. *Gut Microbes*, 5(1), 3-4 (2014).
10. Lushniak BD. Antibiotic resistance: a public health crisis. *Public Health Reports*, 129(4), 314-316 (2014).
11. Cosgrove SE. The relationship between antimicrobial resistance and patient outcomes: mortality, length of hospital stay, and health care costs. *Clinical Infectious Diseases*, 42(Supplement\_2), S82-S89 (2006).
12. Sydnor ER, Perl TM. Hospital epidemiology and infection control in acute-care settings. *Clinical Microbiology Reviews*, 24(1), 141-173 (2011).
13. Yeaman M, Bayer A. Antimicrobial peptides versus invasive infections. *Antimicrobial Peptides and Human Disease*, 111-152 (2006).
14. Mularski A, Separovic F. Atomic force microscopy studies of the interaction of antimicrobial peptides with bacterial cells. *Australian Journal of Chemistry*, 70(2), 130-137 (2017).

15. Golkar Z, Bagasra O, Pace DG. Bacteriophage therapy: a potential solution for the antibiotic resistance crisis. *The Journal of Infection in Developing Countries*, 8(02), 129-136 (2014).
16. Vollmer W, Blanot D, De Pedro MA. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2), 149-167 (2008).
17. Bush K. Antimicrobial agents targeting bacterial cell walls and cell membranes. *Revue Scientifique et Technique*, 31(1), 43-56 (2012).
18. Gordon E, Mouz N, Duee E, Dideberg O. The crystal structure of the penicillin-binding protein 2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance. *Journal of Molecular Biology*, 299(2), 477-485 (2000).
19. Klöckner A, Otten C, Derouaux A, Vollmer W, Buhl H, De Benedetti S, Munch D, Josten M, Molleken K, Sahl HG, Henrichfreise B. AmiA is a penicillin target enzyme with dual activity in the intracellular pathogen *Chlamydia pneumoniae*. *Nature Communications*, 5 (2014).
20. Spratt BG. Properties of the Penicillin-Binding Proteins of *Escherichia coli* K12. *The FEBS Journal*, 72(2), 341-352 (1977).
21. Blake C, Johnson LN, Mair G, North A, Phillips D, Sarma V. Crystallographic studies of the activity of hen egg-white lysozyme. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 167(1009), 378-388 (1967).
22. Gram C. Ueber die isolirte Färbung der Schizomyceten in Schnitt-und Trockenpräparaten. *Fortschritte der Medicin*, 2, 185-189 (1884).
23. Shockman GD, Barren J. Structure, function, and assembly of cell walls of Gram-positive bacteria. *Annual Reviews in Microbiology*, 37(1), 501-527 (1983).
24. Neuhaus FC, Baddiley J. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews*, 67(4), 686-723 (2003).
25. Vollmer W. Structural variation in the glycan strands of bacterial peptidoglycan. *FEMS Microbiology Reviews*, 32(2), 287-306 (2008).
26. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, 32(2), 234-258 (2008).
27. Glauert AM, Thornley MJ. The topography of the bacterial cell wall. *Annual Reviews in Microbiology*, 23(1), 159-198 (1969).

28. Kamio Y, Nikaido H. Outer membrane of *Salmonella typhimurium*: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry*, 15(12), 2561-2570 (1976).
29. Parsons LM, Lin F, Orban J. Peptidoglycan recognition by Pal, an outer membrane lipoprotein. *Biochemistry*, 45(7), 2122-2128 (2006).
30. Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. *Annual Review of Biochemistry*, 71(1), 635-700 (2002).
31. Kovacs-Simon A, Titball R, Michell SL. Lipoproteins of bacterial pathogens. *Infection and Immunity*, 79(2), 548-561 (2011).
32. Hantke K, Braun V. Covalent binding of lipid to protein. *The FEBS Journal*, 34(2), 284-296 (1973).
33. Mitchell P. Approaches to the analysis of specific membrane transport. *Biological Structure and Function*, 2, 581-599 (1961).
34. Elbreki M, Ross RP, Hill C, O'Mahony J, McAuliffe O, Coffey A. Bacteriophages and their derivatives as biotherapeutic agents in disease prevention and treatment. *Journal of Viruses*, 2014 (2014).
35. Ackermann H-W, Prangishvili D. Prokaryote viruses studied by electron microscopy. *Archives of Virology*, 157(10), 1843-1849 (2012).
36. Ashelford KE, Day MJ, Fry JC. Elevated abundance of bacteriophage infecting bacteria in soil. *Applied and Environmental Microbiology*, 69(1), 285-289 (2003).
37. Brüssow H, Kutter E. Phage ecology. *Bacteriophages: Biology and Applications*, 129-163 (2005).
38. Chibani-Chennoufi S, Bruttin A, Dillmann M-L, Brüssow H. Phage-host interaction: an ecological perspective. *Journal of Bacteriology*, 186(12), 3677-3686 (2004).
39. Deresinski S. Bacteriophage therapy: exploiting smaller fleas. *Clinical Infectious Diseases*, 48(8), 1096-1101 (2009).
40. Keary R, McAuliffe O, Ross R, Hill C, O'Mahony J, Coffey A. Bacteriophages and their endolysins for control of pathogenic bacteria. *Méndez-Vilas A. Microbial pathogens and strategies for combating them: science, technology and education, Formatex Research Center, Badajoz, Spain*, 1028-1040 (2013).

41. Sundar MM, Nagananda G, Das A, Bhattacharya S, Suryan S. Isolation of host-specific bacteriophages from sewage against human pathogens. *Asian J Biotechnol*, 1, 163-170 (2009).
42. Marsh P, Wellington E. Phage-host interactions in soil. *FEMS Microbiology Ecology*, 15(1-2), 99-107 (1994).
43. Bergh Ø, BØrsheim KY, Bratbak G, Haldal M. High abundance of viruses found in aquatic environments. *Nature*, 340(6233), 467-468 (1989).
44. Mills S, Shanahan F, Stanton C, Hill C, Coffey A, Ross RP. Movers and shakers: influence of bacteriophages in shaping the mammalian gut microbiota. *Gut microbes*, 4(1), 4-16 (2013).
45. Balcazar JL. Bacteriophages as vehicles for antibiotic resistance genes in the environment. *PLoS pathogens*, 10(7), e1004219 (2014).
46. Weinbauer MG, Agis M, Bonilla-Findji O, Malits A, Winter C. Bacteriophage in the environment. *Bacteriophage: Genetics and Molecular Biology*, 61-92 (2007).
47. Azam F, Fenchel T, Field JG, Gray J, Meyer-Reil L, Thingstad F. The ecological role of water-column microbes in the sea. *Marine Ecology Progress Series*, 257-263 (1983).
48. D'Herelle F. Technique de la recherche du microbe filtrant bactériophage (Bacteriophagum intestinale). *CR Soc. Biol*, 81, 1160-1162 (1918).
49. Wilkinson L. Félix d'Herelle and the origins of molecular biology. *Medical History*, 45(2), 294 (2001).
50. Ruska H. Versuch zu einer Ordnung der Virusarten. *Archives of Virology*, 2(5), 480-498 (1943).
51. Bradley DE. Ultrastructure of bacteriophage and bacteriocins. *Bacteriological Reviews*, 31(4), 230 (1967).
52. King AM, Lefkowitz E, Adams MJ, Carstens EB. *Virus Taxonomy: ninth report of the International Committee on Taxonomy of Viruses* (Elsevier, 2011).
53. Rao VB, Black LW. Structure and assembly of bacteriophage T4 head. *Virology Journal*, 7(1), 356 (2010).
54. Veesler D, Cambillau C. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiology and Molecular Biology Reviews*, 75(3), 423-433 (2011).



55. Fokine A, Rossmann MG. Molecular architecture of tailed double-stranded DNA phages. *Bacteriophage*, 4(2), e28281 (2014).
56. Guttman B, Raya R, Kutter E. Basic phage biology. *Bacteriophages: Biology and Applications*, 4 (2005).
57. Campbell AM. Thirty years ago in genetics: prophage insertion into bacterial chromosomes. *Genetics*, 133(3), 433 (1993).
58. Weinbauer MG. Ecology of prokaryotic viruses. *FEMS Microbiology Reviews*, 28(2), 127-181 (2004).
59. Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann M-L, Brüssow H. Phage as agents of lateral gene transfer. *Current Opinion in Microbiology*, 6(4), 417-424 (2003).
60. Davies J, Davies D. Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, 74(3), 417-433 (2010).
61. Santajit S, Indrawattana N. Mechanisms of antimicrobial resistance in ESKAPE pathogens. *BioMed Research International*, 2016 (2016).
62. Carson L, Gorman SP, Gilmore BF. The use of lytic bacteriophages in the prevention and eradication of biofilms of *Proteus mirabilis* and *Escherichia coli*. *FEMS Immunology & Medical Microbiology*, 59(3), 447-455 (2010).
63. Souza GR, Yonel-Gumruk E, Fan D, Easley J, Rangel R, Guzman-Rojas L, Miller JH, Arap W, Pasqualini R. Bottom-up assembly of hydrogels from bacteriophage and Au nanoparticles: the effect of cis- and trans-acting factors. *PloS one*, 3(5), e2242 (2008).
64. Markoishvili K, Tsitlanadze G, Katsarava R, Glenn J, Sulakvelidze A. A novel sustained-release matrix based on biodegradable poly (ester amide) s and impregnated with bacteriophages and an antibiotic shows promise in management of infected venous stasis ulcers and other poorly healing wounds. *International Journal of Dermatology*, 41(7), 453-458 (2002).
65. McVay CS, Velásquez M, Fralick JA. Phage therapy of *Pseudomonas aeruginosa* infection in a mouse burn wound model. *Antimicrobial Agents and Chemotherapy*, 51(6), 1934-1938 (2007).
66. Soothill J. Treatment of experimental infections of mice with bacteriophages. *Journal of Medical Microbiology*, 37(4), 258-261 (1992).
67. Smith HW, Huggins MB, Shaw KM. The control of experimental *Escherichia coli* diarrhoea in calves by means of bacteriophages. *Microbiology*, 133(5), 1111-1126 (1987).

68. Watanabe R, Matsumoto T, Sano G, Ishii Y, Tateda K, Sumiyama Y, Uchiyama J, Sakurai S, Matsuzaki S, Imai S, Yamaguchi K. Efficacy of bacteriophage therapy against gut-derived sepsis caused by *Pseudomonas aeruginosa* in mice. *Antimicrobial Agents and Chemotherapy*, 51(2), 446-452 (2007).
69. Abedon ST, Kuhl SJ, Blasdel BG, Kutter EM. Phage treatment of human infections. *Bacteriophage*, 1(2), 66-85 (2011).
70. Stone R. Stalin's forgotten cure. *Science*, 298(5594), 728-731 (2002).
71. Moak M, Molineux IJ. Peptidoglycan hydrolytic activities associated with bacteriophage virions. *Molecular Microbiology*, 51(4), 1169-1183 (2004).
72. Hermoso JA, García JL, García P. Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Current Opinion in Microbiology*, 10(5), 461-472 (2007).
73. Rodríguez-Rubio L, Martínez B, Rodríguez A, Donovan DM, García P. Enhanced staphylococcal activity of the *Staphylococcus aureus* bacteriophage vB\_SauS-phiIPLA88 HydH5 virion-associated peptidoglycan hydrolase: fusions, deletions, and synergy with LysH5. *Applied and Environmental Microbiology*, 78(7), 2241-2248 (2012).
74. Young R. Bacteriophage lysis: mechanism and regulation. *Microbiological Reviews*, 56(3), 430-481 (1992).
75. Wang I-N, Smith DL, Young R. Holins: the protein clocks of bacteriophage infections. *Annual Reviews in Microbiology*, 54(1), 799-825 (2000).
76. Young R. Bacteriophage holins: deadly diversity. *Journal of Molecular Microbiology and Biotechnology*, 4(1), 21-36 (2002).
77. Pang T, Savva CG, Fleming KG, Struck DK, Young R. Structure of the lethal phage pinhole. *Proceedings of the National Academy of Sciences*, 106(45), 18966-18971 (2009).
78. Berry J, Rajaure M, Pang T, Young R. The spanin complex is essential for lambda lysis. *Journal of Bacteriology*, 194(20), 5667-5674 (2012).
79. Berry JD, Rajaure M, Young R. Spanin function requires subunit homodimerization through intermolecular disulfide bonds. *Molecular Microbiology*, 88(1), 35-47 (2013).
80. Loessner MJ. Bacteriophage endolysins—current state of research and applications. *Current Opinion in Microbiology*, 8(4), 480-487 (2005).
81. Young R, Wang N, Roof WD. Phages will out: strategies of host cell lysis. *Trends in Microbiology*, 8(3), 120-128 (2000).

82. Ralston DJ, Baer BS, Lieberman M, Krueger AP. Lysis from without of *S. aureus* K1 by the combined action of phage and virolysin. *The Journal of General Physiology*, 41(2), 343-358 (1957).
83. Sonstein SA, Hammel JM, Bondi A. Staphylococcal bacteriophage-associated lysin: a lytic agent active against *Staphylococcus aureus*. *Journal of Bacteriology*, 107(2), 499-504 (1971).
84. Borysowski J, Weber-Dąbrowska B, Górski A. Bacteriophage endolysins as a novel class of antibacterial agents. *Experimental Biology and Medicine*, 231(4), 366-377 (2006).
85. Pirnay J-P, De Vos D, Verbeken G, Merabishvili M, Chanishvili N, Vaneechoutte M, Zizi M, Laire G, Lavigne R, Huyas I, Van den Mooter G. The phage therapy paradigm: pret-a-porter or sur-mesure? *Pharmaceutical Research*, 28(4), 934-937 (2011).
86. Gill JJ, Hyman P. Phage choice, isolation, and preparation for phage therapy. *Current Pharmaceutical Biotechnology*, 11(1), 2-14 (2010).
87. Fischetti VA. 12 The Use of Phage Lytic Enzymes to Control Bacterial Infections. *Bacteriophages: Biology and Applications*, 321 (2004).
88. Fischetti VA. Bacteriophage lytic enzymes: novel anti-infectives. *Trends in microbiology*, 13(10), 491-496 (2005).
89. Stojković EA, Rothman-Denes LB. Coliphage N4 N-acetylmuramidase defines a new family of murein hydrolases. *Journal of Molecular Biology*, 366(2), 406-419 (2007).
90. Takegawa K, Nakoshi M, Iwahara S, Yamamoto K, Tochikura T. Induction and purification of endo- $\beta$ -N-acetylglucosaminidase from *Arthrobacter protophormiae* grown in ovalbumin. *Applied and Environmental Microbiology*, 55(12), 3107-3112 (1989).
91. Höltje J, Mirelman D, Sharon N, Schwarz U. Novel type of murein transglycosylase in *Escherichia coli*. *Journal of Bacteriology*, 124(3), 1067-1076 (1975).
92. Fokine A, Miroshnikov KA, Shneider MM, Mesyanzhinov VV, Rossmann MG. Structure of the bacteriophage  $\phi$ KZ lytic transglycosylase gp144. *Journal of Biological Chemistry*, 283(11), 7242-7250 (2008).
93. Obeso JM, Martínez B, Rodríguez A, García P. Lytic activity of the recombinant staphylococcal bacteriophage  $\Phi$ H5 endolysin active against *Staphylococcus aureus* in milk. *International Journal of Food Microbiology*, 128(2), 212-218 (2008).

94. Cheng X, Zhang X, Pflugrath JW, Studier FW. The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proceedings of the National Academy of Sciences*, 91(9), 4034-4038 (1994).
95. Navarre WW, Schneewind O. Surface proteins of Gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiology and Molecular Biology Reviews*, 63(1), 174-229 (1999).
96. Cheng Q, Nelson D, Zhu S, Fischetti VA. Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrobial Agents and Chemotherapy*, 49(1), 111-117 (2005).
97. Yokoi K-j, Kawahigashi N, Uchida M, Sugahara K, Shinohara M, Kawasaki KI, Nakamura S, Taketo A, Kodaira KI. The two-component cell lysis genes holWMY and lysWMY of the Staphylococcus warneri M phage  $\phi$ WMY: cloning, sequencing, expression, and mutational analysis in Escherichia coli. *Gene*, 351, 97-108 (2005).
98. Pritchard DG, Dong S, Baker JR, Engler JA. The bifunctional peptidoglycan lysin of *Streptococcus agalactiae* bacteriophage B30. *Microbiology*, 150(7), 2079-2087 (2004).
99. Fischetti VA. Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. *International Journal of Medical Microbiology*, 300(6), 357-362 (2010).
100. Yang H, Wang D-B, Dong Q, Zhang Z, Cui Z, Deng J, Yu J, Zhang XE, Wei H. Existence of separate domains in lysin PlyG for recognizing Bacillus anthracis spores and vegetative cells. *Antimicrobial Agents and Chemotherapy*, 56(10), 5031-5039 (2012).
101. Oliveira H, Melo LD, Santos SB, Nobrega FL, Ferreira EC, Cerca N, Azeredo J, Kluskens LD. Molecular aspects and comparative genomics of bacteriophage endolysins. *Journal of Virology*, 87(8), 4558-4570 (2013).
102. Rigden DJ, Jedrzejewski MJ, Galperin MY. Amidase domains from bacterial and phage autolysins define a family of  $\gamma$ -D, L-glutamate-specific amidohydrolases. *Trends in Biochemical Sciences*, 28(5), 230-234 (2003).
103. Layec S, Decaris B, Leblond-Bourget N. Characterization of proteins belonging to the CHAP-related superfamily within the Firmicutes. *Journal of Molecular Microbiology and Biotechnology*, 14(1-3), 31-40 (2008).
104. Bateman A, Rawlings ND. The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. *Trends in Biochemical Sciences*, 28(5), 234-237 (2003).

105. Zou Y, Hou C. Systematic analysis of an amidase domain CHAP in 12 *Staphylococcus aureus* genomes and 44 staphylococcal phage genomes. *Computational Biology and Chemistry*, 34(4), 251-257 (2010).
106. Sáiz JL, López-Zumel C, Monterroso B, Varea J, Arrondo JL, Iloro I, Garcia JL, Laynez JL, Menendez M. Characterization of Ejl, the cell-wall amidase coded by the Pneumococcal bacteriophage Ej-1. *Protein Science*, 11(7), 1788-1799 (2002).
107. Varea J, Monterroso B, Sáiz JL, Lopez-Zumel C, Garcia JL, Laynez J, Garcia P, Menendez M. Structural and thermodynamic characterization of Pal, a phage natural chimeric lysin active against *Pneumococci*. *Journal of Biological Chemistry*, 279(42), 43697-43707 (2004).
108. Romero P, López R, García E. Characterization of LytA-like N-acetylmuramoyl-L-alanine amidases from two new *Streptococcus mitis* bacteriophages provides insights into the properties of the major pneumococcal autolysin. *Journal of Bacteriology*, 186(24), 8229-8239 (2004).
109. Matthews B, Remington S. The three dimensional structure of the lysozyme from bacteriophage T4. *Proceedings of the National Academy of Sciences*, 71(10), 4178-4182 (1974).
110. Schmelcher M, Shabarova T, Eugster MR, Eichenseher F, Tchang VS, Banz M, Loessner MJ. Rapid multiplex detection and differentiation of *Listeria* cells by use of fluorescent phage endolysin cell wall binding domains. *Applied and Environmental Microbiology*, 76(17), 5745-5756 (2010).
111. Loessner MJ, Kramer K, Ebel F, Scherer S. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Molecular Microbiology*, 44(2), 335-349 (2002).
112. Briers Y, Volckaert G, Cornelissen A, Lagaert S, Michiels CW, Hertveldt K, Lavigne R. Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages  $\phi$ KZ and EL. *Molecular Microbiology*, 65(5), 1334-1344 (2007).
113. Adriaenssens E, Mattheus W, Cornelissen A, Shaburova O, Krylov VN, Kropinski AM, Lavigne R. Complete genome sequence of the giant *Pseudomonas* phage Lu11. *Journal of Virology*, 86(11), 6369-6370 (2012).
114. Walmagh M, Briers Y, Dos Santos SB, Azeredo J, Lavigne R. Characterization of modular bacteriophage endolysins from *Myoviridae* phages OBP, 201 $\phi$ 2-1 and PVP-SE1. *PLoS one*, 7(5), e36991 (2012).

115. Garcia E, Garcia JL, Garcia P, Arrarás A, Sánchez-Puelles JM, López R. Molecular evolution of lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Proceedings of the National Academy of Sciences*, 85(3), 914-918 (1988).
116. Vaara M. Agents that increase the permeability of the outer membrane. *Microbiological reviews*, 56(3), 395-411 (1992).
117. Loeffler JM, Nelson D, Fischetti VA. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science*, 294(5549), 2170-2172 (2001).
118. Daniel A, Euler C, Collin M, Chahales P, Gorelick KJ, Fischetti VA. Synergism between a novel chimeric lysin and oxacillin protects against infection by methicillin-resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 54(4), 1603-1612 (2010).
119. Singh PK, Donovan DM, Kumar A. Intravitreal injection of the chimeric phage endolysin Ply187 protects mice from *Staphylococcus aureus* endophthalmitis. *Antimicrobial Agents and Chemotherapy*, 58(8), 4621-4629 (2014).
120. Schuch R, Nelson D, Fischetti VA. A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature*, 418(6900), 884-889 (2002).
121. Pastagia M, Euler C, Chahales P, Fuentes-Duculan J, Krueger JG, Fischetti VA. A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and-sensitive *Staphylococcus aureus* strains. *Antimicrobial Agents and Chemotherapy*, 55(2), 738-744 (2011).
122. Pastagia M, Schuch R, Fischetti VA, Huang DB. Lysins: the arrival of pathogen-directed anti-infectives. *Journal of Medical Microbiology*, 62(10), 1506-1516 (2013).
123. Guariglia-Oropeza V, Helmann JD. *Bacillus subtilis*  $\sigma$ V confers lysozyme resistance by activation of two cell wall modification pathways, peptidoglycan O-acetylation and D-alanylation of teichoic acids. *Journal of Bacteriology*, 193(22), 6223-6232 (2011).
124. Kusuma C, Jadanova A, Chanturiya T, Kokai-Kun JF. Lysostaphin-resistant variants of *Staphylococcus aureus* demonstrate reduced fitness in vitro and in vivo. *Antimicrobial Agents and Chemotherapy*, 51(2), 475-482 (2007).
125. DeHart HP, Heath HE, Heath LS, LeBlanc PA, Sloan GL. The lysostaphin endopeptidase resistance gene (*epr*) specifies modification of peptidoglycan cross bridges in *Staphylococcus simulans* and *Staphylococcus aureus*. *Applied and Environmental Microbiology*, 61(4), 1475-1479 (1995).

126. Ehlert K, Tschierske M, Mori C, Schröder W, Berger-Bächi B. Site-specific serine incorporation by Lif and Epr into positions 3 and 5 of the *Staphylococcal* peptidoglycan interpeptide bridge. *Journal of Bacteriology*, 182(9), 2635-2638 (2000).
127. Schmelcher M, Donovan DM, Loessner MJ. Bacteriophage endolysins as novel antimicrobials. *Future Microbiology*, 7(10), 1147-1171 (2012).
128. Schmelcher M, Waldherr F, Loessner MJ. Listeria bacteriophage peptidoglycan hydrolases feature high thermoresistance and reveal increased activity after divalent metal cation substitution. *Applied Microbiology and Biotechnology*, 93(2), 633-643 (2012).
129. Donovan DM, Dong S, Garrett W, Rousseau GM, Moineau S, Pritchard DG. Peptidoglycan hydrolase fusions maintain their parental specificities. *Applied and Environmental Microbiology*, 72(4), 2988-2996 (2006).
130. Hendrix RW. Bacteriophages: evolution of the majority. *Theoretical Population Biology*, 61(4), 471-480 (2002).
131. Zimmer M, Sattelberger E, Inman RB, Calendar R, Loessner MJ. Genome and proteome of Listeria monocytogenes phage PSA: an unusual case for programmed+ 1 translational frameshifting in structural protein synthesis. *Molecular Microbiology*, 50(1), 303-317 (2003).
132. Sheehan MM, García JL, López R, García P. The lytic enzyme of the *Pneumococcal* phage Dp-1: a chimeric lysin of intergeneric origin. *Molecular Microbiology*, 25(4), 717-725 (1997).
133. Becker SC, Dong S, Baker JR, Foster-Frey J, Pritchard DG, Donovan DM. LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. *FEMS Microbiology Letters*, 294(1), 52-60 (2009).
134. Becker SC, Foster-Frey J, Stodola AJ, Anacker D, Donovan DM. Differentially conserved *Staphylococcal* SH3b\_5 cell wall binding domains confer increased Staphylolytic and Streptolytic activity to a streptococcal prophage endolysin domain. *Gene*, 443(1), 32-41 (2009).
135. Manoharadas S, Witte A, Bläsi U. Antimicrobial activity of a chimeric enzymatic towards *Staphylococcus aureus*. *Journal of Biotechnology*, 139(1), 118-123 (2009).
136. Diaz E, Lopez R, Garcia JL. Chimeric phage-bacterial enzymes: a clue to the modular evolution of genes. *Proceedings of the National Academy of Sciences*, 87(20), 8125-8129 (1990).

137. Diaz E, López R, Garcia JL. Chimeric *Pneumococcal* cell wall lytic enzymes reveal important physiological and evolutionary traits. *Journal of Biological Chemistry*, 266(9), 5464-5471 (1991).
138. Croux C, Ronda C, Lopez R, Garcia J. Interchange of functional domains switches enzyme specificity: construction of a chimeric pneumococcal-clostridial cell wall lytic enzyme. *Molecular Microbiology*, 9(5), 1019-1025 (1993).
139. Donovan D, Becker S, Dong S, Baker J, Foster-Frey J, Pritchard D. Peptidoglycan hydrolase enzyme fusions for treating multi-drug resistant pathogens. *Biotech International*, 21(2), 6 (2009).
140. Rodríguez-Rubio L, Martínez B, Rodríguez A, Donovan DM, Götz F, García P. The phage lytic proteins from the *Staphylococcus aureus* bacteriophage vB\_SauS-phiPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance. *PloS one*, 8(5), e64671 (2013).
141. Becker SC, Foster-Frey J, Donovan DM. The phage K lytic enzyme LysK and lysostaphin act synergistically to kill MRSA. *FEMS Microbiology Letters*, 287(2), 185-191 (2008).
142. Fernandes S, Proença D, Cantante C, Silva FA, Leandro C, Lourenco S, Milheirico C, de Lencastre H, Cavaco-Silva P, Pimentel M, Sao-Jose C. Novel chimerical endolysins with broad antimicrobial activity against methicillin-resistant *Staphylococcus aureus*. *Microbial Drug Resistance*, 18(3), 333-343 (2012).
143. Swift SM, Seal BS, Garrish JK, Oakley BB, HiettK, Yeh HY, Woolsey R, Schegg KM, Line JE, Donovan DM. A thermophilic phage endolysin fusion to a *Clostridium perfringens*-specific cell wall binding domain creates an anti-*Clostridium* antimicrobial with improved thermostability. *Viruses*, 7(6), 3019-3034 (2015).
144. Yang H, Zhang H, Wang J, Yu J, Wei H. A novel chimeric lysin with robust antibacterial activity against planktonic and biofilm methicillin-resistant *Staphylococcus aureus*. *Scientific Reports*, 7 (2017).
145. Cheng Q, Fischetti VA. Mutagenesis of a bacteriophage lytic enzyme PlyGBS significantly increases its antibacterial activity against group B streptococci. *Applied Microbiology and Biotechnology*, 74(6), 1284-1291 (2007).
146. Fujiwara T, Aoki S, Komatsuzawa H, Nishida T, Ohara M, Suginaka H, Sugai M. Mutation analysis of the histidine residues in the glycylglycine endopeptidase ALE-1. *Journal of Bacteriology*, 187(2), 480-487 (2005).
147. Borysowski J, Górski A. Fusion to cell-penetrating peptides will enable lytic enzymes to kill intracellular bacteria. *Medical Hypotheses*, 74(1), 164-166 (2010).



148. Dietz GP, Bähr M. Delivery of bioactive molecules into the cell: the Trojan horse approach. *Molecular and Cellular Neuroscience*, 27(2), 85-131 (2004).
149. Dietz GP. Cell-penetrating peptide technology to deliver chaperones and associated factors in diseases and basic research. *Current Pharmaceutical Biotechnology*, 11(2), 167-174 (2010).
150. Splith K, Neundorf I. Antimicrobial peptides with cell-penetrating peptide properties and vice versa. *European Biophysics Journal*, 40(4), 387-397 (2011).
151. Schwarze SR, Hruska KA, Dowdy SF. Protein transduction: unrestricted delivery into all cells? *Trends in cell biology*, 10(7), 290-295 (2000).
152. Clark E, Nava B, Caputi M. TAT is a multifunctional viral protein that modulates cellular gene expression and functions. *Oncotarget*, 8(16), 27569 (2017).
153. Ryu J, Han K, Park J, Choi SY. Enhanced uptake of a heterologous protein with an HIV-1 Tat protein transduction domains (PTD) at both termini. *Molecules & Cells (Springer Science & Business Media BV)*, 16(3) (2003).
154. Gaeng S, Scherer S, Neve H, Loessner MJ. Gene Cloning and Expression and Secretion of *Listeria monocytogenes* Bacteriophage-Lytic Enzymes in *Lactococcus lactis*. *Applied and Environmental Microbiology*, 66(7), 2951-2958 (2000).
155. Briers Y, Walmagh M, Van Puyenbroeck V, Cornelissen A, Cenens W, Aertsen A, Oliveira H, Azeredo J, Verween G, Pirnay JP, Miller S. Engineered endolysin-based “Artilyns” to combat multidrug-resistant Gram-negative pathogens. *mBio*, 5(4), e01379-01314 (2014).
156. Gerstmans H, Rodríguez-Rubio L, Lavigne R, Briers Y. From endolysins to Artilysin® s: novel enzyme-based approaches to kill drug-resistant bacteria. *Biochemical Society Transactions*, 44(1), 123-128 (2016).
157. Ibrahim HR, Yamada M, Matsushita K, Kobayashi K, Kato A. Enhanced bactericidal action of lysozyme to *Escherichia coli* by inserting a hydrophobic pentapeptide into its C terminus. *Journal of Biological Chemistry*, 269(7), 5059-5063 (1994).
158. Briers Y, Walmagh M, Grymonprez B, Bieble M, Pirnay JP, Defraigne V, Michiels J, Cenens W, Aertsen A, Miller S, Lavigne R. Art-175 is a highly efficient antibacterial against multidrug-resistant strains and persists of *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, 58(7), 3774-3784 (2014).
159. Maher S, McClean S. Investigation of the cytotoxicity of eukaryotic and prokaryotic antimicrobial peptides in intestinal epithelial cells in vitro. *Biochemical Pharmacology*, 71(9), 1289-1298 (2006).

160. Dawson RM, Liu CQ. Cathelicidin peptide SMAP-29: comprehensive review of its properties and potential as a novel class of antibiotics. *Drug Development Research*, 70(7), 481-498 (2009).
161. Zhang H, Bao H, Billington C, Hudson JA, Wang R. Isolation and lytic activity of the *Listeria* bacteriophage endolysin LysZ5 against *Listeria monocytogenes* in soya milk. *Food microbiology*, 31(1), 133-136 (2012).
162. García P, Martínez B, Rodríguez L, Rodríguez A. Synergy between the phage endolysin LysH5 and nisin to kill *Staphylococcus aureus* in pasteurized milk. *International Journal of Food Microbiology*, 141(3), 151-155 (2010).
163. Celia LK, Nelson D, Kerr DE. Characterization of a bacteriophage lysin (Ply700) from *Streptococcus uberis*. *Veterinary Microbiology*, 130(1), 107-117 (2008).
164. Nelson D, Loomis L, Fischetti VA. Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proceedings of the National Academy of Sciences*, 98(7), 4107-4112 (2001).
165. Entenza J, Loeffler J, Grandgirard D, Fischetti V, Moreillon P. Therapeutic effects of bacteriophage Cpl-1 lysin against *Streptococcus pneumoniae* endocarditis in rats. *Antimicrobial Agents and Chemotherapy*, 49(11), 4789-4792 (2005).
166. Rashel M, Uchiyama J, Ujihara T, Ujihara T, Uehara Y, Kuramoto S, Sugihara S, Yagyu KI, Muraoka A, Sugai M, Hiramastu K, Honke K. Efficient elimination of multidrug-resistant *Staphylococcus aureus* by cloned lysin derived from bacteriophage  $\phi$ MR11. *The Journal of Infectious Diseases*, 196(8), 1237-1247 (2007).
167. Jun SY, Jung GM, Yoon SJ, Choi YJ, Koh WS, Moon KS, Kang SH. Preclinical safety evaluation of intravenously administered SAL200 containing the recombinant phage endolysin SAL-1 as a pharmaceutical ingredient. *Antimicrobial Agents and Chemotherapy*, 58(4), 2084-2088 (2014).
168. Amorena B, Gracia E, Monzón M, Leiva J, Oteiza C, Perez M, Alabart JL, Hernandez-Yang J. Antibiotic susceptibility assay for *Staphylococcus aureus* in biofilms developed in vitro. *Journal of Antimicrobial Chemotherapy*, 44(1), 43-55 (1999).
169. Emori TG, Gaynes RP. An overview of nosocomial infections, including the role of the microbiology laboratory. *Clinical Microbiology Reviews*, 6(4), 428-442 (1993).
170. Fux CA, Stoodley P, Hall-Stoodley L, Costerton JW. Bacterial biofilms: a diagnostic and therapeutic challenge. *Expert Review of Anti-Infective Therapy*, 1(4), 667-683 (2003).

171. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. *International Journal of Antimicrobial Agents*, 35(4), 322-332 (2010).
172. Keren I, Kaldalu N, Spoering A, Wang Y, Lewis K. Persister cells and tolerance to antimicrobials. *FEMS Microbiology Letters*, 230(1), 13-18 (2004).
173. Lewis K. Riddle of biofilm resistance. *Antimicrobial Agents and Chemotherapy*, 45(4), 999-1007 (2001).
174. Sass P, Bierbaum G. Lytic activity of recombinant bacteriophage  $\phi$ 11 and  $\phi$ 12 endolysins on whole cells and biofilms of *Staphylococcus aureus*. *Applied and Environmental Microbiology*, 73(1), 347-352 (2007).
175. O'flaherty S, Coffey A, Meaney W, Fitzgerald G, Ross R. The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant *Staphylococci*, including methicillin-resistant *Staphylococcus aureus*. *Journal of Bacteriology*, 187(20), 7161-7164 (2005).
176. Son J-S, Lee S-J, Jun SY, Yoon SJ, Kang SH, Paik HR, Kang JO, Choi YJ. Antibacterial and biofilm removal activity of a podoviridae *Staphylococcus aureus* bacteriophage SAP-2 and a derived recombinant cell-wall-degrading enzyme. *Applied Microbiology and Biotechnology*, 86(5), 1439-1449 (2010).
177. Gutierrez D, Ruas-Madiedo P, Martínez B, Rodríguez A, García P. Effective removal of *Staphylococcal* biofilms by the endolysin LysH5. *PLoS one*, 9(9), e107307 (2014).
178. Domenech M, García E, Moscoso M. In vitro destruction of *Streptococcus pneumoniae* biofilms with bacterial and phage peptidoglycan hydrolases. *Antimicrobial Agents and Chemotherapy*, 55(9), 4144-4148 (2011).
179. Simmons M, Morales CA, Oakley BB, Seal BS. Recombinant expression of a putative amidase cloned from the genome of *Listeria monocytogenes* that lyses the bacterium and its monolayer in conjunction with a protease. *Probiotics and Antimicrobial Proteins*, 4(1), 1-10 (2012).
180. de Vries J, Harms K, Broer I, Kriete G, Mahn A, During K, Wackernagel W. The bacteriolytic activity in transgenic potatoes expressing a chimeric T4 lysozyme gene and the effect of T4 lysozyme on soil- and phytopathogenic bacteria. *Systematic and Applied Microbiology*, 22(2), 280-286 (1999).
181. Ozawa H, Tanaka H, Ichinose Y, Shiraishi T, Yamada T. Bacteriophage P4282, a parasite of *Ralstonia solanacearum*, encodes a bacteriolytic protein important for lytic infection of its host. *Molecular Genetics and Genomics*, 265(1), 95-101 (2001).

182. Yamada T. Bacteriophages of *Ralstonia solanacearum*: their diversity and utilization as biocontrol agents in agriculture. In: *Bacteriophages*. (InTech, 2012)
183. Oey M, Lohse M, Kreikemeyer B, Bock R. Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *The Plant Journal*, 57(3), 436-445 (2009).
184. Fujinami Y, Hirai Y, Sakai I, Yoshino M, Yasuda J. Sensitive Detection of *Bacillus anthracis* Using a Binding Protein Originating from  $\gamma$ -Phage. *Microbiology and Immunology*, 51(2), 163-169 (2007).
185. Kretzer JW, Lehmann R, Schmelcher M, Banz M, Kim KP, Korn C, Loessner MJ. Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. *Applied and Environmental Microbiology*, 73(6), 1992-2000 (2007).
186. Walcher G, Stessl B, Wagner M, Eichenseher F, Loessner MJ, Hein I. Evaluation of Paramagnetic Beads Coated with Recombinant *Listeria* Phage Endolysin-Derived Cell-Wall-Binding Domain Proteins for Separation of *Listeria monocytogenes* from Raw Milk in Combination with Culture-Based and Real-Time Polymerase Chain Reaction-Based Quantification. *Foodborne Pathogens and Disease*, 7(9), 1019-1024 (2010).
187. Sainathrao S, Mohan KVK, Atreya C. Gamma-phage lysin PlyG sequence-based synthetic peptides coupled with Qdot-nanocrystals are useful for developing detection methods for *Bacillus anthracis* by using its surrogates, *B. anthracis*-Sterne and *B. cereus*-4342. *BMC Biotechnology*, 9(1), 67 (2009).
188. Anfinsen CB, Haber E. Studies on the reduction and re-formation of protein disulfide bonds. *The Journal of Biological Chemistry*, 236(5), 1361-1363 (1961).
189. Wijaya EC, Separovic F, Drummond CJ, Greaves TL. Activity and conformation of lysozyme in molecular solvents, protic ionic liquids (PILs) and salt-water systems. *Physical Chemistry Chemical Physics*, 18(37), 25926-25936 (2016).
190. Epstein CJ, Goldberger RF, Anfinsen CB. The genetic control of tertiary protein structure: studies with model systems. In: *Cold Spring Harbor symposia on quantitative biology*. () 439-449 (1963).
191. Levinthal C. How to fold graciously. *Mossbauer Spectroscopy in Biological Systems*, 67, 22-24 (1969).
192. Levinthal C. Are there pathways for protein folding? *Journal de Chimie Physique*, 65, 44-45 (1968).

193. Kim PS, Baldwin RL. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annual Review of Biochemistry*, 51(1), 459-489 (1982).
194. Ptitsyn O. Stages in the mechanism of self-organization of protein molecules. *Doklady Akademii nauk SSSR*, 210(5), 1213 (1973).
195. Gō N. The consistency principle in protein structure and pathways of folding. *Advances in Biophysics*, 18, 149-164 (1984).
196. Mok KH, Kuhn LT, Goetz M, Day IJ, Lin JC, Andersen NH, Hore PJ. A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature*, 447(7140), 106-109 (2007).
197. Fersht AR. Nucleation mechanisms in protein folding. *Current Opinion in Structural Biology*, 7(1), 3-9 (1997).
198. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Structure, Function, and Bioinformatics*, 21(3), 167-195 (1995).
199. Onuchic JN, Luthey-Schulten Z, Wolynes PG. Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry*, 48(1), 545-600 (1997).
200. Wolynes PG. Energy landscapes and solved protein-folding problems. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 363(1827), 453-467 (2005).
201. Shoemaker BA, Wang J, Wolynes PG. Exploring structures in protein folding funnels with free energy functionals: the transition state ensemble. *Journal of Molecular Biology*, 287(3), 675-694 (1999).
202. Panchenko AR, Luthey-Schulten Z, Wolynes PG. Foldons, protein structural modules, and exons. *Proceedings of the National Academy of Sciences*, 93(5), 2008-2013 (1996).
203. Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nature Structural & Molecular Biology*, 4(1), 10-19 (1997).
204. Dobson CM, Karplus M. The fundamentals of protein folding: bringing together theory and experiment. *Current Opinion in Structural Biology*, 9(1), 92-101 (1999).
205. Dyson HJ, Wright PE. Elucidation of the protein folding landscape by NMR. *Methods in Enzymology*, 394, 299-321 (2005).

206. Shukla VK, Singh JS, Vispute N, Ahmad B, Kumar A, Hosur RV. Unfolding of CPR3 gets initiated at the active site and proceeds via two intermediates. *Biophysical Journal*, 112(4), 605-619 (2017).
207. Bhavesh N, Hosur R. Exploring unstructured proteins. In: *Proc. Indian Natl. Sci. Acad. A*, 579-596 (2004).
208. Rajanikanth V, Srivastava SS, Singh AK, Rajyalakshmi M, Chandra K, Arvind P, Sankaranarayanan R, Sharma Y. Aggregation-prone near-native intermediate formation during unfolding of a structurally similar nonlenticular  $\beta\gamma$ -crystallin domain. *Biochemistry*, 51(43), 8502-8513 (2012).
209. Suman SK, Ravindra D, Sharma Y, Mishra A. Association properties and unfolding of a  $\beta\gamma$ -crystallin domain of a *Vibrio*-specific protein. *PLoS one*, 8(1), e53610 (2013).
210. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradović Z. Intrinsic disorder and protein function. *Biochemistry*, 41(21), 6573-6582 (2002).
211. Uversky V. Protein folding revisited. A polypeptide chain at the folding–misfolding–nonfolding cross-roads: which way to go? *Cellular and Molecular Life Sciences CMLS*, 60(9), 1852-1871 (2003).
212. Neudecker P, Lundström P, Kay LE. Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. *Biophysical Journal*, 96(6), 2045-2054 (2009).
213. Kumar A, Srivastava S, Hosur RV. NMR characterization of the energy landscape of SUMO-1 in the native-state ensemble. *Journal of Molecular Biology*, 367(5), 1480-1493 (2007).
214. Kumar A, Srivastava S, Mishra RK, Mittal R, Hosur RV. Residue-level NMR view of the urea-driven equilibrium folding transition of SUMO-1 (1-97): Native preferences do not increase monotonously. *Journal of Molecular Biology*, 361(1), 180-194 (2006).
215. Englander SW, Mayne L, Krishna MM. Protein folding and misfolding: mechanism and principles. *Quarterly Reviews of Biophysics*, 40(4), 1-41 (2007).
216. Krishna MM, Hoang L, Lin Y, Englander SW. Hydrogen exchange methods to study protein folding. *Methods*, 34(1), 51-64 (2004).
217. Dobson CM. Principles of protein folding, misfolding and aggregation. In: *Seminars in Cell & Developmental Biology*. 3-16 (2004).
218. Gehman JD, Separovic F. Solid-state NMR of amyloid membrane interactions. *Protein Folding, Misfolding, and Disease: Methods and Protocols*, 165-177 (2011).

219. Fink AL. Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Folding and Design*, 3(1), R9-R23 (1998).
220. Ranjan P, Ghosh D, Yarramala DS, Das S, Maji SK, Kumar A. Differential copper binding to alpha-synuclein and its disease-associated mutants affect the aggregation and amyloid formation. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1861(2), 365-374 (2017).
221. Rastogi N, Mitra K, Kumar D, Roy R. Metal ions as cofactors for aggregation of therapeutic peptide salmon calcitonin. *Inorganic Chemistry*, 51(10), 5642-5650 (2012).
222. Adler J, Scheidt HA, Krüger M, Thomas L, Huster D. Local interactions influence the fibrillation kinetics, structure and dynamics of A $\beta$  (1–40) but leave the general fibril structure unchanged. *Physical Chemistry Chemical Physics*, 16(16), 7461-7471 (2014).
223. Singh SM, Kongari N, Cabello-Villegas J, Mallela KM. Missense mutations in dystrophin that trigger muscular dystrophy decrease protein stability and lead to cross- $\beta$  aggregates. *Proceedings of the National Academy of Sciences*, 107(34), 15069-15074 (2010).
224. Singh SM, Molas JF, Kongari N, Bandi A, Armstrong GS, Winder SJ, Mallela KM. Thermodynamic stability, unfolding kinetics, and aggregation of the N-terminal actin-binding domains of utrophin and dystrophin. *Proteins: Structure, Function, and Bioinformatics*, 80(5), 1377-1392 (2012).
225. Singh SM, Cabello-Villegas J, Hutchings RL, Mallela KM. Role of partial protein unfolding in alcohol-induced protein aggregation. *Proteins: Structure, Function, and Bioinformatics*, 78(12), 2625-2637 (2010).
226. Salahuddin P. Protein Folding, Misfolding, Aggregation And Amyloid Formation: Mechanisms of A $\beta$  Oligomer Mediated Toxicities. *Journal of Biochemistry and Molecular Biology Research*, 1(2), 36-45 (2015).
227. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*, 19(1), 26-59 (2001).
228. Fink AL. Natively unfolded proteins. *Current Opinion in Structural Biology*, 15(1), 35-41 (2005).
229. Bemporad F, Gsponer J, Hopearuoho HI, Plakoutsi G, Stati G, Stefani M, Taddei N, Vendruscolo M, Chiti F. Biological function in a non-native partially folded state of a protein. *The EMBO Journal*, 27(10), 1525-1535 (2008).

230. Martsev SP, Dubnovitsky AP, Stremovsky OA, Chumanovich AA, Tsybovsky YI, Kravchuk ZI, Deyev SM. Partially structured state of the functional VH domain of the mouse anti-ferritin antibody F11. *FEBS Letters*, 518(1-3), 177-182 (2002).
231. Kjaergaard M, Teilum K, Poulsen FM. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proceedings of the National Academy of Sciences*, 107(28), 12535-12540 (2010).
232. Tan Y-J, Oliveberg M, Davis B, Fersht AR. Perturbed pKa-values in the denatured states of proteins. *Journal of Molecular Biology*, 254(5), 980-992 (1995).
233. Pots AM, de Jongh HH, Gruppen H, Hessing M, Voragen AG. The pH dependence of the structural stability of patatin. *Journal of Agricultural and Food Chemistry*, 46(7), 2546-2553 (1998).
234. Goto Y, Calciano LJ, Fink AL. Acid-induced folding of proteins. *Proceedings of the National Academy of Sciences*, 87(2), 573-577 (1990).
235. Fink AL, Calciano LJ, Goto Y, Kurotsu T, Palleros DR. Classification of acid denaturation of proteins: intermediates and unfolded states. *Biochemistry*, 33(41), 12504-12511 (1994).
236. Juneja J, Bhavesh NS, Udgaonkar JB, Hosur RV. NMR identification and characterization of the flexible regions in the 160 kDa molten globule-like aggregate of barstar at low pH. *Biochemistry*, 41(31), 9885-9899 (2002).
237. Shera JN, Takahashi D, Herrera AI, Prakash O, Sun XS. Morphology and Structural Properties of pH-Responsive Amphiphilic Peptides. *Journal of Nanoscience and Nanotechnology*, 10(12), 7981-7987 (2010).
238. Whitney PL, Tanford C. Solubility of amino acids in aqueous urea solutions and its implications for the denaturation of proteins by urea. *Journal of Biological Chemistry*, 237(5), PC1735-PC1737 (1962).
239. Bhavesh NS, Juneja J, Udgaonkar JB, Hosur RV. Native and nonnative conformational preferences in the urea-unfolded state of barstar. *Protein Science*, 13(12), 3085-3091 (2004).
240. Brandts JF. The thermodynamics of protein denaturation. I. The denaturation of chymotrypsinogen. *Journal of the American Chemical Society*, 86(20), 4291-4301 (1964).
241. Thomas T, Cavicchioli R. Effect of temperature on stability and activity of elongation factor 2 proteins from Antarctic and thermophilic methanogens. *Journal of Bacteriology*, 182(5), 1328-1332 (2000).



242. Barnwal RP, Devi KM, Agarwal G, Sharma Y, Chary KV. Temperature-dependent oligomerization in M-crystallin: Lead or lag toward cataract, an NMR perspective. *Proteins: Structure, Function, and Bioinformatics*, 79(2), 569-580 (2011).
243. Sukthankar P, Whitaker SK, Garcia M, Herrera A, Boatwright M, Prakash O, Tomich JM. Thermally induced conformational transitions in nascent branched amphiphilic peptide capsules. *Langmuir*, 31(10), 2946-2955 (2015).
244. Nick Pace C, Scholtz JM, Grimsley GR. Forces stabilizing proteins. *FEBS Letters*, 588(14), 2177-2184 (2014).
245. Krishna AG, Menon ST, Terry TJ, Sakmar TP. Evidence that helix 8 of rhodopsin acts as a membrane-dependent conformational switch. *Biochemistry*, 41(26), 8298-8309 (2002).
246. Kanuru M, Raman R, Aradhyam GK. Serine protease activity of calnuc regulation by Zn<sup>2+</sup> and G proteins. *Journal of Biological Chemistry*, 288(3), 1762-1773 (2013).
247. Holzwarth G, Doty P. The ultraviolet circular dichroism of polypeptides. *Journal of the American Chemical Society*, 87(2), 218-228 (1965).
248. Greenfield NJ, Fasman GD. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry*, 8(10), 4108-4116 (1969).
249. Strickland EH, Beychok S. Aromatic contributions to circular dichroism spectra of protein. *CRC Critical Reviews in Biochemistry*, 2(1), 113-175 (1974).
250. Kelly SM, Price NC. The use of circular dichroism in the investigation of protein structure and function. *Current Protein and Peptide Science*, 1(4), 349-384 (2000).
251. Greenfield NJ. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nature Protocols*, 1(6), 2527-2535 (2006).
252. Raman R, Ptak CP, Hsieh C-L, Oswald RE, Chang Y-F, Sharma Y. The perturbation of tryptophan fluorescence by phenylalanine to alanine mutations identifies the hydrophobic core in a subset of bacterial Ig-like domains. *Biochemistry*, 52(27), 4589-4591 (2013).
253. Haq SK, Rasheedi S, Khan RH. Characterization of a partially folded intermediate of stem bromelain at low pH. *The FEBS Journal*, 269(1), 47-52 (2002).
254. Hawe A, Sutter M, Jiskoot W. Extrinsic fluorescent dyes as tools for protein characterization. *Pharmaceutical Research*, 25(7), 1487-1499 (2008).

255. Semisotnov G, Rodionova N, Razgulyaev O, Uversky V, Gripas A, Gilmanshin R. Study of the “molten globule” intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers*, 31(1), 119-128 (1991).
256. Singh S, Roy R. The application of absolute quantitative  $^1\text{H}$  NMR spectroscopy in drug discovery and development. *Expert Opinion on Drug Discovery*, 11(7), 695-706 (2016).
257. Bharti SK, Roy R. Quantitative  $^1\text{H}$  NMR spectroscopy. *TrAC Trends in Analytical Chemistry*, 35, 5-26 (2012).
258. Ganguly AK, Ranjan P, Kumar A, Bhavesh NS. Dynamic association of PfEMP1 and KAHRP in knobs mediates cytoadherence during Plasmodium invasion. *Scientific Reports*, 5 (2015).
259. Kashyap M, Ganguly AK, Bhavesh NS. Structural delineation of stem-loop RNA binding by human TAF15 protein. *Scientific Reports*, 5, srep17298 (2015).
260. Sackewitz M, Scheidt HA, Lodderstedt G, Schierhorn A, Schwarz E, Huster D. Structural and dynamical characterization of fibrils from a disease-associated alanine expansion domain using proteolysis and solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 130(23), 7172-7173 (2008).
261. Takahashi D, Hiromasa Y, Kim Y, Anbanandam A, Yao X, Chang KO, Prakash O. Structural and dynamics characterization of norovirus protease. *Protein Science*, 22(3), 347-357 (2013).
262. Keramisanou D, Biris N, Gelis I, Sianidis G, Karamanou S, Economou A, Kalodimos CG. Disorder-order folding transitions underlie catalysis in the helicase motor of SecA. *Nature Structural & Molecular Biology*, 13(7), 594-602 (2006).
263. Ohki S-y, Kainosho M. Stable isotope labeling methods for protein NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 53(4), 208-226 (2008).
264. LeMaster DM, Richards FM. NMR sequential assignment of Escherichia coli thioredoxin utilizing random fractional deuteration. *Biochemistry*, 27(1), 142-150 (1988).
265. Kainosho M, Tsuji T. Assignment of the three methionyl carbonyl carbon resonances in Streptomyces subtilisin inhibitor by a carbon-13 and nitrogen-15 double-labeling technique. A new strategy for structural studies of proteins in solution. *Biochemistry*, 21(24), 6273 (1982).
266. Vuister GW, Kim S-J, Wu C, Bax A. 2D and 3D NMR-study of phenylalanine residues in proteins by reverse isotopic labeling. *Journal of the American Chemical Society*, 116(20), 9206-9210 (1994).

267. Wang H, Janowick DA, Schkeryantz JM, Liu X, Fesik SW. A method for assigning phenylalanines in proteins. *Journal of the American Chemical Society*, 121(7), 1611-1612 (1999).
268. Yamazaki T, Otomo T, Oda N, Kyogoku Y, Uegaki K, Ito N, Ishino Y, Nakamura H. Segmental isotope labeling for protein NMR using peptide splicing. *Journal of the American Chemical Society*, 120(22), 5591-5592 (1998).
269. Staunton D, Schlinkert R, Zanetti G, Colebrook SA, Campbell ID. Cell-free expression and selective isotope labelling in protein NMR. *Magnetic Resonance in Chemistry*, 44(S1) (2006).
270. Tugarinov V, Kay LE. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *Journal of the American Chemical Society*, 125(45), 13868-13878 (2003).
271. Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, Güntert P. Optimal isotope labelling for NMR protein structure determinations. *Nature*, 440(7080), 52-57 (2006).
272. Schubert M, Manolikas T, Rogowski M, Meier BH. Solid-state NMR spectroscopy of 10%  $^{13}\text{C}$  labeled ubiquitin: spectral simplification and stereospecific assignment of isopropyl groups. *Journal of biomolecular NMR*, 35(3), 167-173 (2006).
273. LeMaster DM, LaIuppa JC, Kushlan DM. Differential deuterium isotope shifts and one-bond  $^1\text{H}$ - $^{13}\text{C}$  scalar couplings in the conformational analysis of protein glycine residues. *Journal of Biomolecular NMR*, 4(6), 863-870 (1994).
274. Tzeng S-R, Pai M-T, Kalodimos CG. NMR studies of large protein systems. *Protein NMR Techniques*, 133-140 (2012).
275. Yang L-Q, Sang P, Tao Y, Fu YX, Zhang KQ, Xie YH, Liu SQ. Protein dynamics and motions in relation to their functions: several case studies and the underlying mechanisms. *Journal of Biomolecular Structure and Dynamics*, 32(3), 372-393 (2014).
276. Mittag T, Kay LE, Forman-Kay JD. Protein dynamics and conformational disorder in molecular recognition. *Journal of Molecular Recognition*, 23(2), 105-116 (2010).
277. Ferreira ST, De Felice FG. Protein dynamics, folding and misfolding: from basic physical chemistry to human conformational diseases. *FEBS Letters*, 498(2-3), 129-134 (2001).
278. Mandell JG, Falick AM, Komives EA. Identification of protein-protein interfaces by decreased amide proton solvent accessibility. *Proceedings of the National Academy of Sciences*, 95(25), 14705-14710 (1998).

279. Englander SW, Sosnick TR, Englander JJ, Mayne L. Mechanisms and uses of hydrogen exchange. *Current Opinion in Structural Biology*, 6(1), 18-23 (1996).
280. Bollen YJ, Kamphuis MB, van Mierlo CP. The folding energy landscape of apoflavodoxin is rugged: hydrogen exchange reveals nonproductive misfolded intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, 103(11), 4095-4100 (2006).
281. Takahashi H, Nakanishi T, Kami K, Arata Y, Shimada I. A novel NMR method for determining the interfaces of large protein–protein complexes. *Nature Structural & Molecular Biology*, 7(3), 220-223 (2000).
282. Stilbs P. Molecular self-diffusion coefficients in Fourier transform nuclear magnetic resonance spectrometric analysis of complex mixtures. *Analytical Chemistry*, 53(13), 2135-2137 (1981).
283. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry*, 38(50), 16424-16431 (1999).
284. Yao S, Weber DK, Separovic F, Keizer DW. Measuring translational diffusion coefficients of peptides and proteins by PFG-NMR using band-selective RF pulses. *European Biophysics Journal*, 43(6-7), 331-339 (2014).
285. Dehner A, Kessler H. Diffusion NMR spectroscopy: folding and aggregation of domains in p53. *Chembiochem*, 6(9), 1550-1565 (2005).
286. Cameron KS, Fielding L. NMR diffusion coefficient study of steroid–cyclodextrin inclusion complexes. *Magnetic Resonance in Chemistry*, 40(13) (2002).
287. Jones JA, Wilkins DK, Smith LJ, Dobson CM. Characterisation of protein unfolding by NMR diffusion measurements. *Journal of Biomolecular NMR*, 10(2), 199-203 (1997).
288. Krishnan V. Determination of oligomeric state of proteins in solution from pulsed-field-gradient self-diffusion coefficient measurements. A comparison of experimental, theoretical, and hard-sphere approximated values. *Journal of Magnetic Resonance*, 124(2), 468-473 (1997).
289. Heisel KA, Goto JJ, Krishnan VV. NMR chromatography: molecular diffusion in the presence of pulsed field gradients in analytical chemistry applications. *American Journal of Analytical Chemistry*, 3, 401-409 (2012).

290. Ohnishi M, Urry D. Temperature dependence of amide proton chemical shifts: the secondary structures of gramicidin S and valinomycin. *Biochemical and Biophysical Research Communications*, 36(2), 194-202 (1969).
291. Baxter NJ, Williamson MP. Temperature dependence of  $^1\text{H}$  chemical shifts in proteins. *Journal of Biomolecular NMR*, 9(4), 359-369 (1997).
292. Cierpicki T, Zhukov I, Byrd RA, Otlewski J. Hydrogen bonds in human ubiquitin reflected in temperature coefficients of amide protons. *Journal of Magnetic Resonance*, 157(2), 178-180 (2002).
293. Williamson MP, Hall MJ, Handa BK.  $^1\text{H}$ -NMR assignment and secondary structure of a Herpes simplex virus glycoprotein D-1 antigenic domain. *The FEBS Journal*, 158(3), 527-536 (1986).
294. Baxter NJ, Hosszu LL, Waltho JP, Williamson MP. Characterisation of low free-energy excited states of folded proteins. *Journal of Molecular Biology*, 284(5), 1625-1639 (1998).
295. Williamson MP. Many residues in cytochrome c populate alternative states under equilibrium conditions. *Proteins: Structure, Function, and Bioinformatics*, 53(3), 731-739 (2003).
296. Dunn JJ, Studier FW, Gottesman M. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *Journal of Molecular Biology*, 166(4), 477-535 (1983).
297. Jeruzalmi D, Steitz TA. Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. *The EMBO Journal*, 17(14), 4101-4113 (1998).
298. Moffatt BA, Studier FW. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell*, 49(2), 221-227 (1987).
299. Zhang X, Studier FW. Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. *Journal of Molecular Biology*, 269(1), 10-27 (1997).
300. Inouye M, Arnheim N, Sternglanz R. Bacteriophage T7 lysozyme is an N-acetylmuramyl-L-alanine amidase. *Journal of Biological Chemistry*, 248(20), 7247-7252 (1973).
301. Hattman S, Ives J, Margolin W, Howe MM. Regulation and expression of the bacteriophage mu mom gene: mapping of the transactivation (dad) function to the C region. *Gene*, 39(1), 71-76 (1985).
302. Studier FW. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology*, 219(1), 37-44 (1991).

303. Fuerst TR, Niles EG, Studier FW, Moss B. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences*, 83(21), 8122-8126 (1986).
304. Dunn JJ, Krippel B, Bernstein KE, Westphal H, Studier FW. Targeting bacteriophage T7 RNA polymerase to the mammalian cell nucleus. *Gene*, 68(2), 259-266 (1988).
305. Tabor S, Richardson CC. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Sciences*, 82(4), 1074-1078 (1985).



## Mechanistic insights into pH-dependent structure-function relationship of T7 endolysin

### Abstract

Lysis activity profiles of several of the T7 family members had been known for many years, but the molecular basis for their pH-dependent differential activity is not clear. In this chapter, T7 endolysin (T7L) was expressed, purified, and its pH-induced structural, and activity characteristics were investigated by applying a variety of biophysical techniques, and protein nuclear magnetic resonance (NMR) spectroscopy. Detailed studies established a reversible structural transition of T7L below pH 6 and the formation of a partially folded conformation(s) at low pH. This low-pH conformation exposed its hydrophobic pockets and heterogeneous in nature. A network of buried His residues in a constrained environment is responsible for the observed pH-dependent conformational dynamics and transitions. These results provided insights into pH-dependent structure-function relationship of T7L like family of bacteriophage amidases.

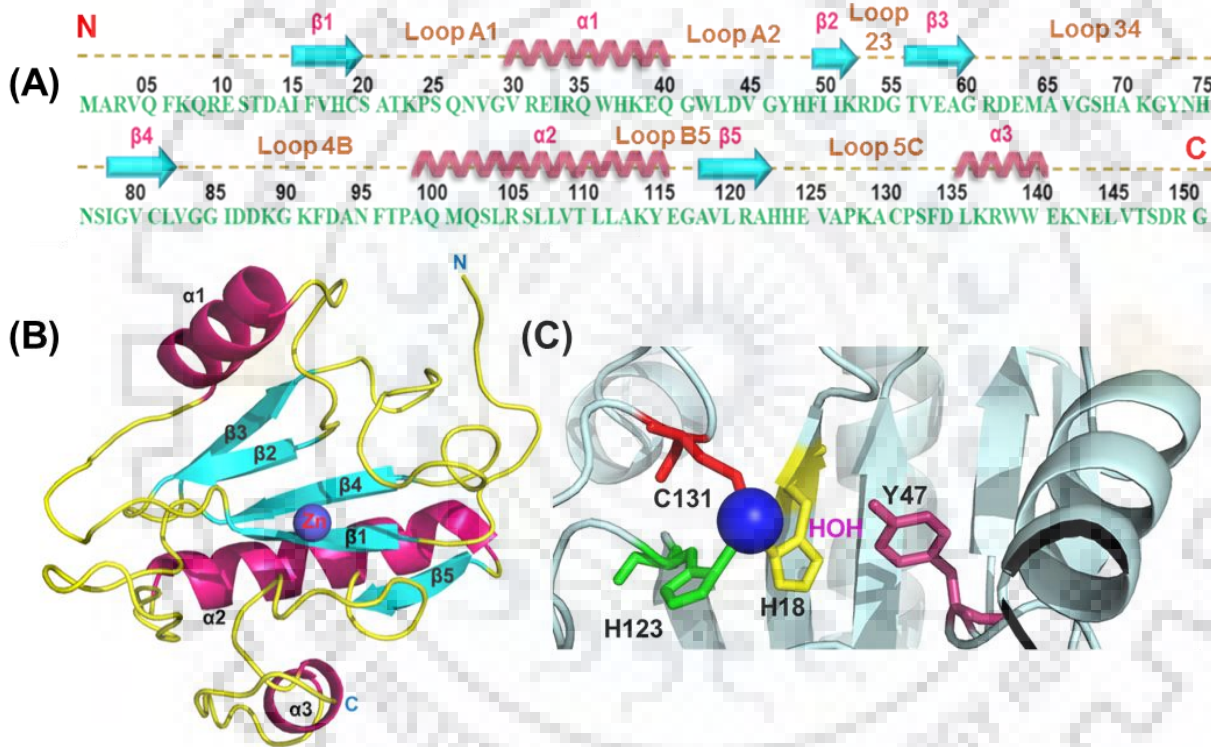
### 2.1 Introduction

Among several classes of endolysins, amidases (also called N-acetylmuramoyl-L-alanine amidases) are known for their rapid lysis activity [1]. Majority of endolysins infecting Gram-negative bacteria are 15–20 kDa single-domain globular proteins consisting of a single catalytic domain [2, 3]. Bacteriophages such as T7, T3, K11, *Thermus scotoductus* phage vB\_Tsc2631, *Thermus scotoductus* MAT2119 bacteriophage Ph2119, *Bacillus anthracis* prophage etc., contain single-domain amidase family endolysins [4-9]. Structurally these proteins share the conserved  $\alpha$ - $\beta$  fold as described for T7L [4, 9].

T7 endolysin (T7L) is a 151 residue long (17 kDa) bifunctional zinc amidase protein with a pI of 8.8 [9, 10] (**Figure 2.1 A**). It has three  $\alpha$ -helices ( $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$ ) and five  $\beta$ -sheets, out of these four are parallel ( $\beta 1$ ,  $\beta 2$ ,  $\beta 4$  and  $\beta 5$ ) and one is antiparallel ( $\beta 3$ ). Loops connected these structural elements are named as A1, A2, 23, 34, 4B, B5 and 5C (**Figure 2.1 B**). A zinc atom ( $\text{Zn}^{+2}$ ) is present in a cleft, and anchored to T7L by connecting with H18, H123 and C131, and Y47 through a water molecule in a tetrahedral geometry (**Figure 2.1 C**). In addition to Zn binding, the amino acids H18,

H123 and C131, K128 and Y47 also establish substrate specific interactions, thus playing a crucial role in amidase activity of protein [4].

Functional studies of these endolysins from different phages such as T7, K11, ph2119, etc., established their maximal lysis activity in the pH range of 6.5–8.5 [7, 9, 11-13]. Further, studies also demonstrated that their functionality is very stringent toward pH changes. All these enzymes are reported to be inactive in lysis assays below pH 6.0. However, their pH-dependent conformational changes have not been understood yet. In this chapter, T7L protein has been expressed, purified and characterized to delineate the pH-dependent structure-function relationship of amidase family of bacteriophage endolysins.



**Figure 2.1:** Structural details of T7L: (A) Complete amino acid sequence with secondary structural elements (B) Three-dimension structural representation (PDB ID: 1LBA) and (C) Amino acids involved in Zn binding. These structures were generated using PyMOL molecular graphics software.

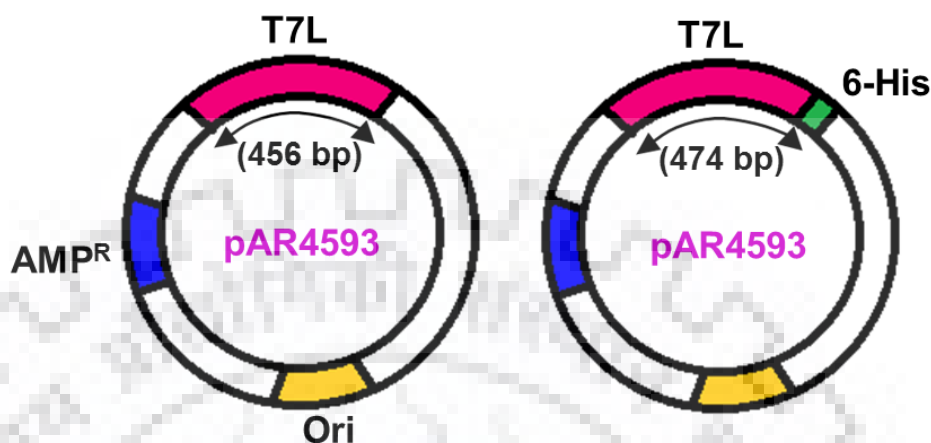
## 2.2 Materials and methods

### 2.2.1 Construct design:

pAR4593 plasmid (pET-3 origin) containing T7L gene sequence (NCBI reference sequence NP\_041973.1) was used as source of wild type T7L expression [14]. To ease the purification process the pAR4593 plasmid was modified by inserting 6-His residues at the C-terminal of T7L sequence



and expressed as His-tag protein. The construct designs of both the vectors are presented in **Figure 2.2**.



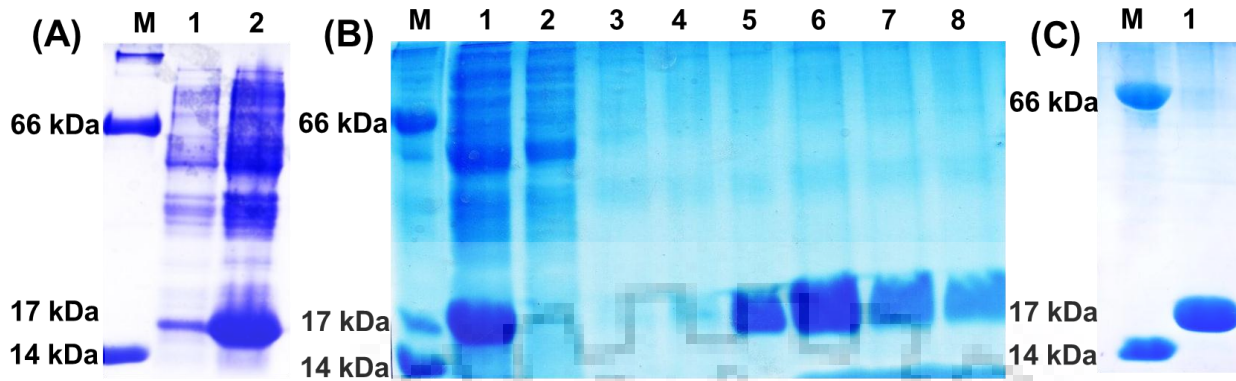
**Figure 2.2:** Schematic representation of vectors used for T7L expression and purifications.

### 2.2.2 Overexpression and purification

*E. coli* BL21 (DE3) cells were transformed with a plasmid containing the T7L gene. For expression of T7L protein(s), the cells were inoculated into 10 mL of Luria-Bertani (LB) medium containing 100 µg/mL ampicillin as an overnight seed culture. The large culture (1000 mL) was grown at 37 °C and 220 rpm to an OD<sub>600</sub> of 0.4–0.5 using the saturated overnight 10 mL seed culture. The culture was then induced with 0.5 mM isopropyl β- D-1-thiogalactopyranoside (IPTG), and incubated at 37 °C and 220 rpm for additional 4 h. Cells were harvested by centrifugation at 6000 rpm for 10 min, resuspended in 50 mL of lysis buffer [50 mM Tris and 500 mM NaCl (pH 8.0)] and treated with 100 mM of phenylmethylsulfonyl fluoride, (PMSF) and 1 mM of β-Mercaptoethanol (BME). Cells were sonicated and then centrifuged at 14000 rpm for 1 h at 4 °C.

#### **Purification of T7L with His tag**

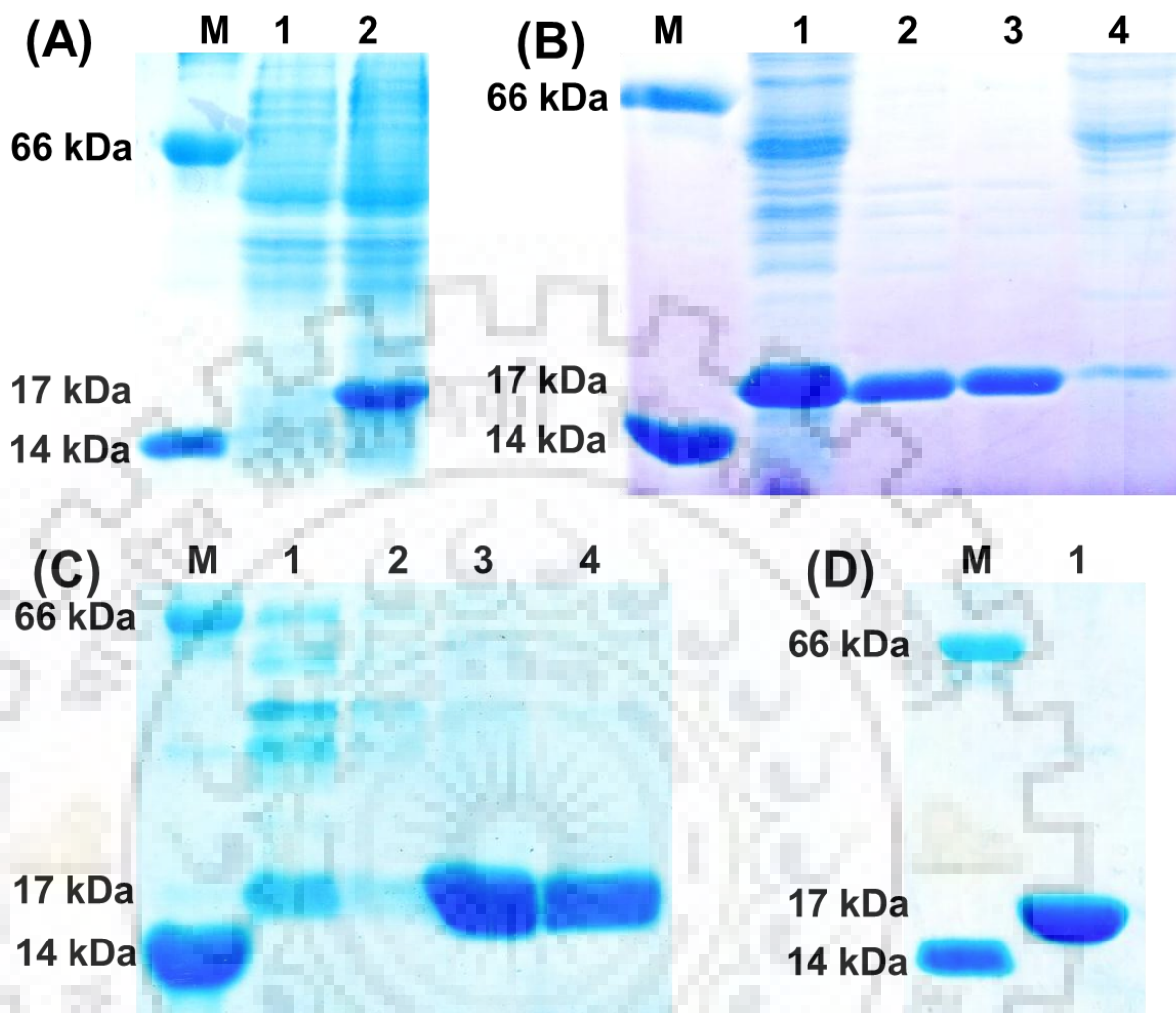
Protein was purified by affinity chromatography using Ni-NTA column. The supernatant was applied directly to the Ni-NTA column, equilibrated with lysis buffer, and passed three to four times through the column to allow the binding of protein. The nonspecific protein impurities were removed by washing the column with an increasing gradient of imidazole in lysis buffer (10, 30, and 50 mM; 50 mL each). The bound protein was eluted with 30–50 mL of 400 mM imidazole. Fractions (10 mL) were collected and analyzed by electrophoresis on a 12 % SDS–PAGE gel (**Figure 2.3 A-B**).



**Figure 2.3:** 12 % SDS-PAGE gel analysis of T7L-His tag purification from affinity chromatography. Lane M represents the marker (BSA=66 kDa, HEL=14 kDa): (A) Induction gel profile, Lane 1=un-induced and Lane 2=induced sample (B) Purification from Ni-NTA column, Lane 1=cell lysate (supernatant), Lane 2=flow through fraction, Lane 3= 30 mM imidazole wash, Lane 4=50 mM imidazole+1 M NaCl wash, Lane 5-8=eluted protein fraction with 400 mM imidazole and (C) Pure protein after gel filtration chromatography.

#### **Purification of T7L without His tag**

Protein was purified by ion-exchange chromatography using DEAE and CM-cellulose columns (Bio-Red). The supernatant was applied directly onto DEAE column equilibrated with buffer containing 20 mM phosphate buffer (pH 8), 1 mM EDTA, 4 % glycerol and 10 mM NH<sub>4</sub>Cl. The flow through (FT) was collected after passing the supernatant for three or four times through the column. Column was then washed with equilibration buffer (W1) and equilibration buffer + 0.5 M NaCl (W2) containing BME. Protein was found to be present in FT and W1 fractions as confirmed by electrophoresis on a 12% SDS-PAGE gel (**Figure 2.4 A and B**). These two protein fractions were then mixed together, pH was adjusted to 6 and centrifuged at 14000 rpm/20 min/4 °C. The supernatant was adsorbed onto CM-Cellulose column pre-equilibrated with 20 mM phosphate buffer (pH 6) and 4 % glycerol. Supernatant was passed three or four times through the column to allow the binding of protein. The bound protein was washed with NaCl (100 mM and 250 mM), and eluted with 20–30 mL of 500 mM NaCl. Protein fractions were collected and analysed by electrophoresis on a 12% SDS-PAGE gel (**Figure 2.4 C**). Untagged T7L exclusively used for the experiments involving Zn for Zn-dependent CD experiments (no EDTA) and cell lysis assays.

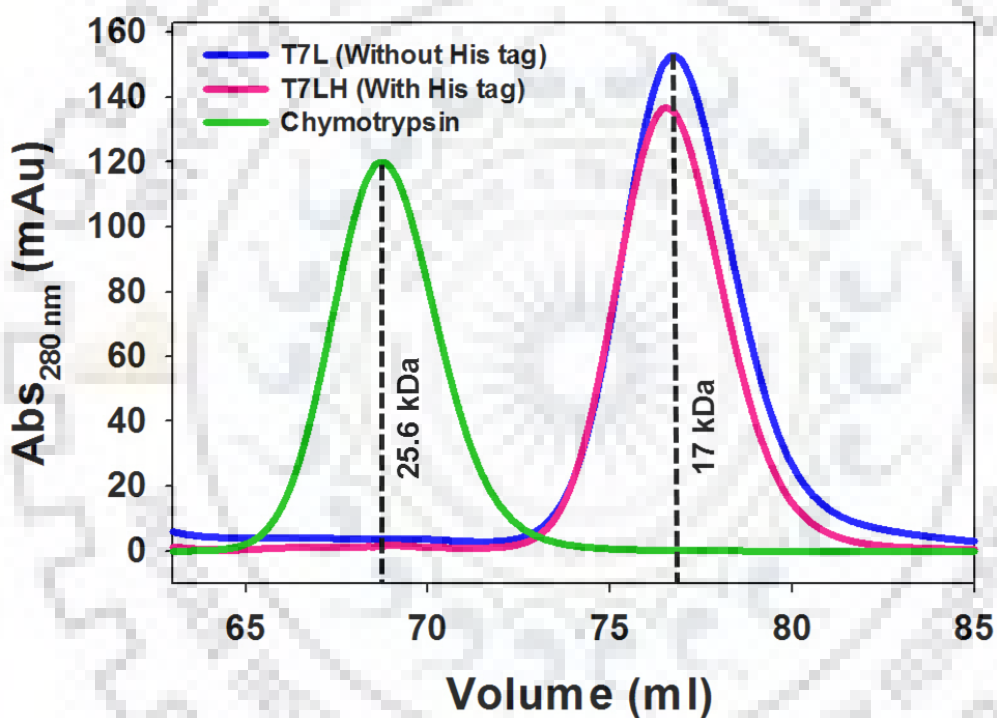


**Figure 2.4:** 12 % SDS-PAGE gel analysis of T7L purification from ion-exchange chromatography. Lane M represents the marker (BSA=66 kDa, HEL=14 kDa): (A) Induction gel profile, Lane 1=un-induced, Lane 2=Induced samples (B) DEAE column purification profile, Lane 1=cell lysate, Lane 2=flow through fraction, Lane 3=wash 1 with equilibration buffer and Lane 4=wash 2 with 500 mM NaCl (C) CM-Cellulose column purification profile, Lane 1=flow through fraction, Lane 2=wash 1 with 100 mM NaCl, Lane 3= wash 2 with 250 mM NaCl and Lane 4=elution with 500 mM NaCl and (D) Pure protein after gel filtration chromatography.

Fractions of T7L with / without His tag were dialyzed overnight in 20 mM Tris-acetate/20 mM glycine-phosphate buffer (pH 7.0) containing 0.1 M NaCl and 5 mM EDTA. For NMR experiments, T7L was grown in M9 minimal medium containing  $^{15}\text{NH}_4\text{Cl}$  and purified as described above.

After affinity/ion-exchange chromatographic purification steps, T7L protein was further purified by size exclusion chromatography (SEC) using a 120 mL HiLoad 16/60 Superdex 75 prep

grade column (GE Healthcare) on an AKTA prime FPLC system (GE Healthcare) in 20 mM Tris buffer containing 0.1 M NaCl, 5 mM EDTA and 1 % glycerol (pH 7) with a flow rate of 1 mL/min (**Figure 2.3 C, Figure 2.4 D**). The Chymotrypsin from bovine pancreas (25.6 kDa, obtained from SRL India) was used as a reference. The FPLC elution of both untagged and His-tagged T7L proteins yields similar profile as a monomeric protein (**Figure 2.5**), suggesting very little influence of His tag on the structural properties of T7L. After performing gel filtration chromatography the protein fractions were concentrated by ultrafiltration using Centricon tubes (10 kDa cut off membrane-Millipore). Concentrations of proteins were measured by UV absorbance at 280 nm using a molar extinction coefficient of  $26470 \text{ M}^{-1} \text{ cm}^{-1}$ .



**Figure 2.5:** FPLC chromatogram showing similar elution profiles of tag and without His-tag T7L proteins.

All the chemicals used in purification are procured from SRL.

### 2.2.3 Cell lysis turbidimetric assay

*E. coli* BL21 (DE3) cells were cultivated in 10 mL of LB medium and allowed to grow at 37 °C until the mid-exponential phase ( $OD_{600}$  of 0.6–0.8). The cells were treated with 0.1 M EDTA for 5 min at room temperature and divided into 1 mL aliquots, pelleted by a brief centrifugation (1 min),

and resuspended in 1 mL of 10 mM potassium phosphate-acetate buffer containing 10  $\mu\text{M}$   $\text{ZnSO}_4$  at different pH (pH 8 to 3). pH-dependent amidase activity of T7L was measured by turbidimetric assay by adding 50  $\mu\text{g}$  of protein into 500  $\mu\text{L}$  of bacterial cell suspensions. As a measure of lysis efficiency, decrease in turbidity of cell suspension was monitored at every 2 min for a period of 12 min at 600 nm using a UV-Vis spectrophotometer. Activity measurements were performed in triplicates. Sensitized bacterial cells without T7L was used as a control.

#### **2.2.4 Circular dichroism (CD) spectroscopy**

CD measurements at pH 8 to 3, (20 mM glycine-phosphate, 5 mM EDTA and 0.1 M NaCl, 25  $^\circ\text{C}$ ) were carried out on a Peltier controlled Jasco J-1500 CD spectrophotometer. Far-UV CD spectra of T7L (20  $\mu\text{M}$ ) were recorded from 190-250 nm using 1 mm quartz cuvette. CD experiments with zinc-loaded T7L were performed in the presence of 100  $\mu\text{M}$   $\text{ZnSO}_4$  (without EDTA). The secondary structural contents of the T7L spectra was calculated using online tool DICHROWEB-K2D (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>) [15]. Near-UV CD measurements of T7L protein (150  $\mu\text{M}$ ) were recorded from 240-320 nm in a 10 mm quartz cuvette.

#### **2.2.5 Fluorescence spectroscopy**

Fluorescence spectra of T7L (20  $\mu\text{M}$ ) from pH 8 to 3 were recorded on a Fluorolog (HORIBA JOBIN YVON) spectrophotometer using a 4 mm path length quartz cuvette with a 5 nm band-pass for both excitation and emission at 25  $^\circ\text{C}$  with an acquisition interval of 1 nm/sec. Intrinsic protein fluorescence spectra of the protein were recorded with tryptophan (Trp) by exciting the protein samples at 295 nm. 8-anilino-1-naphthalenesulfonic acid (ANS-Sigma) was used as an extrinsic fluorophore, and binding measurements were recorded by exciting the samples at 380 nm. The final concentration of ANS was 100  $\mu\text{M}$  with a protein:ANS molar ratio of 1:5.

#### **2.2.6 Fluorescence lifetime measurements**

Time resolved fluorescence decays of T7L (20  $\mu\text{M}$ ) from pH 8 to 3 were measured using a time-correlated single photon counting (TCSPC) spectrofluorometer (Horiba Jobin Yvon Fluoro Cube Fluorescence Lifetime System) equipped with Nano LED as pulse excitation source. The excitation of tryptophan was kept at 295 nm and emission was detected at 350 nm. T7L samples were incubated with 100  $\mu\text{M}$  of ANS were excited at a fixed wavelength of 375 nm and the

fluorescence emission was detected at 472 nm. The decay was measured using a 4 mm quartz cuvette at a scan speed of 150 nm/s with  $\pm 0.5$  nm accuracy. The data points were fitted to tri-exponential model using a re-convolution method and the quality of the best fit was confirmed by chi-square ( $\chi^2$ ) value.

### **2.2.7 Nuclear magnetic resonance (NMR) spectroscopy**

NMR experiments were recorded at Bruker Avance 500 MHz spectrometer equipped with a TXI probe at 25 °C. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR experiments were performed with 48 scans and 256 increments in the indirect dimension using 700  $\mu\text{M}$   $^{15}\text{N}$ -labeled T7L (20 mM glycine-phosphate, 0.1 M NaCl, 5 mM EDTA, and 10 %  $\text{D}_2\text{O}$  (Sigma)). The spectral widths in the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions were 16 and 32 ppm, respectively. For HSQC experiments QSINE window function was used. To determine the reversibility of pH-induced transitions, T7L sample was brought from pH 7 to 3 by adjusting the pH with a microelectrode and then exchanged back to pH 7.

### **2.2.8 Measurement of translational diffusion characteristics of T7L using 2D-DOSY**

For translational diffusion measurements,  $^1\text{H}$  based 2D DOSY (diffusion ordered spectroscopy) experiments for T7L at different pH values (25 °C) were performed on an 800 MHz NMR spectrometer with an actively shielded gradient unit capable of generating a nominal maximum gradient strength of 53 G/cm. The diffusion coefficients were obtained using the widely accepted stimulated echo bipolar gradient pulse program “ledbpgppr2s” (from the Bruker library). For each experiment, a diffusion delay ( $\Delta$ ) of 200–300 ms was used and sine-shaped pulse field gradients (PFGs), with a duration of 3.0 ms ( $\delta/2$ ), followed by a recovery delay of 100  $\mu\text{s}$  were applied with amplitudes varying linearly from 5 to 95% of the maximal strength in 64 equally spaced steps. Each experiment consisted of the accumulation of 32 scans (with a recycle delay of 2 s) and lasted for  $\sim 100$  min. After Fourier transformation and baseline correction, the DOSY spectra were processed using “eddosy” and “dosy2d” analysis modules of the Bruker Topspin software.

Experiments were recorded on T7L protein samples along with the known standard proteins [chicken SH3 domain (SH3, MW  $\sim 7.2$  kDa, obtained from Cambridge Isotopes Limited), hen egg lysozyme (HEL, MW  $\sim 14.3$  kDa, obtained from Sigma), and human serum albumin (HSA, MW  $\sim 65.2$ , obtained from Sigma)] as a reference.

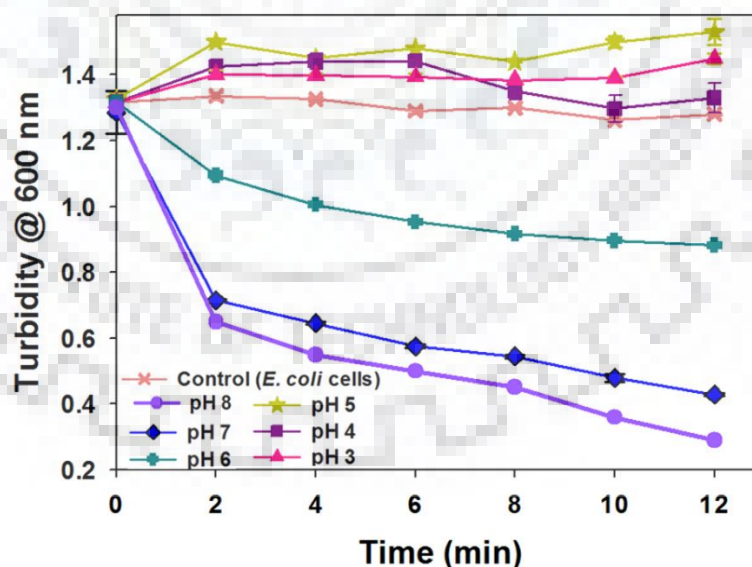
### 2.2.9 Calculation of buried index and pKa Values:

Buried index and pKa values of histidine (His) residues in T7L (PDB ID: 1LBA). The buried index was calculated using online program ASAview (<http://asaview.netasa.org/>) [16]. The pKa values were estimated employing online webserver DEPTH (<http://cospi.iiserpune.ac.in/depth/htdocs/index.html>) [17].

## 2.3 Results

### 2.3.1 pH-dependent enzymatic efficiency of T7L endolysin

Literature reports suggested that the enzymatic activity of T7L drastically drops with decrease in pH, and it shows low/negligible lysis activity below pH 6 [11]. To unravel the functional characteristics of T7L, cell based turbidimetric lysis assay was performed at different pH values (pH 8-3) as defined by the decrease in absorbance of sensitized *E. coli* cells at 600 nm. The suspension of *E. coli* cells alone has been taken as a control. The cells treated with T7L at pH 7 and 8 results in significant decrease in OD compared to pH 6. However no change in absorbance was found at lower pH (5-3) as compared to the control (**Figure 2.6**). These results delineated the pH-dependent activity profile of T7L. The observed results are in line with the previously reported functional studies on bacteriophage amidases [7, 9, 11, 12].

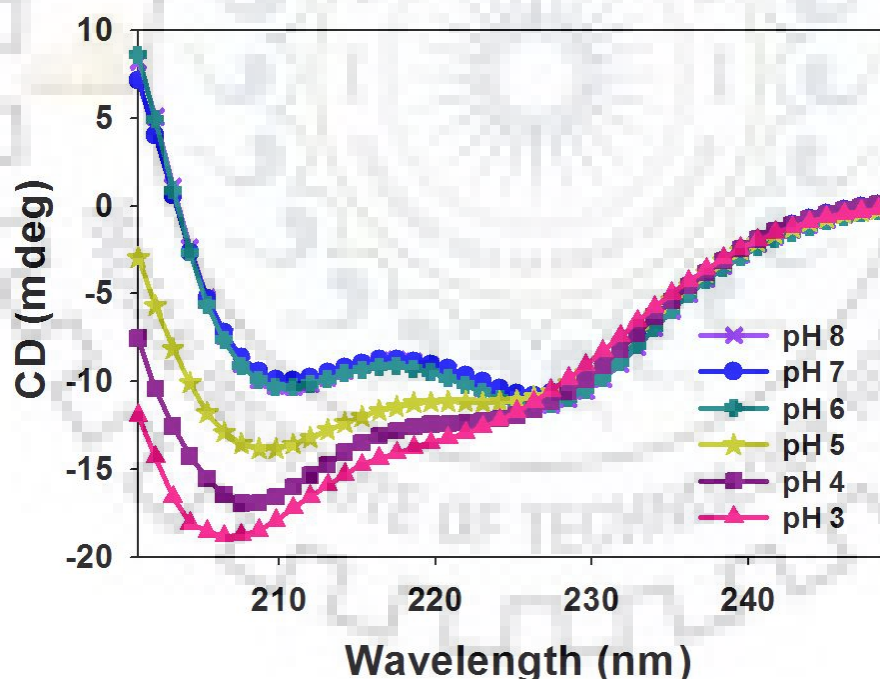


**Figure 2.6:** Amidase activity of T7L measured via the turbidimetric assay. The optical density (OD) measurements (at 600 nm) are presented for T7L at pH 8-3 (Error values are between  $\pm 0.01$  and  $\pm 0.04$ ).

The functional study clearly depicted that the pH alters the lysis activity of T7L. In order to understand the pH-dependent functional characteristics, it is essential to decipher the structural modulations of T7L under same experimental conditions. In the following sections, the pH-dependent conformational changes of T7L were assessed using a variety of biophysical techniques.

### 2.3.2 pH-dependent secondary structural modulations of T7 endolysin

Circular dichroism (CD) spectroscopy is a very sensitive technique for measuring the gross structural features of the proteins and/or biomolecules. Far-UV CD in the range of 190–240 nm provides information about the changes that occurred in the secondary structural elements. Far-UV CD profiles of T7L were analysed at various pH values (pH 8–3) to probe the effect of pH on its secondary structure (**Figure 2.7**). The CD curves at pH 8, 7 and 6 overlaid well on each other, indicating a similar structural preference at these pH values. However, once the pH was decreased to  $\leq 5$ , melting of the secondary structure was observed as indicated by the substantial decrease in the ellipticity value at 222 nm, a measure of the overall helical structure content of the protein.



**Figure 2.7:** Far-UV CD profile of T7L at 25 °C showing the secondary structure transition from pH 8 to 3 in 20 mM glycine-phosphate buffer.

The ellipticity value at 208 nm, a scale for randomness or unfoldedness, had increased significantly. In order to obtain quantitative insights into the extent of structural changes ( $\alpha$ -helices



and  $\beta$ -sheets), the measured CD profiles were analysed and quantified using Dichroweb-K2D (**Table 2.1**). These results revealed that, T7L at pH 8 to 6 contain similar structural content, but decreasing of pH to  $\leq 5$  results in substantial decrease in the helix content ( $\sim 10\%$ ), which is in line with the observed spectral features.

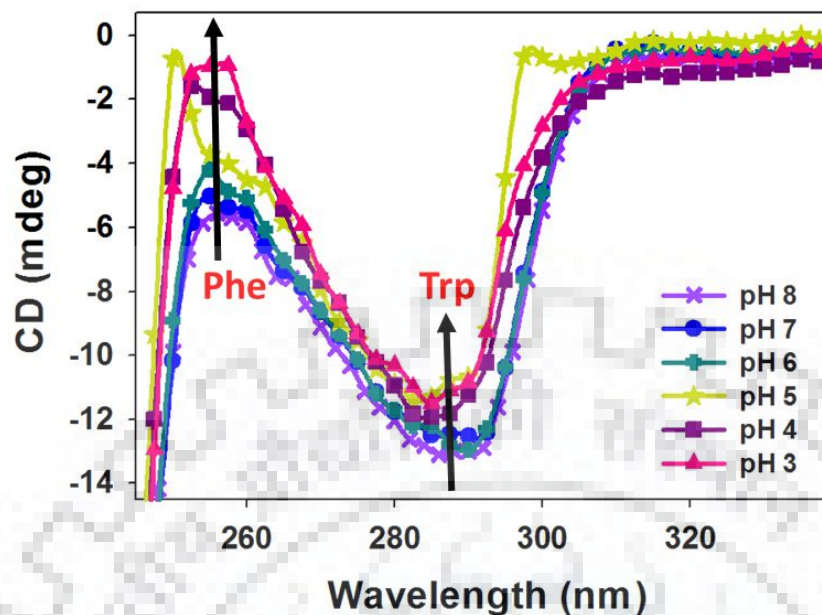
*Table 2.1: Secondary structure composition of T7L at various pH values calculated using DICHROWEB-K2D online software using default setting parameters.*

Category	Helix (%)	Sheet (%)	Coil (%)
pH 8	27	24	49
pH 7	26	24	50
pH 6	26	24	50
pH 5	17	28	55
pH 4	16	27	57
pH 3	17	26	57

### **2.3.3 pH-dependent tertiary structural changes in T7L**

Conformational differences at secondary structural level in general are accompanied with some alterations at tertiary structure level. Near-UV CD and fluorescence measurements can provide promising insights into the tertiary structural changes have been performed.

Near-UV CD studies are very sensitive to the locations of aromatic residues. The near-UV CD spectra showed the characteristic signals around 255 and 290 nm, which are attributed to Phe and Trp amino acids respectively (**Figure 2.8**). Intensities of dichroic signals at 290 nm due to Trp residues are similar at all measured pH values indicating little change in Trp environment. Moreover, such observations were supported by the solvent accessible surface area (SASA) calculations of the T7L, which indicated that three (W36, W42, and W140) out of four Trp residues (W36, W42, W139 and W140) in the T7L are solvent exposed (**Table 2.2**). However, the intensity of the negative band at 255 nm due to Phe varied significantly upon change in pH (**Figure. 2.8**).



**Figure 2.8:** Near-UV CD profile of T7L at 25 °C from pH 8 to 3 in 20 mM glycine-phosphate buffer. Arrows showing transition in ellipticity bands upon pH change.

**Table 2.2:** Solvent accessible surface area (SASA) of Trp and Phe residues calculated by using PDBePISA tool using PDB ID: 1LBA. Buried residues were marked with red color.

Residue	Residue No.	SASA (Å <sup>2</sup> )
Tryptophan	36	75.29
	42	123.11
	139	3.94
	140	75.90
Phenylalanine	6	59.64
	16	0.0
	49	22.77
	92	81.59
	96	12.20
	134	0.0

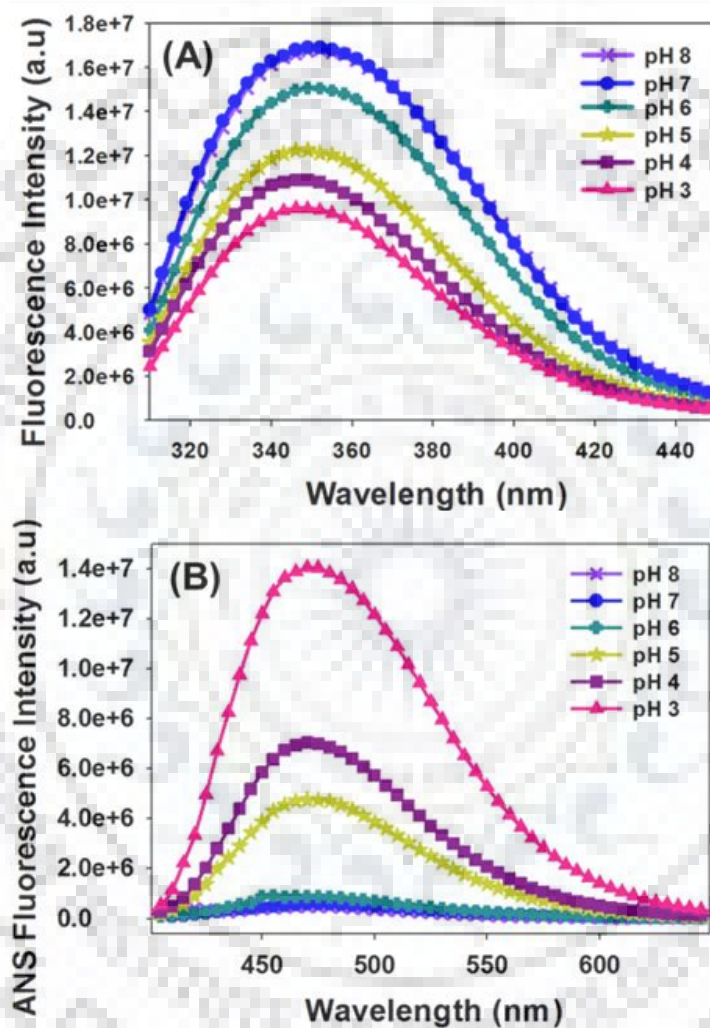
T7L contains six Phe residues (F6, F16, F49, F92, F96 and F134). SASA calculations suggested that except F6 and F92, all the other four Phe residues (F16, F49, F96 and F134) are very much buried in the hydrophobic core of the protein (**Table 2.2**). The decrease in ellipticity upon reducing pH is an indication of exposure of the hydrophobic pockets containing Phe residues.

The combined near-UV CD profiles of Trp and Phe along with their spatial location/SASA evidenced that the pH-dependent structural transition of T7L is responsible for the observed spectral changes. Thus, near-UV CD spectral features clearly distinguished the differences in the aromatic environment between native and low pH conformations. However, it is evident from various studies that partially folded or molten globule state of protein show the lack of rigid tertiary structure formed by the tight packing of amino acid residues [18-21]. Therefore, on the basis of such evidence, and the observed intensity decrease upon reducing the pH from 8 to 5/4/3 depicts the first line of evidence of presence of partially folded (PF) conformations at low pH (pH 5-3). Further, the differences observed in the spectral features of obtained near-UV CD profiles at pH 5-3 hints for a plausible structural heterogeneity among the low pH (5, 4 and 3) PF states of T7L.

#### ***Steady state fluorescence measurements***

In order to substantiate the pH-dependent differential tertiary structural characteristics of T7L, steady state fluorescence experiments have been performed using Trp as an intrinsic fluorophore, and ANS dye as an extrinsic fluorophore. The fluorescence intensity of an intrinsic Trp residue is sensitive to the micro-environment. The Trp fluorescence curves at pH 8 and 7 overlaid well each other and the maximum emission for T7L were observed at 352 nm. A gradual decrease (~ 50 %) of the Trp fluorescence intensity was observed upon change of pH from 8 to 3 (**Figure. 2.9 A**). However, no noticeable wavelength shift has been accompanied with the pH change. Such a fluorescence emission without change in wavelength suggests that the fluorophores are already exposed to solvent as compared to the buried Trp which generally emits fluorescence at ~ 310-320 nm. These observations are in line with the near-UV CD and SASA results as discussed above. Out of the four Trp residues, as only W139 is buried, the Trp fluorescence results establish that, as the pH is reduced from 8 to 3, the environment around the W139 has also opened up due to the collapse of secondary/tertiary structure(s) and hence the decrease in the observed fluorescence intensity.

To probe the nature of collapse, ANS was used as an extrinsic fluorophore, which is very sensitive to the structural changes that are accompanied by the exposure of its hydrophobic surfaces. The quantum yield for ANS is very high when it is bound to hydrophobic patches and fluoresces minimally upon binding to the native/folded and unfolded protein conformations [22-24]. ANS fluorescence profiles of T7L at various pH values are shown in **Figure. 2.9 B**.



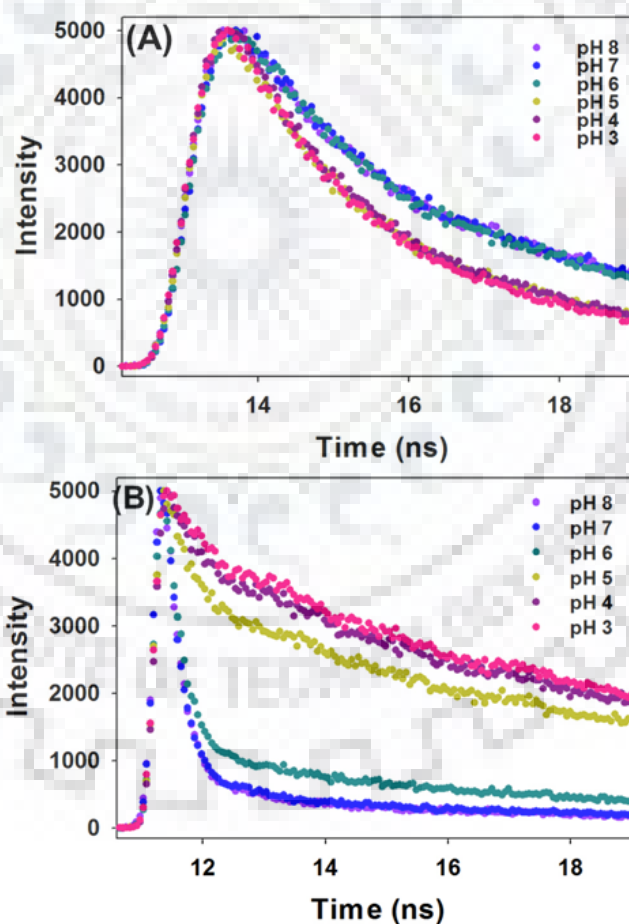
**Figure 2.9:** Fluorescence spectral profiles of T7L from pH 8 to 3 at 25 °C: (A) Tryptophan emission spectra and (B) ANS emission spectra.

A distinguishable ANS fluorescence patterns were observed between the native (pH 8, 7 and 6) and the low pH PF states. ANS binding is comparably weak in the case of pH 8, 7 and 6 native states indicating that the hydrophobic surfaces are buried in the protein core. An immense increase in ANS fluorescence intensity was observed for T7L at pH 3 as compared to that of the pH 8/7 native state, indicating that the hydrophobic surfaces are well-exposed at pH 3. Further, a red shift from

472 towards 508 nm was also observed from the pH 8 to 3 state, thus reinforcing previous conclusion regarding the existence of the partially denatured conformation of T7L at low pH. Among PF states, intensity of ANS fluorescence increases with decrease of pH suggesting for the differential exposure of hydrophobic core. The ANS experiments confirmed the variable structural characteristics of these T7L partially folded states.

### *Fluorescence lifetime measurements*

To probe structural dynamics, Trp and ANS based fluorescence lifetime measurements have been performed using fluorescence lifetime spectroscopy (FLS). Trp fluorescence intensities decay profiles show two distinct sets of decay profiles. One set corresponding to the native conformation (pH 8-6) and the second set correspond to PF conformation (pH 5-3) (**Table 2.3**) (**Figure 2.10 A**).



**Figure 2.10:** Time resolved (A) Tryptophan and (B) ANS fluorescence decay profiles of T7L at pH 8 to 3.

The variable hydrophobic accessibility can also modulate the lifetime of the ANS bound to the T7L protein. Therefore, to probe the differences, ANS based fluorescence life time analysis have been performed (**Figure 2.10 B, Table 2.3**). The results evidenced an increase in the amplitude of the longer lifetime ( $\tau_1$ ) and a decrease in the shorter lifetime ( $\tau_3$ ) upon shifting from native state to PF states. Such a significant change in the lifetime is an indicative of the enhanced ANS binding specificity at hydrophobic patches [25, 26]. Although variations have been observed in the decay profiles of the PF states between pH 5 and 4/3, no significant differences in the lifetimes were observed within the pH 4 and pH 3 PF states.

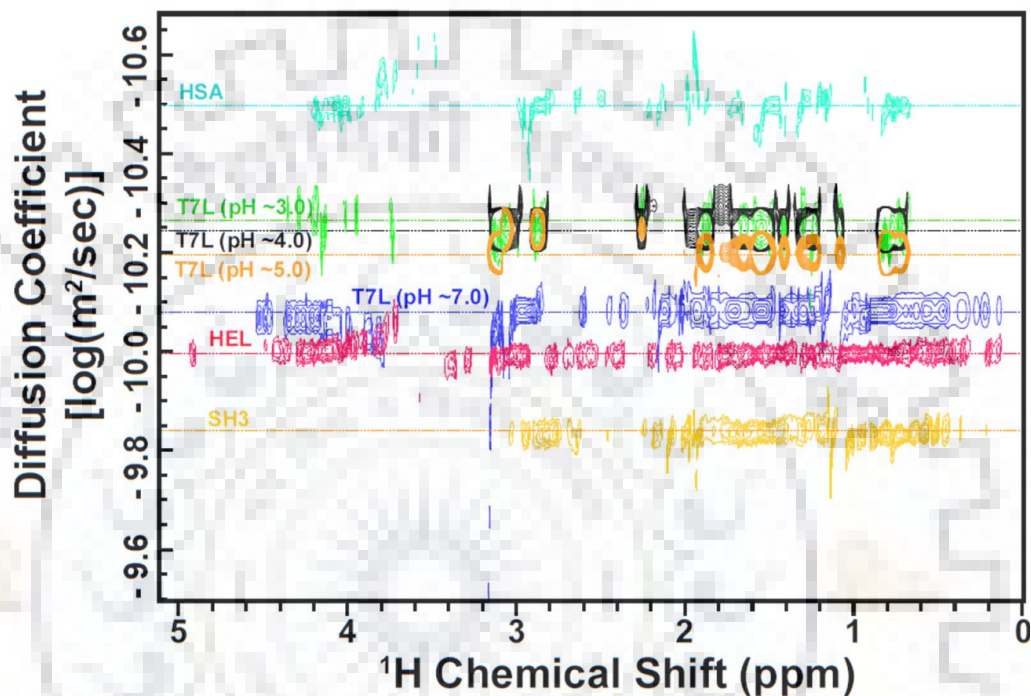
**Table 2.3:** Tryptophan and ANS based time resolved fluorescence decay fitting parameters of T7L.

pH	Tryptophan			ANS		
	$\tau_1$ (ns) (% amp)	$\tau_2$ (ns) (% amp)	$\tau_3$ (ns) (% amp)	$\tau_1$ (ns) (% amp)	$\tau_2$ (ns) (% amp)	$\tau_3$ (ns) (% amp)
8	8.3 (66)	2.6 (24)	0.7 (10)	14.6 (50)	2.8 (16)	0.3 (34)
7	8.3 (66)	2.4 (26)	0.6 (8)	15.2 (53)	2.8 (16)	0.3 (31)
6	9.0 (60)	2.9 (27)	0.8 (12)	16.1 (65)	3.5 (19)	0.3 (16)
5	7.5 (36)	2.5 (48)	0.6 (17)	16.2 (79)	4.8 (19)	0.4 (2.2)
4	5.8 (35)	2.6 (47)	0.6 (19)	16.2 (80)	4.7 (19)	0.38 (1.3)
3	5.0 (41)	2.0 (45)	0.5 (14)	16.1 (81)	4.8 (19)	0.37 (1.0)

### **2D-DOSY translational diffusion**

Variable hydrophobic surface opening of the T7L PF states can result in noticeable changes in diffusion properties as a result of their structural architecture. Literature suggests that the PF states and molten globule states exhibit reduced diffusion constants compared to their native states [27, 28]. To assess such behaviour, NMR based 2D-DOSY translational diffusion experiments have been performed on T7L at pH 7 (representative of native state) and low pH (5, 4 and 3) along with some standard proteins. As expected, the translational diffusion of the PF conformations is slow in comparison to the native T7L at pH 7 (**Figure 2.11**). The observed increase in the extent of translation diffusion for the low-pH conformations can be directly attributed to the protrusions of

the melted structural segments (loss of helical content) thus confirming the presence of the partially denatured conformation. However, a comparative analysis of the diffusion coefficients with standard proteins also established that no higher-order aggregation is present at lower pH. Moreover, it is also evident from the data that the diffusion properties of the pH 5 PF state differ with those of the pH 4/3 PF states, reiterating the existence of structural heterogeneity among the low pH PF states.



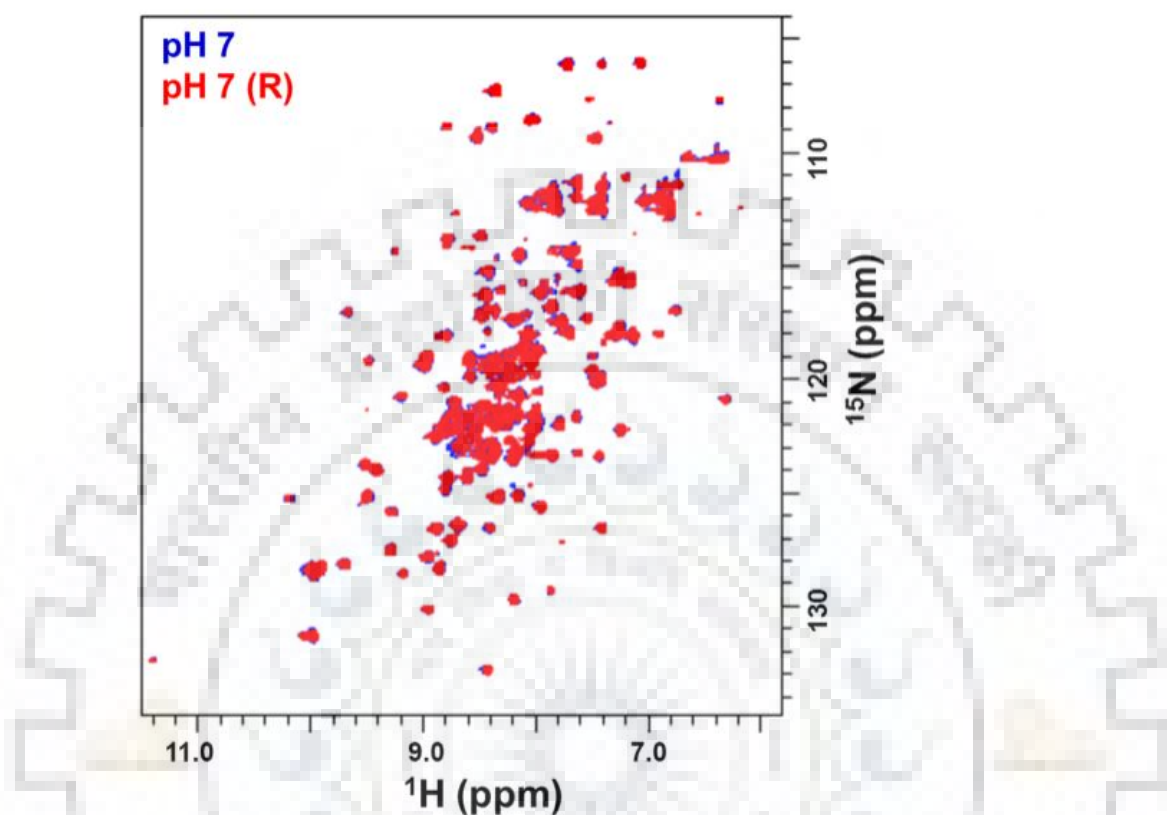
**Figure 2.11:** 2D-DOSY analysis of T7L at pH 7, 5, 4 and 3 along with the reference proteins at 25 °C (SH3, chicken SH3 domain; HEL, hen egg lysozyme; HSA, human serum albumin).

All the above biophysical analysis confirm that T7L exhibits pH-dependent structural transition and forms PF conformation(s) at low pH. These results also establish that, PF conformations at pH 5, 4 and 3 contain differential tertiary structures, structural dynamics and diffusion properties, thus suggesting their structural heterogeneity.

### 2.3.4 Reversibility of the pH-dependent T7L structural transition

To determine the reversible nature of these conformational changes, a <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of T7L was recorded at pH 7. Then the pH of same sample was adjusted to pH 3, exchanged back with pH 7 buffer and the <sup>1</sup>H-<sup>15</sup>N HSQC spectrum [pH 7 (R)] was recorded. The overlay of these two pH 7 HSQC spectra is shown in **Figure 2.12**. Both of the overlaying spectra were identical

to each other with respect to their spectral dispersion, indicating the pH-induced structural transition is a physical phenomenon and is completely reversible.

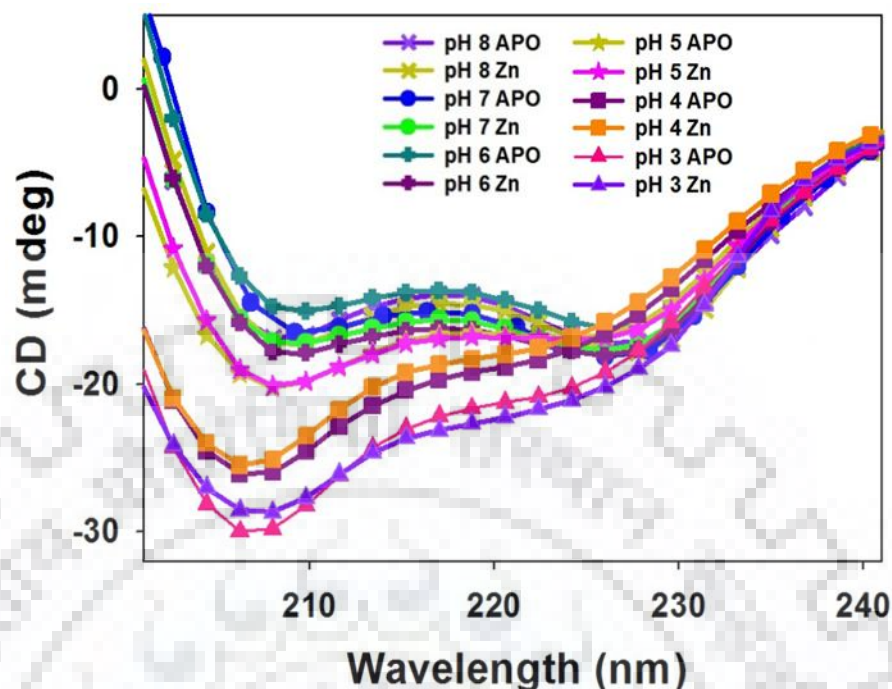


**Figure 2.12:** Overlay of the T7L  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of pH 7 and pH 7 (R) at 25 °C. R stands for reversibility, where the T7L sample is exchanged from pH 7 to 3 and then brought back to pH 7 through buffer exchange.

### 2.3.5 Effect of $\text{ZnSO}_4$ on pH-induced structural transitions

As T7L is a zinc amidase, the role of Zn metal on the observed pH-dependent conformational transition was tested by recording far-UV CD spectra in the absence and presence of Zn (as  $\text{ZnSO}_4$ ) (**Figure 2.13**). Zn containing T7L proteins from pH 8 to 3 yield similar spectral profile as observed for Apo forms. These spectral results confirm that in the presence of metal also, pH-dependent conformational switching of T7L do occur.





**Figure 2.13:** Far-UV CD profiles of T7L at 25 °C showing the secondary structural changes upon changing the pH from 8 to 3 in the presence of 100  $\mu\text{M}$   $\text{ZnSO}_4$  and in absence of zinc (Apo).

## 2.4 Discussion

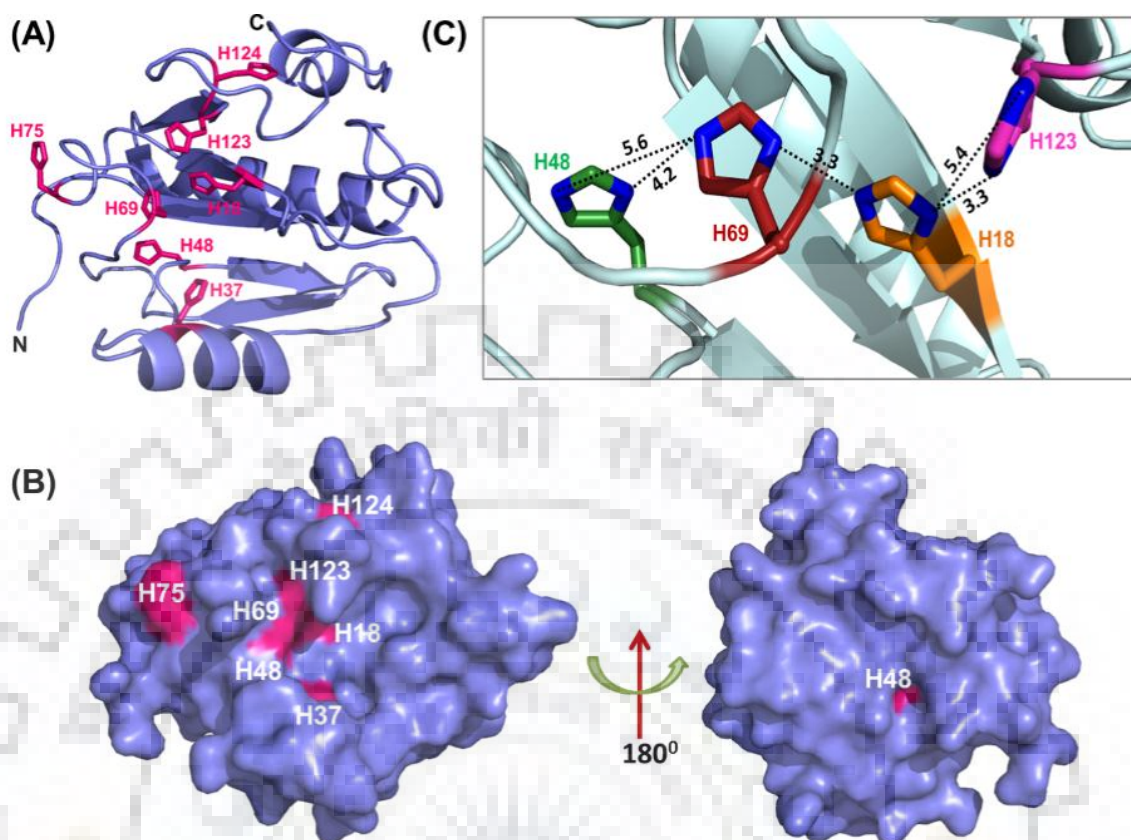
### 2.4.1 Mechanistic insights into the pH-dependent structural transition of T7 endolysin

The functional characteristics of a protein can be directly correlated to its structural integrity. To assess the complete enzymatic activity profile of a protein, it is essential to understand the stability and nature of its conformational transitions. Environmental factors such as temperature, pressure, pH, etc., are known to alter the structure–function relationship of protein to a significant extent [29-33]. Many single-domain amidases exhibit pH-dependent enzymatic profile similar to that of T7, with its highest activity at pH 7.4 [34, 35]. However, no structure–function relationship studies were available for the amidase family of proteins. In the current chapter, comprehensive details regarding the pH-dependent structural intricacy of the bacteriophage T7 endolysin has been presented, a model protein for peptidoglycan amidases [8, 36]. All the biophysical experiments delineated the fact that pH plays a major role in maintaining the structural integrity of the T7L structure. The native state is highly dynamic, and any perturbation in pH toward the acidic side will result in partially folded structure(s) with characteristic exposed hydrophobic surfaces. However,

such a pH-induced conformation essentially preserved its structural characteristics upon reversing the conditions back to the native state.

Proteins exhibit native structures with a global minimum free energy due to the pre-existence of repulsive and attractive electrostatic interactions that provide overall stability to the proteins [37, 38]. Once the pH of the protein switches, the structural and stability changes are caused by the altered ionization states of the amino acid side chains. In the pH range from 8 to 5, the only amino acid that alters the ionization state is histidine (His). The influence of His ionization on the structural stability can be directly correlated to its environment. If the His is completely exposed to solvent, its pKa value ranges from ~7 to 6 and its protonation will not perturb the structure to a large extent as the charge gained is solvent-exposed. However, if His is completely buried, the pKa will be ~4.5-3.5, and such a buried charge can potentially alter the conformation and/or folding of the protein as this additional charge can cause repulsive forces in its vicinity [39, 40]. Further, if the conformational changes originate from His protonation, they are in general exhibit a reversible nature.

As structural changes in the pH range from 6 to 5 for T7L were observed and they are reversible in nature, protonation characteristics of His residues from their surface accessibility features using the crystal structure (PDB entry 1LBA) have been analysed (**Figure 2.14**). T7L consists of seven His residues (H18, H37, H48, H69, H75, H123, and H124) (**Figure 2.14 A**). The computationally measured buried index pKa values for all the His residues are presented in **Table 2.4**. The pKa values and buried index suggest that among the seven His residues, H75 is completely exposed to solvent and H37, H123 and H124 are partially exposed (**Figure 2.14 B**). In the case of *Thermus* phage Ph2119 and Tsc2631 amidases H124 is naturally replaced with Lys and the protein exists in its functional native form [7, 9]. The distance analysis suggested that H37 in T7L is far from the rest of the His and/or positively charged residues and did not show any noticeable side chain repulsive interactions in the surrounding environment (6 Å radius). From all these attributes, H37, H75, and H124 residues have been excluded as the source of the structural collapse. However, such fully and/or partially exposed His residues are the key players that contributed to the observed T7L native-state conformational dynamics.



**Figure 2.14:** (A) Structure of T7L showing the positions of all His residues (B) Surface representation of T7L depicting the buried nature of the His side chains and (C) Close-up on a particular region of the T7L crystal structure (PDB entry 1LBA), to show the His residue (H18, H48, H69, and H123) side chain interactions. The sequential numbering of His residues was given in line with the primary sequence of the protein provided in **Figure 2.1 A**. The atomic distances ( $<6 \text{ \AA}$ ) between the side chain “N” atoms (blue) of His residues are indicated by dotted lines. The structures were generated using PyMOL molecular graphics software.

**Table 2.4:** Buried index and pKa values of His residues calculated by using online webserver DEPTH (as described in Section 2.2.9).

Residue	Buried Index (%)	pKa
H18	100	4.2
H37	43	5.57
H48	76	3.56
H69	90	3.24
H75	0	6.22
H123	46	5.08
H124	43	5.07

The behaviour of the four remaining His residues (H18, H48, H69, and H123) was analyzed. These His residues are highly conserved in members of the T7 family, and their pKa calculations in T7L suggested that these four His residues are completely and/or partially buried in the hydrophobic core of the protein (**Table 2.4**). Moreover, they are present in the proximity of each other in a limited conformational space (**Figure 2.14 C**).

Several of the distances between the N<sup>δ</sup> and N<sup>ε</sup> atoms of each of the His residues with their neighbouring imidazole ring N atoms within a radius of <6 Å were observed (**Figure 2.14 C**). Such short distances and confined orientations between these His imidazole rings will potentially generate a repulsive force between the side chain groups upon protonation and destroy the coupled network of existing stabilizing interactions. Indeed, such conformational transitions due to His protonation were reported in the literature for several other proteins [39-45]. Hence, it can be presumed that the cluster of these four His side chain residues are highly responsible for the observed pH conformational transitions in T7L and hence its differential activity over the pH range.

## **2.5 Concluding remarks**

In summary, T7 endolysin was expressed, purified and the pH-dependent structural transition has been delineated. Studies established that these pH-dependent structural and functional changes of T7L are reversible in nature. The His-His side chain protonation network is the major source of the observed conformational transition and native-state dynamic behaviour. As several of the bacteriophage amidases such as T3, K11, and others contain all these four conserved His residues intact, and display pH-dependent activity profiles similar to that of T7, it is believed that the molecular details elucidated here have provided novel insights into the structure-function paradigm of the single domain amidase family of endolysins.

## **2.6 References**

1. Nelson DC, Schmelcher M, Rodriguez-Rubio L, Klumpp J, Pritchard DG, Dong S, Donovan DM. Endolysins as antimicrobials. *Advance in Virus Research* 83(7), 299-365 (2012).
2. Schmelcher M, Donovan DM, Loessner MJ. Bacteriophage endolysins as novel antimicrobials. *Future Microbiology* 7(10), 1147-1171 (2012).

3. Matthews B, Remington S. The three dimensional structure of the lysozyme from bacteriophage T4. *Proceedings of the National Academy of Sciences* 71(10), 4178-4182 (1974).
4. Cheng X, Zhang X, Pflugrath JW, Studier FW. The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* 91(9), 4034-4038 (1994).
5. Demartini M, Haleboua S, Inouye M. Lysozymes from bacteriophages T3 and T5. *Journal of Virology* 16(2), 459-461 (1975).
6. Junn HJ, Youn J, Suh KH, Lee SS. Cloning and expression of *Klebsiella* phage K11 lysozyme gene. *Protein Expression and Purification* 42(1), 78-84 (2005).
7. Plotka M, Kaczorowska A-K, Morzywolek A, Makowska J, Kozlowski LP, Thorisdottir A, Skirnisdottir S, Hjorleifsdottir S, Fridjonsson OH, Hreggvidsson GO, Kristjansson JK. Biochemical characterization and validation of a catalytic site of a highly thermostable Ts2631 endolysin from the *Thermus scotoductus* phage vB\_Tsc2631. *PLoS One* 10(9), e0137374 (2015).
8. Low LY, Yang C, Perego M, Osterman A, Liddington RC. Structure and lytic activity of a *Bacillus anthracis* prophage endolysin. *Journal of Biological Chemistry* 280(42), 35433-35439 (2005).
9. Plotka M, Kaczorowska A-K, Stefanska A, Morzywolek A, Fridjonsson OH, Dunin-Horkawicz S, Kozlowski L, Hreggvidsson GO, Kristjansson JK, Dabrowski S, Bujnicki JM. Novel highly thermostable endolysin from *Thermus scotoductus* MAT2119 bacteriophage Ph2119 with amino acid sequence similarity to eukaryotic peptidoglycan recognition proteins. *Applied and Environmental Microbiology* 80(3), 886-895 (2014).
10. Moffatt BA, Studier FW. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* 49(2), 221-227 (1987).
11. Kleppe G, Jensen HB, Pryme IF. Purification and characterization of the lytic enzyme n-acetylmuramyl-L-alanine amidase of bacteriophage T7. *European Journal of Biochemistry* 76(2), 317-326 (1977).
12. Alcantara EH, Kim DH, Do SI, Lee SS. Bi-functional activities of chimeric lysozymes constructed by domain swapping between bacteriophage T7 and K11 lysozymes. *Journal of Biochemistry and Molecular Biology* 40(4), 539-546 (2007).
13. Leblanc L, Nezami S, Yost D, Tsourkas P, Amy PS. Isolation and characterization of a novel phage lysin active against *Paenibacillus larvae*, a honeybee pathogen. *Bacteriophage* 5(4), e1080787 (2015).

14. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology* 185 60-89 (1990).
15. Whitmore L, Wallace B. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* 32(suppl\_2), W668-W673 (2004).
16. Ahmad S, Gromiha M, Fawareh H, Sarai A. ASAView: database and tool for solvent accessibility representation in proteins. *BMC Bioinformatics* 5(1), 51 (2004).
17. Tan KP, Nguyen TB, Patel S, Varadarajan R, Madhusudhan MS. Depth: a web server to compute depth, cavity sizes, detect potential small-molecule ligand-binding cavities and predict the pKa of ionizable residues in proteins. *Nucleic Acids Research* 41(W1), W314-W321 (2013).
18. Uversky VN, Karnoup AS, Segel DJ, Seshadri S, Doniach S, Fink AL. Anion-induced folding of *Staphylococcal* nuclease: characterization of multiple equilibrium partially folded intermediates. *Journal of Molecular Biology* 278(4), 879-894 (1998).
19. Ohgushi M, Wada A. 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Letters* 164(1), 21-24 (1983).
20. Dolgikh DA, Gilmanishin RI, Brazhnikov EV, Bychkova VE, Semisotnov GV, Venyaminov SY, Ptitsyn OB.  $\alpha$ -Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Letters* 136(2), 311-315 (1981).
21. Semisotnov G, Rodionova N, Razgulyaev O, Uversky V, Gripas A, Gilmanishin R. Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* 31(1), 119-128 (1991).
22. Ptitsyn O, Uversky V. The molten globule is a third thermodynamical state of protein molecules. *FEBS Letters* 341(1), 15-18 (1994).
23. Gasymov OK, Glasgow BJ. ANS fluorescence: Potential to augment the identification of the external binding sites of proteins. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1774(3), 403-411 (2007).
24. Mukherjee S, Mohan PK, Chary KV. Magnesium promotes structural integrity and conformational switching action of a calcium sensor protein. *Biochemistry* 46(12), 3835-3845 (2007).
25. Uversky VN, Winter S, Löber G. Use of fluorescence decay times of 8-ANS-protein complexes to study the conformational transitions in proteins which unfold through the molten globule state. *Biophysical Chemistry* 60(3), 79-88 (1996).

26. Bushmarina NA, Kuznetsova IM, Biktashev AG, Turoverov KK, Uversky VN. Partially folded conformations in the folding pathway of bovine carbonic anhydrase II: a fluorescence spectroscopic analysis. *ChemBioChem* 2(11), 813-821 (2001).
27. Redfield C. NMR studies of partially folded molten-globule states. *Protein NMR Techniques* 233-254 (2004).
28. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38(50), 16424-16431 (1999).
29. Mohan PK, Barve M, Chatterjee A, Ghosh-Roy A, Hosur RV. NMR comparison of the native energy landscapes of DLC8 dimer and monomer. *Biophysical Chemistry* 134(1), 10-19 (2008).
30. Akasaka K. Probing conformational fluctuation of proteins by pressure perturbation. *Chemical Reviews* 106(5), 1814-1835 (2006).
31. Day R, Bennion BJ, Ham S, Daggett V. Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *Journal of Molecular Biology* 322(1), 189-203 (2002).
32. Jethva PN, Udgaonkar JB. Modulation of the Extent of Cooperative Structural Change During Protein Folding by Chemical Denaturant. *The Journal of Physical Chemistry B* 121(35), 8263-8275 (2017).
33. Feller G. Protein stability and enzyme activity at extreme biological temperatures. *Journal of Physics: Condensed Matter* 22(32), 323101 (2010).
34. Jensen HB, Kleppe K. Effect of ionic strength, pH, amines and divalent cations on the lytic activity of T4 lysozyme. *European Journal of Biochemistry* 28(1), 116-122 (1972).
35. Tsugita A, Inouye M, Terzaghi E, Streisinger G. Purification of bacteriophage T4 lysozyme. *Journal of Biological Chemistry* 243(2), 391-397 (1968).
36. Liepinsh E, Genereux C, Dehareng D, Joris B, Otting G. NMR structure of *Citrobacter freundii* AmpD, comparison with bacteriophage T7 lysozyme and homology with PGRP domains. *Journal of Molecular Biology* 327(4), 833-842 (2003).
37. Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nature Structural Biology* 4(1), 10-19 (1997).
38. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. *Protein Science* 11(4), 739-756 (2002).

39. Sinibaldi F, Howes BD, Piro MC, Caroppi P, Mei G, Ascoli F, Smulevich G, Santucci R. Insights into the role of the histidines in the structure and stability of cytochrome c. *JBIC Journal of Biological Inorganic Chemistry* 11(1), 52-62 (2006).
40. Vila JA, Arnautova YA, Vorobjev Y, Scheraga HA. Assessing the fractions of tautomeric forms of the imidazole ring of histidine in proteins as a function of pH. *Proceedings of the National Academy of Sciences* 108(14), 5602-5607 (2011).
41. Mohan P, Barve M, Chatterjee A, Hosur RV. pH driven conformational dynamics and dimer-to-monomer transition in DLC8. *Protein Science* 15(2), 335-342 (2006).
42. Barrick D, Hughson FM, Baldwin RL. Molecular mechanisms of acid denaturation: The role of histidine residues in the partial unfolding of apomyoglobin. *Journal of Molecular Biology* 237(5), 588-601 (1994).
43. Mohan PK, Chakraborty S, Hosur RV. Hierarchy of local structural and dynamics perturbations due to subdenaturing urea in the native state ensemble of DLC8 dimer. *Biophysical Chemistry* 153(1), 17-26 (2010).
44. Qin Z-L, Zheng Y, Kielian M. Role of conserved histidine residues in the low-pH dependence of the Semliki Forest virus fusion protein. *Journal of Virology* 83(9), 4670-4677 (2009).
45. Cheng CJ, Daggett V. Molecular dynamics simulations capture the misfolding of the bovine prion protein at acidic pH. *Biomolecules* 4(1), 181-201 (2014).



---

## Conformational heterogeneity, dynamics and stability features of T7 endolysin native conformations

### Abstract

Deciphering and characterizing the free energy landscape of low energy states of protein require understanding of structure, stability and dynamic features of native state ensemble. In context, variety of biophysical techniques including CD, fluorescence, size exclusion chromatography, and NMR spectroscopy were applied to study the structural features of T7 endolysin (T7L) native state at pH 8, 7 and 6. Studies suggested that T7L in the native state exhibits a fair amount of conformational dynamics and structural heterogeneity. Broadening of ~20 residues at pH 6 suggest the altered dynamics as a result of conformational exchange. Hydrogen exchange (HX) experiments reveal both the major and minor conformations of  $\alpha$ 2-helix important for the stability of the native state. The dynamic behaviour and the heterogeneity are regulated by the pH; the pH 6 conformation is comparatively more dynamic compared pH 8 and pH 7 native state conformations. Further, stability measurements using chemical denaturant urea and thermal denaturation reveal the differential structural stabilities and unfolding characteristics of pH 6 conformation compared to pH 8 and 7, and formation of insoluble aggregates or precipitates at  $\sim 70$  °C.

### 3.1 Introduction

To accomplish biological functions, proteins must conserve its native state with considerable degree of structural flexibility. Indeed, many proteins with identical structural fold display different dynamic features which in turn govern their unique functional behaviours [1,2]. The structural stability, dynamics and functional properties of proteins are determined by their unique structural organization [3]. These properties are extremely difficult to elucidate at atomic level due to their high level of complexity. The arrangement of hydrophobic residues within the core and presence of secondary and tertiary structure are common factors that are operative in native state [4]. The native state occupies the lowest energy level down to the folding funnel, and comprises of an ensemble of conformations existing in equilibrium under various physiological conditions [4-7]. The ensemble involves one predominant and small population of other fluctuating low-lying energy conformations that are surrounded by the free energy minima [8-11]. These fluctuations are described as occasionally visited conformations

and have also been implicated in catalytic functions, ligand binding and allosteric regulations [10,12,13].

The native state is the most stable state of protein to perform biological functions, and is governed by the surrounding environmental factors such as pH, temperature, solvent and ionic strength due to their effect on inter- and intramolecular interactions that are crucial for stability and integrity of proteins [14-19]. Failure of protein to survive in such environmental conditions produces biological inactivation, and/or misfolding of proteins that can be associated with various fatal diseases [20-23]. Characterization of low energy states/alternative conformations separated by very low kinetic barrier can be done by NMR and X-ray diffraction studies, and more rigorously inferred using NMR spin relaxations [24-30]. Identification and characterization of such low energy states under mild external perturbation such as pH, urea and temperature unravels the native state heterogeneity [12,31,32].

This chapter provides the detailed study on structure, stability and dynamic features of T7L native conformations at pH 8, 7 and 6. Various biophysical and NMR studies elucidated the structural heterogeneity, which is also confirmed by assigning the multiple conformations of T7L residues. Studies suggested that the heterogeneity and stability of native conformations varies upon small change in pH.

## **3.2 Materials and methods**

### **3.2.1 Protein expression and purification**

The unlabeled and isotopically labelled ( $^{15}\text{N}$  /  $^{15}\text{N}$ - $^{13}\text{C}$ ) T7L protein was produced following the protocol described in **Section 2.2.2 of Chapter 2**.

### **3.2.2 Circular dichroism (CD) spectroscopy**

CD measurements at pH 8, 7 and 6 (20 mM glycine-phosphate, 5 mM EDTA and 0.1 M NaCl) were performed as described in **Section 2.2.4 of Chapter 2**. Temperature dependent far-UV CD of 20  $\mu\text{M}$  T7L was measured from 20 to 90  $^{\circ}\text{C}$  at regular intervals of 10  $^{\circ}\text{C}$ , with a temperature gradient of 1  $^{\circ}\text{C}/\text{min}$ , and an incubation time of 5 min at every resting temperature. Thermal reversibility has been assessed by gradually cooling (1  $^{\circ}\text{C}/\text{min}$ ) back the samples and re-recording the spectra at 20  $^{\circ}\text{C}$ .

Chemical denaturation studies were performed by dissolving the urea in phosphate buffer, and the stock concentrations were calculated using the refractive index method. Far-UV

CD spectra of T7L was recorded in the range of 0 M to 10 M urea using pre-equilibrated (~12 h) protein samples.

### ***Free energy calculation***

Experimental data at 222 nm thus obtained was smoothened by five point averaging, and then normalized using the following equation.

$$F_{App} = \frac{S_{obs} - S_F}{S_U - S_F} \quad (3.1)$$

where,  $F_{App}$  is apparent unfolding fraction,  $S_F$  and  $S_U$  represent the spectral values of folded and unfolded states.  $S_{obs}$  is the observed signal at a given urea concentration.

The thermodynamic parameters were obtained from fitting the CD and fluorescence based denaturation curves to a two-state unfolding mechanism ( $N \leftrightarrow U$ ) using the following equation [33]. All the fittings were done with Sigma Plot 12.0 software.

$$A_{obs} = \frac{A_0 + A_U \exp - \frac{(\Delta G_1^0 - m_1[U])}{RT}}{1 + \exp - \frac{(\Delta G_1^0 - m_1[U])}{RT}} \quad (3.2)$$

where, R is the gas constant, T is the absolute temperature,  $\Delta G^0$  is the free energy change at each point,  $m_1$  is the dependence of  $\Delta G$  on urea concentration, [U] is the urea concentration and A's are the specific contribution of each species to the signal.

### **3.2.3 Steady state fluorescence measurements**

Fluorescence measurements of T7L (20  $\mu$ M) at pH 8, 7 and 6 were done as described in **Section 2.2.5 of Chapter 2**. Fluorescence based urea denaturation experiments of T7L-ANS samples were performed under identical conditions as described in above CD section.

### **3.2.4 Size exclusion chromatography (SEC)**

SEC experiments were performed as described in **Section 2.2.2 of Chapter 2**. 70  $\mu$ M of T7L samples at pH 8, 7 and 6 were loaded separately onto a column pre-equilibrated with buffers of respective pH (20 mM phosphate, 0.1 M NaCl, 5 mM EDTA, and 1 % glycerol).

### **3.2.5 Nuclear magnetic resonance (NMR) spectroscopy**

2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were recorded at 1 mM T7L at pH 8, 7 and 6 as described in **Section 2.2.7 of Chapter 2**. All the native state triple-resonance, relaxation, and

HX experiments were conducted in 20 mM Tris, 0.2 M NaCl, 5 mM EDTA buffers (pH 8, 7 and 6), where the protein is found to be more stable over longer periods of time.

### **3.2.5.1 Backbone resonance assignment**

The three-dimensional NMR experiments were conducted at pH 7 T7L using a triple-channel Bruker Avance III 800 MHz instrument equipped with a TXI cryoprobe, pulse shaping, and pulse field gradient capabilities. For backbone resonance assignments, a series of standard 3D NMR experiments such as HNCOC, HNCA, intra HNCA, HNCACB and CBCA(CO)NH were recorded at 25 °C using [<sup>13</sup>C, <sup>15</sup>N] T7L (~1 mM) at pH 7. The <sup>13</sup>C carrier frequency was set at 54 ppm for HNCA/intra HNCA and hNCAnH, 42 ppm for HNCACB and CBCA(CO)NH, and 174 ppm for HNCOC. All the NMR data were processed using Bruker software TOPSPIN version 3.2, and the resultant spectra were analysed using CARRA [34].

### **3.2.5.2 Hydrogen-exchange (HX) experiment**

The native-state hydrogen exchange experiment was initiated by dissolving the lyophilized T7L protein sample (pH 7 and 25 °C) in 100 % D<sub>2</sub>O. The sample was loaded on a pretuned and shimmed NMR spectrometer at 25 °C. The dead time (time from the addition of D<sub>2</sub>O to the start of the HSQC spectrum) was 10 min. The HX spectrum was recorded with 16 scans and consisted of 128 complex increments in the indirect <sup>15</sup>N dimension (total experiment time of 80 min).

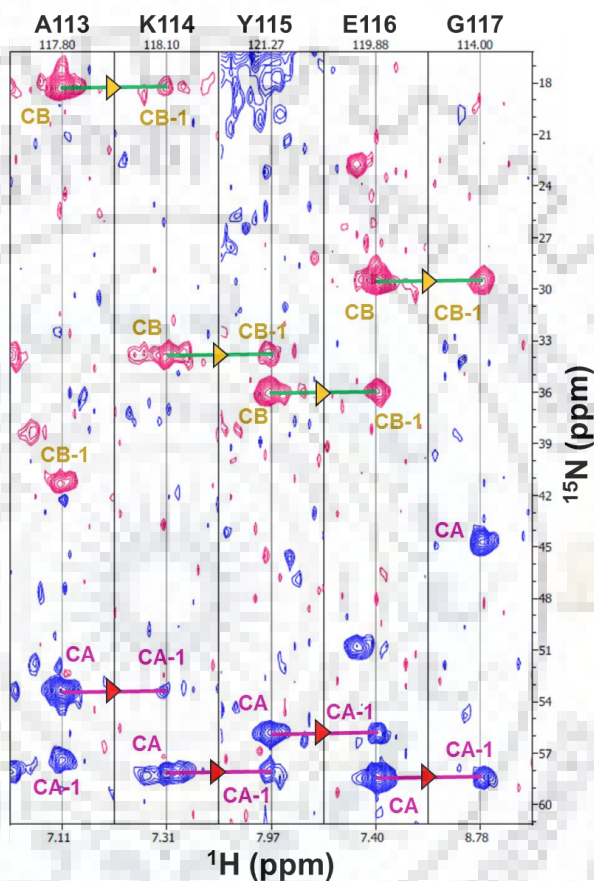
### **3.2.5.3 NMR relaxation experiments**

The relaxation experiments were performed at pH 8, 7 and 6 using the pulse sequences described by Peng and Wagner and Farrow et al., [35,36]. A complete set of backbone amide R<sub>1</sub>, R<sub>2</sub>, and <sup>1</sup>H-<sup>15</sup>N NOE relaxation data sets were acquired on the Bruker Avance 800 MHz spectrometer. The R<sub>1</sub> and R<sub>2</sub> data sets were collected with a recycle delay of 2.5 s. For T<sub>1</sub> (R<sub>1</sub>) measurements, the following relaxation delays were used: 10, 50, 90, 110, 150, 250, 350, 550, 770, and 990 ms. For T<sub>2</sub> (R<sub>2</sub>) measurements, the following relaxation delays were used: 18.56, 37.12, 55.68, 74.24, 92.8, 111.36, and 129.92 ms. The Sigma Plot 12.0 software was used to calculate the R<sub>1</sub> and R<sub>2</sub> values and their uncertainties by fitting the measured peak intensities to a single-exponential decay curve. Steady-state <sup>1</sup>H-<sup>15</sup>N heteronuclear NOE measurements were taken with a proton saturation time of 3.0 s and a relaxation delay of 3.0 s. Steady-state <sup>1</sup>H-<sup>15</sup>N NOEs were calculated as a ratio of intensities of the peaks with and without proton saturation. The errors in the NOEs were obtained as described by Farrow et al. [35].

### 3.3 Results

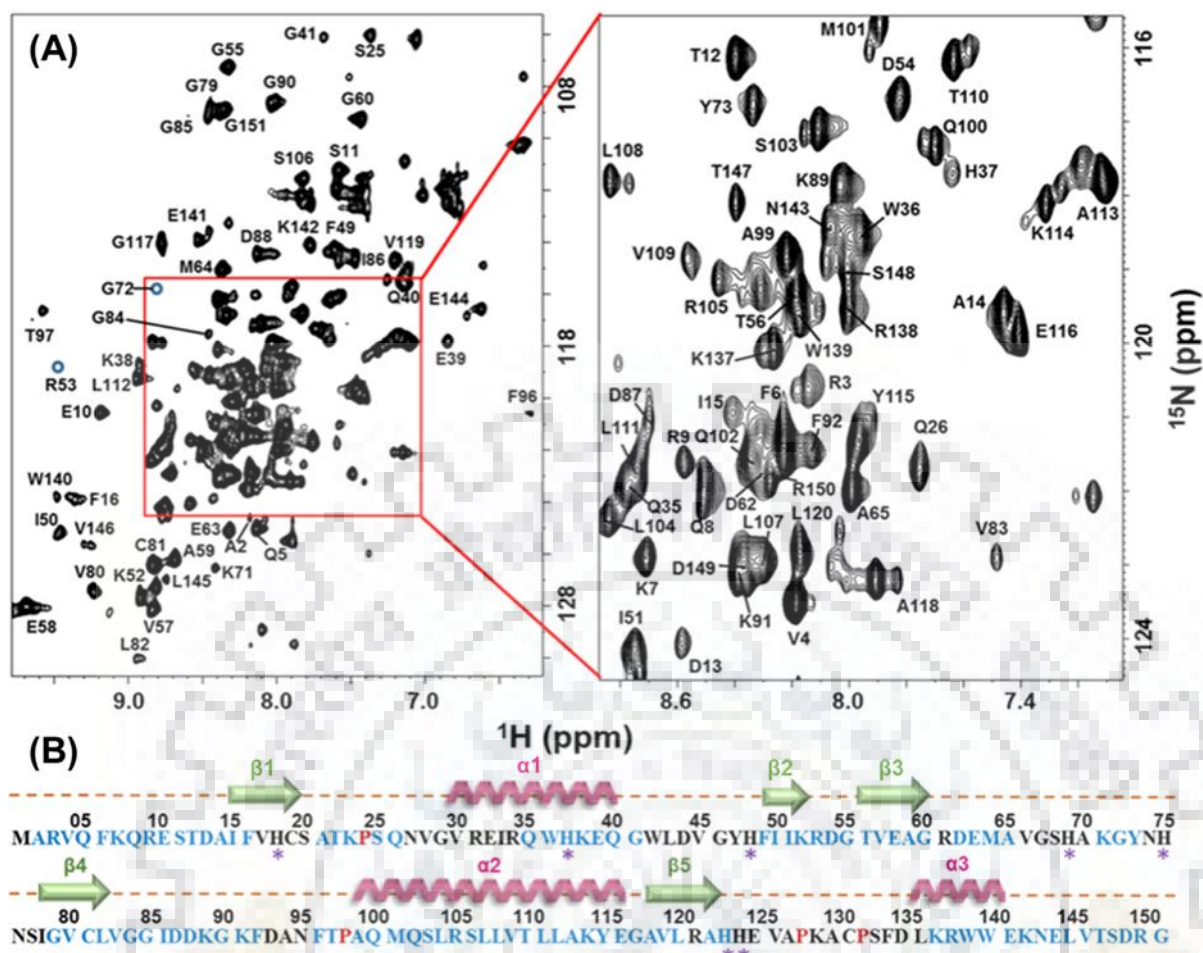
#### 3.3.1 Backbone resonance assignment of T7L

Assignment of NH cross-peaks of T7L at pH 7 in the HSQC spectrum was achieved by using a set of conventional triple-resonance experiments. The HNCACB strip is shown in **Figure 3.1** to describe the connectivity of C $\alpha$  (CA) and C $\beta$  (CB) resonances for the stretch of A113-G117.



**Figure 3.1:** Representative HNCACB strip plot showing sequential connectivities for residues A113 to G117. Purple lines with red triangles are showing connectivity of CA resonances whereas, green lines with yellow triangles showing connectivity of CB resonances.

It is evident from HSQC spectrum (**Figure 3.2 A**), that the intensities of the NH resonances are not uniform indicating an inherent conformationally dynamic nature. Of the 140 observed amino acid residues in the HSQC spectrum,  $\sim 25$ – $30$  of them do not show any sequence specific connection to the preceding residues even in highly sensitive 3D experiments such as HNCA/HNCO. Overall  $> 70$  % of the resonance assignments were obtained. The backbone NH assignments for the assigned residues are labelled in **Figure 3.2 A**, and the summary of the assignments is given in **Figure 3.2 B**. The chemical shift details of the assigned residues are provided in **Appendix I**.

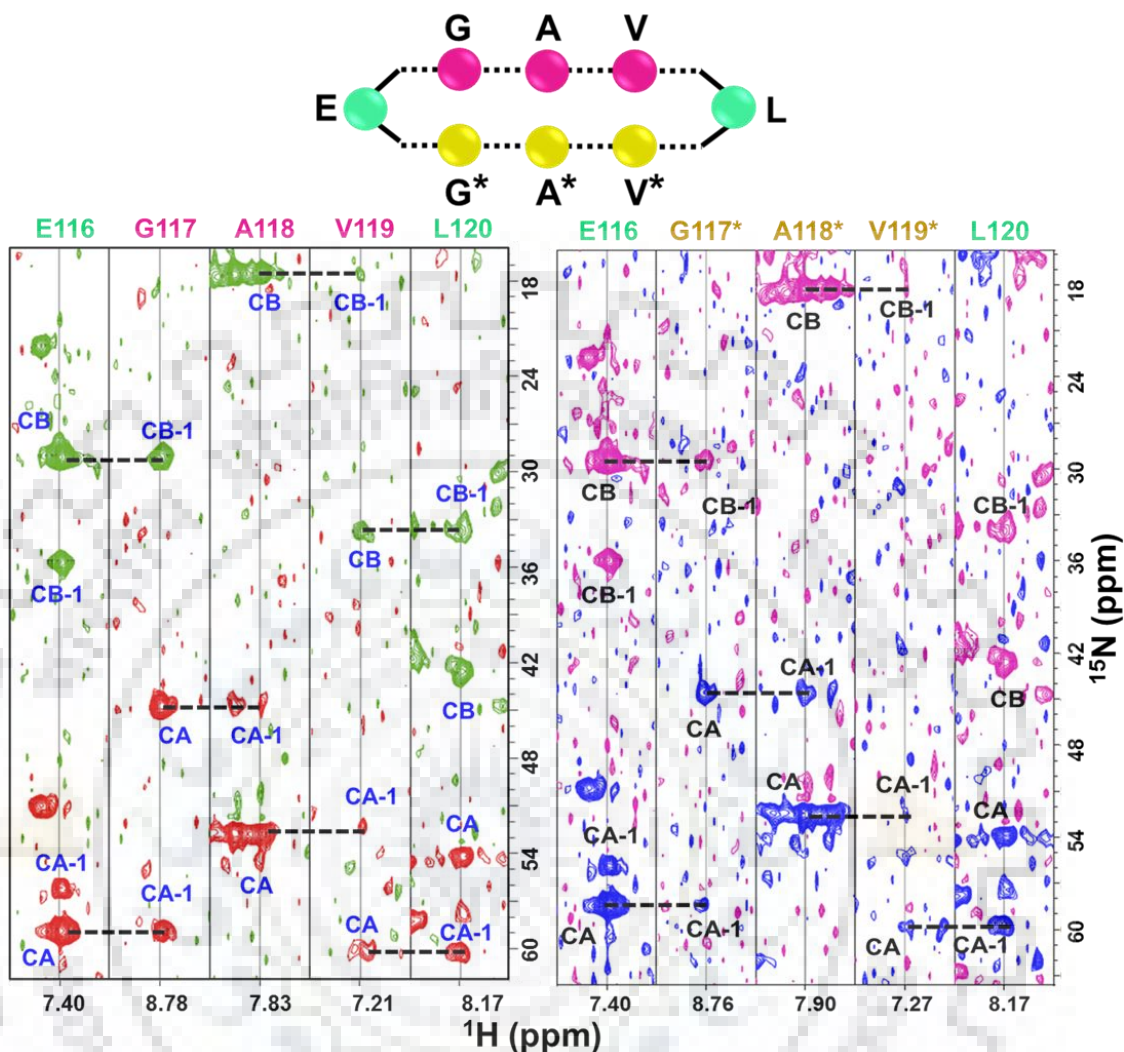


**Figure 3.2:** Backbone NMR assignment of T7L at pH 7 and 25 °C: (A)  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum of the T7L assignment with NH cross-peaks indicated by residue name and number, with a few peaks with lower intensity in the HSQC spectrum, are shown with circles. The close-up is the expansion of the central region of the T7L spectrum (B) Summary of residue-wise NMR assignments of T7L (blue, assigned; black, unassigned). Proline residues are colored red, and the histidine residues are highlighted with asterisk (\*) for clarity. The corresponding residue numbers and secondary structural elements are shown at top of the sequence.

### 3.3.2 Existence of multiple conformations in native state ensemble

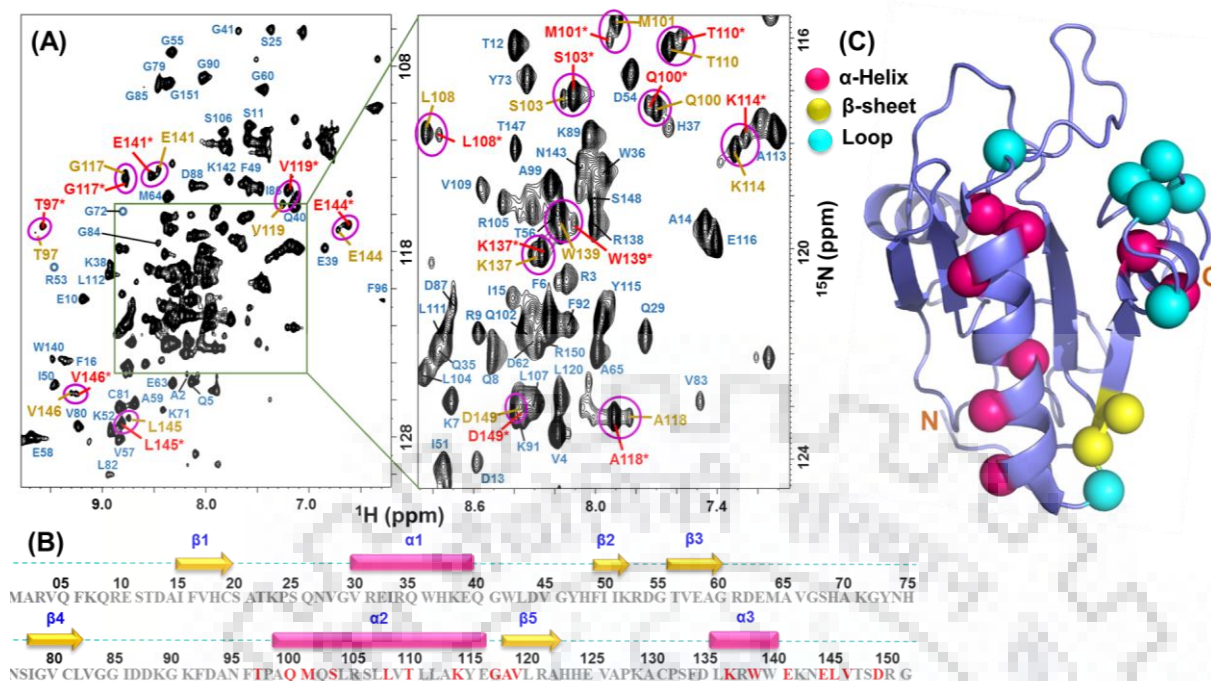
After assigning most of the available sequential connectivities in the HSQC spectrum of T7L native state at pH, there are still several resonances present which do not correspond to any new stretch or missing unassigned residues. Such resonances draw a serious attention towards the assignment/sequence-specific location of those residues. The assignments of these extra observed peaks were accomplished by using the same approach as described above. It was found that majority of these peaks belong to some of the already assigned residues, and corresponded to a major and a minor conformation. The resonance assignments of these alternative stretches of T7L clearly establish that some of the protein segments are accessing the alternative conformation in the slow exchange regime of the NMR timescale. The

alternative sequential connectivity of residues G117, A118 and V119 which are present in two conformations is shown in **Figure 3.3** using HNCACB strip plots for the stretch of E116-L120.



**Figure 3.3:** Representative HNCACB strip plot showing alternative sequential connectivities for CA and CB resonances of G117, A118 and V119 residues in the stretch of E116-L120.

Approximately, 17 residues out of total assigned amino acids show the presence of multiple conformations at pH 7 (**Figure 3.4 A**). Residues exhibiting the alternative conformations belong to four different regions of the polypeptide chain (**Figure 3.4 B and C**). Six residues are located in  $\alpha$ 2-helix, two residues each in  $\beta$ 5-sheet and  $\alpha$ 3-helix region, followed with five residues in C-terminal loop. It clearly indicates that  $\alpha$ 2-helix and C-terminal loop are highly populated with multiple conformation accessing residues. All the data suggests that C-terminal half/segment contribute extensively to the structural heterogeneity of T7L in its native conformation.

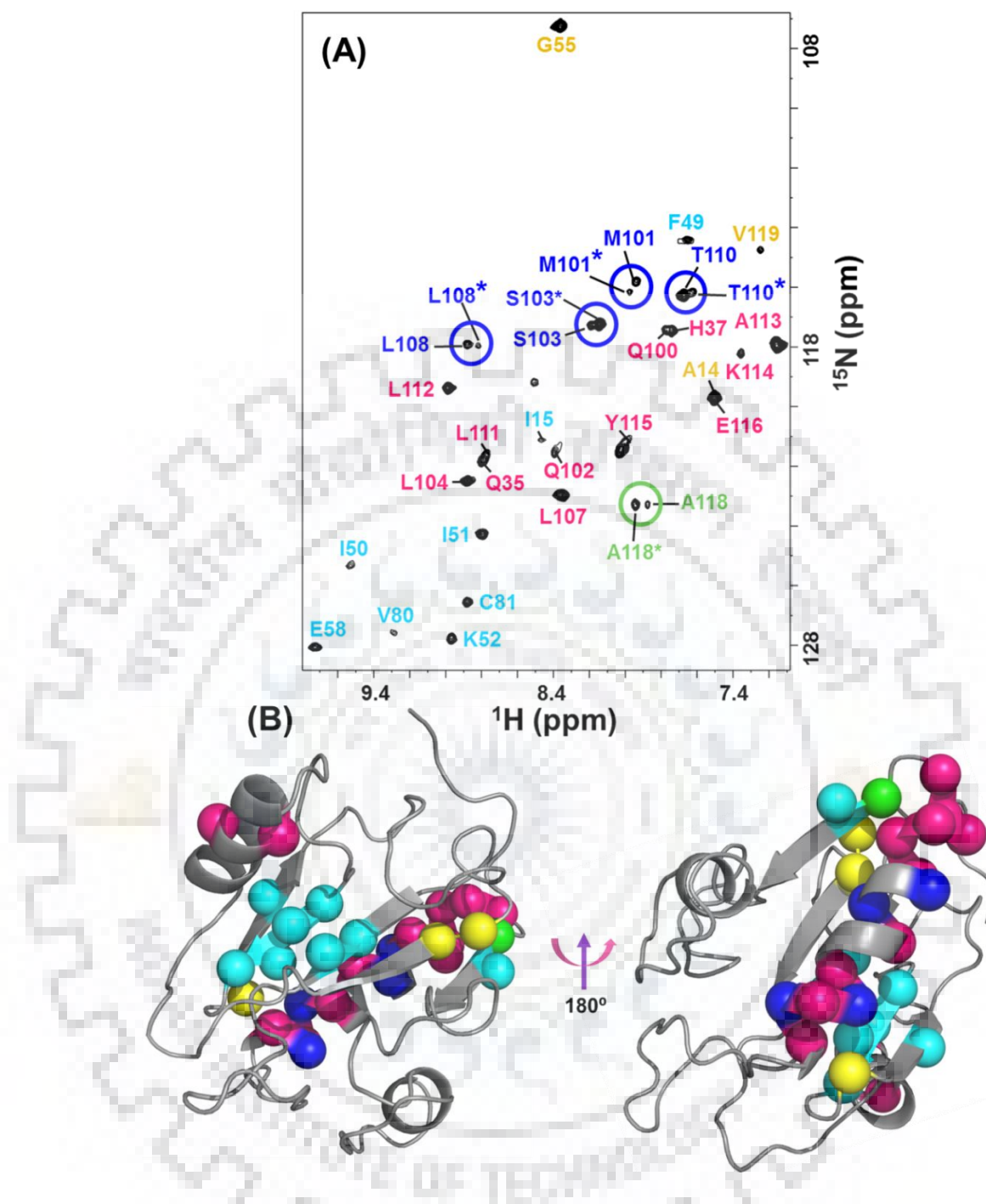


**Figure 3.4:** Backbone NMR assignment of T7L multiple conformations at pH 7 and 25 °C: (A)  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum of the T7L assignment. Circles are showing NH cross-peaks of residues having two conformations (with and without asterisk) (B) Amino acid sequence indicating the summary of total residues showing multiple conformations highlighted with red color and (C) Structural localization of assigned residues with multiple conformations onto crystal structure of T7L (PDB ID:1LBA) generated by using PyMOL molecular graphics software.

### 3.3.3 Residue level insights into the T7L native state stability

As native state of T7L presents conformational heterogeneity, it is interesting to know which of the structural elements contribute to its stability. In order to probe the conformational stability of T7L in its native-state ensemble, hydrogen exchange experiment was performed. NMR-based native-state hydrogen/deuterium exchange is a powerful technique for calculating residue-wise stabilities of proteins. In principle, the exchange rates of backbone amide protons depend on their accessibility to the solvent deuterons, which can be correlated to the stability of secondary structural elements and/or tertiary and/or quaternary interactions. The HX spectrum of T7L is shown in **Figure 3.5 A**, and the structural locations of the protected residues are marked in **Figure 3.5 B**. Of 151 amino acids, only ~32 amino acids were protected and are major contributors to the structural elements (**Figure 3.5 B**). Among the protected residues, helix  $\alpha$ 2 makes a major contribution with 14 residues followed by sheets  $\beta$ 2 and  $\beta$ 4 accounting for six residues, suggesting that these structural elements make major contributions to the stability of the native state.





**Figure 3.5:** (A) Hydrogen–Deuterium (H/D) exchange HSQC spectrum of T7L. The disappearance of NH peaks from the  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectrum is monitored at 25 °C, by re-dissolving lyophilized T7L (pH 7) in  $\text{D}_2\text{O}$ . Circles are showing appearance of minor conformations from  $\alpha$ 2-helix region (B) Residues present in the HSQC spectrum after H/D exchange mapped on the T7L structure (PDB ID:1LBA) generated by using PyMOL molecular graphics software.

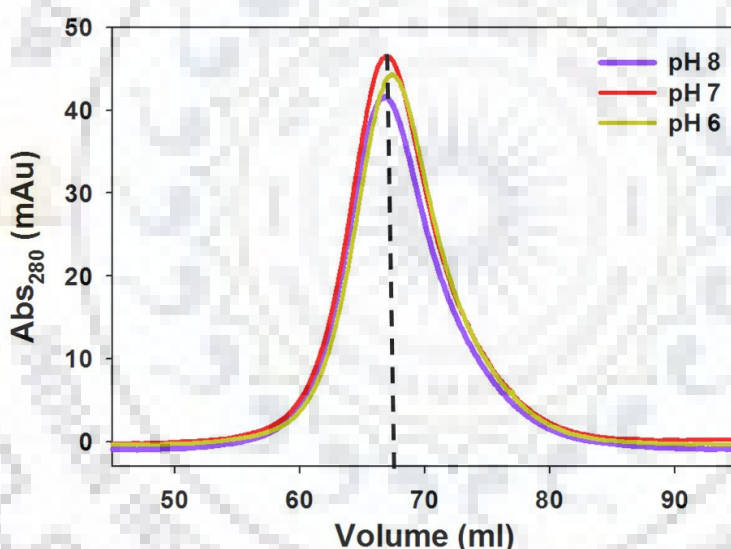
Interestingly, some of the minor NH resonances belonging to the  $\alpha$ 2-helix were also appeared with noticeable intensities in the HX spectra. Appearance of these minor conformations of  $\alpha$ 2-helical residues establish that both these T7L dynamic conformations that

are in slow exchange at NMR time scale likely contributes to the structural stability of protein (Figure 3.5 A, enclosed in circles; Figure 3.5 B, blue, green spheres).

### 3.3.4 Assessing the pH dependent conformational heterogeneity in native state ensemble

To elucidate the heterogeneity of T7L native state ensemble at pH 7, it has been subjected to mild perturbation using pH (pH 8 and pH 6), where the protein remains in its native conformation.

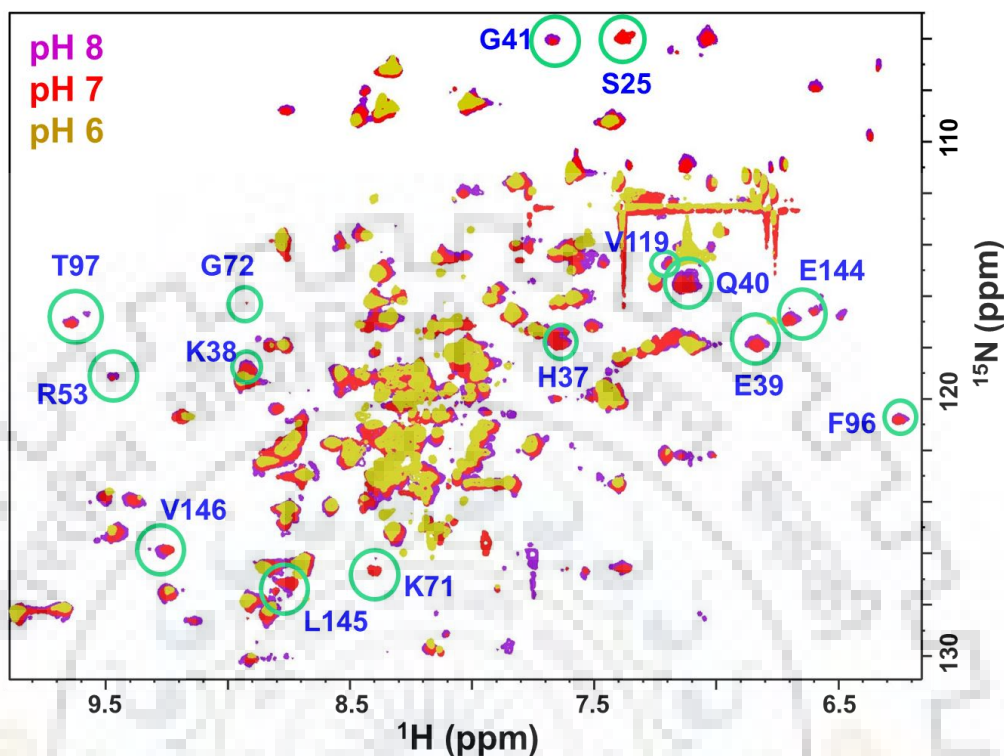
SEC experiments have been performed to assess the conformational state of T7L at pH 8, 7 and 6. Equimolar samples of T7L at these three pH values were subjected separately to a Hi-load 16/60 Superdex-75 column. The FPLC elution profiles of the proteins are shown in Figure 3.6. As depicted, the elution peaks of for T7L at pH 8, 7 and 6 look very much similar establishing that T7L exists in its native state in the chosen pH range.



**Figure 3.6:** FPLC chromatogram showing similar elution profiles of T7L native state at pH 8, 7 and 6.

In order to assess the native state heterogeneity at different pH, NMR based  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were performed. The  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts are sensitive probes for the environment of a given residue in a protein structure. It is evident from the spectra that at both pH 8 and 6, T7L shows very good chemical shift dispersion along the proton (7–11 ppm) and nitrogen axes, indicating that the protein is in its folded conformation. As the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra at these two pH values are very similar to pH 7, the transfer of assignment is readily possible. The assignments at pH 8 and 6 were obtained by direct transfer of assignments using pH 7 assignments. The number of peaks present in the HSQC spectra at pH 8, 7 and 6 are

140, 140 and 120 respectively. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectral overlay of T7L at pH 8, 7 and 6 is shown in **Figure 3.7**.



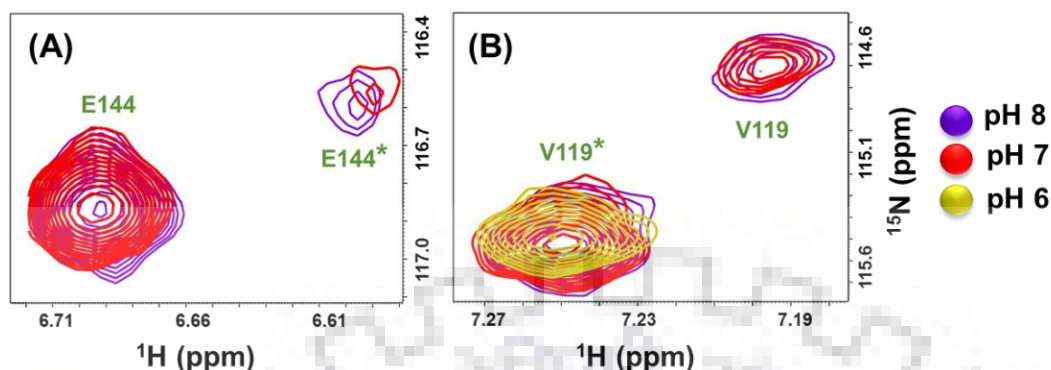
**Figure 3.7:** The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectral overlay of T7L at pH 8, 7 and 6 and 25 °C. Circles are showing assigned residues which are disappeared at pH 6.



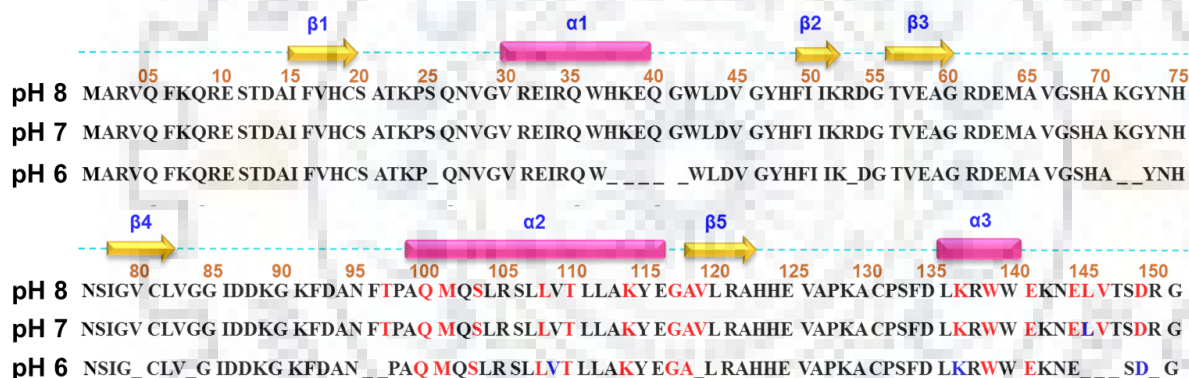
**Figure 3.8:** Summary of broadened residues at pH 8, 7 and 6. The corresponding residue numbers and secondary structural elements are shown at top of the sequence.

The loss of ~20 resonances at pH 6 indicates that these peaks were broadened because of their altered dynamics as a result of conformational exchange, although the protein is in its folded conformation [37]. All the assigned residues broadened due to such exchange at pH 6 are shown in **Figure 3.8**, and are marked with red colour on the primary sequence of the

protein. The broadened residues are majorly located on  $\alpha$ 1-helix (amino acid 37-41) and on the C-terminal (amino acid 145-151) of the protein.



**Figure 3.9:** The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectral overlay of T7L at pH 8, 7 and 6 and 25 °C. (A) Residue E144 present at pH 8 and 7 in both the major and minor (\*) conformations but disappeared at pH6 and (B) V119 present at pH 8 and 7 in both the major and minor (\*) conformations but show only major conformation at pH6.



**Figure 3.10:** The primary sequence of amino acids showing the residues accessing multiple conformations marked as red at pH 8, 7 and 6. Residues marked with blue are devoid of minor conformations, and ‘\_’ indicating disappeared residues. The corresponding residue numbers and secondary structural elements are shown at the top of the sequence.

Further, identification of the residues accessing alternative conformations in the native states present at pH 8, 7 provides a conscious view of the landscape of the native state ensemble. Comparison of HSQC spectra revealed that all the residues are present at pH 8 with both their major and minor conformations except, L145 minor conformation. Whereas, two situations were observed in the case pH 6 on the HSQC overlay with pH 7; (a) resonances corresponding to both major and minor conformations completely disappeared at pH 6; which includes T97, E144, L145 and V146 and, (b) residues are present in their major conformation and the minor conformation was vanished; these residues include V119 and K137 (**Figure 3.9**).

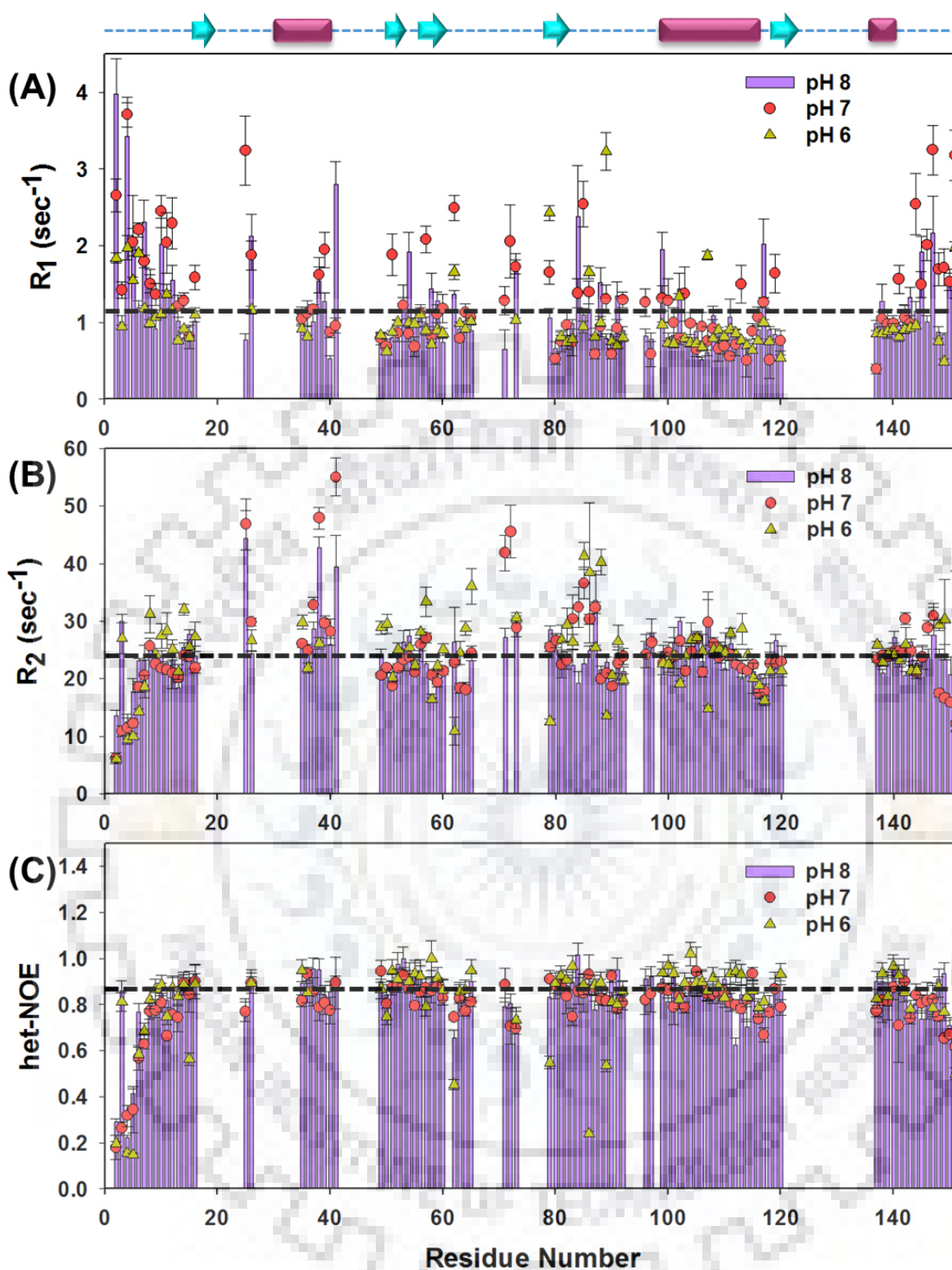
The locations of these residues are marked on the native structure of the monomeric protein in **Figure 3.10**.

Results show that the observed dynamic behaviour and the heterogeneity are regulated by the pH, and the differences in the heterogeneity detected at pH 6 as compared to those at pH 7 and 8 must be attributed to the changes in the protonation states of the various side chains. It clearly indicates that the low pH state (pH 6) is comparatively more dynamic than the higher pH (8, 7) native state conformations.

### **3.3.5 Structural dynamics of T7L native conformations**

Dynamics measurements using NMR relaxation experiments provide detailed sequence specific information along the polypeptide chain at various time scales. In general, the longitudinal relaxation rate ( $R_1$ ), the transverse relaxation rate ( $R_2$ ), and the steady-state heteronuclear NOE measurements (het-NOE) provide motions that are sensitive to various time scales. The values for the longitudinal relaxation rate,  $R_1$ , are sensitive to both low- to high-frequency motions (nanosecond to picosecond time scale motions). However, the longitudinal relaxation ( $R_1$ ) rate by itself does not discriminate effectively between faster and slower motions, whereas the heteronuclear  $[^1\text{H}]-^{15}\text{N}$  NOEs are typically most sensitive to higher-frequency motions (picosecond time scale motions) of the backbone. Their lower values indicate the increased local flexibility of the polypeptide [38,39]. Similarly, the transverse relaxation ( $R_2$ ) rates are more sensitive to the contributions from slower millisecond to microsecond exchange processes to the observed conformational dynamics.

To assess the dynamic characteristics of T7L, relaxation experiments have been performed. The relaxation rate parameters,  $R_1$ ,  $R_2$ , and NOE, were reliably obtained for ~ 96 residues of major conformation, and a summary of the results is given in **Figure 3.11 A–C**.



**Figure 3.11:** Summary of the residue-wise backbone  $^{15}\text{N}$  relaxation parameters of T7L recorded on 800 MHz NMR instrument: (A) longitudinal relaxation rates ( $R_1$ ), (B) transverse relaxation rates ( $R_2$ ), and (C) steady-state heteronuclear NOE (het-NOE). The corresponding secondary structure elements are also indicated on the top. The black horizontal line represents the average value of the particular relaxation parameter.

From the data, it is evident that T7L shows a significant chain dependent conformational dynamics in its  $R_1$  and  $R_2$  values. Further detailed analysis of  $R_2$  values at pH 8,

7 and 6 provide information about the residue specific conformational exchange. The average values  $R_2$  obtained are  $24.3 \pm 1.2$ ,  $24.3 \pm 1.0$  and  $24.1 \pm 1.5$  s<sup>-1</sup> at pH 8, 7 and 6, respectively. At pH 8 a group of many residues such as S25, H37–G41, K71, D87, T147 and G151 showed large  $R_2$  values (**Figure 3.11 B**). At pH 7 more number of residues showed enhanced  $R_2$  values such as S25, H37–G41, K71, G72, V83–D87, L107, K142, V146 and T147. In case of pH 6, apart from the broadened residues detected earlier using HSQC measurements, Q8-F16, F49 I50, T56 V57 E63-A65 G85, I86, D88, A113, S148 and D149 have shown large  $R_2$  values than the average indicating that these residues are experiencing conformational exchange.

As  $R_2$  values serve as useful monitors of local conformational transitions occurring on the millisecond-to-microsecond time scale, and the conformational exchange present at particular residue site results in conspicuously enhanced  $R_2$  values, the residues and/or segments described at pH 8/7 are regarded as the source of the T7L dynamic behaviour. Comparative analysis of the peak intensities of the HSQC spectra of pH 7 and 6 (**Figure 3.7**), along with the observed  $R_2$  pattern at pH 7 and 6 suggested that several of the broadened residues at pH 6 are identified as being in the same group or vicinity of residues that showed enhanced  $R_2$  values at pH 7, thus establishing the enhanced dynamics of pH 6 native state conformation.

The steady-state NOEs at pH 8, 7 and 6 show a constant pattern over the polypeptide chain with average values of  $0.83 \pm 0.01$ ,  $0.82 \pm 0.03$  and  $0.8 \pm 0.04$  at pH 8, 7 and 6, respectively (**Figure 3.11 C**), thus evidencing a characteristic folded structural conformation. The NOE patterns clearly revealed that the N-terminus is more flexible than its C-terminus and attain greater flexibility upon decreasing pH. However, in case of pH 6 some of the residues D62, G79, I86 and D88 etc., show very less NOE values indicating the enhanced dynamic flexibility at these sites in faster time scales.

Overall, the observed relaxation results confirm the dynamic heterogeneity of native states and revealed that residues present at pH 6 experienced more fluctuations in  $R_1$ ,  $R_2$  and NOE values as compared to pH 8 and 7.

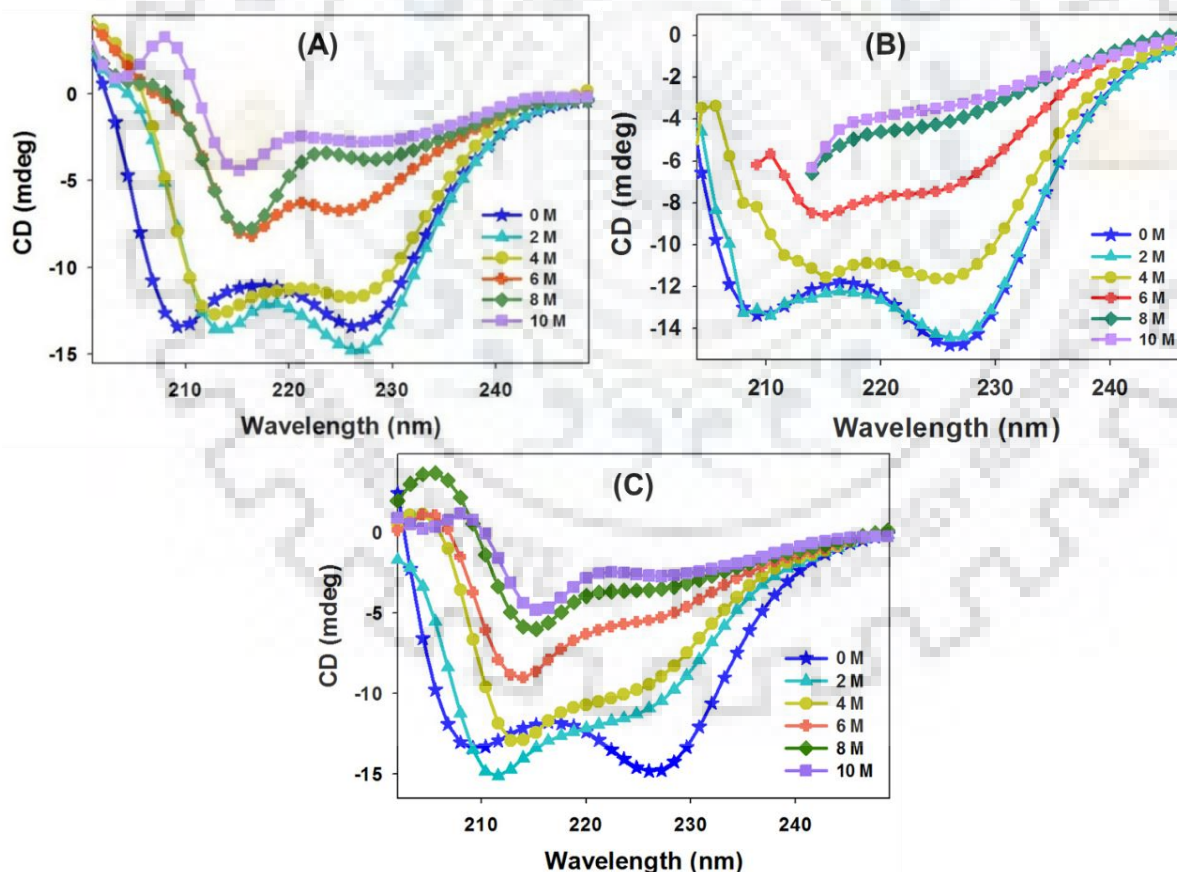
### **3.3.6 Assessing the structural stability features of T7L native state conformations**

Structural stability of native fold has regulatory consequences on the function of proteins which is determined by the overall balance of forces that stabilize the native state and forces that destabilize it. Perturbation in these interactions useful to investigate the stability features of well-defined 3D structure. CD and fluorescence spectroscopic techniques were used

to characterize stability of T7L native conformations during unfolding with chemical denaturant urea and temperature.

### 3.3.6.1 Chemical denaturation of T7L native conformations

Far-UV CD studies were carried out to investigate the effect of protein denaturant urea on structural stability and conformational changes of T7L native states. Urea induced denaturation profiles of T7L native conformations at pH 8, 7 and 6 obtained from far-UV spectroscopy are shown in **Figure 3.12 A–C**. The magnitude of ellipticity bands use as an index of protein unfolding due to their complete dependency on structural conformation of proteins. It is evident from CD profile that at 0 M urea T7L exhibits native conformations at all these three pH values. Increasing the urea concentration from 0 to 10 M associated with gradual melting of secondary structural elements indicated by decrease in ellipticity at 222 nm. However, the melting profile of pH 6 seems to be different to those of at pH 8 and 7, as CD signal at 222 nm for pH 6 conformation decreased sharply.

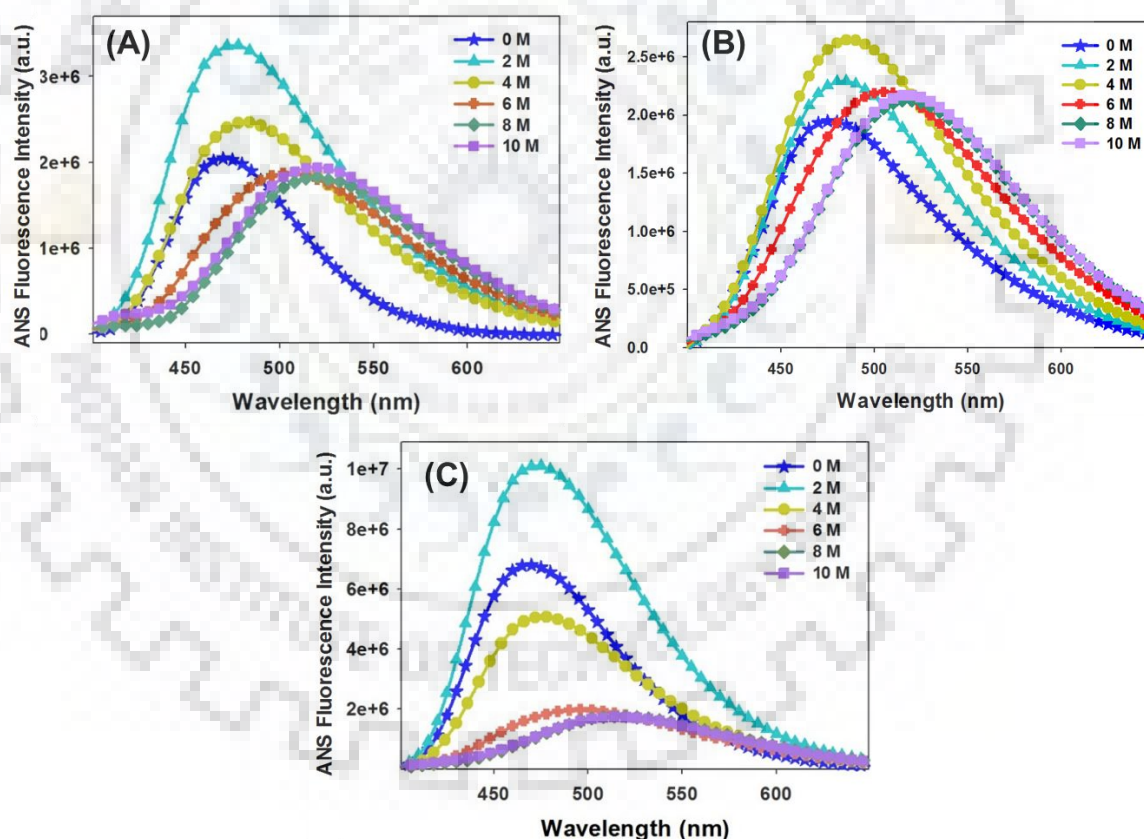


**Figure 3.12:** Far-UV CD profiles of urea induced chemical denaturation of T7L at (A) pH 8 (B) pH 7 and (C) pH 6.



**Steady-state fluorescence measurements**

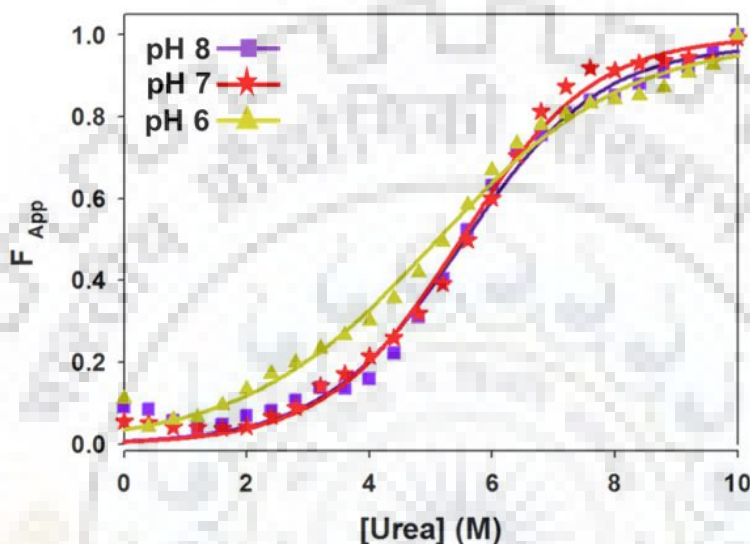
The exposure of any hydrophobic region during urea induced denaturation has been studied by ANS binding to the T7L. The respective changes in the emission intensity and wavelength ( $\lambda_{\max}$ ) have been monitored as a function of urea concentration (**Figure 3.13 A–C**). The spectral parameters such as intensity, position and intensity of fluorescence emission spectrum provides the information of structural and dynamic properties of proteins [40]. ANS binding to pH 8 and 7 states did not show any significant change in the fluorescence intensity in the range of 0 to 10 M urea. However, in pH 6 state at 2 M urea the maximum intensity value increased by two-fold, and then decreased upon increasing urea concentration, thus indicating its hydrophobic opening around 2 M urea concentration. pH 8 and 7 states show fluorescence emission maximum at 480 nm at 0 M urea and shifts to 520 nm with very less change in intensity ( $I_{\max}$ ) at 10 M urea.



**Figure 3.13:** ANS fluorescence emission profiles of urea induced chemical denaturation of T7L at (A) pH 8 (B) pH 7 and (C) pH 6.

Urea dependent unfolding experiments based on CD and fluorescence evidenced for differential structural stabilities and unfolding characteristics of pH 6 conformation compared to pH 8 and 7. Further, to determine the Gibbs free energy changes, and the transition midpoints

of the urea unfolding process, transition curves were analysed using ellipticity values at 222 nm. Decreasing the ellipticity with increasing urea concentration shows a sigmoidal shaped two-state transition curve (**Figure 3.14**). The transition curve at pH 6 show considerable difference as compared to pH 8 and 7. The fitted parameters are described in **Table 3.1**. Although, the mid-point of transitions are almost same for all the three native conformations, the free energy values clearly indicated that pH 6 state is comparatively less stable.



**Figure 3.14:** Normalized denaturation curves monitored by far-UV CD ellipticity at 222 nm as a function of urea concentration.

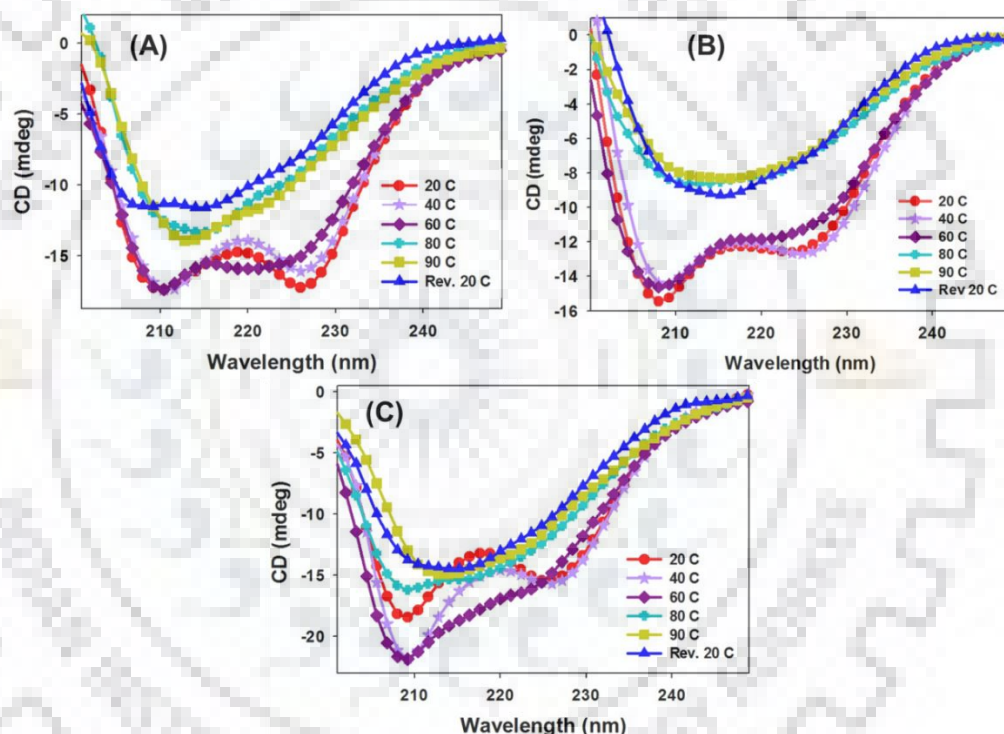
**Table 3.1:** Unfolding free energies ( $\Delta G$ ) and transition midpoints ( $C_m$ ) of T7L at different pH values for urea denaturation process

pH	$\Delta G$ (kcal/mole)	$C_m$
8	$-6.3 \pm 0.4$	5.6
7	$-6.1 \pm 0.3$	5.6
6	$-4.3 \pm 0.2$	5.4

### 3.3.6.2 Thermal stability and reversibility of low pH conformations

Temperature-dependent far-UV CD measurements were performed at pH 8, 7 and 6 to analyze the thermal stability and reversibility of T7L native conformations. Far-UV CD spectra were recorded as a function of temperature from 20 to 90 °C at each of the pH values (**Figure 3.15 A–C**). The far-UV denaturation profiles of pH 8 and 7 were found to be identical. The

thermal denaturation curves suggested that at pH 8 and 7, no significant changes in ellipticity were observed between 20 and 40 °C (**Figure 3.15 A and B**). Extensive changes were observed once the temperature was increased beyond 40 °C. The gradual decrease in the CD values at 222 nm above 50 °C is indicative of the thermal melting of the secondary structure. In contrast, at pH 6 (**Figure 3.15 C**) ellipticity values at 222/208 nm had shown continuous changes from 20 °C onwards, thus indicating a continuous thermal transition / temperature induced structural change. However, the melting behavior of native conformations at pH 8, 7 and 6 were found to be irreversible above 70 °C, as the protein started forming soluble and insoluble aggregates leading to loss of CD signals. This limited the thermodynamic analysis of the thermal denaturation process of T7L native states.



**Figure 3.15:** Thermal stability of T7L native conformations monitored by far-UV CD from 20 °C to 90 °C in 20 mM glycine-phosphate buffer; (A) pH 8 (B) pH 7 and (C) pH 6. Spectra of the protein samples that were cooled back to 20 °C (Rev 20 °C) after heating to 90 °C were also shown to elucidate the thermal irreversibility/ aggregation properties.

### 3.4 Discussion

#### 3.4.1 Heterogeneity and differential structural stability features of T7L native state ensemble

Deciphering the structure, stability and dynamic properties of proteins provide useful mechanistic insights into the free energy landscape of its native state ensemble. At pH 7 T7L

exhibits in its native conformation which is characterized by inherent conformational dynamics nature and heterogeneity. Under the natively folded conditions, the structure of T7L is populated with multiple conformations in the slow exchange regime of NMR time scale as the amino acid residues accessed two different conformations (major and minor conformations on the basis of their NH peak intensities). Mild perturbation in pH (pH 8 and 6), and exposure to chemical denaturant and temperature helped in elucidating the differential conformational heterogeneity, and stability properties of native T7L native state ensemble. This indeed enabled in characterization of the plasticity of the T7L native energy landscape. Most of these heterogeneous residues belonged to the C-terminal half of the protein and majority of them are in the  $\alpha$ 2-helix and C-terminal loop of T7L polypeptide chain.

The HX studies indicated that the stability of T7L native state majorly lies around  $\alpha$ 2-helix, as both major and minor conformations are present after addition of D<sub>2</sub>O. Detailed investigations on the unfolding behaviour of the T7L using urea and temperature have provided valuable insights on the differential stability characteristics. The chemical denaturation profile monitored by far-UV CD reveals that T7L native conformations unfold by a simple two state process. pH 6 state shows less stability towards chemical and thermal denaturation as compared to pH 7 and 8 states. The thermal denaturation profile monitored by far-UV CD reveals that at all pH values T7L forms soluble/insoluble aggregates above 70 °C.

The conformational dynamics of T7L is very much sensitive toward pH change, and thus responsible for the difference in the numbers of multiple conformation accessing residues at lower pH (pH 6). These studies provided insights into the sensitivity of the conformational fluctuations due to the environmental perturbations. The roughness of energy landscape and fluctuations in the native state ensemble of a protein is a function of mutual interactions of different amino acid side chains. Changes at amino acid side chains caused by pH are responsible for the observed conformational fluctuations, as the protonation state of a side chain causes a local change in the electrostatic potential. Upon close observation, it was noticed that the residues missing in the sequence specific resonance assignments are the stretches that are either side of the His residues. T7L comprises seven His residues with varied surface exposure. The conformational exchange of the protonated and deprotonated states of these His residues depending on their side chain pK<sub>a</sub> values (**Table 2.4, Chapter 2**) can potentially contribute to an enhanced conformational exchange, which in turn resulted in the variable intensities and/or broadening of the HSQC peaks as well as the loss of sequence specific connections in the 3D spectrum. Such perturbations are also responsible for partial loss of

amino acid population of multiple conformations and function (**Section 2.3.1, Chapter 2**) at pH 6. The sensitivity of conformational dynamics can be attributed to the partial protonation of His side chains. H37 (pKa 5.57), H75 (pKa 6.22), H123 (pKa 5.08) and H124 (pKa 5.07) are the major contributor to the observed differences in the fluctuations of native state in the range of pH 6-7. All these results suggest that, although at pH 8, 7 and 6 T7L presents in its native state, they exhibit differential structural stability and dynamics properties which play a crucial role in the conformational heterogeneity of native state ensemble and their pH-dependent lysis activity as elucidated in **Chapter 2**.

### **3.5 Concluding remarks**

In summary, the current chapter deciphered that the native state of T7L exhibits conformational heterogeneity with presence of multiple conformations. The native conformations present at pH 8, 7 and 6 have differential structural stability and dynamic behaviours. As decoding the information of native ensemble and presence of multiple conformations of proteins is important for understanding the relationship between molecular flexibility and function, therefore such an outcome can be useful to study the structure-stability relationships of other endolysins which share common structural features like T7L. Implications of environmental factors that dictate the stability are useful to design the protein with improved stability and functionality. Furthermore, elucidating the folding and dynamic characteristics of proteins are indeed essential to understand the shallowness of their native energy landscapes.

### **3.6 References**

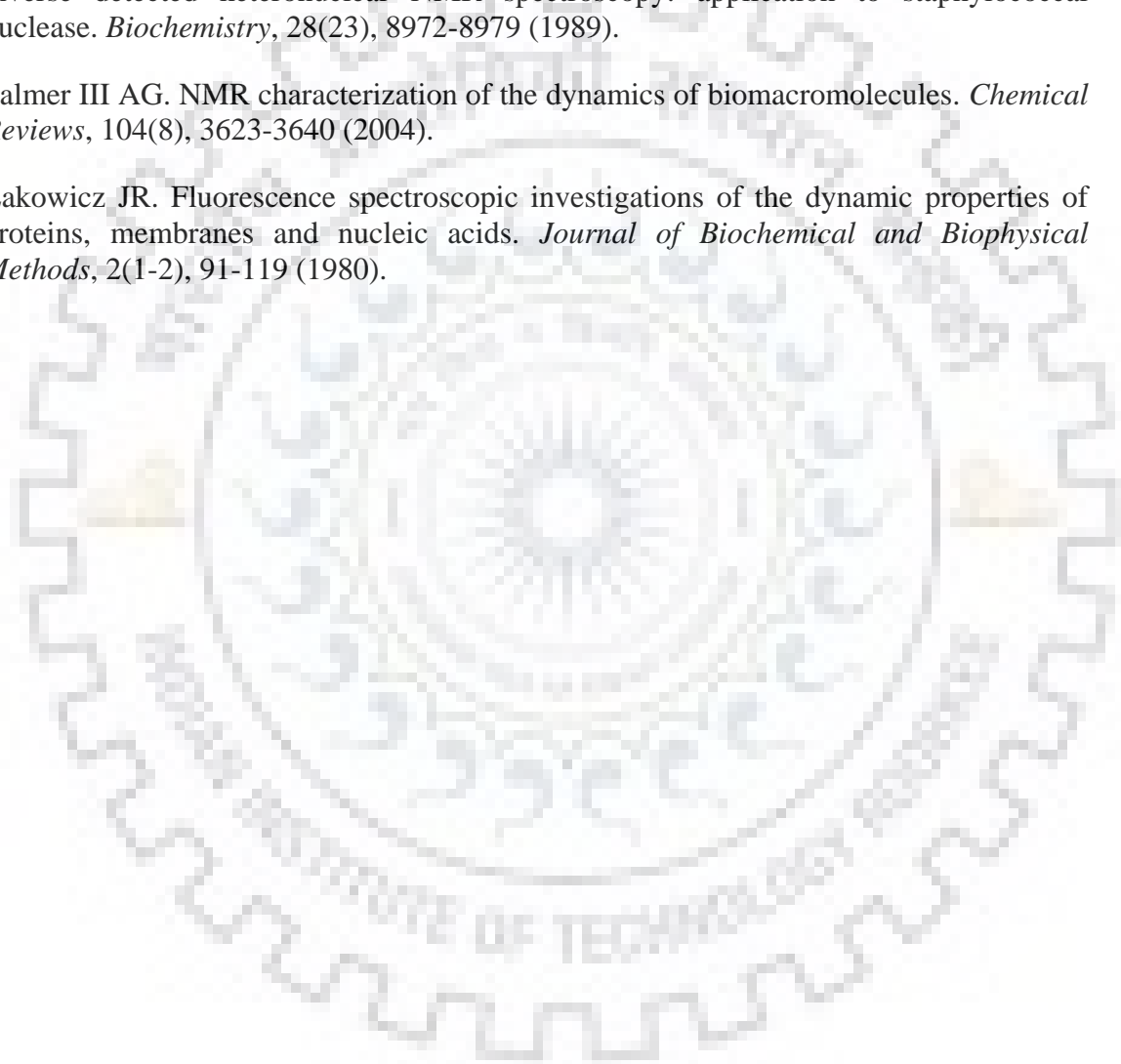
1. Di Iorio EE, Tavernelli I, Yu W. Dynamic properties of monomeric insect erythrocrucorin III from *Chironomus thummi-thummi*: Relationships between structural flexibility and functional complexity. *Biophysical Journal*, 73(5), 2742-2751 (1997).
2. Szilágyi A, Závodszky P. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure*, 8(5), 493-504 (2000).
3. Ansari A, Berendzen J, Bowne SF, Frauenfelder H, Iben IE, Sauke TB, Shyamunder E, Young RD. Protein states and proteinquakes. *Proceedings of the National Academy of Sciences*, 82(15), 5000-5004 (1985).
4. Tavernelli I, Cotesta S, Di Iorio EE. Protein dynamics, thermal stability, and free-energy landscapes: a molecular dynamics investigation. *Biophysical Journal*, 85(4), 2641-2649 (2003).

5. Bertini I, Giachetti A, Luchinat C, Parigi G, Petoukhov MV, Pierattelli R, Ravera E, Svergun DI. Conformational space of flexible biological macromolecules from average data. *Journal of the American Chemical Society*, 132(38), 13553-13558 (2010).
6. Bernadó P, Blackledge M. Structural biology: proteins in dynamic equilibrium. *Nature*, 468(7327), 1046-1048 (2010).
7. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Structure, Function, and Bioinformatics*, 21(3), 167-195 (1995).
8. Smock RG, Gierasch LM. Sending signals dynamically. *Science*, 324(5924), 198-203 (2009).
9. Swain JF, Gierasch LM. The changing landscape of protein allostery. *Current Opinion in Structural Biology*, 16(1), 102-108 (2006).
10. Kumar S, Ma B, Tsai C-J, Sinha N, Nussinov R. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Science*, 9(1), 10-19 (2000).
11. Boehr DD, McElheny D, Dyson HJ, Wright PE. Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. *Proceedings of the National Academy of Sciences*, 107(4), 1373-1378 (2010).
12. Kumar A, Srivastava S, Hosur RV. NMR characterization of the energy landscape of SUMO-1 in the native-state ensemble. *Journal of Molecular Biology*, 367(5), 1480-1493 (2007).
13. Boehr DD, McElheny D, Dyson HJ, Wright PE. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science*, 313(5793), 1638-1642 (2006).
14. Jaenicke R. Protein stability and molecular adaptation to extreme conditions. *The FEBS Journal*, 202(3), 715-728 (1991).
15. Stigter D, Alonso D, Dill KA. Protein stability: electrostatics and compact denatured states. *Proceedings of the National Academy of Sciences*, 88(10), 4176-4180 (1991).
16. Tang KE, Dill KA. Native protein fluctuations: the conformational-motion temperature and the inverse correlation of protein flexibility with protein stability. *Journal of Biomolecular Structure and Dynamics*, 16(2), 397-411 (1998).
17. Mohan PK, Barve M, Chatterjee A, Ghosh-Roy A, Hosur RV. NMR comparison of the native energy landscapes of DLC8 dimer and monomer. *Biophysical Chemistry*, 134(1), 10-19 (2008).
18. Pace CN, Trevino S, Prabhakaran E, Scholtz JM. Protein structure, stability and solubility in water and other solvents. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 359(1448), 1225-1235 (2004).
19. Maulik V, Jennifer S, Teruna J. The role of thiols and disulfides in protein chemical and physical stability. *Current Protein & Peptide Science*, 10, 614-625 (2009).

20. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry.*, 75, 333-366 (2006).
21. Quist A, Doudevski I, Lin H, Azimova R, Ng D, Frangione B, Kagan B, Ghiso J, Lal R. Amyloid ion channels: a common structural link for protein-misfolding disease. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30), 10427-10432 (2005).
22. Soto C. Unfolding the role of protein misfolding in neurodegenerative diseases. *Nature Reviews Neuroscience*, 4(1), 49-60 (2003).
23. Herczenik E, Gebbink MF. Molecular and cellular aspects of protein misfolding and disease. *The FASEB Journal*, 22(7), 2115-2133 (2008).
24. Dyson HJ, Wright PE. Elucidation of the protein folding landscape by NMR. *Methods in Enzymology*, 394, 299-321 (2005).
25. Dyson HJ, Wright PE. Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Methods in Enzymology*, 339, 258-270 (2000).
26. Dyson HJ, Ewright P. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Advances in Protein Chemistry*, 62, 311-340 (2002).
27. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. *Chemical Reviews*, 104(8), 3607-3622 (2004).
28. Baldwin AJ, Kay LE. NMR spectroscopy brings invisible protein states into focus. *Nature Chemical Biology*, 5(11), 808-814 (2009).
29. Peng JW, Wagner G. [20] Investigation of protein motions via relaxation measurements. *Methods in Enzymology*, 239, 563-596 (1994).
30. Palmer AG, Williams J, McDermott A. Nuclear magnetic resonance studies of biopolymer dynamics. *The Journal of Physical Chemistry*, 100(31), 13293-13310 (1996).
31. Mishra P, Jha SK. An alternatively packed dry molten globule-like intermediate in the native state ensemble of a multidomain protein. *The Journal of Physical Chemistry B*, 121(40), 9336-9347 (2017).
32. Sridevi K, Lakshmikanth G, Krishnamoorthy G, Udgaonkar JB. Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble. *Journal of Molecular Biology*, 337(3), 699-711 (2004).
33. Chatterjee A, Mohan PK, Prabhu A, Ghosh-Roy A, Hosur RV. Equilibrium unfolding of DLC8 monomer by urea and guanidine hydrochloride: Distinctive global and residue level features. *Biochimie*, 89(1), 117-134 (2007).
34. Keller R, Wuthrich K. Computer-aided resonance assignment (CARA). *Verlag Goldau, Cantina, Switzerland*, (2004).
35. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. Backbone dynamics of a free and a

phosphopeptide-complexed Src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry*, 33(19), 5984-6003 (1994).

36. Peng JW, Wagner G. Mapping of the spectral densities of nitrogen-hydrogen bond motions in Eglin c using heteronuclear relaxation experiments. *Biochemistry*, 31(36), 8571-8586 (1992).
37. Mohan P, Barve M, Chatterjee A, Hosur RV. pH driven conformational dynamics and dimer-to-monomer transition in DLC8. *Protein science*, 15(2), 335-342 (2006).
38. Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, 28(23), 8972-8979 (1989).
39. Palmer III AG. NMR characterization of the dynamics of biomacromolecules. *Chemical Reviews*, 104(8), 3623-3640 (2004).
40. Lakowicz JR. Fluorescence spectroscopic investigations of the dynamic properties of proteins, membranes and nucleic acids. *Journal of Biochemical and Biophysical Methods*, 2(1-2), 91-119 (1980).





---

## **Deciphering the differential structural stability and dynamics of partially folded conformations of T7 endolysin**

### **Abstract**

Characterization of partially collapsed protein conformations at atomic level is a daunting task due to their inherent flexibility and conformational heterogeneity. T7 endolysin (T7L) exhibits a pH-dependent structural transition from native state to partially folded (PF) conformation. In the pH range 5-3, T7L PF states display differential 8-anilino-1-naphthalenesulfonic acid (ANS) binding characteristics. The structure, stability and dynamics relationships of pH driven T7L PF conformations have been investigated using CD, fluorescence and NMR spectroscopy. Structural studies indicated a partial loss of secondary/tertiary structures compared to its native state. The loss in the tertiary structure and the hydrophobic core opening increases upon decrease of pH from 5 to 3. Urea dependent unfolding features of PF state at pH 5 and 4 evidenced for a collapsed conformation at intermediate urea concentrations. Further, thermal denaturation experiments delineated that the pH 5 conformation is thermally irreversible in contrast to pH 3 indicating that hydrophobic core opening is essential for thermal reversibility. Residue level studies revealed that  $\alpha$ 1-helix and  $\beta$ 3- $\beta$ 4 segment of T7L are the major contributors for such a structural collapse and inherent dynamics. All these results suggest that the low pH PF states of T7L are heterogeneous and exhibits differential structural, unfolding, thermal reversibility, and dynamic features. Unraveling these biophysical features of single domain bacteriophage endolysins is essential for designing of novel chimeric and engineered endolysins as broadband antimicrobial agents over a diverse pH range.

### **4.1 Introduction**

Folding of a protein from a disordered denatured state to the precisely folded native state involves a complex network of elementary interactions. From a denatured state, proteins can fold rapidly to their native state via a simple two-state process or by multi-state process, which comprises of intermediates, partially folded and molten globule states [1-5]. Global organization of a protein into specific three dimensional fold is crucial for its stability, cellular localization and functional activity. Environmental factors such as pH, solvent, temperature and chemical denaturants

differentially contribute to a protein's structure-stability-dynamics paradigm, and henceforth their functionality [6-9]. Among these environmental factors, pH induced conformational changes and stability characteristics are closely relevant with respect to biological perspective, as pH change alters the ionization state of charged side chain amino acids [10-12]. Moreover, the mechanism of pH induced structural changes can be specifically correlated with its primary sequence composition. pH-dependent conformational changes have been demonstrated for a wide variety of protein families [13-16]. Changes in pH are shown to regulate; (a) oligomerization of proteins [17-19], (b) formation of aggregates and amyloid fibers [20-23], (c) domains swapping/merging [24,25], (d) folding-unfolding transitions [26,27], (e) interactions of proteins with ligands/biomolecules [28-30] and, (f) structure-activity relationships [31-33].

T7 endolysin (T7L) access its native state at pH 7, and forms partially unfolded/collapsed conformation below pH 6. Such a conformational transition of T7L explained the loss of functionality of the bacteriophage amidases at low pH. However, the details related to the conformational stability, and dynamics of these low pH (5-3) partially folded (PF) states are not yet deciphered. This chapter deals with the structure-stability-dynamics relationship of T7L low pH partially folded conformations. Detailed biophysical and NMR studies revealed that these PF states differ in their structural and stability characteristics, and exhibit differential dynamics. The observed heterogeneity in the PF states evidence the role of environmental perturbations in regulating the fine-tuned protein folding intermediates across the protein's energy landscape.

## **4.2 Materials and methods**

### **4.2.1 Protein expression and purification**

The unlabeled and isotopically labeled ( $^{15}\text{N}$  /  $^{15}\text{N}$ - $^{13}\text{C}$ ) T7L protein was produced following the protocol described in **Section 2.2.2 of Chapter 2**.

### **4.2.2 Circular dichroism (CD) spectroscopy**

Temperature and chemical denaturant (urea) based CD measurements of 20  $\mu\text{M}$  T7L at pH 7, 5, 4 and 3 were carried out as described in **Section 2.2.4 of Chapter 2** and **Section 3.2.2 of Chapter 3**. The thermodynamic parameters were obtained by normalizing the data using equation 1 and fitting the CD denaturation curve to a two-state unfolding mechanism using equation 2 as described in **Section 3.2.2 of Chapter 3**.

### 4.2.3 Steady state fluorescence measurements

Fluorescence measurements of thermal and chemical denaturation of 20  $\mu$ M T7L at pH 7, 5, 4 and 3 were performed as described in **Section 2.2.5 of Chapter 2** and **Section 3.2.3 of Chapter 3**.

At each pH value, buffer containing ANS and ANS+T7L were measured and the resultant fluorescence spectra were used to assess the spectral properties. For urea denaturation studies, ANS fluorescence experiments recorded on T7L-ANS samples were performed under identical conditions as discussed in **Section 3.2.3 of Chapter 3**. ANS emission intensities at 470 nm ( $I_{\max}$ ) were used to analyze denaturation curves. The curves were normalized by using the following equation.

$$F_{App} = \frac{S_F - S_{Obs}}{S_F - S_U} \quad (4.1)$$

where,  $F_{App}$ ,  $S_F$ ,  $S_U$  and  $S_{Obs}$  represent the same as described in equation (1).

The thermodynamic parameters were obtained by fitting the fluorescence based denaturation curves to a two-state unfolding mechanism ( $N \leftrightarrow U$ ) using equation 3.2 as described in **Section 3.2.2 of Chapter 3**.

### 4.2.4 Cell lysis turbidimetric assay

pH- and temperature dependent amidase activity of T7L was measured by turbidimetric assay as described in **section 2.2.3** of chapter 2. As a measure of lysis efficiency, decrease in turbidity of cell suspension was monitored at every 2 min for a period of 12 min using a UV-Vis spectrophotometer. Activity measurements were performed in triplicates. Sensitized bacterial cells without T7L were used as a negative control, and the lysis profile of native T7L at pH 7 was considered as positive control. For temperature dependent reversibility based lysis assay, protein samples (0.1 mM) at pH 7, 5, 4 and 3 were gradually heated and incubated at 90  $^{\circ}$ C for 20 min. After incubation, the samples were gradually cooled to room temperature, centrifuged, and then supernatants were recovered. Protein samples at pH 5, 4 and 3 were adjusted back to 7 using a microelectrode and their respective concentrations in the supernatant were re-calculated using UV-Vis spectroscopy.

#### **4.2.5 Nuclear magnetic resonance (NMR) spectroscopy**

2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were recorded at pH 5, 4 and 3 as described in **Section 2.2.7** of **Chapter 2**.

To perform pH-dependent  $^1\text{H}$ - $^{15}\text{N}$  HSQC titration, the 700  $\mu\text{M}$   $^{15}\text{N}$ -labeled T7L sample was taken from pH 7 to 3 by adjusting the pH with a microelectrode. The completely unfolded spectrum of [ $^{15}\text{N}$ ] T7L was obtained by dissolving the lyophilized protein sample in 9.5 M urea buffer (pH 3). The sample at pH 3 was lyophilized, and 500  $\mu\text{L}$  of saturated urea buffer ( $\sim 10.5$  M, pH 3) was added to the lyophilized powder along with 10%  $\text{D}_2\text{O}$ .

For backbone resonance assignments, a series of standard 3D NMR experiments such as HNCO, HNCA, HNCACB, and CBCA(CO)NH, were recorded using 1 mM T7L at pH 3 [34]. The  $^{13}\text{C}$  carrier frequency was set at 52 ppm for HNCA; at 41 ppm for HNCACB, CBCA(CO)NH; and at 173 ppm for HNCO.

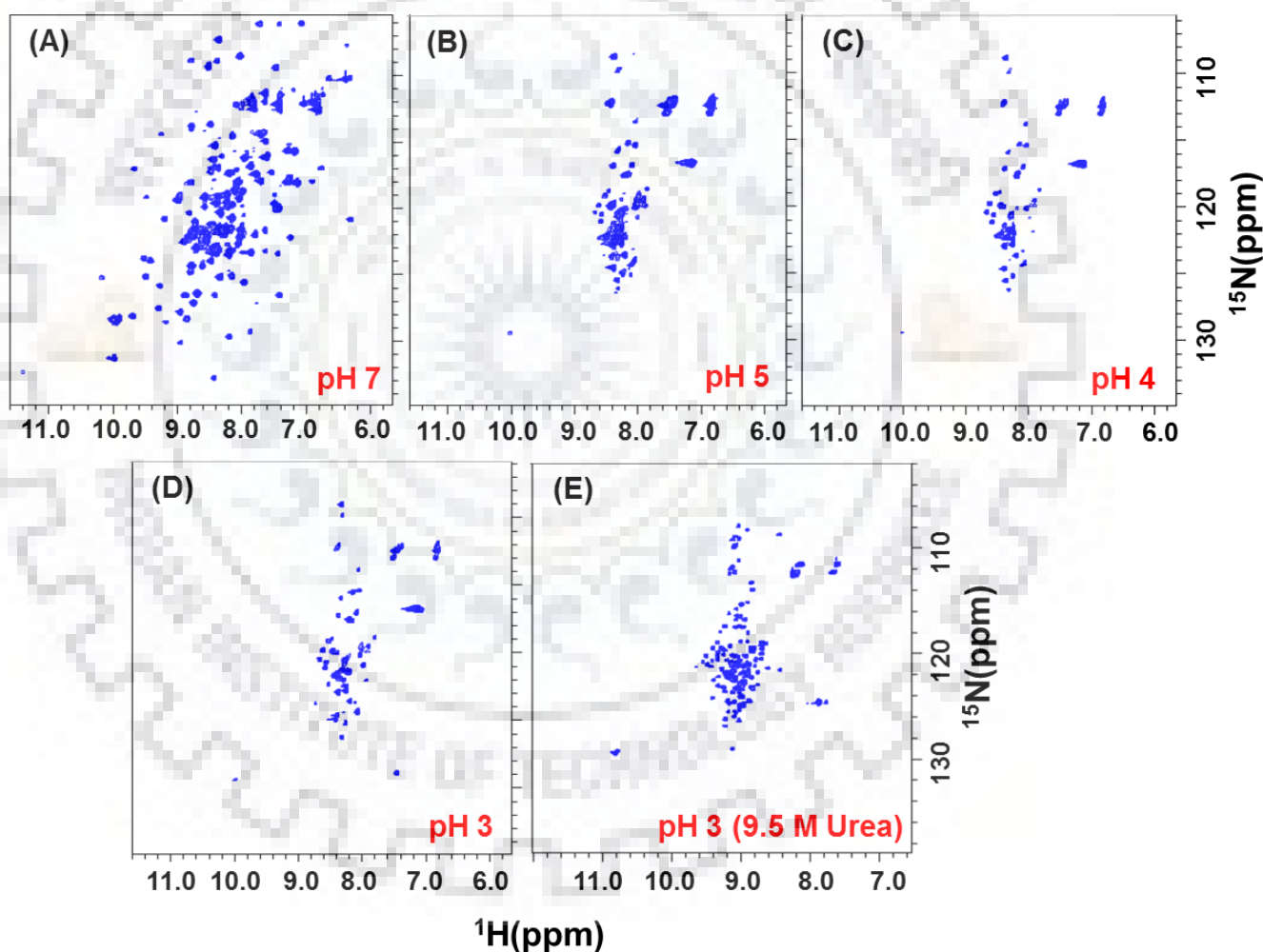
For urea based T7L HSQC experiments, 10.5 M stock solutions of urea were made, adjusted the pH as required, and added to lyophilized T7L (0.25 mM). The samples were diluted with respective buffers to prepare NMR samples with a final urea concentration from 0 to 9.5 M. The spectral widths in the  $^1\text{H}$  /  $^{15}\text{N}$  dimensions are 16 ppm / 32 ppm at pH 7, and 11 ppm / 24 ppm at pH 5, 4 and 3 respectively.

Temperature dependent  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of T7L at pH 5, 4 and 3 were acquired in the temperature range 20 to 70  $^{\circ}\text{C}$  at regular intervals of 10  $^{\circ}\text{C}$ . Thermal reversibility of the protein has been assessed by recording a HSQC spectrum, after cooling the sample back to 20  $^{\circ}\text{C}$ . Amide proton ( $^{\text{N}}\text{H}$ ) temperature coefficients ( $\Delta\delta_{\text{HN}}/\Delta T$ ) were analyzed by fitting the chemical shift changes to a linear equation.  $^{15}\text{N}$  transverse relaxation experiments ( $T_2$ ) and steady-state  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE measurements were obtained using the pulse sequences described by Farrow *et. al.*, [35] and performed as described in **Section 3.2.5.3** of **Chapter 3**.  $T_2$  measurements were done using the following relaxation delays: 10, 30, 50, 70, 90, 130, 170 and 230 ms. het-NOE experiments were performed using the same parameters as described in **Section 3.2.5.3** of **Chapter 3**.

## 4.3 Results

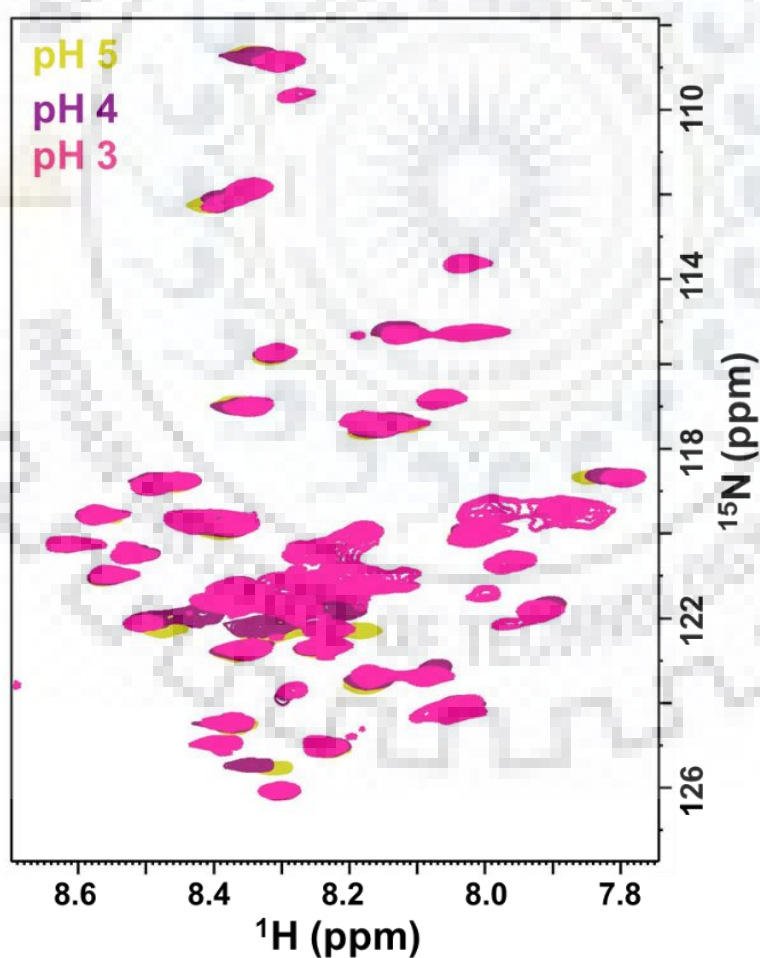
### 4.3.1 Structural heterogeneity of T7L partially folded conformations

In **Chapter 2** it has been observed that T7L exists in partially folded conformation below pH 6. Such a PF conformation is established by the difference in the spectral resolution and the total number of peaks in the two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (**Figure 4.1 A-D**). However, once the pH is lowered below 6 (i.e., between pH 5 and 3), the amide shift dispersion of T7L deteriorates in both dimensions with peaks clustering around 8.3 ppm (NH frequency) suggesting a huge conformational change.



**Figure 4.1:** (A–D)  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of T7L at 25 °C at different pH values and (E) pH 3 denatured state using 9.5 M urea.

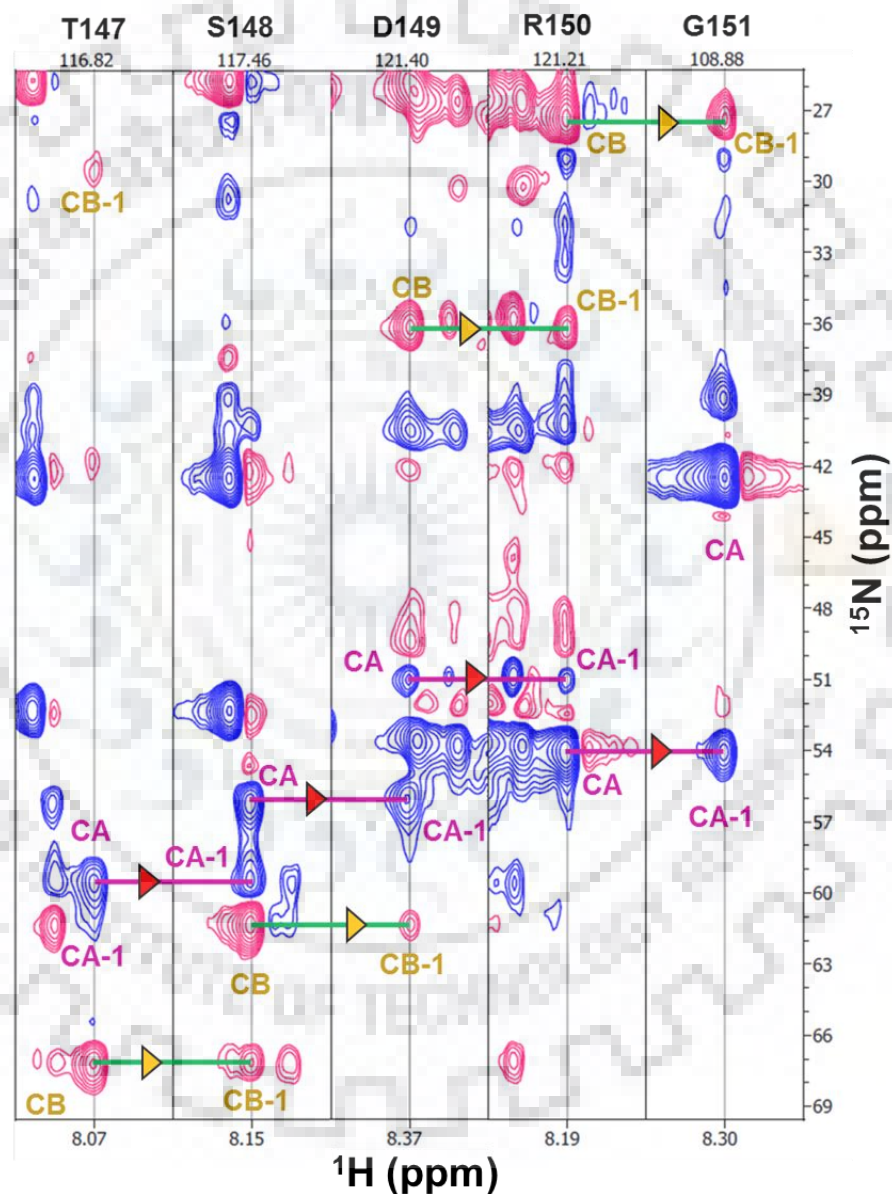
HSQC spectra of these PF states suggest a highly dynamic conformation(s) in the intermediate exchange of NMR time scale, which is evident from the total number of peaks present in the spectrum (~ 42 out of 147 resonances). In the case of the unfolded and/or denatured state, although the protein loses its dispersion in its amide proton axis, all the peaks do appear in the spectrum. Therefore, to rule out the possibility of the unfolded conformation, the HSQC spectra of T7L in a 9.5 M urea denaturant was recorded (**Figure 4.1 E**). A total of ~130 peaks with a good dispersion on the  $^{15}\text{N}$  axis were observed, a signature of the denatured protein spectrum thus clearly ruling out the denatured conformation at low pH. Although, the NMR HSQC spectra differentiated the native and the PF states of T7L, the HSQC spectra by itself did not provide any evidence for the structural heterogeneity, as the position and number of resonances at pH 5, 4 and 3 are almost identical (**Figure 4.2**). Hence, to gain insights into the differential global features of these PF states, optical spectroscopic (CD and fluorescence techniques) and NMR studies have been performed.



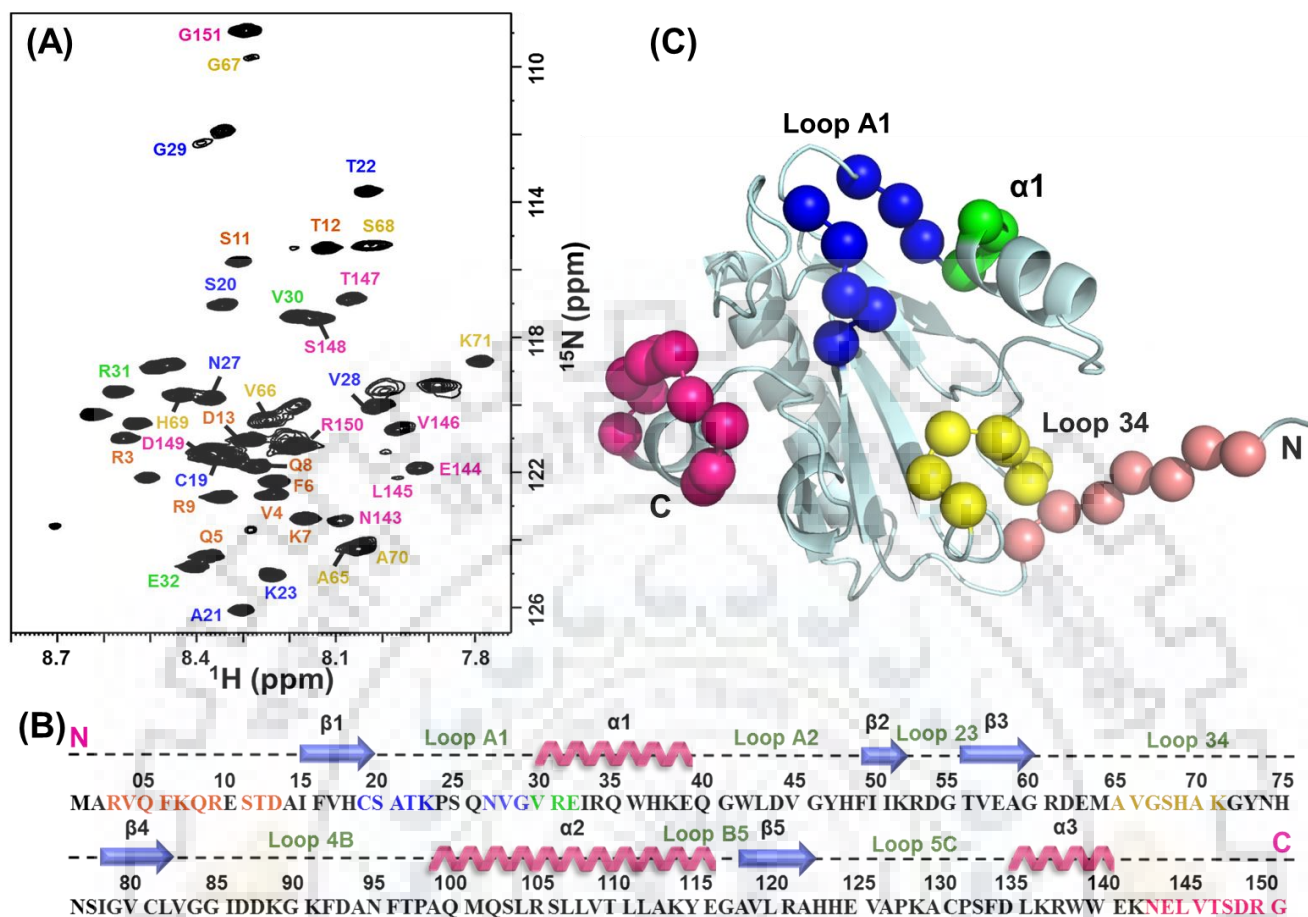
**Figure 4.2:** A  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectral overlay of the T7L at pH 5, 4 and 3, at 25 °C.

### 4.3.2 Residue level insights into structure and dynamics of T7L partially folded states

In order to unravel the structural characteristics of T7L low pH PF states, backbone resonance assignments of the observed NH resonances were obtained by analyzing triple resonance NMR experiments. The HNCACB strip is shown in **Figure 4.3** to describe the connectivity of C $\alpha$  (CA) and C $\beta$  (CB) resonances for the stretch of T147-G151.



**Figure 4.3:** Representative HNCACB strip plot showing sequential connectivities for residues T147 to G151. Purple lines with red triangles are showing the connectivity of CA resonances and green lines with yellow triangles are showing the connectivity of CB resonances.



**Figure 4.4:** Backbone NMR assignment of T7L at pH 3 and 25 °C (A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of assigned backbone amide cross-peaks labeled with one letter amino acid code and residue number (B) Amino acid sequence of T7L showing assignment summary and (C) Structural localization of assigned residues onto crystal structure of T7L (PDB ID: 1LBA) generated by using PyMOL molecular graphics software.

The resonance assignments were obtained for pH 3 PF state and were transferred directly on to the spectra at pH 5 and 4, as all the PF states represent highly overlapping spectral features with almost same number of NH resonances (**Figure 4.4 A**). The non-overlapping nature of observed NH resonances was checked using highly sensitive 3D HNCOC spectrum. Clearly, for all the residues present in the spectrum only one  $i-1$  peak of carbonyl carbon was observed, indicating that multiple amino acids/residues were not present under a given NH peak. A total of 37 residues were assigned out of  $\sim 42$  observable peaks (**Figure 4.4 B and C**). It is evident from the assignment that most of the residues either belong to the N, C-terminal and some interconnecting loops. Moreover, a few



residues at the termini of  $\beta$ 1 sheet and  $\alpha$ 1 helix were also assigned. The corresponding chemical shift values of assigned amino acid residues are listed in **Appendix II**.

### **4.3.3 Differential stability characteristics of T7L partially folded states**

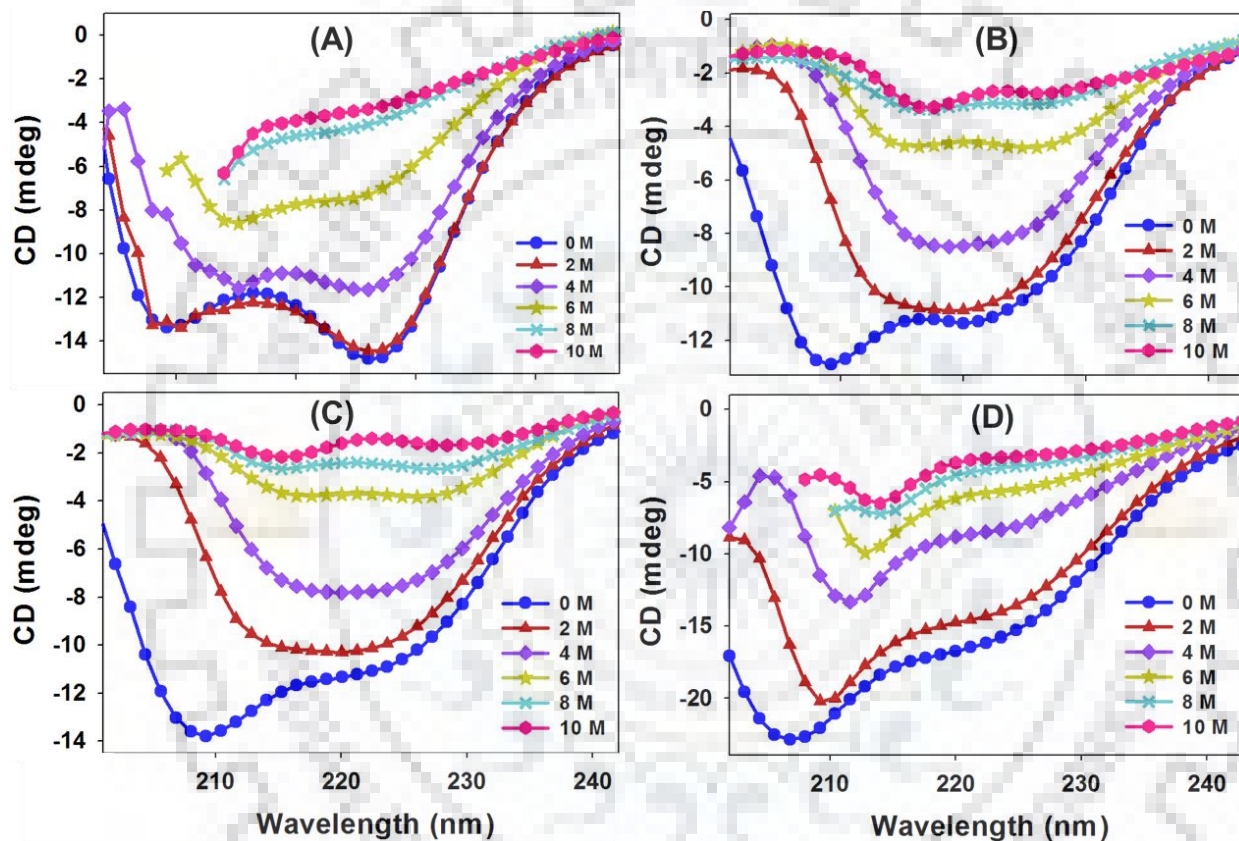
Structural modulations and conformational transitions greatly influence the stability of the proteins. In order to elucidate the accompanying stability changes associated with pH-dependent conformational transition of T7L, a comparative analysis of its stability profiles at different pH values have been performed using chemical denaturant urea and temperature.

#### **4.3.3.1 Chemical denaturation profiles of low pH conformations**

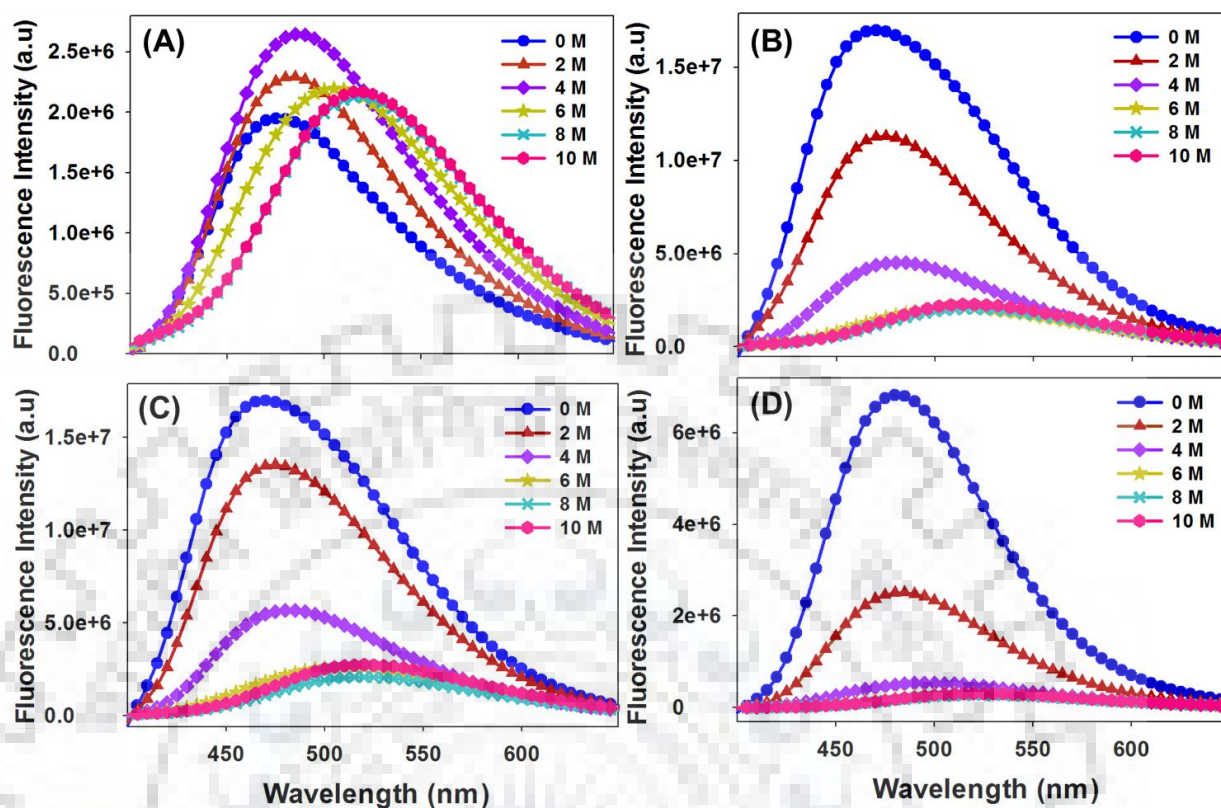
Urea induced denaturation profiles were obtained using far-UV CD and ANS fluorescence experiments (**Figure 4.5 A-D**). At 0 M urea, for native and PF states, the far-UV CD evidenced a characteristic double well spectrum, although the ratio of the structured (222 nm) to denatured protein (208 nm) varied significantly. Addition of urea to T7L has resulted in specific spectral patterns. In the case of pH 7 native conformation and pH 3 PF conformation, T7L unfolding followed a conventional melting of structural elements upon increasing of urea concentration (0 M to 10 M), which is evidenced by the gradual decrease in CD signal at 222 nm (**Figure 4.5 A and D**). However, in the case of pH 5 and pH 4 PF states, the unfolding curves at 2 M and 4 M urea evidenced for an existence of molten globule like structure, as the double well feature has coalesced into a broad single well (**Figure 4.5 B and C**). A further increase of urea concentrations above 6 M resulted in attaining back their double well spectral features with decreased ellipticity value at 222 nm, indicating the denaturation phenomenon.

The associated changes in the denaturation process for the native state and the PF state were also monitored using ANS fluorescence experiments (**Figure 4.6 A-D**). In line with the CD experiments, the ANS fluorescence experiments showed a continuous decrease of the ANS intensities upon increasing of urea concentrations. As no such specific hydrophobic pockets are exposed in the native state, the urea unfolding transition evidenced for a shift in wavelength from 480 nm to 520 nm with minimal intensity changes (**Figure 4.6 A**). Whereas, all the PF states experienced both intensity loss and wavelength shift phenomenon during the denaturation process (**Figure 4.6 B-D**). Initially, the PF states resulted in sharp decrease of ANS intensity indicating the conformational change at their hydrophobic patches. For pH 5 and pH 4 conformations sharp

intensity loss was observed until 6 M urea; and for pH 3 PF state it was 4 M urea, suggesting that the latter is a less stable conformation (**Figure 4.6 B-D**). The gradual decrease of intensity around 2 M and 4 M urea also supports the partially collapsed spectral features that are observed in the CD spectra for the pH 5 and pH 4 PF states. Further increase of urea concentration above 6 M urea resulted in complete unfolding of the protein, which is evidenced by the bathochromic shift of ANS from 470 to 520 nm.

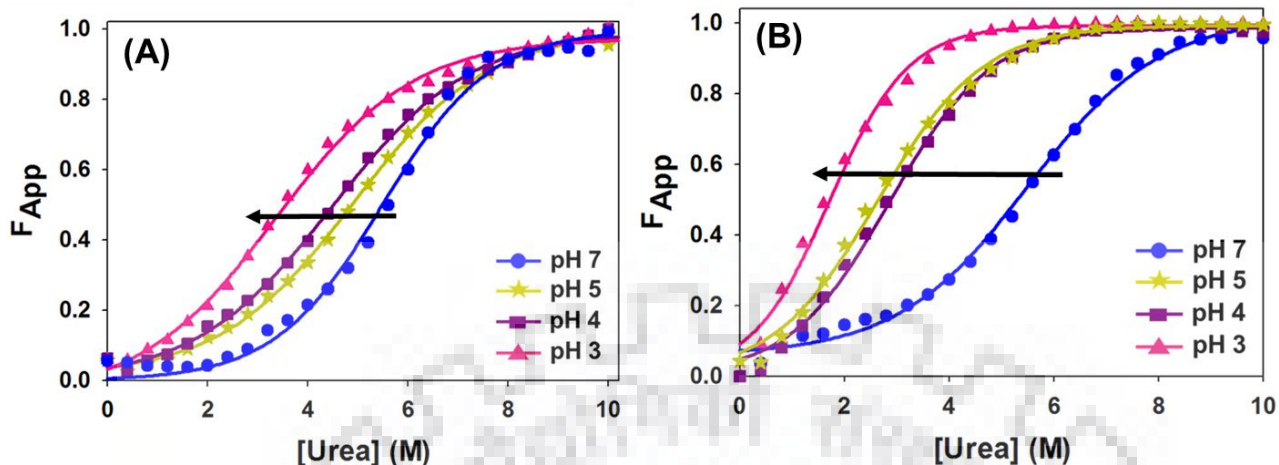


**Figure 4.5:** (A-D) Far-UV CD profiles of urea induced chemical denaturation of T7L: (A) pH 7 (B) pH 5 (C) pH 4 and (D) pH 3.



**Figure 4.6:** ANS fluorescence emission profiles of urea induced chemical denaturation of T7L: (A) pH 7 (B) pH 5 (C) pH 4 and (D) pH 3.

In order to characterize the conformational stabilities and the transition midpoints of the urea unfolding processes, all the CD and ANS fluorescence profiles were analyzed using a two-state unfolding transition (**Figure 4.7 A and B, Table 4.1**). The free energy values ( $\Delta G$ ) and the transition midpoints ( $C_m$ ) undoubtedly indicated that the PF states are less stable compared to its native state. Moreover, these results suggested that, among the PF states, the stability pattern follows pH 5 > pH 4 > pH 3. Such a stability profile is completely justified considering the extent of protonation of ionized side chains of a protein (T7L), and the extent of hydrophobic opening upon lowering pH (**Figure 2.8 B, Chapter 2**). The differences in the obtained  $\Delta G$  and  $C_m$  values were also observed between the CD and fluorescence results. The underestimated thermodynamic parameters in the ANS measurements can be attributed to the following: (a) hydrophobic opening preceding the complete secondary structural melting and/or; (b) inability of ANS to assess the global unfolding phenomenon due to its localized binding in hydrophobic patches.

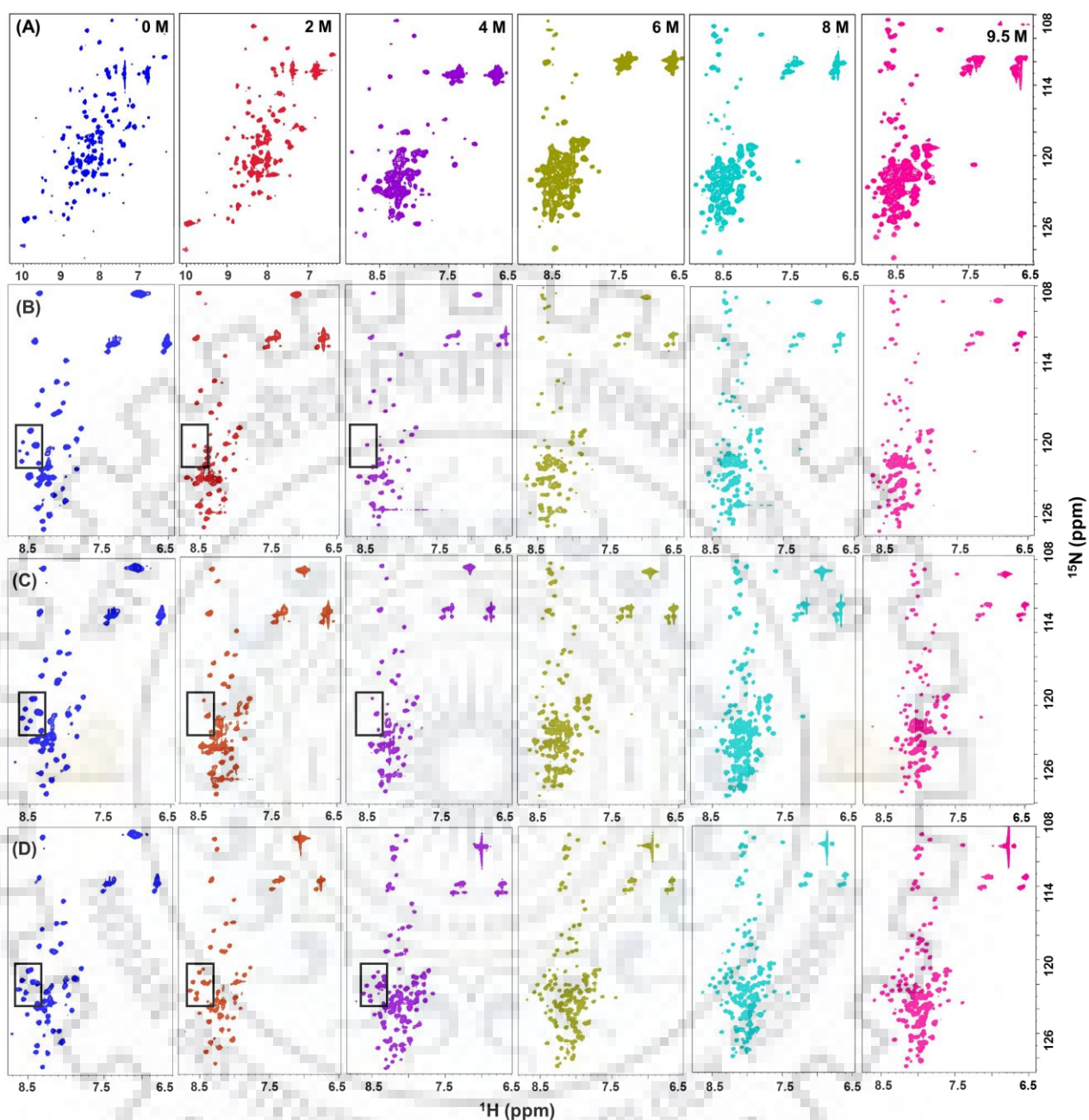


**Figure 4.7:** Normalized denaturation curves monitored by (A) far-UV CD ellipticity at 222 nm and (B) ANS fluorescence as a function of urea concentration.

**Table 4.1:** Unfolding free energies ( $\Delta G$ ) and transition midpoints ( $C_m$ ) of T7L at different pH values for urea denaturation process.

pH	Circular dichroism		Fluorescence	
	$\Delta G$ (kcal/mole)	$C_m$	$\Delta G$ (kcal/mole)	$C_m$
7	$-6.1 \pm 0.3$	5.6	$-4.5 \pm 0.3$	5.5
5	$-3.7 \pm 0.1$	4.9	$-2.3 \pm 0.1$	2.6
4	$-3.3 \pm 0.1$	4.6	$-2.0 \pm 0.1$	2.4
3	$-2.5 \pm 0.2$	3.4	$-1.0 \pm 0.1$	1.0

To further visualize the urea unfolding characteristics at residue-level,  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of T7L were acquired at various urea concentrations ranging from 0 to 9.5 M. At pH 7, population of native state is predominant up to 2 M urea as confirmed by the wide chemical shift dispersion in both the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions (**Figure 4.8 A**). At 4-6 M urea, significant reduction in number of resonances, and loss in chemical shift dispersion clearly confirms it as the unfolding transition zone. HSQC spectra at 8 M urea and above resulted in appearance of all the resonance with a dispersion loss, thus evidencing the complete denaturation of T7L. In the case of pH 5 and 4 PF states, the number of resonances almost remained same in number at 2 M and 4 M urea, although some of the 0 M residues disappeared/broadened upon addition of urea (**Figure 4.8 B-C**).



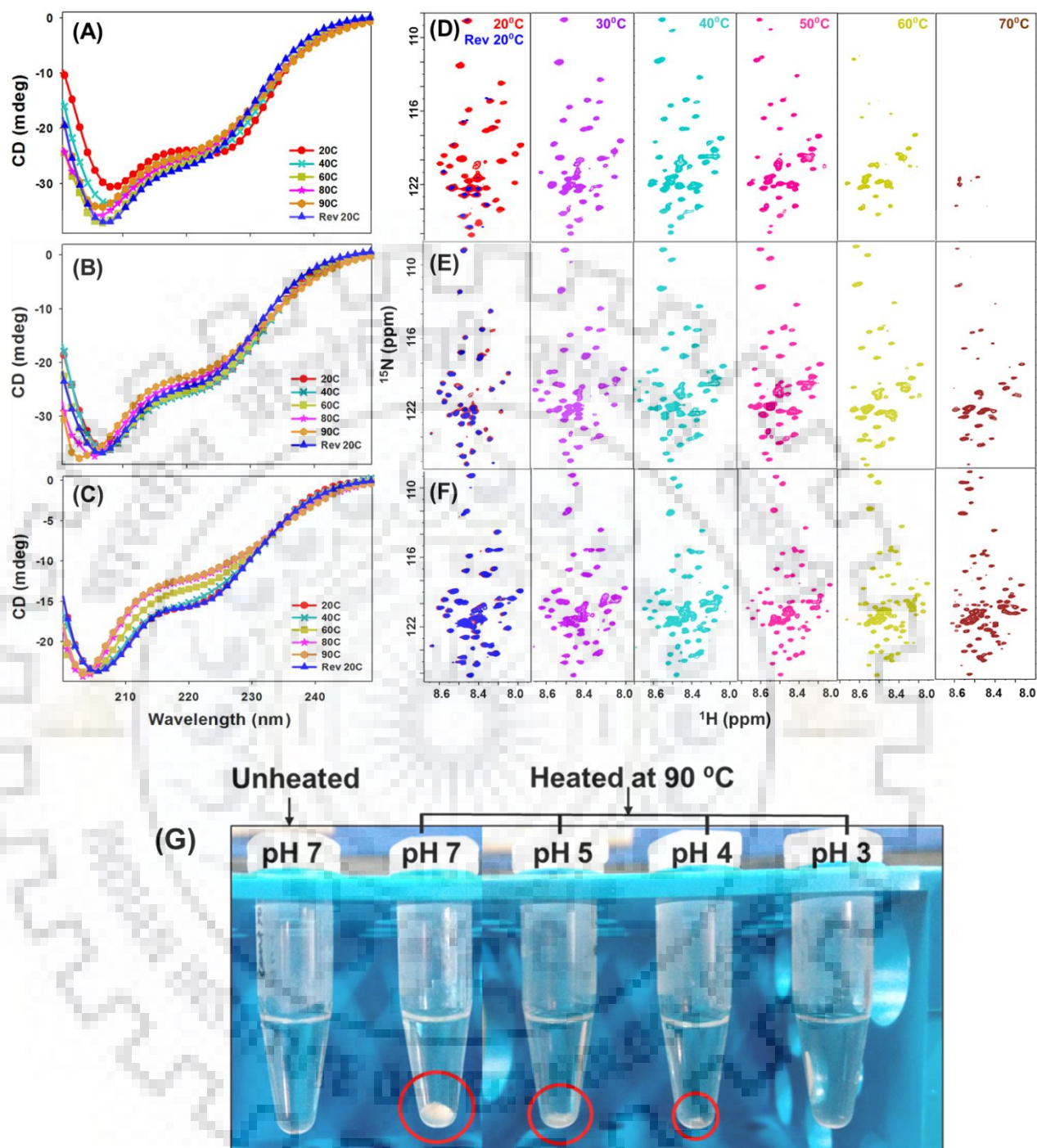
**Figure 4.8:**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of T7L at 0 to 9.5 M urea concentration: (A) pH 7 (B) pH 5 (C) pH 4 and (D) pH 3. The black rectangle enclosing few NH resonances in 0 M, 2 M and 4 M urea spectra depicts the broadened peaks.

Few new resonances popped up with very little intensities. Such variability in the peaks clearly indicates the structural changes within the PF states and indeed supports the observed partially collapsed states around 2 M - 4 M urea CD spectra. Upon increasing the urea  $\sim$  6 M and above, the protein is completely unfolded (**Figure 4.8 B-C**). At pH 3 no such broadening of the

peaks were observed in the range of 2 M to 4 M urea (**Figure 4.8 D**). A progressive melting of the structure upon increase of urea concentration was evident from the total number of resonances. The pH 3 PF state almost unfolded at 4 M, as no major changes have been noticed thereafter (**Figure 4.8 D**).

#### **4.3.3.2 Thermal stability and reversibility of low pH conformations**

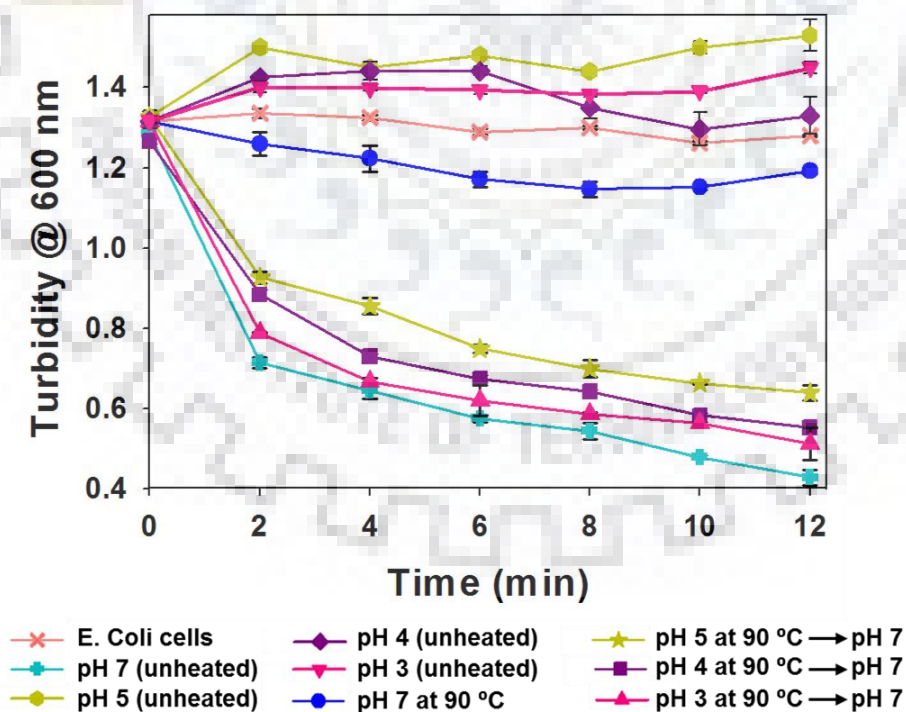
Native state conformation of T7L is not stable at higher temperatures ( $> 60$  °C) as it forms irreversible soluble and insoluble aggregates. The thermal stability characteristics of PF states were measured in the temperature range 20 °C to 90 °C using far-UV CD and  $^1\text{H}$ - $^{15}\text{N}$ HSQC NMR spectroscopy (**Figure 4.9**). These T7L low pH conformations exhibited variable thermal stability/reversibility features. The pH 5 conformation shows reasonable thermal stability up to 50 °C and forms completely irreversible soluble/insoluble aggregates above this temperature (**Figure 4.9 A and D**). The pH 4 conformation is comparatively more stable and thermally reversible to that of pH 5 conformation, as most of the resonances appeared back upon cooling the sample back to 20 °C (**Figure 4.9 B and E**), although a certain extent of insoluble aggregates at high temperature was observed. In contrast to the pH 5 and pH 4 PF states, the pH 3 conformation is highly thermally stable and is completely reversible (**Figure 4.9 C and F**). No detectable feature of aggregation was observed in the pH 3 sample. All the peaks were present even at 70 °C in addition to some minor unfolded conformations, moreover all the peaks reappeared upon reversing the temperature back to 20 °C (**Figure 4.9 F**). The formation of visible aggregates/precipitates of T7L protein at different pH values upon systematic heating and cooling of the samples is quantified using a temperature regulated water bath experiment (**Figure 4.9 G**). Fair amount of precipitate was observed in pH 7, 5 and 4 heated samples but not in pH 3. Concentration measurements of the recovered supernatants suggested a recovery of 60 % and 80 % of the protein in soluble forms at pH 5 and 4, whereas total of 100 % is recovered at pH 3. All these thermal stability experiments establish the differential thermal stability and aggregation characteristics of T7L PF states.



**Figure 4.9:** Thermal stability of T7L PF conformation monitored by far-UV CD from 20 °C to 90 °C; (A) pH 5 (B) pH 4 and (C) pH 3.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of T7L from 20 °C to 70 °C at (D) pH 5 (E) pH 4 and (F) pH 3. Spectra of the protein samples that were cooled back to 20 °C (Rev 20 °C) after heating to 70 °C to determine thermal reversibility were also shown. (G) Temperature regulated water bath experiment showing the amount of precipitation (enclosed in red circles).

#### 4.3.4 pH and temperature dependent amidase activity

Functional studies conducted in **section 2.3.1** of Chapter 2 suggested that T7L shows low/negligible peptidoglycan cleaving activity below pH 6, and the pH-dependent conformational transition of T7L between pH 7 and pH 3 is completely reversible. In this context, understanding the functional reversibility of pH/temperature dependent T7L low pH conformations is essential in order to comprehensively delineate their functional characteristics, and also to know the activity profiles of natively folded/aggregated conformation in the supernatant of thermally challenged protein samples. To assess the functional behaviour of these low pH conformations, pH and temperature dependent amidase activity of T7L was measured by performing turbidimetric assay using sensitized *E. coli* cell suspensions (**Figure 4.10**). Aliquots of bacterial cell suspensions treated with T7L at pH 7 were taken as a positive control, and *E. coli* cells alone as a negative control. Two sets of experiments were performed; (a) in the first experiment, T7L at its native or low pH state i.e., cell lysis assay was performed at different pH values (pH 5, 4 and 3). (b) In the second experiment, the lysis assay was performed exclusively at pH 7, but the samples at different pH were thermally challenged to 90 °C, cooled back, and the pH of the supernatants were adjusted back to pH 7.



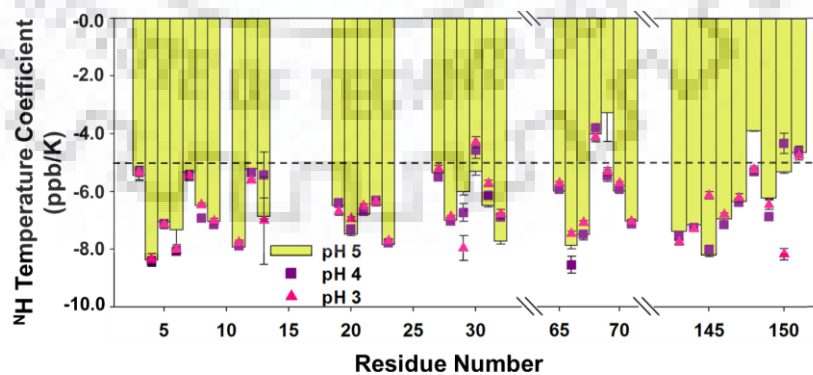
**Figure 4.10:** Turbidimetric assay depicting the pH and temperature dependent amidase activity profiles of T7L.



As described in **Chapter 2**, the low pH experiments performed according to (a) did not show any cell lysis activity, indicating that the protein is completely inactive in the partially folded state at all the three pH values (**Figure 2.6, Figure 4.10**). However, a differential pattern of cell lysis activity was observed in experiential set (b). The cell lysis activity of thermally challenged samples yielded the following order pH 3 > pH 4 > pH 5, although cell suspensions were incubated in equal amount of protein (50  $\mu$ g) at all pH values (**Figure 4.10**). The observed trend in the lysis activity of these low pH conformations can be attributed to the presence of 100 % native conformation for pH 3 sample and certain extent of inactive conformations (soluble aggregates) for pH 4 and pH 5 samples subjected to same treatment (b). Such an explanation is in line with the results of thermal stability experiments presented in the previous section.

#### 4.3.5 Effect of temperature on amide proton chemical shifts

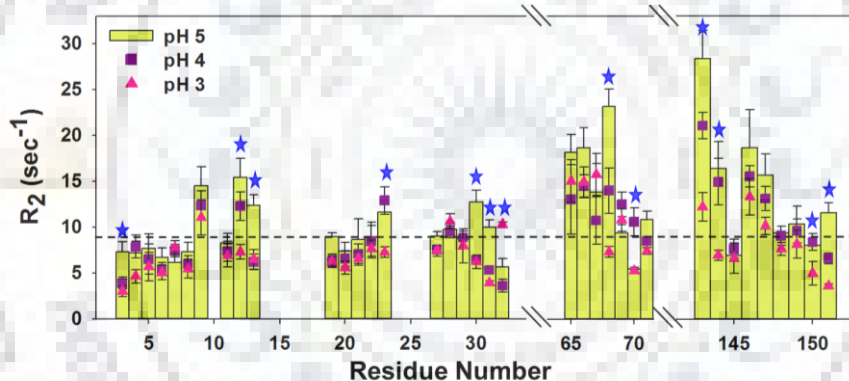
Temperature dependence of the amide proton chemical shifts were measured by recording,  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra as a function of temperature to substantiate the random coil conformations of the assigned chemical shifts of PF states (**Figure 4.11**). Data revealed that most of the amide protons had temperature coefficient values in the range of - 6 ppb/K to - 10 ppb/K, which are much higher than the usual temperature coefficients range (< - 4.5 ppb/K) for structural elements, thus establishing that the assigned residues are devoid of H-bonding and are present in the unstructured elements [36,37]. Few residues such as S68, H69 (Loop 34) and S148 (C-terminal) are notable as their temperature coefficients are < - 4 ppb/K at pH 5 and are > - 5 ppb/K at pH 4 and 3 suggesting that these residues might be involved in some local H-bonding interactions at pH 5 that melted upon lowering the pH (4/3).



**Figure 4.11:** Overlay of sequence specific amide proton temperature coefficients of T7L at pH 5, 4 and 3.

#### 4.3.6 Comparative studies on structural- dynamics of PF conformations

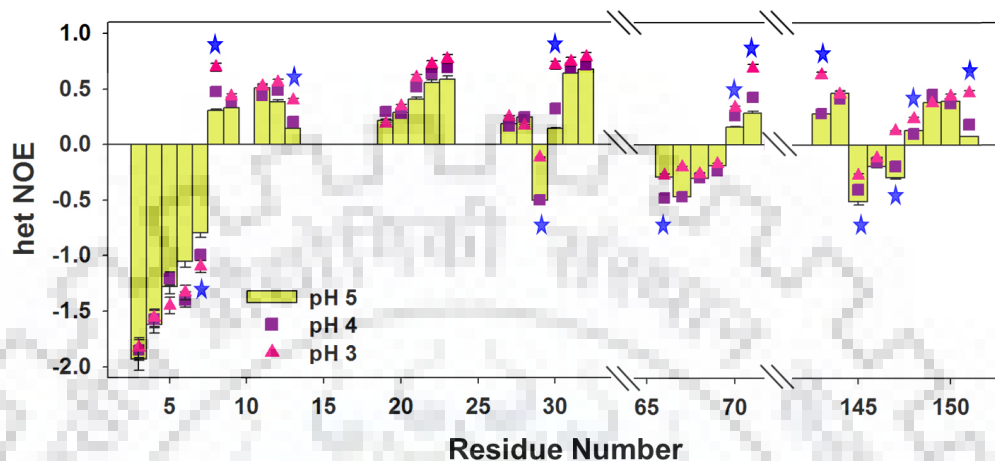
As partially folded proteins are highly flexible and exhibit heterogeneous molecular motions, backbone  $^{15}\text{N}$  transverse relaxation rate measurements ( $R_2$ ) were used to probe their dynamics and conformational exchange behavior in  $\mu\text{s}$  to  $\text{ms}$  time scale.  $R_2$  values obtained at all the three pH values are shown in **Figure 4.12**. The  $R_2$  values showed significant differences at the measured pH values. Several residues exhibited enhanced  $R_2$  values compared to their average value ( $\sim 9.3 \pm 1.3 \text{ s}^{-1}$ ), indicating the conformational exchange behavior at that sites. These sites are majorly concentrated in the loop 34 and C-terminal region. It is interesting to note that at pH 5, the number of residues exhibiting conformational exchange are greater compared to pH 4, which indeed is higher than at pH 3. In fact, most of these exchangeable sites at pH 3 obtained a mean  $R_2$  value suggesting a dynamically uniform PF conformation. All the residues that showed differential  $R_2$  values upon pH changes are highlighted with “\*” in **Figure 4.12**.



**Figure 4.12:** Overlay of sequence specific  $^{15}\text{N}$  transverse relaxation rates ( $R_2$ ) of PF conformations at pH 5, 4 and 3. Residues that showed differential  $R_2$  values upon pH change are highlighted with “\*”.

Further, to substantiate the differential dynamics, the steady state het-NOE experiments was also performed, which monitor dynamics on a faster (ps–ns) time scale **Figure 4.13**. Overall the observed het-NOE values are  $\leq 0.5$  for most of the residues indicating that the PF forms are highly dynamic in the faster time scale motions. It is worth noting that, sequence specific, and pH dependent variations have also been observed in the het-NOE values. Several segments have shown very low / negative NOE values. They include; N- and C-terminal of protein and the residues present around the  $\alpha 1$ -helix and the loop 34. It is worth noting that the C-terminal and loop 34 are also dynamic in the  $\mu\text{s}$ - $\text{ms}$  time scale as evidenced by enhanced  $R_2$  values. The magnitude of het-NOE values is

comparatively higher for several of the residues at pH 3 conformation in comparison to pH 5 and 4, thus pointing towards its restricted motion. All these results establish that the low pH PF forms of T7L are structurally and dynamically variable.

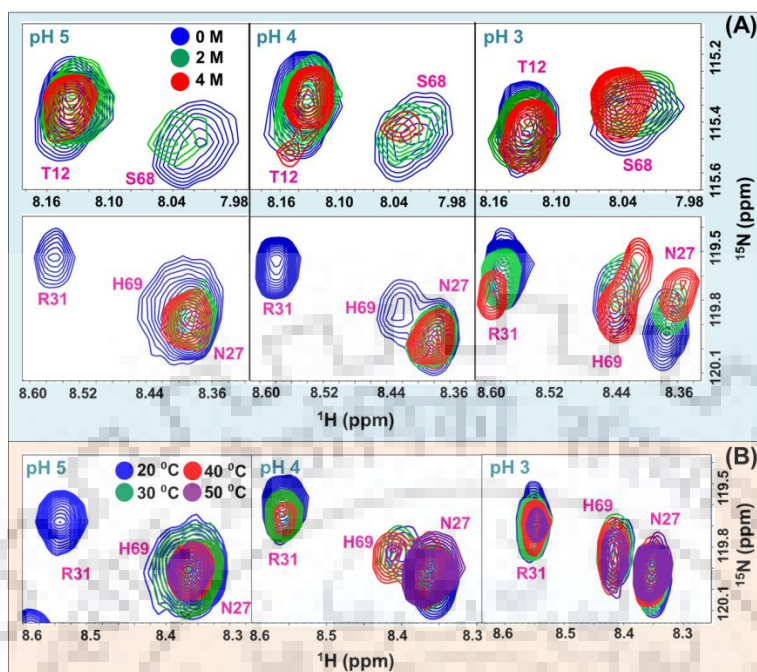


**Figure 4.13:** Overlay of sequence specific steady-state heteronuclear NOEs (het-NOE) of PF conformations at pH 5, 4 and 3. Residues that showed differential NOEs upon pH change are highlighted with “\*”.

#### 4.4 Discussion

##### 4.4.1 Mechanistic insights into differential biophysical characteristics of T7L PF conformations

Characterization of partially folded forms is highly challenging and an exciting research topic in the areas of molecular biophysics and structural biology. Most of the biophysical techniques yield very limited information in characterizing even one particular conformation of partially collapsed /molten globule structure at atomic level. The experimental studies involved in this chapter delineated the structural, stability and dynamic differences of low pH partially collapsed structures of T7L with the aid of several structural techniques. A recent pH-dependent computational study on T7L reported that upon decrease in pH,  $\alpha$ 1-helix unwinds partially [38]. Indeed, the NMR measurements on the low pH PF forms presented in the current study established that the helical segment (V30-R31-E32) of  $\alpha$ 1-helix was unwound upon pH perturbation (**Figure 4.4**). Such a partially unwinding will disorient the catalytic residues (H18 and Y47) and gate keeping residues (H37) thus making the enzyme inactive at low pH conformation. Indeed, lysis experiments evidenced the complete loss of activity of the T7L at pH 5 and below (**Figure 4.10**).



**Figure 4.14:** Overlay of a part of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of some T7L hotspot residues illustrating loss of resonance peaks from (A) 0 M to 4 M urea and (B) 20 to 50 °C at pH 5, 4 and 3.

The present study also provided insights into the hierarchy of the structural changes that are responsible for differential stability and dynamics behavior of T7L PF states. It is evident from both the steady state and time resolved ANS fluorescence experiments that T7L exposed its hydrophobic pockets upon structural collapse, and the extent of hydrophobic opening further increased upon lowering of pH. The lack of NMR signals from almost all the structural elements and presence of significant amount of secondary structure in CD experiments at all the three pH values depict that the PF conformations contain most of the essential structural elements intact in its globule. The differential stability features assessed by chemical denaturation and thermal perturbation studies as measured by integrative biophysical tools depicted the vulnerable sites/hot spots of T7L opening. Both these experiments suggested that the  $\alpha$ 1-helix and segment connecting the  $\beta$ 3- $\beta$ 4 sheets are the key hot spots for structural opening transitions upon chemical and thermal denaturation (**Figure 4.14**). Dynamic measurements including  $R_2$  and het NOE values provided concrete evidences for enhanced dynamics in both ps-ns and  $\mu$ s-ms time scales in these structural regions, which can be directly correlated to their structural fragility. Further opening of partially unwound  $\alpha$ 1-helix resulted in differential melting of the low pH conformation at pH 5/4 at 2 M and 4 M urea, which resulted in further loss of resonance peaks in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (**Figure 4.14 A**). Moreover,

the same sites were found to be vulnerable in temperature induced aggregation/precipitation at pH 5 (**Figure 4.14 B**).

Interestingly both these vulnerable sites are present in the N-terminal segment of the protein, and they do accommodate all the nearby buried His residues (H18, H48, and H69) whose pKa values are  $\leq 4.5$  (**Table 2.4, Chapter 2**) thus establishing that His protonation is the major source of observed structural fragility. All these evidences suggest that  $\alpha 1$ -helix plays a regulatory role in the pH-dependent structural transition of T7L endolysin. Its unwinding and conformational exchange with respect to  $\beta 3$ - $\beta 4$  sheets, along with other interacting/ionizable residues are responsible for the observed variable stability and dynamics features of T7L PF states.

#### **4.5 Concluding remarks**

In summary, the present study deciphered the variable biophysical properties of T7L low pH conformations. Detailed biophysical investigations of low pH conformations under different denaturation conditions using optical techniques depicted that the low pH PF states indeed possess differential structural stabilities and dynamic characteristics. Partial loss of conformation in the  $\alpha 1$ -helix resulted in complete loss of functionality. The fragile  $\alpha 1$ -helix along with the  $\beta 3$ - $\beta 4$  segments comprises the structurally vulnerable sites for further opening of the protein's hydrophobic core, thermal irreversibility behavior and in formation of partially collapsed structures at low concentrations of urea at pH 5 and 4. All these results undoubtedly establish the presence of structural and dynamic heterogeneity among the T7L low pH PF states.

#### **4.6 References**

1. Dobson CM. Protein folding and misfolding. *Nature*, 426(6968), 884-890 (2003).
2. Onuchic JN, Wolynes PG. Theory of protein folding. *Current opinion in structural biology*, 14(1), 70-75 (2004).
3. Fink AL. Compact intermediates states in protein folding. In: *Proteins: Structure, Function, and Engineering*. (Springer, 1995) 27-53.
4. Hoerner JK, Xiao H, Kaltashov IA. Structural and dynamic characteristics of a partially folded state of ubiquitin revealed by hydrogen exchange mass spectrometry. *Biochemistry*, 44(33), 11286-11294 (2005).

5. Kuwajima K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Structure, Function, and Bioinformatics*, 6(2), 87-103 (1989).
6. Kishore D, Kundu S, Kayastha AM. Thermal, chemical and pH induced denaturation of a multimeric  $\beta$ -galactosidase reveals multiple unfolding pathways. *PLoS One*, 7(11), e50380 (2012).
7. Mohan PK, Chakraborty S, Hosur RV. Hierarchy of local structural and dynamics perturbations due to subdenaturing urea in the native state ensemble of DLC8 dimer. *Biophysical chemistry*, 153(1), 17-26 (2010).
8. Dubey VK, Jagannadham M. Differences in the unfolding of procerain induced by pH, guanidine hydrochloride, urea, and temperature. *Biochemistry*, 42(42), 12287-12297 (2003).
9. Yadav SC, Jagannadham MV. Complete conformational stability of kinetically stable dimeric serine protease milin against pH, temperature, urea, and proteolysis. *European Biophysics Journal*, 38(7), 981-991 (2009).
10. Yang A-S, Honig B. On the pH dependence of protein stability. *Journal of molecular biology*, 231(2), 459-474 (1993).
11. Pace CN, Grimsley GR, Scholtz JM. Protein ionizable groups: pK values and their contribution to protein stability and solubility. *Journal of Biological Chemistry*, 284(20), 13285-13289 (2009).
12. Bycroft M, Fersht AR. Stabilization of protein structure by interaction of  $\alpha$ -helix dipole with a charged side chain. *Nature*, 335(6192), 740-743 (1988).
13. Harrison JS, Higgins CD, O'Meara MJ, Koellhoffer JF, Kuhlman BA, Lai JR. Role of electrostatic repulsion in controlling pH-dependent conformational changes of viral fusion proteins. *Structure*, 21(7), 1085-1096 (2013).
14. Weng J, Tan C, Shen JR, Yu Y, Zeng X, Xu C, Ruan K. pH-induced conformational changes in the soluble manganese-stabilizing protein of photosystem II. *Biochemistry*, 43(16), 4855-4861 (2004).
15. Mohan P, Barve M, Chatterjee A, Hosur RV. pH driven conformational dynamics and dimer-to-monomer transition in DLC8. *Protein science*, 15(2), 335-342 (2006).
16. Langella E, Improta R, Barone V. Checking the pH-induced conformational transition of prion protein by molecular dynamics simulations: effect of protonation of histidine residues. *Biophysical journal*, 87(6), 3623-3632 (2004).

17. Bhat V, Kurouski D, Olenick MB, McDonald CB, Mikles DC, Deegan BJ, Seldeen KL, Lednev IK, Farooq A. Acidic pH promotes oligomerization and membrane insertion of the BclXL apoptotic repressor. *Archives of biochemistry and biophysics*, 528(1), 32-44 (2012).
18. Singh J, Udgaonkar JB. Unraveling the molecular mechanism of pH-induced misfolding and oligomerization of the prion protein. *Journal of molecular biology*, 428(6), 1345-1355 (2016).
19. Cabezon E, Butler PJG, Runswick MJ, Walker JE. Modulation of the oligomerization state of the bovine F1-ATPase inhibitor protein, IF1, by pH. *Journal of Biological Chemistry*, 275(33), 25460-25464 (2000).
20. Perico N, Purtell J, Dillon TM, Ricci MS. Conformational implications of an inversed pH-dependent antibody aggregation. *Journal of pharmaceutical sciences*, 98(9), 3031-3042 (2009).
21. Petkova A, Buntkowsky G, Dyda F, Leapman R, Yau W-M, Tycko R. Solid state NMR reveals a pH-dependent antiparallel  $\beta$ -sheet registry in fibrils formed by a  $\beta$ -amyloid peptide. *Journal of molecular biology*, 335(1), 247-260 (2004).
22. Srinivasan R, Jones EM, Liu K, Ghiso J, Marchant RE, Zagorski MG. pH-dependent amyloid and protofibril formation by the ABri peptide of familial British dementia. *Journal of molecular biology*, 333(5), 1003-1023 (2003).
23. Tipping KW, Karamanos TK, Jakhria T, Iadanza MG, Goodchild SC, Tuma R, Ranson NA, Hewitt EW, Redford SE. pH-induced molecular shedding drives the formation of amyloid fibril-derived oligomers. *Proceedings of the National Academy of Sciences*, 112(18), 5691-5696 (2015).
24. Shingate P, Warwicker J, Sowdhamini R. Energetic calculations to decipher pH-dependent oligomerization and domain swapping of proteins. *PloS one*, 10(6), e0127716 (2015).
25. Kumar DP, Tiwari A, Bhat R. Effect of pH on the stability and structure of yeast hexokinase a acidic amino acid residues in the cleft region are critical for the opening and the closing of the structure. *Journal of Biological Chemistry*, 279(31), 32093-32099 (2004).
26. Horng J-C, Cho J-H, Raleigh DP. Analysis of the pH-dependent folding and stability of histidine point mutants allows characterization of the denatured state and transition state for protein folding. *Journal of molecular biology*, 345(1), 163-173 (2005).
27. Sato S, Raleigh DP. pH-dependent stability and folding kinetics of a protein with an unusual  $\alpha$ - $\beta$  topology: The C-terminal Domain of the Ribosomal Protein L9. *Journal of molecular biology*, 318(2), 571-582 (2002).

28. Hom RA, Vora M, Regner M, Subach OM, Cho W, Verkhusha VV, Stahelin RV, Kutateladze TG. pH-dependent binding of the Epsin ENTH domain and the AP180 ANTH domain to PI (4, 5) P 2-containing bilayers. *Journal of molecular biology*, 373(2), 412-423 (2007).
29. Re F, Sesana S, Barbiroli A, Bonomi F, Cazzaniga E, Lonati E, Bulbarelli A, Masserini M. Prion protein structure is affected by pH-dependent interaction with membranes: A study in a model system. *FEBS letters*, 582(2), 215-220 (2008).
30. Olson LJ, Hindsgaul O, Dahms NM, Kim J-JP. Structural insights into the mechanism of pH-dependent ligand binding and release by the cation-dependent mannose 6-phosphate receptor. *Journal of Biological Chemistry*, 283(15), 10124-10134 (2008).
31. Stănciuc N, Aprodu I, Ioniță E, Bahrim G, Răpeanu G. Exploring the process–structure–function relationship of horseradish peroxidase through investigation of pH-and heat induced conformational changes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 147, 43-50 (2015).
32. Castellano A, Barteri M, Castagnola M, Bianconi A, Borghi E, Dellalonga S. Structure-function relationship in the serotransferrin: the role of the pH on the conformational change and the metal ions release. *Biochemical and biophysical research communications*, 198(2), 646-652 (1994).
33. Ivery MT, Gready JE. Structure-activity relationships and pH dependence of binding of 8-alkyl-N5-deazapterins to dihydrofolate reductase. *Journal of medicinal chemistry*, 37(24), 4211-4221 (1994).
34. Permi P, Annala A. Coherence transfer in proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 44(1), 97-137 (2004).
35. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry*, 33(19), 5984-6003 (1994).
36. Baxter NJ, Williamson MP. Temperature dependence of <sup>1</sup>H chemical shifts in proteins. *Journal of biomolecular NMR*, 9(4), 359-369 (1997).
37. Cierpicki T, Otlewski J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *Journal of biomolecular NMR*, 21(3), 249-261 (2001).
38. Borkotoky S, Murali A. A computational assessment of pH-dependent differential interaction of T7 lysozyme with T7 RNA polymerase. *BMC structural biology*, 17(1), 7 (2017).



## Concluding remarks and future perspectives

### 5.1 Concluding remarks

The thesis work presents the structural stability, functional and dynamic features of T7 bacteriophage endolysin T7L. Functional studies established that the lysis characteristics of T7L are pH-dependent, and it possesses maximum activity around pH 8. All the biophysical experiments demonstrated that pH plays a major role in maintaining the structural integrity of T7L structure. Once the pH is lowered below 6, a conformational transition occurs from native state to a partially folded (PF) state. These PF states at pH 5, 4 and 3 are characterized by the substantial secondary structure, absence of tightly packed tertiary structure and exposed hydrophobic pockets as compared to native state conformations (pH 8, 7 and 6). These PF states also display structural heterogeneity and differential exposure of hydrophobic surfaces as confirmed by fluorescence and translational diffusion experiments. Further, studies established that these pH-dependent structural transitions of T7L are reversible in nature.

NMR studies evidenced that the native state of T7L is conformationally heterogeneous as two different conformations were detected in the HSQC spectrum. For several residues, the alternate conformations were assigned. Results suggested that the C-terminal half of the protein including the long  $\alpha$ 2-helix access alternate conformation. All the NMR relaxation experiments unravelled that the native state is highly dynamic, and several residues experiencing enhanced conformational exchange. HX experiments revealed that the stability of T7L majorly lies in the structural elements  $\alpha$ 2-helix and  $\beta$ 2,  $\beta$ 4-sheets of T7L. In the native state ensemble, the pH 6 conformation is more dynamic and less stable as evidenced by NMR relaxation parameters and temperature / urea based unfolding studies.

Further exploring the structure and dynamics properties of partially folded (PF) of T7L suggested that these PF states are heterogeneous in nature and exhibit differential thermal and chemical stability. Urea induced unfolding studies evidenced that pH 3 state is found to be chemically less stable, and the unfolding pathway of pH 5 and 4 states involve the formation of molten globule sort of structures at intermediate concentration of urea. In contrast to chemical stability, pH 3 state shows extreme thermal stability and exhibits complete reversibility even at 90 °C. Residue level insights using NMR studies on these low pH

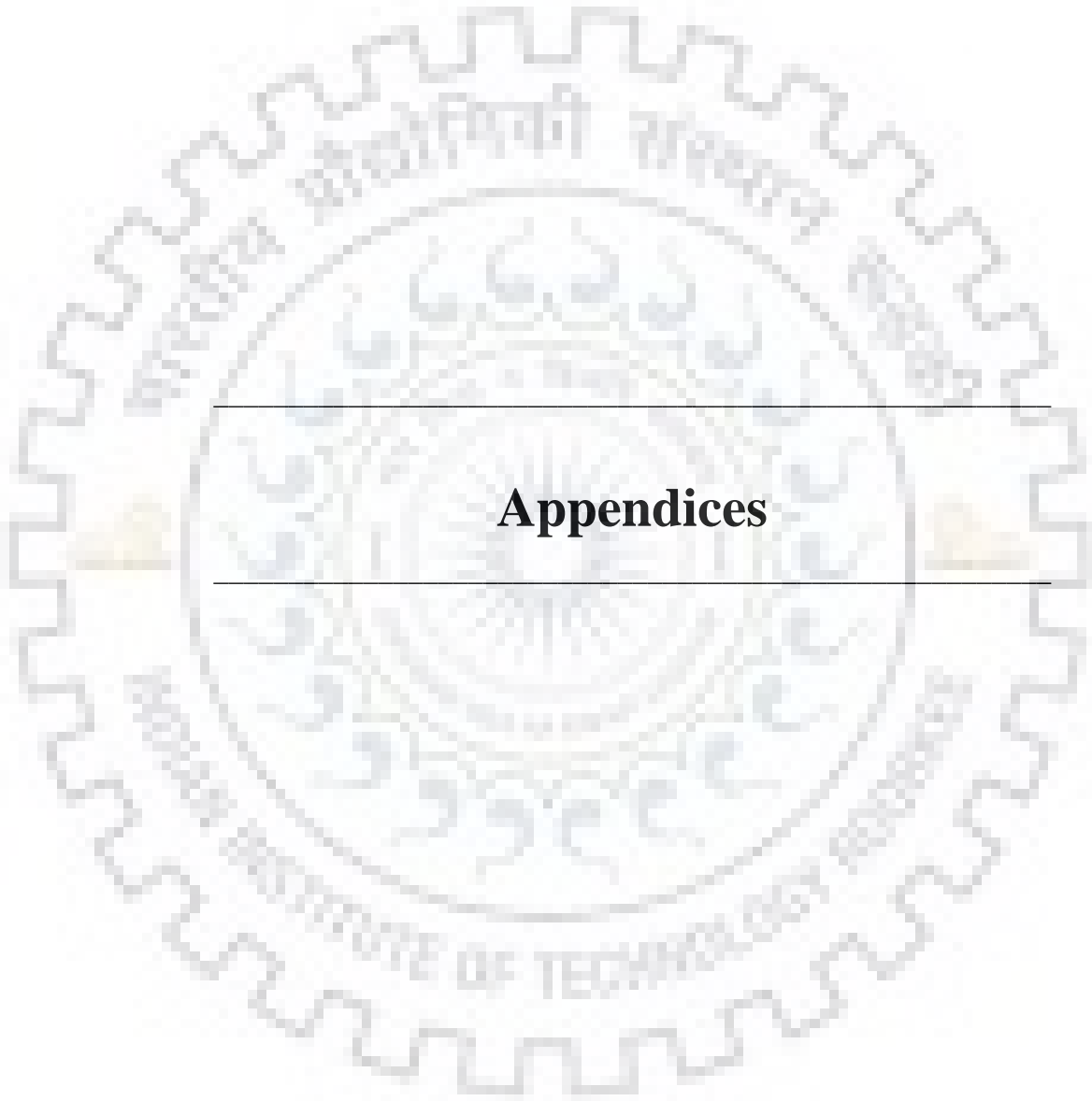
conformations of T7L revealed that  $\alpha$ 1-helix and  $\beta$ 3- $\beta$ 4 segments are highly dynamic and are responsible for such variable structural collapse.

The peculiar nature of various structural transitions upon varying the pH can be interpreted on the basis of protonation states of the ionisable side chains. The structural analysis of T7L suggested that a coupled network of buried His residues (H18, H48, H69 and H123) is the major source of observed pH induced structural transition. Surface accessibility and protonation characteristics of His residues established that, protonation of these buried His residues generated repulsive forces between the side chains and triggered the T7L structural collapse. Future mutational studies of all these His residues are essential to exactly underpin the atomic details of the observed transition.

## **5.2 Future perspectives**

Unravelling the structure-function relationships of single domain bacteriophage endolysins using T7L under different environmental conditions is crucial to apply these proteins as enzybiotics. As T7L is a metal containing protein, effect of various metal ions can be exploited to study structure-function relationship of endolysins. The studies performed here provided insights into the molecular basis of the stability and structure-function relationship of endolysins, which in turn can be used to design a robust enzyme with specific and enhanced biotechnological applications. As T7 endolysin belongs to the family of single-domain proteins, understanding the structural and function constraints of T7L will further aid in fusing it with CBD's (Cell binding domains) and/or CHAPS (Cysteine and histidine dependent amidohydrolase/peptidase) domains, and other functionally relevant motifs to produce next generation artificial lysins (artilysin) and chimeric lysins (chimeolysins) as novel enzybiotic formulations with improved stability and lytic activity against multiple drug resistance pathogens over a broad spectrum of hosts. Fusion of specific pathogenesis related proteins with T7L, and their expression in plant cells will helpful in providing resistance from bacterial infections in transgenic plant.

The study outcome can be applied on other endolysin families and other regulatory proteins/enzymes involved physiological and pathological functionalities to decipher various conformations and structural transitions across the energy landscape. Exploring the plausibility of multiple conformations in other similar proteins will be significant advance in terms of structural biology research, as decoding the structure-function relationship is crucial to engineering proteins with specific and enhanced functionalities.



---

## Appendices

---



## Appendix-I

$^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  (CA, CB & CO) chemical shift (CS) values (ppm) for the assigned residues in T7L sequence at pH 7.

S. No.	Residue	Residue No.	$^1\text{H}$ -CS	$^{15}\text{N}$ -CS	CA-CS	CB-CS	CO-CS
1	A	2	8.179	124.597	52.041	19.303	173.713
2	R	3	8.146	120.532	56.107	30.753	171.267
3	V	4	8.187	123.498	62.394	32.386	170.171
4	Q	5	8.137	125.031	54.838	30.071	169.981
5	F	6	8.230	121.150	57.409	42.028	171.281
6	K	7	8.714	122.910	56.707	32.926	170.759
7	Q	8	8.518	121.983	55.004	29.822	172.674
8	R	9	8.582	121.627	57.653	31.462	172.783
9	E	10	9.192	120.552	57.478	30.468	171.523
10	S	11	7.578	111.128	57.673	64.832	167.821
11	T	12	8.401	116.119	62.197	69.306	167.559
12	D	13	8.589	123.987	54.258	42.671	171.586
13	A	14	7.464	119.549	50.801	22.719	170.041
14	I	15	8.421	120.931	60.117	39.395	169.279
15	F	16	9.387	123.847	56.775		
16	A	21	8.732	124.429	53.470	16.561	169.801
17	T	22	8.312	119.261	59.187	70.308	169.040
18	K	23	7.057	106.115	55.294	32.294	
19	P	24	--	--	66.385	31.761	170.977
20	S	25	7.369	105.927	58.429	62.726	170.106
21	Q	26	7.754	121.724	55.037	29.288	
22	Q	35	8.779	121.925	59.726	29.613	174.024
23	W	36	7.964	118.557	60.703	28.069	172.348
24	H	37	7.646	117.664	64.893	37.521	172.435
25	K	38	8.939	118.647	60.777	37.973	172.761
26	E	39	6.845	117.764	59.018	28.880	173.414
27	Q	40	7.114	115.583	55.432	27.918	170.781
28	G	41	7.691	105.982	44.732	--	
29	F	49	7.616	114.227	57.320	44.952	168.713
30	I	50	9.489	125.258	59.637	40.979	169.127
31	I	51	8.752	124.135	60.441	37.571	172.087
32	K	52	8.923	127.737	58.070	33.514	173.306
33	R	53	9.472	119.003	59.302	28.685	172.043

*Appendices*

34	D	54	7.822	116.661	53.132	39.416	172.109
35	G	55	8.321	107.155	44.677	--	169.323
36	T	56	8.195	119.345	65.340	68.753	168.365
37	V	57	8.833	128.144	62.191	31.944	171.042
38	E	58	9.687	128.041	54.474	32.556	170.803
39	A	59	8.684	126.178	52.370	18.844	172.435
40	G	60	7.425	109.166	43.770	--	168.691
41	D	62	8.323	121.716	55.480	42.548	172.674
42	E	63	8.312	125.039	60.594	31.089	172.827
43	M	64	8.370	114.993	54.192	31.318	171.412
44	A	65	8.000	121.958	52.192	21.078	
45	K	71	8.410	126/555	59.065	32.166	172.631
46	G	72	8.825	116.007	44.921	--	169.584
47	Y	73	8.338	116.664	57.132	39.826	
48	G	79	8.476	109.073	44.302	--	167.690
49	V	80	9.237	127.456	60.787	33.955	168.300
50	C	81	8.844	126.438	56.750	27.821	168.648
51	L	82	8.898	130.083	54.156	44.858	171.434
52	V	83	7.489	122.930	63.634	31.724	170.890
53	G	84	8.465	117.322	44.736	--	170.258
54	G	85	8.597	109.022	45.366	--	169.910
55	I	86	7.471	114.588	58.866	41.761	169.780
56	D	87	8.700	121.102	51.571	41.867	172.979
57	D	88	8.138	114.363	56.730	40.798	172.457
58	K	89	8.039	117.824	55.090	32.129	172.304
59	G	90	8.036	108.559	45.452	--	168.844
60	K	91	8.399	123.084	54.566	32.405	171.760
61	F	92	8.133	121.449	58.386	38.188	
62	F	96	6.260	120.652	57.029	39.304	171.281
63	T	97	9.650	116.905	60.270	68.897	
64	P	98	--	--	65.985	31.495	175.003
65	A	99	8.221	118.781	54.951	18.106	176.135
66	Q	100	7.700	117.319	60.204	30.191	172.892
67	M	101	7.895	115.605	56.010	28.874	175.069
68	Q	102	8.352	121.522	59.129	28.390	174.549
69	S	103	8.161	117.052	61.462	62.756	171.760
70	L	104	8.848	122.366	58.151	40.253	172.827
71	R	105	8.463	119.095	61.835	29.546	172.566
72	S	106	7.813	111.435	61.844	62.810	172.239

*Appendices*

73	L	107	8.305	122.899	57.786	41.907	173.719
74	L	108	8.845	117.880	58.426	41.169	173.915
75	V	109	8.571	118.825	67.267	31.667	174.829
76	T	110	7.634	116.165	66.710	68.712	172.500
77	L	111	8.741	121.692	58.070	42.922	173.980
78	L	112	8.954	119.220	57.557	41.169	173.741
79	A	113	7.112	117.801	53.392	17.830	173.915
80	K	114	7.309	118.096	58.111	33.143	171.869
81	Y	115	7.967	121.273	55.900	36.009	169.627
82	E	116	7.404	119.876	58.525	29.361	173.589
83	G	117	8.776	114.004	44.795	--	169.671
84	A	118	7.832	123.208	52.491	18.291	171.956
85	V	119	7.213	114.614	59.668	33.789	170.193
86	L	120	8.169	122.799	53.884	42.553	
87	A	122	7.980	123.179	52.430	18.345	171.869
88	H	123	7.266	115.440	59.848		
89	K	137	8.268	120.074	60.085	32.251	173.458
90	R	138	8.003	119.521	58.821	30.468	174.742
91	W	139	8.178	119.665	62.377	27.424	174.024
92	W	140	9.484	123.816	59.641	30.310	173.806
93	E	141	8.534	113.840	58.586	30.805	173.349
94	K	142	7.764	114.191	55.128	33.512	171.783
95	N	143	8.074	118.391	53.635	37.756	169.323
96	E	144	6.728	116.720	54.715	32.479	170.215
97	L	145	8.740	126.907	54.568	42.876	170.171
98	V	146	9.308	125.606	60.389	35.381	170.302
99	T	147	8.398	117.993	62.208	69.491	169.714
100	S	148	8.015	118.646	58.472	63.861	169.322
101	D	149	8.372	122.940	54.240	40.966	171.586
102	R	150	8.277	121.764	56.332	30.191	172.130
103	G	151	8.371	108.916	45.327	--	

## Appendix-II

<sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>C (CA, CB & CO) chemical shift values (ppm) of T7L at pH 3.

S. No.	Residue	Residue No.	<sup>1</sup> H-CS	<sup>15</sup> N-CS	CA-CS	CB-CS	CO-CS
1	R	3	8.558	121.01	53.5	27.831	173.392
2	V	4	8.232	122.616	59.621	29.704	173.025
3	Q	5	8.372	124.491	52.927	26.854	172.637
4	F	6	8.229	122.234	55.107	36.914	172.777
5	K	7	8.17	123.351	53.494	30.279	173.313
6	Q	8	8.274	121.852	53.341	26.485	173.204
7	R	9	8.349	122.71	53.646	27.807	
8	S	11	8.311	115.749	55.827	61.379	172.181
9	T	12	8.126	115.314	59.556	67.013	171.714
10	D	13	8.289	121.018	50.681	35.764	
11	C	19	8.354	121.486	53.058	26.025	173.432
12	S	20	8.345	117.018	55.849	61.092	171.376
13	A	21	8.306	126.069	49.896	16.367	175.041
14	T	22	8.038	113.623	59.272	67.415	171.555
15	K	23	8.24	125.06	51.575	29.646	
16	N	27	8.367	119.794	50.506	35.912	172.419
17	V	28	8.012	119.989	60.079	29.556	174.088
18	G	29	8.389	112.211	42.52		171.217
19	V	30	8.187	117.369	52.469	28.68	171.446
20	R	31	8.573	119.606	52.621	26.33	170.909
21	E	32	8.399	124.77	53.908	25.530	
22	A	65	8.054	124.291	50.07	16.309	174.952
23	V	66	8.256	120.433	53.952	29.934	174.413
24	G	67	8.29	109.67	42.504		171.694
25	S	68	8.032	115.274	56.023	61.289	171.982
26	H	69	8.438	119.692	52.512	25.704	171.396
27	A	70	8.078	124.14	49.983	16.194	174.614
28	K	71	7.787	118.686	58.64	35.682	
29	N	143	8.1	123.451	54.715	36.627	172.727
30	E	144	7.92	121.868	59.599	29.876	173.184
31	L	145	7.965	122.14	52.774	39.362	174.684
32	V	146	7.971	120.737	59.861	29.589	173.829
33	T	147	8.065	116.815	59.425	67.243	171.982
34	S	148	8.147	117.462	55.849	61.309	171.644
35	D	149	8.367	121.4	50.746	36.109	172.955
36	R	150	8.187	121.213	53.973	27.289	174.195
37	G	151	8.297	108.878	42.504		