Perturbation of Amyloidogenic Conditions of Hen Lysozyme, Inhibition thereof by Naturally Occurring Osmolytes and Numeric Code of Amino Acids

A Thesis Submitted towards Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

By

Nividh Chandra

Under the Supervision of

Prof. R. Swaminathan

Department of Biosciences & Bioengineering

Indian Institute of Technology Guwahati

ASSAM, INDIA

June 2018

DEPARTMENT OF BIOSCIENCES & BIOENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DECLARATION

I do hereby declare that the matter embodied in this thesis entitled, "Perturbation of Amyloidogenic Conditions of Hen Lysozyme, Inhibition thereof by Naturally Occurring Osmolytes and Numeric Code of Amino Acids" is the result of investigations carried out by me in the Department of Biosciences & Bioengineering, Indian Institute of Technology Guwahati, India under the supervision of Prof Rajaram Swaminathan. In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on findings of other investigators. I also declare that this work has not been submitted elsewhere for award of any degree at any other academic organization.

Nividh Chandra IIT Guwahati

Roll Number: 07610605

DEPARTMENT OF BIOSCIENCES & BIOENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI, INDIA

CERTIFICATE

It is certified that the work described in this thesis, entitled, "Perturbation of Amyloidogenic Conditions of Hen Lysozyme, Inhibition thereof by Naturally Occurring Osmolytes and Numeric Code of Amino Acids" done by Mr. Nividh Chandra for the award of degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences & Bioengineering, Indian Institute of Technology Guwahati, India.

Prof. Rajaram Swaminathan
Department of Biosciences & Bioengineering
Indian Institute of Technology Guwahati

Dedicated to

My Sister Pallavi who was my engine for motivation, encouragement and meditation

My Mentor & Best Friend Dr Bhaskar Sarma who supported by continuously showering life changing advices

My Wife, Dr Gunjan Saxena who sacrificed her fledgling career (hopefully not irreversibly) to be with me and support in all possible ways in all my endeavours

ACKNOWLEDGMENT

I thank my supervisor Prof. R. Swaminathan for helping me by giving me a chance to work in his Lab, arranging for logistics, and the help rendered at the time of Thesis writing.

I thank my Doctoral Committee Members Dr A. Ramesh, Dr Vikash Kumar Dubey and Dr Aditya Panda. I am especially thankful to Dr Ramesh for nearly being my Co-supervisor by going way out of his mandate.

I thank my lab seniors: Dr Lopamudra Homchaudhari for discovering alkaline pH condition of amyloidogenesis which I followed on and Dr Vijay Kumar Ravi who demonstrated exceptional experimental skills.

In thank Lakshminath Bezbaroa Central Library for vast collection of books and online journals. I thank my friend Ajay Ravella who sent PDF's of hundreds of papers often in a matter of hours.

I thank staff at Bani Mandir for all my Printing/Xeroxing needs, supply of wide variety of stationery items and Books. They are a good example of how to run a MSME. I am also grateful to them for lending me a huge credit.

I thank Hostel Affairs Board of IIT Guwahati and especially Mr Narayan Kalita and Dr. K. V. Krishna for allowing me to stay in Married Scholar's Hostel much beyond permissible period. The efficiency with which Staff of IIT Guwahati worked was heart-warming. Off course the best amongst them was our Medical Section. Having worked in Health Sciences myself, it was amazing to see the professionalism with which they worked. I think a B.Tech./ M.Tech. or PhD degree in Biotechnology is incomplete until courses on Physiology, Pharmacology & Medical Lab Technology are taught. The very qualified team of doctors we have could be persuaded to fulfil this lacuna by teaching these courses.

I thank Deans of Academic Affairs for granting numerous extensions well beyond permissible duration.

I thank MHRD for giving me scholarship.

My PhD journey is unusually long and spanned through several institutes. During this journey I came across several mentors who honed my research skills: Prof P. Balaram (IISc), Prof Uday Maitra (IISc), Prof S. Chandrasekaran (IISc), Prof Iain D. Campbell (University of Oxford), Dr David Staunton (University of Oxford) and Dr Ram Vishvakarma (National Institute of Immunology).

I thank my Friends: Drs Himanshu Jha (especially helping me with getting me my new job), Bhunaeshwar Mandal (for all his concern and pointing out that numeric code can have application with Amino Acid Scales and not merely Memorization), Naresh Hanchaten, Anoop, Shampa Sen, Bishwanath Mohanty Sir, Shadab Ahmed, P. Sarvanan, Ganesh Thapa, Mousam Handique, Peter Symonds, Arvind Gupta, Indrajeet, Sirajul Faizi, Praveen Tiwari for keeping a stimulating and supportive atmosphere around me.

I have no words to thank God, my Parents, My Nani Ji and my Parents-in-law for patience and continuously reposing faith in this long journey of PhD and basically EVERYTHING.

I thank various food and coffee/tea providers inside and in the vicinity of IITG. In particular, I thank Krishna ji and Ganesh ji for serving delicacies.

Table of Contents

Chapter 3: Experimental Methodology used in this Thesis

Chapter 4: Results: Perturbations of Amyloidogenic Conditions of Lysozyme

Chapter 5: Results: Concomitant use of Osmolytes with Amyloidogenic conditions of Lysozyme

Four Important Abbreviations

These Abbreviations Are Important because they have been used Extensively in this Thesis (especially to document the outcome of Experiments in the Figures)

HEWL Hen Egg-White Lysozyme

REA Residual Enzymatic Activity (of Lysozyme)

r_{ss} Steady State Fluorescence Anisotropy of HEWL sample doped with Dansyl-labelled HEWL (excitation at 380 nm and Emission at 444 nm)

Th-T Thioflavin-T fluorescence intensity of HEWL sample at 482 nm when excited at 450 nm

Other Abbreviations Used in This Thesis

1.	α-LA	Alpha-lactalbumin
2.	AD	Alzheimer's disease
3.	Aβ40/42	β-amyloid 40/42
4.	ASA	Water Accessible Surface Area
5.	BA	Benzyl alcohol
6.	BSA	Bovine serum albumin
7.	CRABP	Cellular retinoic acid binding protein
8.	GB	Glycine betaine
9.	G-CSF	Granulocyte colony stimulating factor
10.	GdnHCl	Guanidinium hydrochloride
11.	Gly-Gly	Glycine-Glycine
12.	GPC	Glycerophosphocholine
13.	GST-GFP	Glutathione S-Transferase-Green Fluorescent Protein
14.	HD	Huntington's disease
15.	HTAB	Hexadecyltrimethylammonium bromide
16.	IAPP	Islet amyloid polypeptide
17.	IDP	Intrinsically disordered proteins
18.	LEA Proteins	Late Embryogenesis Abundant Proteins
19.	MJD	Machado-Joseph disease
20.	NAG	N-Acetyl Glucosamine
21.	NAM	N-Acetyl Muramic Acid
22.	OPMD	Oculopharyngeal muscular dystrophy
23.	PABPN	PolyA binding protein nuclear 1

24. rhGH Recombinant human growth hormone

25. rhGCSF Recombinant Human Granulocyte colony stimulating factor

26. rhIL-1RA Recombinant Human Interleukin-1 receptor antagonist

27. rhIFN-γ Recombinant human interferon- gamma

28. RNase A Ribonuclease A

29. SDS Sodium dodecyl sulphate

30. TMAO Trimethylamine oxide

Abstract of Thesis

This thesis starts by recognizing seminal discovery of amyloid made by point mutations in human Lysozyme and confidence exuded by its discoverers expressed in lines "as the structures of human and Hen egg-white Lysozymes are known to atomic resolution and their folding and structure-function relationships have been extensively analysed, our observations should provide a powerful model for understanding amyloidogenesis" [1]. However same doesn't seem to have been fully appreciated despite discovery of many *in vitro* conditions for making Lysozyme amyloid which are biophysically appealing and complementary.

First part of this thesis dwells around *in vitro* amyloidogenic conditions of Lysozyme. Barring one condition (acidic pH/elevated temperatures) which has become almost ubiquitously used despite assertion that the biophysical parameters themselves could be amyloidogenic rather than Lysozyme itself. All conditions deserve attention if we have to reach to the core of mechanistic insight. This state of affairs has been reviewed in Chapter-1. In this endeavour it was felt that systematically perturbing biophysical parameters which were classified as a) temperature, b) physicochemical force of either high or low pH or amount of denaturant Ethanol or Guandinium Hydrochloride, c) fluid mechanical parameter and d) seeding. Another crucial aspect of this thesis was the chosen methodology using simple instruments but each peeking into crucial events of this multistep process that too in form of 'kinetics' [a pillar of Physical Chemistry but something which seems to has been overlooked by workers of the field of Protein Aggregation]. These are; a) extent of unfolding by enzyme kinetics; b) extent of oligomerization from fluorescence steady state anisotropy kinetics and c) extent of amyloid by Thioflavin-T fluorescence kinetics). The same was carried out and forms Chapter-4 of this thesis, bringing out the rather critical onset and very fast growth thereafter under the alkaline condition. Surprises came out of ethanol and Guandinium Hydrochloride work because they were nearly abandoned.

Second part of thesis dwells around osmolytes which derive their name from their role to help cells counter osmotic pressure but which have been shown to be useful for helping proteins retain their structures and functions in extreme conditions. Surprisingly importance of osmolytes for proteins were pointed out as late as 1982 [2]. It is felt that our understanding of the roles which osmolytes play to help proteins not only in extreme

conditions but our normal cellular physiology is highly underestimated and understudied, but an impression has been made otherwise by putting theories like 'osmophoic effect' and 'preferential hydration'. Apart from relevance to osmolytes, this has close connection with Hofmeister series for which there is no satisfactory theory yet. This has been reviewed in Chapter-2. Osmolytes have been sparingly used in amyloid research, that too on different proteins making cross-comparison difficult. It was reasoned that studying many different osmolytes for a single well studied protein like Lysozyme will help us test theories like 'osmophoic effect' and 'preferential hydration'; compare their efficacy and offer mechanistic insights both for protein amyloidogenesis and the mode of osmolyte action. This was experimentally carried out in Chapter-5 using three methods previously mentioned [and in 'kinetics' format] using wide variety of naturally occurring osmolytes and their analogues, viz., Arginine, Argininamide, Guanidino Propionoic Acid, Ornithine, Polyamines (Putrescine, Spermidine, Spermine & Cadaverine), functional group variants and structural homologues of Putrescine, Trehalose, Taurine, TMAO (the most well studied osmolyte), Betaine, Ectoine & Hydroxy-Ectoine. Another feature of this study was the use of multiple concentrations of osmolyte (a requirement if kinetics and mechanism is aimed at). Not surprisingly all these osmolytes behaved in different ways at different steps which was in stark contrast to the 'osmophoic' effect' and 'preferential hydration'. These osmolytes were also attempted at acidic pH/elevated temperatures but could not be studied further due to problems of gelation and precipitation (except Trehalose and Taurine). Results with Trehalose were similar both at acidic as well as alkaline pH. All osmolytes led to inhibition of amyloidogenesis. Finally, an experiment was done wherein two, three, or four osmolytes were used together. It emerged that their effect was more than additive apart from 100% abolition of amyloid with Trehalose, if other osmolytes are present. Considering that cytoplasm consists of thousands of osmolytes (and same is likely to be the case for extracellular fluid), this is a promising result because even small-small concentrations of individual osmolytes when taken together can prevent amyloid formation. Such studies are blossoming in the field of Metabolomics. Since these molecules are all natural, it is hoped that a tonic can be developed containing most potent osmolytes which can be given to diagnosed or susceptible patients.

In Chapter-6 of the thesis, a new numeric code of amino acids is proposed for visual bioinformatics analysis. In it amino acids are represented as numbers (ranging from 0 to 9 and their underlined counterparts 0 to 9. This scheme has advantage, that it can allow

bioinformatics analysis to be done without computers. Since brain is by far the best computer in doing abstract thinking so it is expected that we can find patterns which a computer can't [until we incorporate it in computer's logic]. Thus such analysis can be forerunner to bioinformatics programming. Its use lies whenever an amino acid scale is used. One example of such scale is hydropathy plot. Whenever such plots are made, they mask identity of amino acids. This numeric code has been applied for transmembrane tendency of amino acids in protein Rhodopsin. Another advantage of numeric code is that it helps with facile memorization and recall of primary and secondary structure of a small protein.

References:

1 Pepys, M. B. *et al.* Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature* 362, 553-557, doi:10.1038/362553a0 (**1993**).

2 Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. & Somero, G. N. Living With Water-Stress - Evolution Of Osmolyte Systems. *Science* 217, 1214-1222, doi:10.1126/science.7112124 (**1982**).

Chapter-1

Introduction to Lysozyme Amyloidogenesis

1.1 PROTEIN AGGREGATION

In a broad sense, protein aggregation refers to the process by which protein molecules tend to adopt non-native conformations in unfavourable environmental conditions and polymerize to form stable complexes of different sizes called protein aggregates. These aggregates are highly ordered polymeric structures formed by alternative or off-pathway folding steps which may involve self-assembly of partially unfolded intermediates accumulated due to their relative high kinetic stability. The rate and extent of aggregation is influenced by innumerable factors which are categorized as extrinsic and intrinsic factors. The intrinsic factors include different structural levels of the protein: primary, secondary, tertiary and quaternary. Changes in protein structure while unfolding or refolding of the protein during stress conditions lead to conformational changes which might induce aggregation. The extrinsic factors include the immediate environment in which the proteins are present i.e. pH of the solution, ionic strength, concentration of co-solutes, etc. and the processing conditions.

1.11 Mechanisms of protein aggregation: For any given protein, there can be multiple aggregation pathways, and each pathway can involve multiple intermediate steps. Therefore, it is essential to know the mechanism of aggregation and nature of aggregates formed before designing any strategy to combat the process of protein aggregation. Different mechanisms involved in protein aggregation are briefly explained here as three major pathways: aggregation through unfolding protein intermediates, aggregation through protein self-association or chemical linkages, and aggregation through chemical degradations [1].

1.111 Aggregation through unfolding protein intermediate:

For a protein in solution, the three forms of the protein: native form, unfolding intermediates and completely unfolded form are in equilibrium with each other under normal environmental conditions. In stress conditions like heat or shear, the partially unfolded intermediates serve as the precursors for the process of aggregation. These intermediates have properties and characteristics which are significantly different from the native state of protein as the exposed hydrophobic regions in these molecules are more than the other two forms of the protein [2], whereas in case of completely folded or unfolded species, the hydrophobic side chains are either mostly buried inside the protein structures, or randomly scattered, such that they are not in contact with water molecules. Thus, the hydrophobic interactions in the partially unfolded intermediates and solvent molecules lead to protein aggregation. The aggregates formed may have disordered or ordered structures, such as amorphous precipitates /inclusion bodies) or amyloid fibrils, respectively. Also the aggregates formed do not dissociate easily, as the aggregation is effectively irreversible.

1.112 Aggregation through self-association or chemical linkages (Physical aggregation):

Protein molecules may also associate physically with each other under the influence of various non-covalent interactions such as electrostatic, hydrophobic or both and van der Waals forces. The aggregates are formed directly from the native state, so the intermediates are not involved in the process of aggregation. Reversible association of the monomers is the rate-limiting step where protein-protein interactions are favored over protein-solvent interactions for the minimization of charge. This behavior of the molecules results in the

colloidal instability of proteins, in contrast to the conformational stability in the aggregation pathway through unfolded intermediates.

Cross linking of the protein chains may also lead to protein aggregation which involves intermolecular disulfide bond formation/exchange, formaldehyde-mediated cross-linking, di-tyrosine formation, oxidation, and Maillard type reactions. Disulfide bond formation or exchange between cysteines located on the surface is the most frequent chemical linkage leading to protein aggregation.

Native state proteins may also form aggregates by isoelectric precipitation when the net charge on the protein molecule is zero [2]. These aggregates dissolve readily in the buffer solution when retained to their normal physiological conditions and thus, are different from the pathological aggregates which dissociate only in presence of strong destabilizers or denaturants.

1.113 Chemical degradation-induced protein aggregation:

Chemical degradation process like oxidation/auto-oxidation, dimerization, de-amidation, hydrolysis and glycation have been observed to induce aggregation of many proteins. It often results in the change in properties of the protein such as hydrophobicity, secondary/tertiary structures, association tendency or thermodynamics of unfolding.

Proteins may also form aggregates in their denatured form by undergoing chemical degradation either directly or indirectly. Protein misfolding may also induce aggregation of the refolded molecules[3].

The kinetics of aggregation is described by several models, of which the widely accepted is the nucleated growth model. According to this model the process of aggregation involves three phases: nucleation, polymerization and saturation (Invernizzi et al., 2012). In the

nucleation step, the soluble monomers of protein associate to form nuclei. This thermodynamically disfavored step is the lag phase of the process and influences the overall kinetics of aggregation. The second phase is the exponential phase which involves the polymerization of these nuclei to form larger aggregates. Finally, the exhaustion of the monomeric species leads to the saturation phase resulting in the formation of aggregates of different morphology.

Table 1.1: Examples of different proteins and their aggregation mechanisms [4]

Mechanism 1	REVERSIBLE ASSOCIATION OF THE NATIVE MONOMER
Mechanism 2	AGGREGATION OF CONFORMATIONALLY-ALTERED MONOMER
Mechanism 3	AGGREGATION OF CHEMICALLY-MODIFIED PRODUCT
Mechanism 4	NUCLEATION-CONTROLLED AGGREGATION
Mechanism 5	SURFACE INDUCED AGGREGATION

1.12 PROTEIN AGGREGATION: Mostly unwanted

It must be stated that protein aggregation can have useful consequences too (e.g. biofilms in bacteria) and so is not always undesirable [5,6]. Protein aggregation is a frequently encountered adversity in the field of recombinant protein production, drug development and their storage. During the process of manufacturing and formulation development of proteins, aggregates may be formed at any stage, for example, fermentation/expression, unfolding/refolding,

purification/concentration, freeze-thaw, shaking/shearing, pressurization, drying, analytical procedures, lyophilization / storage and other miscellaneous processes [1]. The change in environmental factors like temperature, pH, and solution conditions is the primary cause of aggregation during industrial production of commercially important proteins. Biological activity of the protein aggregates is highly reduced, making them undesirable for further application. Moreover, in spite of changes in the characteristics of the protein monomers, some native secondary and tertiary structures are intact in the aggregates. Such conformational changes make them immunogenic as compared to the native protein. Such aggregation-prone drug candidates have an adverse effect on protein activity and function. They can trigger severe immune response in some patients.

1.13 Diseases associated with protein aggregation – Aggregation of some physiologically important proteins such as βamyloid- peptides, insulin, lysozyme and many more causes fatal degenerative diseases grouped under the name amyloid disorders or protein conformational disorders (PCD) [7]. They are characterized by the deposition of aggregated fibrils of proteins either as intracellular inclusion bodies or extracellular plaques in different tissues or organs [8]. When the protein aggregates are deposited in brain or neurons causing the degeneration of the neural cells, it is known as neurodegenerative disorder [9]. Some of the aggregation involving disorders caused by protein aggregation [10] are discussed below.

I. Alzheimer's Disease: AD is a late onset dementing disorder caused by aggregation of two types of proteins, Aβ peptide and tau. Aβ peptides are derived from proteolytic degradation of amyloid precursor protein and aggregate to form extracellular neuritic plaques rich in β-sheets [11]. Tau is microtubule-associated protein which aggregates to form intracellular aggregates called neurofibrillary tangles [12].

II. Parkinson's Disease: PD is characterized by the presence of Lewy bodies in the cytoplasm of neurons or aggregates in neuritis, which are referred to as Lewy neuritis. The Lewy bodies are majorly composed of the aggregates of α-Synuclein. The cause of protein aggregation is either point mutations in the Synuclein gene (autosomal PD) or genes encoding parkin, DJ-1, PINK1 (recessive PD).

III. Huntington's Disease: HD is a progressive neurodegenerative disorder caused by polyglutamine (polyQ) expansion in the N-terminus of huntingtin protein due to expansion of CAG repeat in its gene. The mutated protein forms ordered aggregates that contain fibers and β-sheets similar to amyloid plaques in AD.

IV. Spino-Cerebral Ataxia: SCA is also caused by the polyglutamine expansion in the ataxin-3 protein which results into the aggregation of the protein.

V. Oculopharyngeal Muscular Dystrophy: OPMD is an autosomal, progressive disease caused by abnormal expansion of GCG repeat in the coding region of *papbn1* gene [13]. The protein with an expanded polyalanine tract forms aggregates consisting of tubular filaments within the nuclei of skeletal muscle fibers, which are toxic and pivotal to OPMD pathology.

VI. Prion diseases: Prion diseases are caused by point mutation in the genes encoding prion proteins. The altered proteins form aggregates and amyloids similar to that formed in AD. For example, Creutzfeld–Jakob Disease is caused by the aggregation of prion proteins.

More details on protein aggregation and its adverse effects are available elsewhere [5,14-18].

Table 1.2: Different diseases involving Protein aggregation

DISEASE	PROTEINS ASSOCIATED
Alzheimer's Disease	βamyloid-peptides, tau
Parkinson's	α- Synuclein
FrontoTemporal Dementia	Tau
OPMD	PAPBN1 (polyA expansion)
Huntington Disease	Huntingtin (polyQ expansion)
Spino-Cerebellar Ataxia/ MJD	Ataxin proteins (polyQ expansion)
Primary Systemic Amyloidosis	Ig light chain
Lysozyme Systemic Amyloidosis	Lysozyme
Insulin- Related Amyloid	Insulin
Diabetes Mellitus II	IAPP
Creutzfeld– Jakob Disease	Prion proteins

1.2 LYSOZYME AGGREGATION and AMYLOIDOGENESIS

1.21 Systemic amyloidosis in human lysozyme

However, the protein aggregation is not only limited to the neuropathies but it also causes some of the systemic disorders (Soto et al, 2001). Lysozyme amyloidosis is one of them and is a non-neuropathic, fatal disorder. Amyloidoses are a heterogeneous group of diseases characterized by deposition of protein aggregates in an abnormal fibrillar form. At a time when we understand little about brain and also because of no intervention in real time, the neuropathic amyloid diseases can hardly be employed as a model to study the amyloidosis.

On the other hand, the localized or systemic ones can be easily studied. In 1993, in a seminal discovery it was shown that I56T and D67H point mutants of human lysozyme led to amyloid deposition in huge quantities in viscera (systemic) which was fatal (Pepys et al, 1993). Since then, eight natural point mutations have been reported in wild type LYZ gene which are pathogenic and cause accumulation of large quantities of amyloid proteins in visceral organs such as liver, kidney, muscles, bone marrow, gall bladder, lymph vessels and gastrointestinal tract (Valleix et al, 2002; Merlini et al, 2005; Granel et al, 2006; Jean et al, 2014). The deposition of protein amyloids in non-native fibrillary conformations leads to disruption of tissues or potential organ dysfunction. The core of these fibrils structure consists of β-sheet with strands perpendicular to the long axis of the fiber (Booth et al, 1997). The amyloidosis involving lysozyme is an autosomal dominant hereditary non-neuropathic systemic disorder. The studies done by Canet et al, have shown that the partially folded intermediates of the lysozyme variants are more stable than the native protein itself (Canet et al, 1999). The patients with lysozyme amyloidosis are clinically heterogeneous and the disease is very penetrant. The symptoms related to the condition are different for the patients with different manifestations and generally involve gastrointestinal haemorrhage, hepatic rupture, sicca syndrome, purpura and petechiae, lymphadenopathy, or renal dysfunction (Sattianayagam et al, 2011). Since lysozyme amyloidosis shares common symptoms with the acquired monoclonal immunoglobin light chain amyloidosis, better understanding is required to diagnose the disease efficiently (Jean et al, 2014; Granel et al, 2006). In 2006, Granel et al, highlighted the need of better knowledge of the disease in order to provide efficient prevention and treatment for the cases of lysozyme amyloidosis (Granel et al, 2006).

Lysozyme, serendipitously discovered by Alexander Fleming in 1922 is a 130-residue protein found in secretions like saliva, sweat and mucus and more generally in kidneys, liver, salivary glands, lactating mammary tissues, lymph nodes, bone marrow and leukocytes (Mason and Taylor, 1975). As the name suggests, this enzyme has capacity to lyse the bacterial cells. It preferentially hydrolyses β-1,4 glycosidic linkages between N-acetylmuramic acid and N-acetylglucosamine in the mucopeptide of bacterial cell wall (Chipman and Sharon, 1969).

1.22 Human and Hen Lysozyme structure similarity

Apart from human tissues, chicken egg white is also rich in lysozyme. The crystal structure of human lysozyme and hen egg white lysozyme (abbreviated as HEWL) are shown in Figure 1.1. Since the 3D structures are actually more conserved than sequence, comparing 3D structures allows us to look even further back into the biological prehistory of the molecules. The most common method of comparing 3D structures is superimposing the crystal structures of the proteins obtained from PDB. HEWL, 129-residue protein, is homologous to human lysozyme with 61% identity and more than 76% similarity in sequence and structure respectively as compared by SuperPose web server (shown in Figure 1.2). It uses a combination of sequence alignment, difference distance matrix comparison and quaternion eigenvalue superposition. The difference distance matrix is also one of the visual representations of structural similarity between two protein molecules. The proteins are scored on the basis of increasing dissimilarity in their structures and these scores are represented by colors in the distance matrix. More the dissimilarity in structures, higher will be the score and darker will be the color. The difference distance matrix of HEWL and

human lysozyme is shown in Figure 1.3. The light colored regions clearly suggest that the structures of both the enzymes are not much different

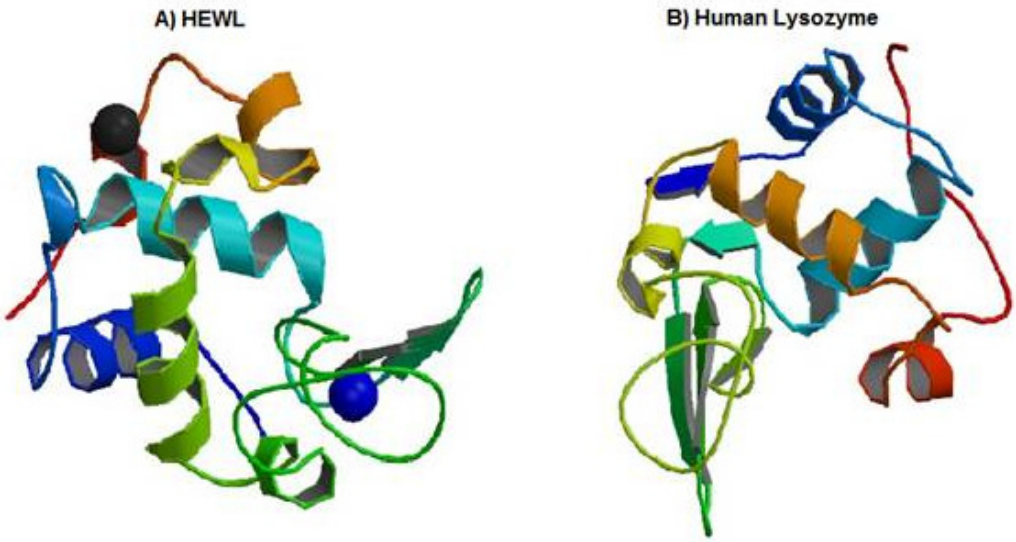


Figure 1.1: Crystal Structure of A) HEWL and B) human lysozyme (PDB ID-193L & 1LZ1 respectively).

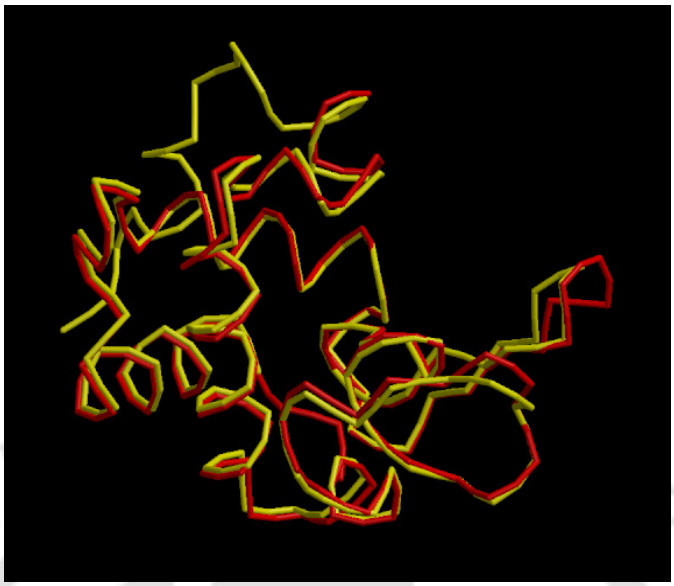


Figure 1.2: Superposed image for HEWL (yellow chain) and Human lysozyme (red chain) (image created using SuperPose version 1.0 and PDB files: 193L and 1LZ1).

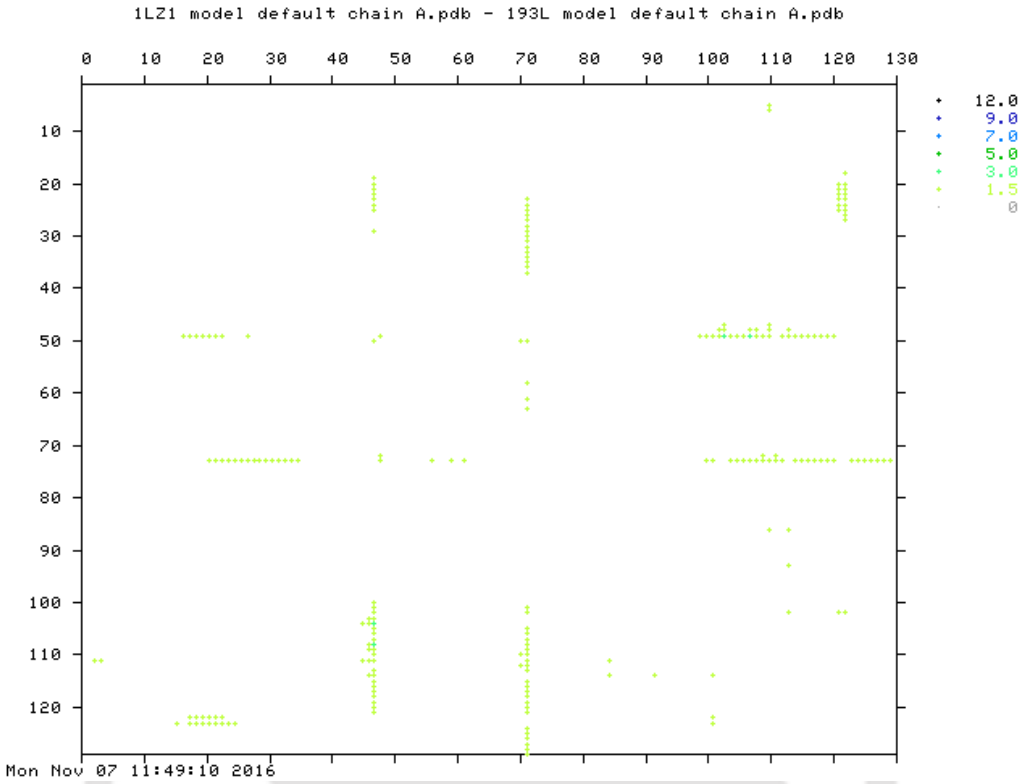


Figure 1.3: Difference Distance matrix of HEWL and human lysozyme (image created using SuperPose version 1.0 and PDB files: 193L and 1LZ1).

1.23 Hen lysozyme aggregation previous literature

HEWL is examined as a model system for protein folding/misfolding because it is relatively small, inexpensive and has excellent solubility in aqueous medium (Blake et al, 1977). After the discovery of amyloidosis causing mutant variants of human lysozyme by Pepys and his coworkers, in-vitro studies involving lysozyme gained a lot of attention. Since HEWL shares more than 60% structural and sequence similarity with human lysozyme as described above and its structure is known to atomic resolutions, it has also been used as a model for studies

involving lysozyme amyloidosis for so many years. It has been extensively studied in terms of structure, mechanism, and aggregation kinetics earlier under different which include: acidic pH at elevated temperature, concentrated solutions of ethanol, moderate concentrations of Guandinium chloride at moderate temperature, alkaline pH at room temperature, etc. The rich information available on lysozyme coupled with the multiple conditions that successfully induced its aggregation in vitro makes lysozyme an ideal model protein to investigate amyloidogenesis. Analysis of primary sequence and in vitro studies have resulted in the prediction of the regions in lysozyme sequences (both Human lysozyme and HEWL) which are hot-spots for amyloid formation. In-vitro studies carried out by Frare and his coworkers at pH 2.0 and 65 ℃ show that amyloid fibrils of HEWL mainly consists of fragments of proteins corresponding to following residues: 49-100, 49-101, 53-100 and 53-101 (Frare et al, 2004). On the other hand, Groot et al analyzed the mutant variants and wild type sequences of human lysozyme and predicted three regions in the wild type sequence: 20-34, 50-62 and 73-104 (Groot et al, 2005). The result of equivalent studies on HEWL predicted the three hot-spot regions as: 24-34, 50-62 and 76-98. It has been accounted that each of these hot-spots fall into hydrophobic cluster regions (Trexler and Nilsson, 2007). Krebs and his coworkers have reported the fibrillation of HEWL full length protein, its peptide fragment of 49-64 residues and a mutant peptide of same region at different experimental conditions like incubation at acidic pH (2.0 and 4.0), neutral pH (7.4) with subsequent addition of trifluoroethanol, elevated temperature (37°C and 65°C) and rapid heating and cooling followed by incubation at 37°C (Krebs et al, 2000). It was reported that the addition of small aliquots of pre-formed fibrils (seeding) accelerated the fibril formation. Vernaglia and his team has shown the effect of guanidine hydrochloride on fibrillation of

HEWL at pH 6.3 (Vernaglia et al, 2004) [19]. Fibrillation was induced at moderate concentrations of GdnHCl (2 – 5M) but at low and high concentrations it failed to produce HEWL fibrils. The explanation behind this disparative effect is that at low concentrations, native structure of the protein is stable, whereas at high concentrations of the denaturant, it dissolves and dis-aggregates the formed fibrils. In 2005, studies by Arnaudov and Vries suggested the involvement of some monomolecular nucleation step in aggregation process at acidic pH and high temperatures of 57°C and above and also stated that the slow change occurs in the protein conformation preceding the aggregation (Arnaudov and Vries, 2005). The fibrils thus formed mainly consisted of full-length protein. In 2006, Homchaudhuri et al looked upon the behavior of HEWL at alkaline pH and found out that the aggregates were formed spontaneously at pH 12.2 (Homchaudhuri et. al, 2006). Later, various studies involving aggregation of HEWL and human lysozyme at different conditions of pH, solvent, temperature etc. have been nicely reviewed in the following literature: Swaminathan et al, 2011; Trexler and Nilsson, 2007; Dumoulin et al, 2007. Studies done by S. Bhattacharya and coworkers suggested that the HEWL fibrils grown at different pH (7 and 11) have different structure but β-sheet content is very similar in them (S. Bhattacharya et al, 2013). They induced the fibril formation at both the pH in the presence of 80% ethanol (v/v) at 60˚C for 6h. According to their studies, the β-sheet content decreased by 10% in the presence of Cu (II) metal ions at both the pH. The aggregation mechanism of HEWL and the morphology of the aggregates at pH 12.2 was further explored by Ravi and coworkers. They observed that the HEWL aggregates follow isodesmic mechanism of aggregation at alkaline pH and lack lag phase (Ravi et al, 2014; Kumar and Udgaonkar, 2010).

References

1 Wang, W., Nema, S. & Teagarden, D. Protein aggregation-Pathways and influencing factors. *International Journal of Pharmaceutics* 390, 89-99, doi:10.1016/j.ijpharm.2010.02.025 (**2010**).

2 Fink, A. L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Folding and Design* 3, R9-R23, doi:10.1016/S1359-0278(98)00002-9 (**1998**).

3 Roberts, C. J. Therapeutic protein aggregation: Mechanisms, design, and control. *Trends in Biotechnology* 32, 372-380, doi:10.1016/j.tibtech.2014.05.005 (**2014**).

4 Philo, J. S. & Arakawa, T. Mechanisms of protein aggregation. *Current Pharmaceutical Biotechnology* 10, 348-351, doi:10.2174/138920109788488932 (**2009**).

5 Chiti, F. & Dobson, C. M. in *Annual Review of Biochemistry*, 75, 333-366 (**2006**).

6 Knowles, T. P. J. & Buehler, M. J. Nanomechanics of functional and pathological amyloid materials. *Nature Nanotechnology* 6, 469-479, doi:10.1038/nnano.2011.102 (**2011**).

7 Soto, C. Protein misfolding and disease; protein refolding and therapy. *FEBS Letters* 498, 204-207, doi:10.1016/S0014-5793(01)02486-3 (**2001**).

8 Stefani, M. Protein misfolding and aggregation: New examples in medicine and biology of the dark side of the protein world. *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1739, 5-25, doi:10.1016/j.bbadis.2004.08.004 (**2004**).

9 Soto, C. Unfolding the role of protein misfolding in neurodegenerative diseases. *Nature Reviews Neuroscience* 4, 49-60, doi:10.1038/nrn1007 (**2003**).

10 Ross, C. A. & Poirier, M. A. Protein aggregation and neurodegenerative disease. *Nature Medicine* 10, S10-S17, doi:10.1038/nm1066 (**2004**).

11 Liu, R., Barkhordarian, H., Emadi, S., Chan, B. P. & Sierks, M. R. Trehalose differentially inhibits aggregation and neurotoxicity of beta-amyloid 40 and 42. *Neurobiology of Disease* 20, 74-81, doi:10.1016/j.nbd.2005.02.003 (**2005**).

12 Wang, Y. & Mandelkow, E. Tau in physiology and pathology. *Nature Reviews Neuroscience* 17, 5-21, doi:10.1038/nrn.2015.1 (**2016**).

13 Davies, J. E., Sarkar, S. & Rubinsztein, D. C. Trehalose reduces aggregate formation and delays pathology in a transgenic mouse model of oculopharyngeal muscular dystrophy. *Human Molecular Genetics* 15, 23-31, doi:10.1093/hmg/ddi422 (**2006**).

14 Morris, A. M., Watzky, M. A. & Finke, R. G. Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature. *Biochimica et Biophysica Acta - Proteins and Proteomics* 1794, 375-397, doi:10.1016/j.bbapap.2008.10.016 (**2009**).

15 Aguzzi, A. & O'Connor, T. Protein aggregation diseases: Pathogenicity and therapeutic perspectives. *Nature Reviews Drug Discovery* 9, 237-248, doi:10.1038/nrd3050 (**2010**).

16 Invernizzi, G., Papaleo, E., Sabate, R. & Ventura, S. Protein aggregation: Mechanisms and functional consequences. *International Journal of Biochemistry and Cell Biology* 44, 1541-1554, doi:10.1016/j.biocel.2012.05.023 (**2012**).

17 Ross, C. A. & Poirier, M. A. What is the role of protein aggregation in neurodegeneration? *Nature Reviews Molecular Cell Biology* 6, 891-898, doi:10.1038/nrm1742 (**2005**).

18 Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nature Reviews Molecular Cell Biology* 15, 384-396, doi:10.1038/nrm3810 (**2014**).

19 Vernaglia, B. A., Huang, J. & Clark, E. D. Guanidine hydrochloride can induce amyloid fibril formation from hen egg-white lysozyme. *Biomacromolecules* 5, 1362-1370, doi:10.1021/bm0498979 (**2004**).

Chapter-2

Introduction to Interaction of Osmolytes with Proteins

Osmolytes are naturally occurring, low molecular weight organic molecules (see Table 1) or inorganic ions (e.g. K^+, Na^+) which are abundantly present in cells of different organisms. They play a vital role in maintaining the osmotic balance of the cell during unfavorable environmental conditions. The concentrations of these osmotically active solutes in a cell vary greatly depending upon the environmental water stress in the cell such as, fluctuating salinity, desiccation or freezing [1].

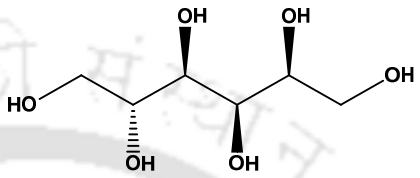
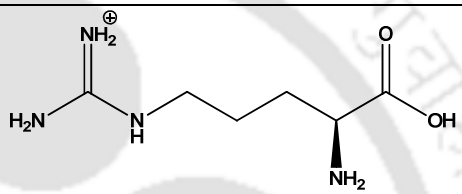
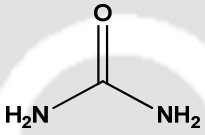
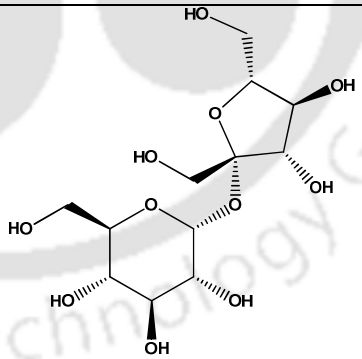
Many inorganic ions, known as Hofmeister ions perform different functions during stress which account for the specific interactions with the proteins affecting their stability [2]. Changes in the concentration of these ions such as K^+ and Na^+ disturbs the ionic strength of the physiological systems and thus, results in perturbations in the activity of macromolecules like enzymes and proteins. Therefore, cells cannot rely solely on inorganic ions to cope up with the osmotic stress. Organic osmolytes serve as an appropriate alternative instead of ions during water stress. These co-solutes do not interfere with the physiological processes and have favorable or non-perturbing effects on macromolecule-solvent interactions.

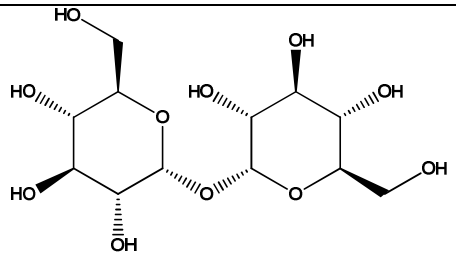
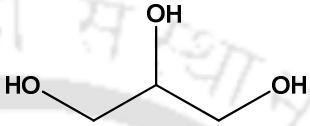
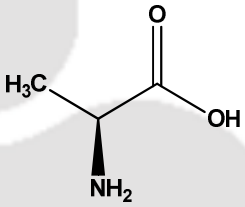
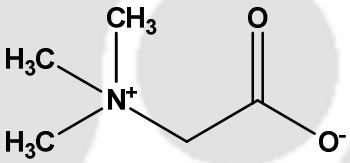
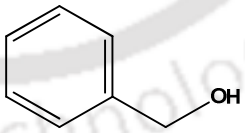
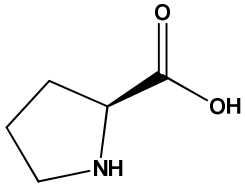
Mainly four classes of organic osmolytes are found in all organisms during osmotic stress:

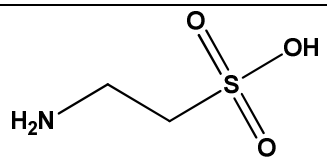
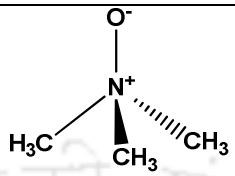
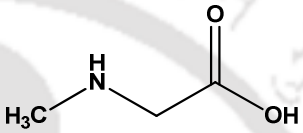
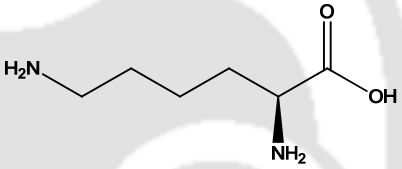
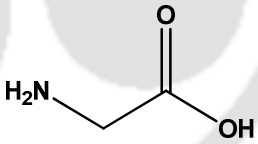
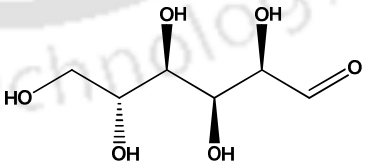
1. Polyhydric alcohols (polyols – glycerol, benzyl alcohol, sorbitol)
2. Sugars (sucrose, glucose, fructose)
3. Free amino acids and their derivatives (alanine, Proline, taurine)
4. Combinations of urea and methylamine compounds.

Structures of common osmolytes are given in Table 2.1

Table 2.1: Common osmolytes and their structures

#	Osmolyte	Structure
1.	Sorbitol	
2.	Arginine	
3.	Urea	
4.	Sucrose	
5.	Trehalose	

		
6.	Glycerol	
7.	Alanine	
8.	Glycine betaine	
9.	Benzyl alcohol	
10.	Proline	

11.	Taurine	
12.	Trimethylamine oxide(TMAO)	
13.	Sarcosine	
14.	Lysine	
15.	Glycine	
16.	Glucose	

Most of them are not obligatory for metabolism and are known to stabilize the protein structure such as polyols (glycerol/sorbitol), amino acids (glycine/ proline/ taurine) and methylamines (TMAO/betaine); thus, are termed as compatible solutes [3]. Some of the free amino acid

osmolytes like arginine, lysine etc. affect the activity of enzymes whose functions are critical for metabolism. Urea also has strong perturbing effects on macromolecules, especially on proteins and is counteracted by other osmolytes: methylamines-TMAO, betaine and sarcosine in ratio 2:1 of Urea: Methylamine [1]. The mechanism of stabilizing and destabilizing nature of these osmolytes depends on the protein backbone. Preferential exclusion of compatible osmolytes from the surface of the proteins is the basis of their protecting nature, as a result of which they shift the equilibrium towards native state of the molecule. On the other hand, the non-stabilizing osmolytes accumulate at the surface of the proteins and favor the denatured/unfolded state of the molecule. Osmolytes are generally classified in two categories of stabilizing and non-stabilizing as described in Table 2.2.

Table 2.2: Classification of osmolytes into stabilizing and non-stabilizing

Stabilizing or non-perturbing osmolytes	Polyols (Glycerol, sorbitol, inositol),
	Methylamine (TMAO, Sarcosine, Betaine)
	Sugars (Trehalose, Sucrose, Glucose, Fructose) Amino Acids(Proline, Taurine, Glycine, Alanine, β-alanine)
Non-stabilizing or perturbing osmolytes	Amino acids (Arginine)
	Urea

2.1 Physiological occurrence and Biological roles of osmolytes:

Occurrence of different osmolytes in different physiological systems of prokaryotes, plants and animals is well described in a review by Burg and Ferraris [4]. Many different osmolytes may occur even in the same organism or cell, and also can replace each other in the times of scarcity.

In Bacteria and Archaea (e.g. *E. coli, Salmonella typhimurium)*, the osmotic balance is initially regulated by uptake of large amounts of K^+ via turgor-responsive transport systems which is counteracted by glutamate synthesis. They also accumulate the organic osmolytes available in the surroundings, namely trehalose, proline, Ectoine, glycine betaine, and carnitine, out of which betaine and proline are preferentially accumulated by uptake. If the exogenous osmolytes are not available, trehalose or glycine betaine are synthesized and accumulated endogenously in order to maintain the osmotic balance in the cells.

Yeasts (*S. cerevisiae)* exclusively employ glycerol as an osmolyte, although, it also accumulates glycerophosphocholine (GPC) at levels much lower to that of glycerol. Other yeasts and fungi synthesize and/or accumulate different polyols such as erythritol, ribitol, arabinitol, xylitol, sorbitol, mannitol, and galactitol. Trehalose is also found especially in baker's and brewer's yeast [5].

Plants accumulate wide variety of organic osmolytes in cytoplasm of their cells, such as proline, valine, isoleucine, Ectoine, aspartic acid, betaine, glycine betaine, sucrose, glucose, fructans, mannitol, pinitol, and inositol.

In insects, trehalose is the basic sugar, found in high amount in their hemolymph.

Marine animals accumulate osmolytes depending upon their habitat in the depths of water. Mollusks and crustaceans exist at shallow depths and they accumulate different osmolytes such as betaine, taurine, TMAO, glycine, alanine, proline, homarine, and arginine. Deep sea

invertebrates have high levels of scyllo-inositol, β-alanine, betaine, hypotaurine, and N-methyltaurine. Oysters have been shown to contain osmolytes like glycine, proline, taurine, glutamate and alanine to protect them against fluctuating extracellular osmolality [6]. Hagfishes produce large volumes of slime to protect themselves when attacked by gill breathing fish predators [7]. Hagfish slime exudates have been shown to be rich in osmolytes like glucose, dimethyl glycine, TMAO, inositol apart from inorganic ions like Na^+, K^+, Mg^{2+} and chloride [8]. In elasmobranchs, urea and methylamines counteract each other in ratio 2:1 in their tissues and extracellular fluid to balance osmolality.

Mammals are subjected to high osmotic imbalance in impure blood, thus, renal medullary cells accumulate high amount of organic osmolytes such as sorbitol, betaine, inositol, taurine and GPC. These osmolytes accumulate in response to hypertonicity in the renal medullary cells. GPC displays exclusive behavior as it accumulates in response to both urea and hypertonicity and is known to counteract the perturbing effect of urea. Cells in other tissues also accumulate osmolytes, for example, brain contains amino acids, choline, creatine, inositol, and taurine; hepatocytes accumulate betaine, inositol, and taurine. Table 2.3 describes what osmolytes are there in different proteins.

Table 2.3: Physiological occurrence of osmolytes in different organisms

(Burg and Ferraris, 2008)

BACTERIA/ARCHAEA	Trehalose, Proline, Ectoine, Glycine Betaine, Carnitine, Glutamate, K^+
YEAST/FUNGI	Glycerol, Trehalose, GPC, Erythritol, Ribitol, Arabinitol, Xylitol, Sorbitol, Mannitol, and Galactitol
PLANTS	Proline, Valine, Isoleucine, Ectoine, Aspartic acid, Betaine, Glycine Betaine, Trehalose, Sucrose, Glucose, Fructans, Mannitol, Pinitol and Inositol.
INSECTS	Trehalose
MARINE ANIMALS	Betaine, Taurine, TMAO, Glycine, Alanine, Proline, Arginine, Scyllo-Inositol, β-alanine, Hypotaurine, Urea and N-methyltaurine
MAMMALS	Sorbitol, Betaine, Inositol, Taurine, Choline, Creatine and GPC

2.2 HOW OSMOLYTES INFLUENCE PROTEIN AGGREGATION?

Since protein aggregation is a prominent hurdle in various fields like recombinant protein production and pathophysiology, it is extremely important to tackle this problem. Osmolytes serve as excellent excipients for this purpose due to several advantages: natural macromolecules, comparatively smaller size and efficacy in obstructing protein aggregation. Some of the well-studied osmolytes, their effects on aggregation of some important model proteins and suggested mechanisms are discussed below.

2.2.1 Different osmolytes have different impact on the protein aggregation-:

Behavior of osmolytes against aggregation varies from protein to protein. Some osmolytes are found to induce protein aggregation and others inhibit the aggregation of same protein. Also same osmolyte may have distinctive effects on aggregation of different proteins depending upon the structural specificities of the proteins [9]. This conflicting nature of osmolytes makes it essential to study the effect of each osmolytes on different proteins separately. Effect of different osmolytes on different proteins and whether they induced or suppressed aggregation has been mentioned in Table 2.4.

Table 2.4: Different Osmolytes and their Effect on Proteins

	OSMOLYTE	AGGREGATION INDUCED	AGGREGATION SUPPRESSED
1.	Arginine		RNase A, HEWL, Carbonic Anhydrase II, IL-6, mAb
2.	Trehalose	α-Synuclein	Aβ-40/42, HEWL, RNase A, PAPBN1, mutant Huntingtin's protein, polyQ proteins
3.	Benzyl alcohol	rhIL-1RA, rhGCSF, rhGH, rhIFN-γ	HEWL(75°C)
4.	TMAO	tau	IAPP, mutated ataxin-3, HEWL
5.	Urea	HEWL	IAPP
6.	Betaine	GST-GFP	IAPP, Insulin
7.	Glycine Betaine	RNase A, α-LA	Tetra-cys CRABP
8.	Glycerol		Mutated ataxin-3, HEWL, α-globulin, Citrate Synthase
9.	Sorbitol		Nucleocapsid protein of rhabdovirus, Citrate Synthase
10.	Taurine	β-amyloids	Lysozyme, glucagon
11.	Proline	Huntingtin1, Glucagon	HEWL, CRABP, Insulin
12.	Sucrose		rhIFN- γ, rhIL-1RA, rhGCSF, HEWL

2.2.2 ARGININE: Arginine is one of the most extensively studied osmolyte against protein aggregation. It is a positively charged amino acid with a complex Guandinium group in its side chain which is a shared property with GdnHCl. As reported by Ghosh and co-workers [10], it acts as destabilizing agent like GdnHCl or urea as well as a stabilizing agent like betaine or sucrose. Arginine and GdnHCl show similar effect on the protein aggregation but exhibit different kinetic behavior.

A possible mechanism for Arginine as a suppressor of protein aggregation relative to GdnHCl has been discussed by Arakawa and co-workers [11]. Both Arginine and GdnHCl suppress protein aggregation but vary in their mechanisms of action. These additives interact with the proteins with different approaches:

i. *Interaction with amino acid side chains, thereby affecting their solubility*: Arginine interacts favorably with most of the amino acids side chains and the peptide bonds prefer to interact with GdnHCl.

ii. *Interaction with proteins (preferential interaction)*: The net preferential binding of Arginine with proteins is dissimilar from that of GdnHCl, for example preferential interaction of Arginine with lysozyme is negative.

iii. *Effect on water structure (surface tension):* Arginine is known to strongly increase the surface tension of water unlike GdnHCl.

Thus, at the amino acid level Arginine and GdnHCl are similar but when comes to proteins, they have different preferential interactions [11]. Arginine is not a conventional protein stabilizer as it increases the solubility of the protein instead of causing local conformational changes like other protein stabilizer, whereas, GdnHCl produces conformational changes in

the structure of the proteins. The anti-aggregation effect of arginine at different pH suggests that Guanidino group is responsible for binding to the protein as it retains charge at both neutral and alkaline pH [12].

The influence of Arginine on aggregation has been studied on different proteins. It proves to be a potent inhibitor of BSA aggregation as it contracts the native state of protein[10]. Also its interaction with the aromatic side chains and efficient binding to the partially unfolded intermediates destabilizes the protein. Arginine reduces the average molecular weight in carbonic anhydrase II aggregates, by promoting the native structure of the protein and, GdnHCl, on contrary is used to dissolve the aggregates and had little impact on size of aggregates [13]. It also suppresses aggregation of IL-6 and mAb at a wide range of concentrations by increasing the melting temperature of both the proteins [14]. Aggregation of RNase A and HEWL during thermal unfolding was also suppressed by different concentrations of Arginine at both neutral and alkaline pH [12,15]

2.2.3 TREHALOSE: Trehalose is a non-reducing disaccharide of glucose which is essential in many organisms like yeasts, plants and insects for its abilities to enhance desiccation tolerance and provide cold protection. It is a natural, non-toxic osmoprotectant which suppresses the aggregation of denatured proteins by two mechanisms, direct and indirect interactions which work side by side for stabilizing the proteins against aggregation [5].

i. *Direct interaction or water replacement theory*: Trehalose directly interacts with the protein molecules via hydrogen bonding with amino acid side chains or peptide backbone and reduces the protein-solvent hydrogen bonds, replacing

the water molecules at protein surface. Also it has strong unfavourable interactions with aliphatic carbon. Therefore, it stabilizes the protein conformations and its activity making the process of aggregation less favourable. For example, the addition of 1.5M trehalose induces the formation of alpha-helical and some tertiary structures in HEWL at 100°C at pH 6.0, which are more compact than the random coils of denatured lysozyme. Thus, it results in suppression of aggregation of the protein [16].

ii. *Indirect interaction*: Trehalose is known to increase the denaturation temperature of proteins by depressing the unfavourable intermolecular interactions, increasing their stability and inhibiting propensity for aggregation. For example, it increases the stability of RNase A by indirect interaction and inhibits its aggregation [5].

Trehalose suppresses the aggregation of many proteins involved in various diseases like polyglutamine diseases, Huntington's, Alzheimer's, OPMD and so on.

Trehalose is fairly effective in inhibiting the aggregation of Aβ40/42 peptides responsible for Alzheimer's disease in dose-dependent manner. It dissociates preformed Aβ40 aggregates by direct interaction and attenuates their toxicity [17]. But in case of Aβ42 trehalose is less efficient. The explanation behind differential effect of trehalose is that Aβ42 is more hydrophobic than Aβ40, consequently, has higher entropy gain during the association of molecules and aggregation process. Thus, trehalose is not able to compensate for the free energy change for Aβ42 to isolate the water molecules from its surroundings.

Trehalose reduces aggregation of mutant PABPN1, poly-alanine rich protein and promotes its clearance in both cell models and mouse models of OPMD, alleviating the disease symptoms [18].

Trehalose inhibits nucleation and aggregation of the polyQ containing β-sheet proteins by inhibiting intermolecular H-bonding and thus preferentially accumulating at backbone amide groups as a result of its favorable interactions with amide oxygen [19].

The aggregation of poly glutamine-containing proteins, truncated Huntingtin protein and MbGln35 is decreased upon addition of trehalose in a dose-dependent manner in mammalian cells. The mechanism suggested behind is the direct binding of trehalose with expanded glutamine stretches [20].

Trehalose is known to stabilize the proteins in their native state. But in case of α-Synuclein, an IDP, it favors rapid conversion of native state protein to semi-ordered structure which then interacts with other partially folded structures. Therefore, trehalose shifts the equilibrium of natively disordered protein towards a compact structure, inducing aggregation of the α-Synuclein [21].

Effectiveness of trehalose against P39A tetra-cys CRABP aggregation is marginal as compared to other osmolytes studied [22].

2.2.3.1 LEA PROTEINS and TREHALOSE [23]: Late embryogenesis abundant (LEA) protein initially identified in plants are now documented in animals also. These are extremely hydrophilic proteins that are intrinsically unstructured but acquire their native conformation during drying and provide desiccation tolerance. They are abundantly found in different organisms ranging from bacteria and fungi to nematodes and rotifers etc.

Different LEA proteins have different effects on protein aggregation, e.g. Ar-LEA1A from Adineta ricciae inhibits desiccation induced aggregation of CS, whereas Ar-LEA1B increases its aggregation. In many animals, LEA proteins carry out their function of desiccation tolerance synergistically with small solutes like **Trehalose** and protect the macromolecules and biological structures. Evidences show the effect of trehalose and LEA proteins in protection of mitochondria during freezing.

2.2.4 BENZYL ALCOHOL: Benzyl alcohol, being the least toxic preservative, can serve as a potent osmolytes against protein aggregation. It has been shown that it stabilizes hen egg white lysozyme against thermal denaturation and aggregation at 75°C. BA suppresses the aggregation of HEWL by preferentially binding to the partially unfolded molecules of protein in the nucleation step during thermal stress [24]. As a result, it slows down the rate of conversion of partially unfolded form to fully unfolded form of peptide.

Studies involving different proteins have reported that benzyl alcohol induces aggregation of different proteins like rhGH [25], rhIFN-γ [26].

Results by Zhang et al support that BA favours the formation of aggregation-prone intermediates of protein rhIL-1RA (β-sheet protein), accelerating the process of protein aggregation. It binds to the hydrophobic surface of partially unfolded intermediates and favors aggregation-prone protein species in the molecular population thus inducing the aggregation of the rhIL-1RA [27]. BA also accelerated aggregation and precipitation of rhGCSF (α-helical protein) at pH 7.0, converting the native α-helical structure into non-native intermolecular β-sheets which are highly aggregation-prone molecules [28]. It increased the population of

aggregation-prone intermediates which are basically partially unfolded molecules having perturbed tertiary structure and some native secondary structures. The aggregation-favouring effect of BA on protein of two different secondary structures suggests that it may affect most of the proteins in similar manner.

2.2.5 TMAO: TMAO is a widely present in marine animals as an osmotic pressure regulator in cell cytoplasm and is known to counteract the effect of urea in ratio 2:1 of urea and TMAO [4].

It is also one of the well-studied osmolyte which acts as a structure maker for proteins and stabilizes the proteins by the mechanism of preferential exclusion from the hydrophobic surface of the protein. These unfavorable interactions result from the osmophoic interaction of TMAO with the protein backbone and amino acid side chains [29]. Moreover, TMAO interacts strongly with water by making stronger H-bonds with water molecules than urea, which are also shorter and more ordered than those in pure water [30]. It affects the protein structure even at low concentrations via indirect mechanism i.e. ordering the solvent water molecules surrounding the protein.

TMAO induces structural changes in HEWL by the indirect interaction mechanism mentioned above [30]. At high concentrations (400 mM) of protonated form of TMAO, the amyloidogenesis of HEWL is suppressed to a great extent at 60^{o}C when incubated for 72h. The protein undergoes irreversible denaturation but well-defined oligomeric aggregates are formed. At 250 mM TMAO, the amyloids formed are much shorter and straight than the ones formed in absence of TMAO which are wavy in shape [31]. This suggests that the morphology of amyloid is highly dependent on the concentration of the osmolyte.

TMAO affects the aggregation kinetics of the protein IAPP but not the morphology of the fibrils formed. It decreases the growth rate of fibrils, while lag phase remains unchanged. When both TMAO and urea are studied for their influence on fibrillation, it was found that TMAO fully counteracts the effect of urea on the protein in ratio 1:2 [3]

TMAO suppresses the aggregation of the mutated ataxin- 3 protein in the concentration range 0mM to 150mM, but decreased the frequency of cell death at 150mM only in both the cell lines BHK-21 and neuro-2a [32].

The C-terminal fragment of tau is structurally reorganized in presence of TMAO. TMAO is found to accelerate both aggregation and fiber formation of tau and tau-dependent microtubule assembly in presence of heparin and tubulin. The filaments thus formed have similar morphology to that present in AD [33].

2.2.6 UREA: Urea is a well-known denaturing agent**.** In contrast to TMAO, it is water structure breaker and makes weak hydrogen bonds with water molecules which are similar to those formed in pure water. Although urea forms more H-bonds but they have decreased energy as compared to TMAO. It preferentially interacts with the polar side chains on the protein surface but not with the non-polar groups in the backbone, destroying the internal network of the H-bond in the protein due to repulsive forces between the side chains. As a result of these interactions, the water molecules are allowed to penetrate into the protein interior destabilizing the protein's native structure [30]. Another possible mechanism suggested by Canchi and Garcia explains the direct interactions of urea molecules with the peptide backbone. Urea being structurally similar to

peptide backbone competes with intra-backbone H-bonds and causes protein unfolding [29].

Urea destabilizes the structure of HEWL by preferential interaction with the protein surface at high concentrations [30].

It also prolongs the lag phase of IAPP aggregation process and shifts the equilibrium towards the unfolded state of the protein. Thus, it retards the nuclei formation of IAPP and its aggregation. On the other hand it accelerates the aggregation of insulin protein [3].

2.2.7 BETAINE: Betaine is another osmolyte known to stabilize the proteins during intracellular stress conditions. The possible mechanism behind the betaine-protein interaction was discussed by Natalello [34], and Choudhary [35] as mentioned below:

i. Osmophoic effect involving the unfavourable interaction of the osmolyte with the peptide backbone of protein which stabilizes the protein.
ii. Direct interaction of the osmolyte with the amino acid side chains on the protein surface which destabilizes the protein.
iii. Hydrophobic interactions also play an important role while stabilizing the protein by strengthening intramolecular and inhibiting intermolecular H-bonding.

Betaine may inhibit or induce protein aggregation depending upon its concentration [36]. At very low concentrations of the solute (1-3 mM), betaine had no effect on the structure of model protein: GST-GFP which exhibits aggregation during thermal stress [34]. At concentration 5-7.5 mM of betaine, it induces the formation of white coloured

aggregates of the protein which have amyloid-like features. At very high concentration of the betaine (10-20mM), it dissolves the large aggregates to smaller soluble aggregates which resist further aggregation during thermal stress and are also stable over longer incubation time. In case of GST-GFP, the direct interaction between betaine and the protein seems to overcome the osmophoic effect.

Betaine has been found to reduce the extent of insulin fibrillation but is not able to delay the onset of aggregation [37]. It also prevents Ig-light chain fibrillation by favouring the native state of protein [36].

When the effect of betaine on human IAPP was studied, it was found to affect the aggregation kinetics and hinders the growth of fibrils without changing the morphology of the fibrils already formed [3]. In the combination of betaine and urea, betaine counteracts the denaturing effect of urea to a lesser extent as compared to TMAO.

2.2.8 GLYCINE BETAINE: GB is the stress counteracting solute which acts as osmoprotectant in both prokaryotic and eukaryotic cells. It is a quaternary ammonium compound. It increases the thermal stability of the proteins and may also suppress the irreversible aggregation of the proteins to restore their function and activities.

The anti-aggregating properties of GB are dependent on extent of unfolding and the hydrophobicity of the proteins [38]. It exclusively shields the exposed hydrophobic residues on protein surface and also destabilizes the aggregation prone intermediates [22], thus influences the aggregation and refolding of the protein.

GB retards the aggregation of tetra-cys CRABP protein, by destabilizing the intermediate monomeric forms of the protein [22].

But in case of the model proteins: RNAse A and alpha-lactalbumin, GB induces aggregation of the proteins depending on their hydrophobicity in presence of HTAB and SDS respectively [38]. RNase A being less hydrophobic than α-LA, shows least aggregation at different pH, whereas massive aggregates of α-LA are observed which are directly proportional to betaine concentration.

2.2.9 PROLINE: Proline is an osmoprotectant which inhibits aggregation of proteins *in vivo* as well as *in vitro*. The mechanism of the stabilizing nature of proline can also be explained via preferential exclusion of the solvent and favorable/unfavorable interactions of the osmolyte with amino acid side chains.

In vivo studies conclude that proline disfavors the aggregation of slow folding mutant of CRABP protein by destabilization of the aggregation prone intermediates and solubilization of the small aggregated molecules when added prior to aggregation. In-vitro studies have also shown the similar results [39].It is found to stabilize the native state and destabilize aggregation-prone species. On the other hand; proline when added at later stages of aggregation doesn't affect the preformed large aggregates and thus, has diminished effect on aggregation [22].

Proline was also found to inhibit the aggregation of HEWL and induce the aggregation of mutant Huntingtin exon1 and fibrillation of glucagon [9].

Proline not only suppresses but also delays the process of fibril formation of insulin when added prior to the onset of fibrillation [37].

2.2.10 SUCROSE- Sucrose is another natural osmolyte which efficiently stabilizes proteins. The mechanism behind the anti-aggregation properties of sucrose is that it is preferentially excluded from the surface of the protein molecule exposed to the solvent due to hydrophobic effect, thus favours the native state of protein over the partially unfolded state, inhibiting the aggregation. Since sucrose effects on protein stability against aggregation are independent of properties of the protein, it can be employed to prevent preservative-induced protein aggregation.

Sucrose inhibits benzyl alcohol-induced aggregation of rhIL-1RA [27] and rhGCSF [28].

It is also observed that the addition of 1.5M sucrose induces the formation of alpha-helical and some tertiary structures in HEWL at 100°C at pH 6.0, which are more compact than the random coils of denatured lysozyme. Therefore, the addition of sucrose diminishes unfavourable intermolecular interactions and increases the solubility of unfolded lysozyme, reducing the aggregation and chemical reaction of the protein [16].

Sucrose shifts the equilibrium within the ensemble of rhIFN-γ native conformations to favor the most compact native species over the unfolded intermediates, thus stabilizing rhIFN-γ against aggregation [40].

2.2.11 GLYCEROL: Glycerol is one of the most widely used polyols against protein aggregation in formulation of biopharmaceutics and in food industry. It affects both native and non-native protein aggregation.

Glycerol's effect on the protein stability depends on the nature of protein as well as the physicochemical properties of the aqueous solution [41]. As the protein aggregation is mainly

influenced by the presence of surface exposed hydrophobic residues, thus, by molecular dynamic simulations it is hypothesized that the amphiphilic orientation of the glycerol molecules at the protein surface is favoured by the preferential solvation of the hydrophobic regions. It results in the increase in preferential interaction coefficient of aggregation-prone intermediates as compared to the native state, and exclusion of the water molecules from the protein surface [42]. As a consequence, glycerol molecules stabilize the partially unfolded intermediates and inhibit the protein aggregation. (Preferential interaction coefficient- Γ_{XP} is measure for the excess number of co-solvent molecules near the protein surface).

This mechanism explains the glycerol induced inhibition of HEWL aggregation by rearrangement of hydrophobic surface regions into the interior of the protein [42].

Glycerol also inhibits the aggregation of α-globulin in concentration dependent manner. It reduces the unfavourable environment due to hydrophobic residues in the protein and affects the hydrophobic interactions of globulin with water in a polynomial manner [43].

Glycerol was also found to be highly effective in suppressing the aggregation of thermally denatured citrate synthase enzyme and restoration of its activity. But at high concentrations, it may lead to off-pathway aggregation of the enzyme [44].

It also had negative impact on aggregation of mutated ataxin-3 protein in dose-dependent manner in BHK-21 and Neuro-2a cell lines [32]. Glycerol is also non-toxic to the cells, thus, it increased the cell viability.

2.2.12 SORBITOL: Sorbitol follows same mechanism of preferential hydration as other polyols like glycerol but it is a better stabilizer than glycerol [41]. The studies also show that the co-solute size and concentration also influences the

activity of the osmolytes against protein aggregation. For example, sorbitol has stronger effect on aggregation than Glycerol at same concentration [45].

It inhibits the aggregation of nucleocapsid protein of rhabdovirus when expressed in E. coli. Also it is shown to increase the refolding yield of citrate synthase in concentration dependent manner, by preferential exclusion from the protein surface. It is likely to interact with the denatured or intermediate state of the protein [44].On the other hand, addition of sorbitol is least effective against insulin fibrillation [37].

2.2.13 TAURINE: Taurine is a free amino acid abundantly found in mammalian cells. Under the crowding conditions in the cell, taurine's interactions with the lysozyme molecules are higher than beta-alanine in both folded and denatured states. Since aggregation occurs through the denatured form of protein, it is hypothesized that taurine might interfere with the formation of aggregates due to its favourable interaction with the denatured or unfolded state [46].

Taurine favors the aggregation of β-amyloids but it slows the fibrillation of glucagon [9]. The following Table (Table 2.5) describes in detail the effects of different osmolytes on different proteins with information such as concentration of osmolyte, putative mechanism and whether suppression or induction of protein aggregation.

Table 2.5: Summary of the Effect of Different Osmolytes on Proteins

Osmolyte	Concentrati-on	Protein studied	Interaction	Effect on Aggregation	Reference
Benzyl Alcohol	0.5%, 0.9%, 2% v/v	HEWL	Preferential binding to partially unfolded molecules	Suppresses	Goyal et al, 2009
	0.08M (0.9% w/v)	rhIL-1RA	Binds to partially unfolded intermediates	Favors	Zhang et al, 2004
	0.08M (0.9% w/v)	rhGCSF	Converts α-helices to β-sheets	Favors	Thiruman-galathu et al, 2006
	10mg/mL	rhGH	Binds to partially unfolded intermediates	Favors	Maa et al, 1996
	0.9% w/v	rhIFN-γ[26]	Binds to partially unfolded intermediates	Favors	Lam et al, 1997

Sucrose	0.5M	rhIL-1RA	Preferential exclusion from protein surface	Suppresses	Zhang et al, 2004
	1.0M	rhGCSF	Preferential exclusion from protein surface	Suppresses	Thiruman-galathu et al, 2006
	1.5M	HEWL	Increases the solubility of unfolded protein	Suppresses	Ueda et al, 2001
	300mM	Insulin	Preferential exclusion	Delays the nucleation process during fibrillation	Nayak et al, 2009
	1M	rhIFN-γ	Favors the formation of compact native state	Suppresses	Kendrick et al, 1998
Glycine Betaine	500mM	Tetra-cys CRABP	Destabilizes the	Suppresses	Ignatova et al, 2007

			intermediate monomeric form		
	4.0mol dm^{-3}	RNase A	Hydrophobic interactions	Induces	Misra et al, 2012
	4.0mol dm^{-3}	α-LA	Hydrophobic interactions	Induces	Misra et al 2012
Arginine	500mM	BSA	Contraction of native state	Inhibits	Ghosh et al, 2009
	0.5M	Carbonic anhydrase II	Slower the protein-protein association and accelerates dissociation	Reduces the size of aggregates	Baynes et al, 2005
	1.1M	IL-6	Increases the solubility of protein	Suppresses	Arakawa et al, 2006
	1.65M	mAb	Increases the solubility of protein	Suppresses	Arakawa et al, 2006

	1M	RNase A	Binding of gaunidino group with the protein	Suppresses	Arakawa et al 2003
	0.9M	HEWL	Binding of guanidino group with the protein	Suppresses	Arakawa et al, 2003; Homchaudh-uri et al, 2006
Trehalose	1.5M	HEWL	Hydrogen bonding	Suppresses	Ueda et al, 2001
	Not available	RNase A	Increases the denaturation temperature	Inhibits	Ohtake et al, 2011
	50mM	Aβ40/42	Hydrophobic interaction	Inhibits	Liu et al, 2005
	100mM	PABPN1	Direct interaction	Inhibits	Davies et al, 2006
	300mM	Insulin	Preferential exclusion	Delays the nucleation process during fibrillation	Nayak et al, 2009

	0—2 %	Huntingtin protein	Direct binding with polyglutamine stretches	Suppresses	Tanaka et al, 2004
	0.5M	α-synuclein	Interaction with partially unfolded intermediates	Induces	Naik et al, 2016
	500mM	Tetra-cys CRABP	Hydrophobic interaction	Marginal suppression	Ignatova et al, 2007
TMAO	1.0-4.0 mol dm^{-3}	HEWL	Orders the water molecules surrounding the protein	Suppresses amyloidogenesis	Panuszko et al, 2009, Wawer et al, 2014
	1-2M	IAPP	Preferential exclusion	Decreases the growth rate of fibrils formed	Seeliger et al, 2013
	150mM	Mutated ataxin-3	Preferential exclusion	Suppresses	Yoshida et al, 2002

	1M	Tau	Direct interaction	Induces	Scaramozz-ino et al, 2006
Urea	>1.5mol dm^{-3}	HEWL	Direct interaction	Destabilizes protein structure	Panuszko et al, 2009
	1-2M	IAPP	Preferential interaction	Suppresses	Seeliger et al, 2013
	1-2M	Insulin	Direct interaction with peptide bonds	Accelerates	Seeliger et al, 2013
Betaine	5-20mM	GST-GFP	Direct interaction with amino acid side chain	Dissolves large aggregates at high concentration(10-20mM) but induces aggregation at low	Natalello et al, 2009

				concentration(5-7.5mM)	
	50mM, 100mM, 250mM	Insulin		Reduces fibrillation	Choudhary et al, 2015
	Not available	Ig-light chain	Native state stabilization	Prevents fibrillation	Melo et al, 2010
	1-2M	IAPP		Reduces aggregation	Seeliger et al, 2013
Proline	Intracellular concentration > 0.4M	Tetra-cys CRABP	Destabilization of aggregation prone intermediates	Suppresses	Ignatova et al, 2006
	125-400mM	HEWL	Preferential exclusion	Inhibits	Macchi et al, 2012
	50-250mM	Insulin	Hydrophobic interactions	Suppresses fibrillation	Choudhary et al, 2015
	125-400mM	Glucagon		Induces	Macchi et al, 2012
Glycerol	30% v/v	HEWL	Hydrophobic interactions	Inhibits	Vagenende et al, 2009

	5-30% w/v	α-globulin	Hydrophobic interactions	Inhibits	Sahu et al, 2008
	1-9M	Citrate synthase	Interaction with partially unfolded intermediates	Suppresses	Mishra et al, 2005
	2% v/v	Mutated ataxin-3		Inhibits	Yoshida et al, 2002
Sorbitol	1-9M	Citrate synthase	Interaction with partially unfolded intermediates	Inhibits	Mishra et al, 2005
	250mM	Insulin	Preferential exclusion	Marginal effect against fibrillation	Choudhary et al, 2015
Taurine	0.2M	HEWL	Favorable interaction with denatured protein	Suppresses	Abe et al, 2015
	125-400mM	Glucagon		Slowers fibrillation	Macchi et al, 2012

2.3 CONCLUSION

Osmolytes are small molecules which accumulate intracellularly in response to stress conditions like heat shock or salinity. They exert stabilizing or destabilizing effect on proteins depending on the interactions between amino acid constituents of proteins and the osmolyte. There are different theories proposed for the mechanism followed by the osmolytes like preferential hydration theory and water replacement theory [47]. The universally accepted theory suggests that there is direct interaction between the osmolyte molecules and peptide backbone of the protein. According to this theory, free energy transfer of the peptide backbone takes place from pure water to the osmolyte solution which is negatively correlated with the fractional polarity of the osmolyte [48]. These direct interactions influence the stabilization of the protein structure. For stabilizers, their local concentration on the protein backbone is much lower than in the bulk solution. For destabilizers, concentration distribution is the other way.

2.4 FUTURE PERSPECTIVE and CHALLENGES

Osmolytes are used for formulation of innumerable biopharmaceutical or recombinant protein products. For example, trehalose is found in monoclonal antibody products**:** Herceptin, Avastin and Lucentis which are very important drugs developed by Genentech. A recombinant protein, Advate was developed by Baxter [5] similarly. Other osmolytes can also be used as co-solutes in the production of different proteins depending upon their specific interaction with the proteins and their effect on protein structure. Suppression of aggregation of many disease-related proteins by sugars makes them a desirable candidate for drug development. The stabilization effect of sugars on the aggregation process is in the order: tri-saccharides> disaccharides> monosaccharide depending upon the free energy change of the nucleation step

for the native protein [49]. But, this aspect is still to be worked upon because oral administration of sugars results into their degradation into their respective monomers and absorption in small intestine. Trehalose, being non-reducing sugar cannot be digested in mammals. Therefore, an osmolyte-based drug can be developed using trehalose or other non-toxic osmolytes.

Different mechanisms of the osmolyte-induced suppression or enhancement of aggregation are being proposed on the basis of experimental studies but the exact molecular basis of the phenomena of is still lacking. The lack of understanding is due the disparity in the functional behavior of the osmolytes with different proteins. Thereby, the most important question is whether the effect of individual osmolyte is protein specific or can it be generalized for all the proteins of same family or structure.

REFERENCES

1 Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. & Somero, G. N. Living with water stress: Evolution of osmolyte systems. *Science* 217, 1214-1222 (**1982**).

2 Baldwin, R. L. How Hofmeister ion interactions affect protein stability. *Biophysical Journal* 71, 2056-2063 (**1996**).

3 Seeliger, J., Estel, K., Erwin, N. & Winter, R. Cosolvent effects on the fibrillation reaction of human IAPP. *Physical Chemistry Chemical Physics* 15, 8902-8907, doi:10.1039/c3cp44412k (**2013**).

4 Burg, M. B. & Ferraris, J. D. Intracellular organic osmolytes: Function and regulation. *Journal of Biological Chemistry* 283, 7309-7313, doi:10.1074/jbc.R700042200 (**2008**).

5 Ohtake, S. & Wang, Y. J. Trehalose: Current use and future applications. *Journal of Pharmaceutical Sciences* 100, 2020-2053, doi:10.1002/jps.22458 (**2011**).

6 Hosoi, M., Kubota, S., Toyohara, M., Toyohara, H. & Hayashi, I. Effect of salinity change on free amino acid content in Pacific oyster. *Fisheries Science* 69, 395-400, doi:10.1046/j.1444-2906.2003.00634.x (**2003**).

7 Fudge, D. S., Schorno, S. & Ferraro, S. in *Annual Review of Biochemistry*, 84, 947-967 (**2015**).

8 Herr, J. E., Clifford, A. M., Goss, G. G. & Fudge, D. S. Defensive slime formation in Pacific hagfish requires Ca(2+) and aquaporin mediated swelling of released mucin vesicles. *Journal of Experimental Biology* 217, 2288-2296, doi:10.1242/jeb.101584 (**2014**).

9 Macchi, F., Eisenkolb, M., Kiefer, H. & Otzen, D. E. The effect of osmolytes on protein fibrillation. *International Journal of Molecular Sciences* 13, 3801-3819, doi:10.3390/ijms13033801 (**2012**).

10 Ghosh, R., Sharma, S. & Chattopadhyay, K. Effect of arginine on protein aggregation studied by fluorescence correlation spectroscopy and other biophysical methods. *Biochemistry* 48, 1135-1143, doi:10.1021/bi802065j (**2009**).

11 Arakawa, T. *et al.* Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects. *Biophysical Chemistry* 127, 1-8, doi:10.1016/j.bpc.2006.12.007 (**2007**).

12 Arakawa, T. & Tsumoto, K. The effects of arginine on refolding of aggregated proteins: Not facilitate refolding, but suppress aggregation. *Biochemical and Biophysical Research Communications* 304, 148-152, doi:10.1016/S0006-291X(03)00578-3 (**2003**).

13 Baynes, B. M., Wang, D. I. C. & Trout, B. L. Role of arginine in the stabilization of proteins against aggregation. *Biochemistry* 44, 4919-4925, doi:10.1021/bi047528r (**2005**).

14 Arakawa, T., Kita, Y., Ejima, D., Tsumoto, K. & Fukada, H. Aggregation suppression of proteins by arginine during thermal unfolding. *Protein and Peptide Letters* 13, 921-927, doi:10.2174/092986606778256171 (**2006**).

15 Homchaudhuri, L., Kumar, S. & Swaminathan, R. Slow aggregation of lysozyme in alkaline pH monitored in real time employing the fluorescence anisotropy of covalently labelled dansyl probe. *FEBS Letters* 580, 2097-2101, doi:10.1016/j.febslet.2006.03.012 (**2006**).

16 Ueda, T., Nagata, M. & Imoto, T. Aggregation and chemical reaction in hen lysozyme caused by heating at pH 6 are depressed by osmolytes, sucrose and trehalose. *Journal of Biochemistry* 130, 491-496 (**2001**).

17 Liu, R., Barkhordarian, H., Emadi, S., Chan, B. P. & Sierks, M. R. Trehalose differentially inhibits aggregation and neurotoxicity of beta-amyloid 40 and 42. *Neurobiology of Disease* 20, 74-81, doi:10.1016/j.nbd.2005.02.003 (**2005**).

18 Davies, J. E., Sarkar, S. & Rubinsztein, D. C. Trehalose reduces aggregate formation and delays pathology in a transgenic mouse model of oculopharyngeal muscular dystrophy. *Human Molecular Genetics* 15, 23-31, doi:10.1093/hmg/ddi422 (**2006**).

19 Hong, J., Gierasch, L. M. & Liu, Z. Its Preferential Interactions with Biopolymers Account for Diverse Observed Effects of Trehalose. *Biophysical Journal* 109, 144-153, doi:10.1016/j.bpj.2015.05.037 (**2015**).

20 Tanaka, M. *et al.* Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease. *Nature Medicine* 10, 148-154, doi:10.1038/nm985 (**2004**).

21 Naik, V., Kardani, J. & Roy, I. Trehalose-Induced Structural Transition Accelerates Aggregation of α-Synuclein. *Molecular Biotechnology*, 1-5, doi:10.1007/s12033-016-9923-4 (**2016**).

22 Ignatova, Z. & Gierasch, L. M. in *Methods in Enzymology*, 428 355-372 (**2007**).

23 Hand, S. C., Menze, M. A., Toner, M., Boswell, L. & Moore, D. in *Annual Review of Physiology*, 73 115-134 (**2011**).

24 Goyal, M. K., Roy, I., Banerjee, U. C., Sharma, V. K. & Bansal, A. K. Role of benzyl alcohol in the prevention of heat-induced aggregation and inactivation of hen egg white lysozyme. *European Journal of Pharmaceutics and Biopharmaceutics* 71, 367-376, doi:10.1016/j.ejpb.2008.09.012 (**2009**).

25 Maa, Y. F. & Hsu, C. C. Aggregation of recombinant human growth hormone induced by phenolic compounds. *International Journal of Pharmaceutics* 140, 155-168, doi:10.1016/0378-5173(96)04577-2 (**1996**).

26 Lam, X. M., Patapoff, T. W. & Nguyen, T. H. The effect of benzyl alcohol on recombinant human interferon. *Pharmaceutical Research* 14, 725-729, doi:10.1023/A:1012190120061 (**1997**).

27 Zhang, Y. *et al.* Mechanism for benzyl alcohol-induced aggregation of recombinant human interleukin-1 receptor antagonist in aqueous solution. *Journal of Pharmaceutical Sciences* 93, 3076-3089, doi:10.1002/jps.20219 (**2004**).

28 Thirumangalathu, R., Krishnan, S., Brems, D. N., Randolph, T. W. & Carpenter, J. F. Effects of pH, temperature, and sucrose on benzyl alcohol-induced aggregation of recombinant human granulocyte colony stimulating factor. *Journal of Pharmaceutical Sciences* 95, 1480-1497, doi:10.1002/jps.20619 (**2006**).

29 Canchi, D. R. & García, A. E. in *Annual Review of Physical Chemistry*, 64, 273-293 (**2013**).

30 Panuszko, A., Bruździak, P., Zielkiewicz, J., Wyrzykowski, D. & Stangret, J. Effects of urea and trimethylamine-N-oxide on the properties of water and the secondary structure of hen egg white lysozyme. *Journal of Physical Chemistry B* 113, 14797-14809, doi:10.1021/jp904001m (**2009**).

31 Wawer, J. *et al.* Inhibition of amyloid fibril formation of hen egg white lysozyme by trimethylamine N-oxide at low pH. *International Journal of Biological Macromolecules* 70, 214-221, doi:10.1016/j.ijbiomac.2014.06.057 (**2014**).

32 Yoshida, H., Yoshizawa, T., Shibasaki, F., Shoji, S. & Kanazawa, I. Chemical chaperones reduce aggregate formation and cell death caused by the truncated Machado-Joseph disease gene product with an expanded polyglutamine stretch. *Neurobiology of Disease* 10, 88-99, doi:10.1006/nbdi.2002.0502 (**2002**).

33 Scaramozzino, F. *et al.* TMAO promotes fibrillization and microtubule assembly activity in the C-terminal repeat region of tau. *Biochemistry* 45, 3684-3691, doi:10.1021/bi052167g (**2006**).

34 Natalello, A., Liu, J., Ami, D., Doglia, S. M. & De Marco, A. The osmolyte betaine promotes protein misfolding and disruption of protein aggregates. *Proteins: Structure, Function and Bioinformatics* 75, 509-517, doi:10.1002/prot.22266 (**2009**).

35 Choudhary, S. & Kishore, N. Addressing mechanism of fibrillization/aggregation and its prevention in presence of osmolytes: Spectroscopic and calorimetric approach. *PLoS ONE* 9, doi:10.1371/journal.pone.0104600 (**2014**).

36 Melo, E. P. *et al.* Compacting proteins: Pros and cons of osmolyte-induced folding. *Current Protein and Peptide Science* 11, 744-751, doi:10.2174/138920310794557727 (**2010**).

37 Choudhary, S., Kishore, N. & Hosur, R. V. Inhibition of insulin fibrillation by osmolytes: Mechanistic Insights. *Scientific Reports* 5, doi:10.1038/srep17599 (**2015**).

38 Misra, P. P. & Kishore, N. Glycine betaine: A widely reported osmolyte induces differential and selective conformational stability and enhances aggregation in some proteins in the presence of surfactants. *Biopolymers* 97, 933-949, doi:10.1002/bip.22110 (**2012**).

39 Ignatova, Z. & Gierasch, L. M. Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant. *Proceedings of the National Academy of Sciences of the United States of America* 103, 13357-13361, doi:10.1073/pnas.0603772103 (**2006**).

40 Kendrick, B. S., Carpenter, J. F., Cleland, J. L. & Randolph, T. W. A transient expansion of the native state precedes aggregation of recombinant human interferon-γ. *Proceedings of the National Academy of Sciences* 95, 14142-14146, doi:10.1073/pnas.95.24.14142 (**1998**).

41 Tiwari, A. & Bhat, R. Stabilization of yeast hexokinase A by polyol osmolytes: Correlation with the physicochemical properties of aqueous solutions. *Biophysical Chemistry* 124, 90-99, doi:10.1016/j.bpc.2006.06.003 (**2006**).

42 Vagenende, V., Yap, M. G. S. & Trout, B. L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry* 48, 11084-11096, doi:10.1021/bi900649t (**2009**).

43 Sahu, R. K. & Prakash, V. Mechanism of prevention of aggregation of proteins: A case study of aggregation of α-globulin in glycerol. *International Journal of Food Properties* 11, 613-623, doi:10.1080/10942910701580367 (**2008**).

44 Mishra, R., Seckler, R. & Bhat, R. Efficient refolding of aggregation-prone citrate synthase by polyol osmolytes: How well are protein folding and stability aspects coupled? *Journal of Biological Chemistry* 280, 15553-15560, doi:10.1074/jbc.M410947200 (**2005**).

45 Sukenik, S. & Harries, D. Insights into the disparate action of osmolytes and macromolecular crowders on amyloid formation. *Prion* 6, 26-31, doi:10.4161/pri.6.1.18132 (**2012**).

46 Abe, Y., Ohkuri, T., Yoshitomi, S., Murakami, S. & Ueda, T. Role of the osmolyte taurine on the folding of a model protein, hen egg white lysozyme, under a crowding condition. *Amino Acids* 47, 909-915, doi:10.1007/s00726-015-1918-0 (**2015**).

47 Arakawa, T. & Timasheff, S. N. The stabilization of proteins by osmolytes. *Biophysical Journal* 47, 411-414 (**1985**).

48 Patel, K. A., Sethi, R., Dhara, A. R. & Roy, I. Challenges with osmolytes as inhibitors of protein aggregation: Can nucleic acid aptamers provide an answer? *International Journal of Biological Macromolecules*, doi:10.1016/j.ijbiomac.2016.05.014 (**2016**).

49 Nayak, A., Lee, C. C., McRae, G. J. & Belfort, G. Osmolyte controlled fibrillation kinetics of insulin: New insight into fibrillation using the preferential exclusion principle. *Biotechnology Progress* 25, 1508-1514, doi:10.1002/btpr.255 (**2009**).

Thesis Objectives

One aspect of this study was to employ many different Osmolytes against Amyloidogenesis of a single protein. Indeed, different Osmolytes have been used as amyloid inhibitor but such studies suffer from two major drawbacks.

1. Different Osmolytes have been used against different proteins thus making a comprehensive comparison is difficult.
2. Mostly a single concentration osmolyte has been used. In most of the studies this concentration is in molar range (rarely less than 500mM). But if want to determine the 'Order' of the process, we need to use many concentrations.
3. All studies have only looked the effect of osmolyte either on final amyloid or on the extent of unfolding.
4. There are few studies where more than one osmolytes have been used in a single study (though studies of animal tissues shows presence of more than on osmolyte in a single tissue).

   Thus the purpose of this study was to overcome above limitations one by one.

So we chose HEWL as a representative protein against which different Osmolytes were pitted one by one. Many different osmolytes were used. However, to make results free of non-natural amyloidogenic conditions we decided to use many different *in vitro* conditions of lysozyme Amyloidogenesis. They are

a. Alkaline pH
b. Acidic pH at elevated temperature
c. Ethanol-Water mixture
d. In presence of Guanidine Hydrochloride at elevated temperature

Sheer biophysical elegance of these conditions led us to probe them in deep before introducing osmolyte(s) in them. This formed Chapter-4 of this thesis, viz., the mechanistic exploration of these conditions. Clue was taken from Physical Organic Chemistry. One by one, the parameters which define amyloidogenic conditions were systematically perturbed by small changes, and their effect on three key steps of unfolding, oligomerization and amyloid content was measured. These parameters are physicochemical force of pH or denaturant concentration, temperature force, effect of mechanical agitation.

Then in Chapter-5, Osmolytes were introduced. For each osmolyte 3-4 concentrations were used. Their effect was measured at three crucial steps of their effect on three key steps of unfolding, oligomerization and amyloid amount. Finally, an experiment was done wherein two, three of four osmolytes were used concomitantly. Though in this thesis we focussed on only a handful osmolytes but tissues have thousands of them. To study them all with many more techniques and with more protein and at more time points will yield so much data which will be difficult to assimilate when presented with pictures. While pondering over this issue, we came up with an idea of a Matrix Formalism.

It became obvious in the beginning of my experimental work with Lysozyme that I would be working exclusively on this protein during entire tenure of my PhD. As a curiosity it occurred to me that If I am going to spend so many years on a single protein which is not very long, why not memorize its primary sequence and secondary structure. The idea was that it could help in some way to analyse the results. This has been shown that Lysozyme at acidic pH/elevated temperature starts breaking up into pieces [1]. This was done at a single time point but it can be expected that multiple time points would yield several overlapping pieces of increasingly smaller sizes. The same can be expected at alkaline pH. In the presence of different osmolytes, this breaking down patterns can be different. This sort of breakdown analysis was a research plan which we wanted to but could not be pursue [and now it is a future plan]. But had it been done, it would have been facilitated with memorization of sequence. Off course one can take recourse to bioinformatics software which everyone does. But we reasoned that having imprinted sequence in our mind will help us do the thing which a computer can't. I started the efforts by beginning with alphabetic code. But it turned out to be extremely difficult. Then came the idea of numeric code. Once the idea sparked, we found out that it is not merely useful in memorizing the primary sequence but also has applications in other domain of 'Amino Acid Scale'. All this has been covered in Chapter-6.

Reference:

1 Frare, E. *et al.* Identification of the core structure of lysozyme amyloid fibrils by proteolysis. *Journal of molecular biology* 361, 551-561, doi:10.1016/j.jmb.2006.06.055 (**2006**).

Chapter-3

Experimental Methodology Used in this Thesis

3.1 Materials used

All Osmolytes (at least 98% pure), HEWL (95% pure, remaining being buffer salts), Human Lysozyme, Thioflavin-T and Micrococcus *lysodeikticus* [as a powder] were purchased from Sigma-Aldrich. Absolute ethanol was from Fluka. PD-10 desalting column was purchased from GE Health Care Pvt. Ltd. Dansyl chloride (2-dimethyl aminonaphthalene-6-sulfonyl Chloride was purchased from Invitrogen, Molecular Probes™, USA. Sodium dihydrogen phosphate dihydrate, Potassium Chloride, Sodium Bicarbonate, Sodium Hydroxide and DMF were purchased from Merck Limited, Mumbai and were of GR grade.

3.2 Buffers

Buffers were made in the following way. pH was measured with Sartorius pH meter (P-20). The 50 mM phosphate pH 7.0 buffer was made using KH_2PO_4 with pH adjusted using 5-M and 1-M NaOH / HCl. The pH 8.5 buffer (20 mM) was made in Gly-Gly buffer [for Thioflavin-T experiments]. The pH 12.2 [except in above case where a series of pH ranging from 11.0 to 12.80] was made by starting with 50 mM Potassium diphosphate buffer pH 7.0 and pH altered by adding NaOH solution. To make pH 2.0 buffer, 70 mM KCl solution was made, and pH adjusted to 2.0. The pH meters used were pre-calibrated using pH 7.0 and pH 10.0 standards [when working with alkaline solutions] just before measurements. They were calibrated using pH 7.0 and pH 4.0 standards when working with acidic solutions. To prevent microbial contamination, sodium azide was added to all buffers to a final concentration of 0.02% (w/v).

3.3 HEWL stocks

The HEWL stocks were made in Milli-Q water with ~40 mg/mL concentrations and their concentrations were cross-checked with UV Spectrophotometer using extinction coefficient of 38,400 M-1 cm-1 at 280 nm. To prevent microbial contamination, sodium azide was added to a final concentration of 0.02% (w/v). These stocks were stored at 4°C and used within a month.

3.4 Preparation of Osmolyte Stocks and Final Reaction Mixtures for Alkaline and Acidic Conditions

Stocks were made at 1.0 M or 500 mM depending upon highest concentrations used. They were almost all soluble (in pH 12.2) except some cases which have been mentioned in thesis at appropriate places. Trehalose was only exception. Because Trehalose has been used at final concentration of 1.0 M, so its stock was made at 1.2 M. To make stocks of osmolytes, they were added to pH 12.2 buffer, the pH of resulting solution(s) were re-adjusted back to pH 12.2. Then, sodium azide was added to a final concentration of 0.02% (w/v). These stocks were then stored at 4^{o}C. Because of these protective measures apart from the fact that pH was harsh (12.2), these osmolyte stocks could be used for up to 3-months without any problem. To make final assay with different concentrations of osmolyte, appropriate amount of these stocks were diluted into pH 12.2 buffer. Just before starting the reaction; HEWL in Milli-Q was added. HEWL stock was typically diluted at least 20-fold [for alkaline pH]. Procedure of osmolyte stock making for acidic pH (2.0) was similar. The only difference was that for acidic case due to high concentration of HEWL (600 µM), it was not added as a Milli-Q solution but directly as HEWL powder.

Since amount of stock solutions made with osmolytes was small keeping cost in mind (of the order of 5-10 mls); therefore, a micro pH glass electrode [Make: Sartorius] was employed.

3.5 Samples for Ethanol condition

For making ethanol samples [ethanol/water mixture of different v/v], it was noticed that if we try to dissolve HEWL powder in pre-mixed ethanol/water mixture, it fails to dissolve despite sonication or mild heating. Therefore, firstly HEWL powder was dissolved in water. Then this was added to ethanol [ethanol used was absolute ethanol from Fluka] and topped up with remaining water to achieve desired v/v of Ethanol/Water and 2 mg HEWL/ml. The way we conducted our ethanol experiments was different from its discoverers. They also used salt [NaCl at such low concentrations as 10 mM] in addition to ethanol and obtained amyloid precipitates [1] . However, their claim of using salt-free HEWL is wrong since they [like us] have used HEWL from Sigma which has 5% buffer salts in it. But we reasoned that precipitates are very difficult to study. Moreover, perturbation of salts is something we would like to do later and in this thesis focussed on using no exogenous salt [though 5% buffer salts are there]. Reason for not using salt [exogenous] was that we were interested

in soluble amyloid [manifested as clear solution to naked eye though exhibiting Tyndall effect when viewed with 650 nm, 5 mW Laser pointer in orthogonal direction]. Certainly soluble amyloid is far from ideal solution wherein particle size is supposed to be tending to zero. By soluble amyloid we mean that particle size is very small [not visible to naked eye. It is to be noted that vision acuity for adults is ~30 µm] and remains suspended with Brownian motion thus appearing homogeneous and mono-phasic. Though far from ideal solution, it is nevertheless a solution and hence extremely well developed principles of solution state physical chemistry can be applied [and many techniques working for solution state can be made use of]. On the contrary our knowledge for biphasic solution-solid equilibria is limited and so are techniques at hand. A very good and relevant example is applicability of steady state fluorescence anisotropy (r_{ss}). We could use r_{ss} only because it was a solution. Had it been solid or biphasic suspension, r_{ss} was ruled out due to problems of scattering. Detailed study with osmolytes under ethanol condition was not carried out and this remains future plan. Such studies are likely to be marred by the fact that many of the molecules used in this thesis are salts which when used with ethanol/water mixture are known to give precipitates (which could be amyloid but difficult to study and as of now we are interested in soluble amyloids). An elementary solubility experiment did however show that TMAO and Spermine [in free base form, not in salt form] did dissolve easily up to ~500 mM concentration in absolute ethanol. Whereas Ectoine did not dissolve much. Of course here we are talking about solubility in absolute ethanol. Chances are that Ectoine and other osmolytes could dissolve more in 80-90% Ethanol/Water [v/v] mixture OR [just as we circumvented solubility problem with HEWL as described above] first we dissolve them in pure water and dilute in ethanol OR even if that fails, we might opt to work with lower concentration solutions. In this thesis we have only changed ethanol concentration and perturbation of temperature still remains on drawing board. Once that is done and result in amyloidogenesis at higher temperatures, we can make use of the same and it is likely that many osmolytes dissolve at such high temperatures. One perturbation in the form of HEWL [2,3] concentration especially at lower concentration is also desired and it is very much possible that lower HEWL concentration might well tolerate osmolytes without causing precipitates.

3.6 Guandinium Hydrochloride Samples

A stock of 6.0-M Guandinium Hydrochloride was made in pH 7.0 buffer. This was diluted appropriately into pH 7.0 buffer to make 2.0 M, 3.0 M, 4.0 M and 5.0 M samples. Just before starting reaction, 120 µM HEWL [final concentration] was added [which itself was made in Milli-Q water and was not more than one month old].

3.7 Heating and Cooling of Samples: All 25°C (alkaline pH and ethanol) samples were kept in an incubator with both heating and cooling mechanism activated. 57°C samples (used with acidic pH samples) were kept in hot air oven). Samples in Guandinium Hydrochloride (45°C) were kept in a water bath. Samples at 4°C were kept in refrigerator. Samples at 15°C were kept in cold julabo.

3.8 General Methodology used in this thesis with Rationale

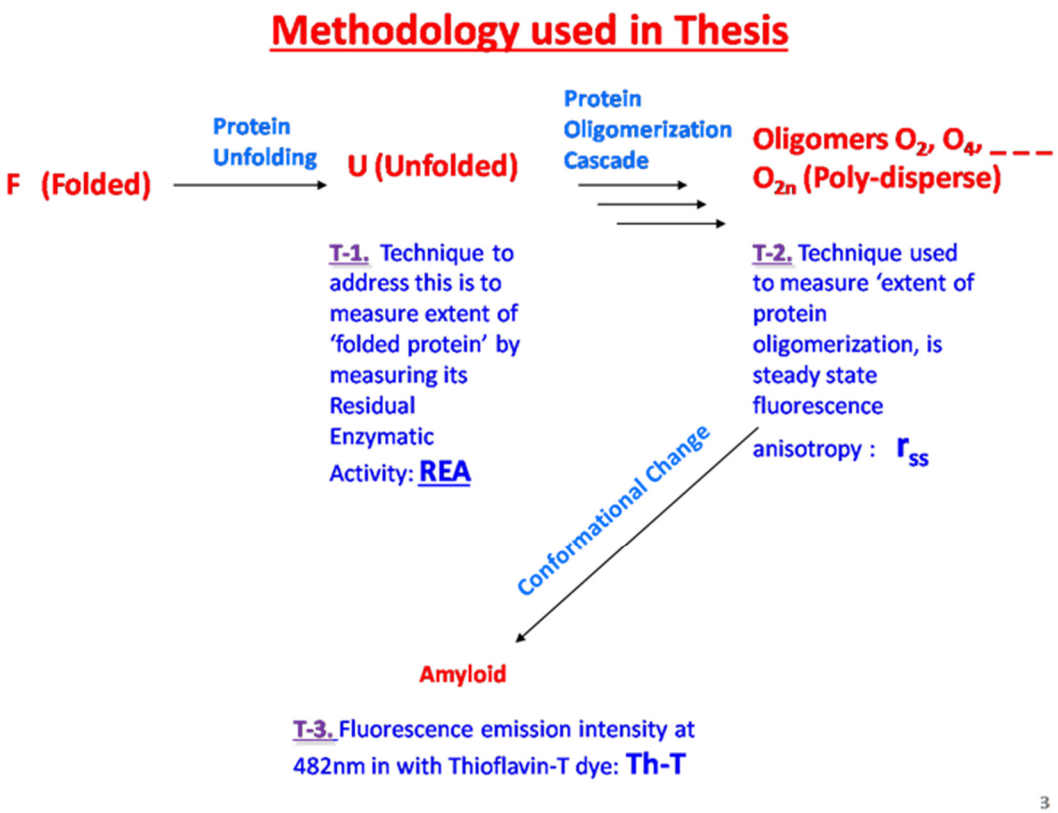


Figure 3.1: Methodology used in Thesis

3.9.1 Mechanism of Lysozyme Action

Lysozyme was the first enzyme for which the X-ray structure was determined at high resolution. This was achieved in 1965 by David Phillips [4]. Phillips went on to propose a mechanism for lysozyme action that was based principally on structural data. The Phillips mechanism has since been borne out by experimental evidence.

****

Following mechanistic details have been taken from the website: http://www.open.edu/openlearn/science-maths-technology/science/biology/proteins/content-section-6.3

The reason to go in this greater detail than what is usual for Methodology chapter is to document few observations we have about the method. We have listed them in later part of this section.

Lysozyme catalyses the hydrolysis of glycosidic bonds that link *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG) in polysaccharides of bacterial cell walls. In doing so, it damages the integrity of the cell wall and thereby acts as a bactericidal agent. The NAM–NAG bond is represented in Figure-3.2, with the site of cleavage by lysozyme indicated.

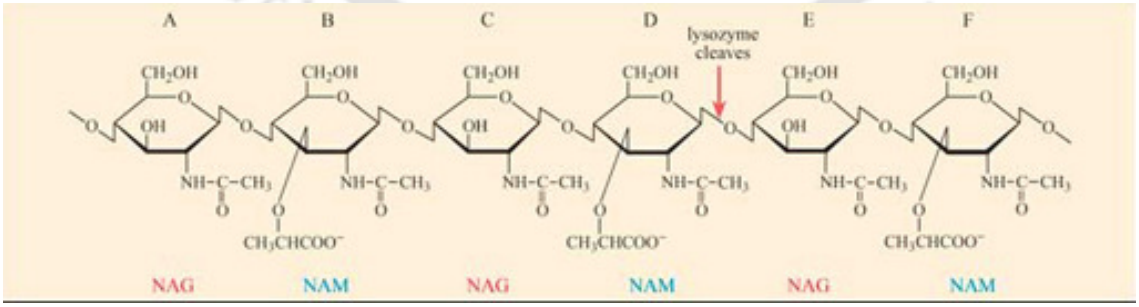


Figure-3.2 Part of the polysaccharide component of bacterial cell walls, showing the alternating *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG) residues. This polysaccharide is a substrate for lysozyme, which hydrolyses the glycosidic bond at the position indicated. (For clarity, and to permit a linear representation of the molecule, some of the bonds are shown in a zig-zag form.)

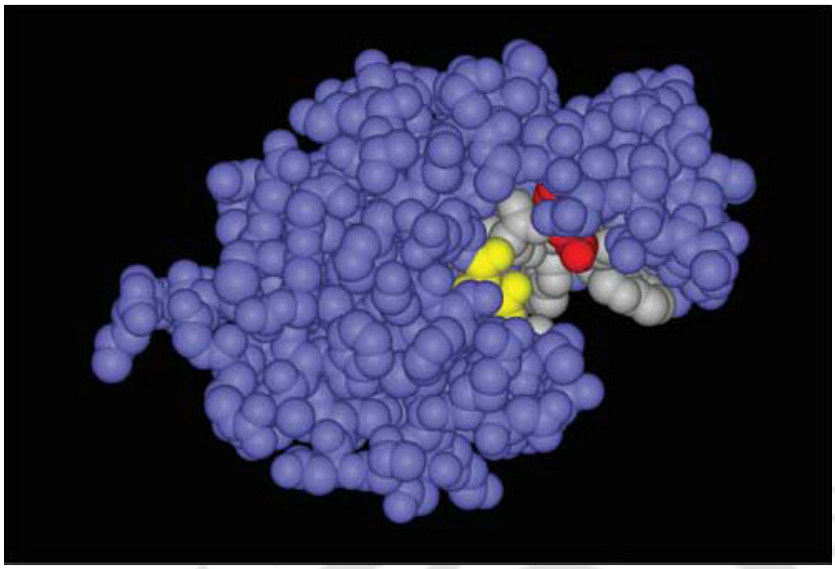


Figure-3.3 A space-filling model of hen egg white lysozyme in which key residues have been highlighted. Asp 52 is in red; Glu 35 is in yellow; some of the residues lining the substrate binding pocket are shown in grey.

The active site of lysozyme is a long groove that can accommodate six sugars of the polysaccharide chain at a time (Figure-3.3). On binding the polysaccharide, the enzyme hydrolyses one of the glycosidic bonds. If the six sugars in the stretch of polysaccharide are identified as A–F, the cleavage site is between D and E, as indicated in Figure-3.2. The two polysaccharide fragments are then released. Figure-3.4 depicts the stages of this reaction, which are also described in detail below.

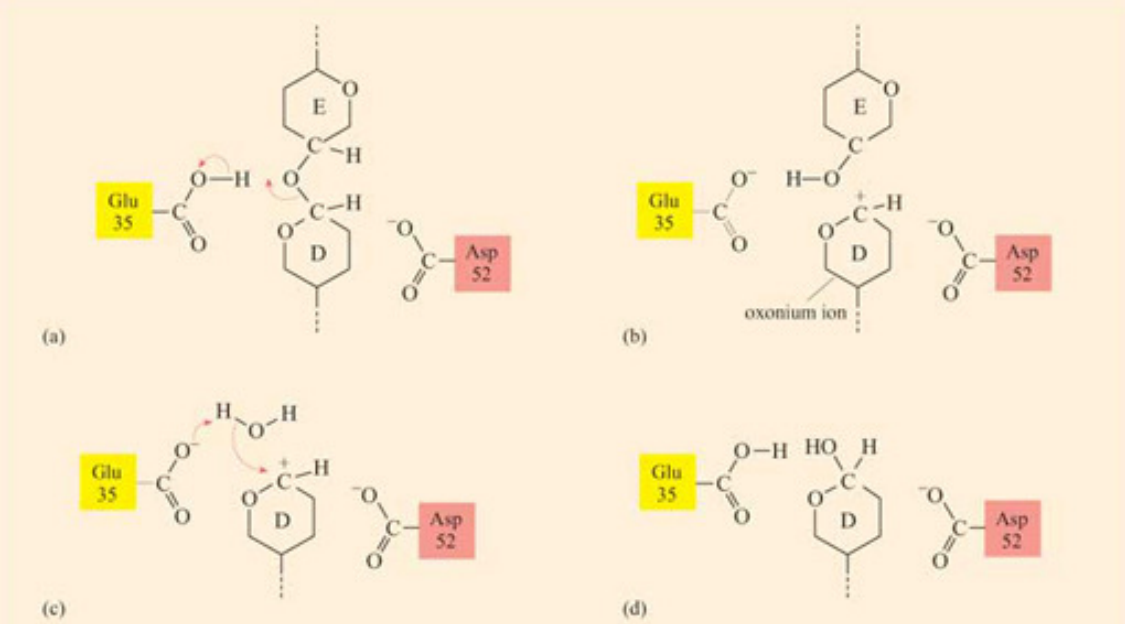


Figure-3.4 The catalytic mechanism of lysozyme. Note that only key residues involved in catalysis (Glu 35 and Asp 52) are shown. The stages are described in detail in the text. [5]

1. On binding to the enzyme, the substrate adopts a strained conformation. Residue D is distorted (not shown in the diagram) to accommodate a $–CH_2OH$ group that otherwise would make unfavourable contact with the enzyme. In this way, the enzyme forces the substrate to adopt a conformation approximating to that of the transition state.
2. Residue 35 of the enzyme is glutamic acid (Glu 35) with a proton that it readily transfers to the polar O atom of the glycosidic bond. In this way, the C–O bond in the substrate is cleaved (Figure-3.4a and Figure-3.4b).

Residue D of the polysaccharide now has a net positive charge; this reaction intermediate is known as an *oxonium ion*). The enzyme stabilises this intermediate

3. in two ways. Firstly, a nearby aspartate residue (Asp 52), which is in the negatively charged carboxylate form, interacts with the positive charge of the oxonium ion. Secondly, the distortion of residue D enables the positive charge to be shared between its C and O atom. (Note that this sharing of charge between atoms is termed *resonance* in the same way as the sharing of electrons between the atoms of the peptide group.) *Thus the oxonium ion intermediate is the transition state*. Normally,

such an intermediate would be very unstable and reactive. Asp 52 helps to stabilise the oxonium ion, but it does not react with it. This is because, at 3 Å distance, the reactive groups are too far apart.

4. The enzyme now releases residue E with its attached polysaccharide, yielding a glycosyl-enzyme intermediate. The oxonium ion reacts with a water molecule from the solvent environment, extracting a hydroxyl group and re-protonating Glu 35 (Figure-3.4c and Figure-3.4d).
5. The enzyme then releases residue D with its attached polysaccharide and the reaction is complete.

- The catalytic mechanism of lysozyme involves both general acid and general base catalysis.

The Phillips mechanism for lysozyme catalysis, as outlined above, is supported by a number of experimental observations. In particular, the importance of Glu 35 and Asp 52 in the process has been confirmed by **site-directed mutagenesis (SDM)** experiments.

****

3.9.2 Enzyme Activity as a measure of extent of Folded Lysozyme

As we just saw, Lysozyme catalyses hydrolysis of 1,4-beta-linkages between N-Acetyl Muramic acid (NAM) and N-Acetyl-D-Glucosamine (NAG) in peptidoglycan of cell wall. Gram-positive bacteria are quite susceptible to this hydrolysis as their cell walls have a high proportion of peptidoglycan. The activity of lysozyme is a function of both pH and ionic strength. The enzyme is active over a broad pH range (6.0 to 9.0). At pH 6.2, maximal activity is observed over a wider range of ionic strengths (0.02 to 0.100 M) [6]. The next two figures have been quoted from this reference.

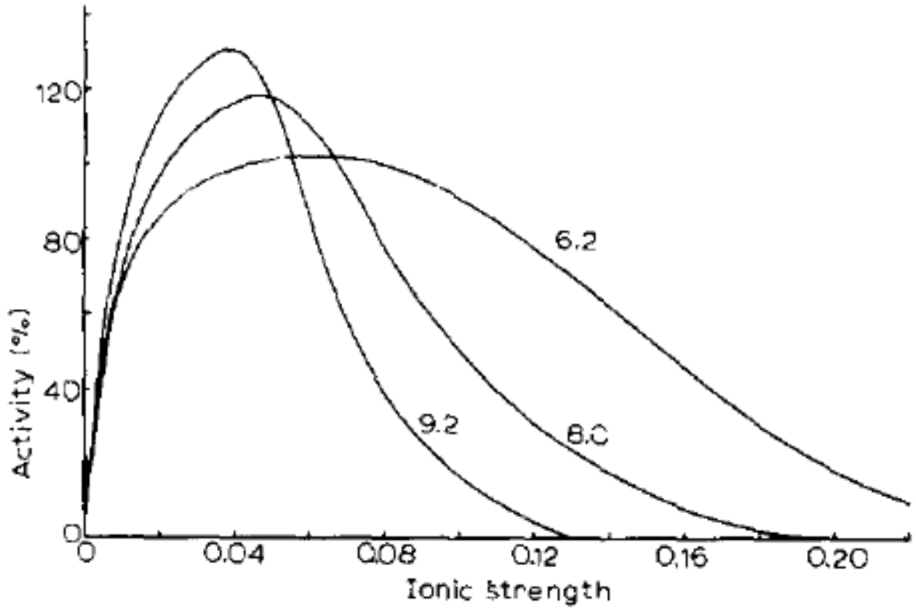


Figure-3.5. Lytic activity of hen's egg-white lysozyme as a function of the ionic strength of the solution at pH 6.2, 8.0 and 9.2. Activities are expressed as a percentage of that observed in 0.05 M phosphate (pH 6.2, I = 0.067).

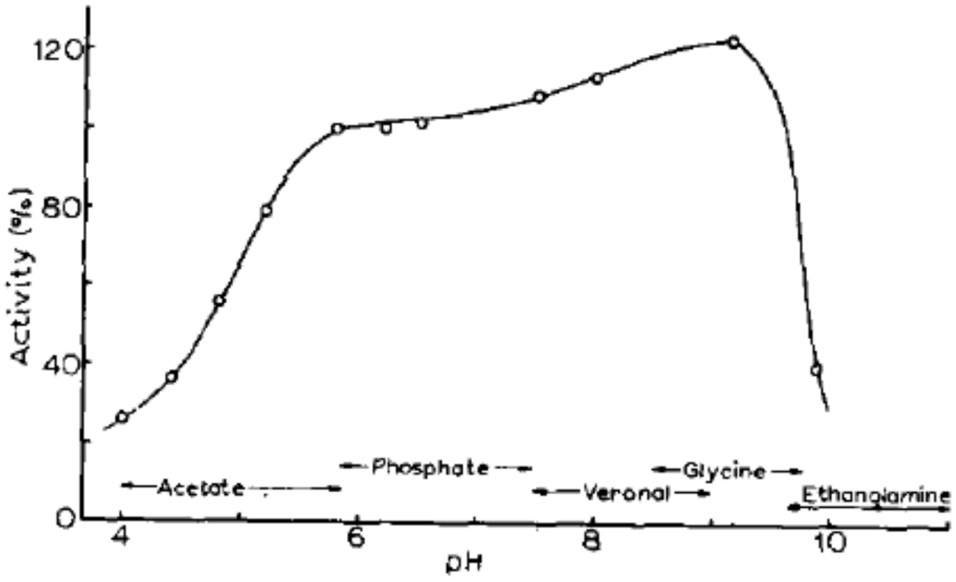


Figure- 3.6. Lytic activity of lysozyme as a function of pH at ionic strength 0.04-0.05. Activity is recorded as a percentage of the activity at pH 6.2. Buffers (0.01 M) were used with NaC1 (0.04 M)

Due to this lysis bacteria dies and bacterial suspension becomes clearer. This is measured by noting down absorbance at 450 nm.

Unit definition of HEWL: one unit will produce a change in A_{450} of 0.001 per minute at pH 6.24 at 25ºC using a suspension of Micrococcus lysodeikticus as a substrate in 2.6 ml assay (1 cm light path).

3.9.3 Observations and Concerns We Have

We however have concerns as to how bacteria survives at pH 9.2 and whether lysis due to such high pH could be simultaneously occurring thus leading to higher lysis at pH 9.2 as compared to pH 6.2. An alternate explanation of higher activity at pH 9.2 could be that at this pH, protein is mostly dimer [7,8] and also from our studies. This dimer is likely to be more than doubly active than monomer (because it punctures bacteria twice at same place, so likely to have more lethal effect than if puncturing at distant places] [this could be a future plan wherein we cross-link dimer at ~pH 10.0 and study its activity]. This hypothesis is supported by the dependence on the ionic strength. Higher ionic strength is likely to dissociate dimer and hence sharper ionic strength dependence profile at pH 9.2. Another concern about such high activity at pH 9.2 comes from mechanism itself. Glu-35 which displays unusually high pKa of ~6.5 (rather than ~4.0) and acts as general acid is unlikely to be available at such high pH of 9.2 (as a General Acid) so in a way established mechanism fails. In nutshell we feel that Lysozyme mechanism is far from settled. Anyways the ionic strength dependence is not explained by Phillips mechanism.

3.9.4 Measurement of Residual Enzymatic Activity

For this purpose, lysis of Micrococcus *lysodeikticus* was carried out [as in Fleming's original paper] using a protocol by Davies [6] . To this end a Micrococcus *lysodeikticus* stock was made in de-ionized water and later diluted in pH 7.0 buffer such that the absorbance at 450 nm was ~0.40. To this was added 700 nM HEWL solution in pH 7.0 phosphate buffer (made by double dilution starting from 120 µM) at 25ºC

120 µM → 7 µM → 700 nM which when used in

Micrococcus *lysodeikticus* (900 µL) + 100 µL of HEWL stock (700 nM) in final assay [HEWL] became 70 nM.

In the beginning (zero time) the HEWL was fully folded (thus its concentration being 70 nM) but as unfolding progressed the amount of folded species decreased and this is what was termed as "Residual Enzymatic Activity (REA)

The way to measure Residual Enzymatic Activity (REA) is described below: as lysis proceeds, absorbance of Micrococcus *lysodeikticus* suspension at 450 nm (A_{450}) decreases. This drop is monitored for 30 seconds. This drop in absorbance is plotted w.r.t. time and yields a straight line with a very good 'R' value. The slope of this line is noted down and this is termed REA [Residual Enzymatic Activity]. REA has no unit. Spectrophotometer used was Cary-100 [Make: Varian] which is a double beam spectrophotometer.

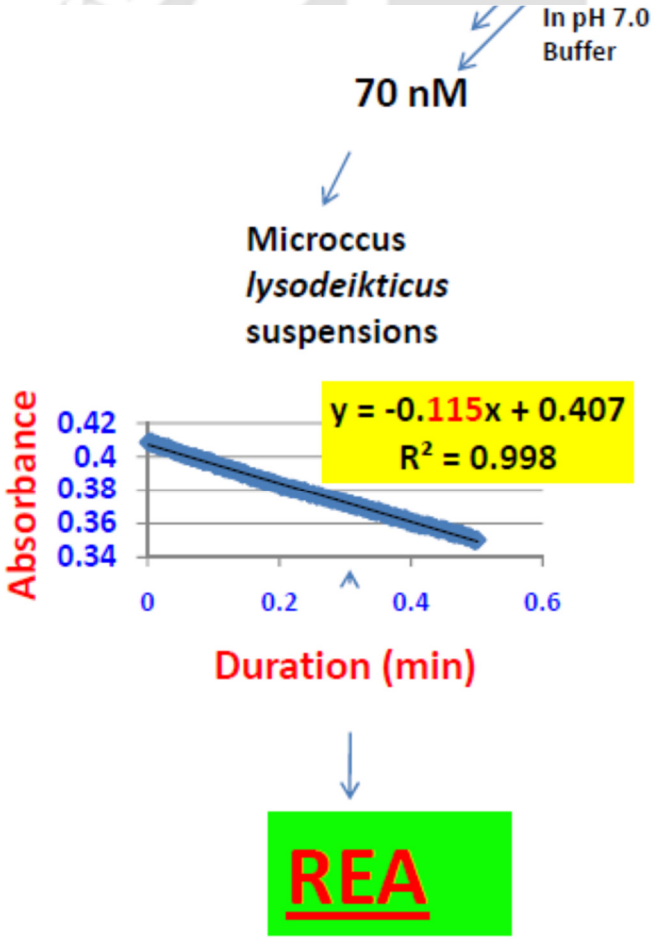


Figure-3.7 Measurement of Residual Enzymatic Activity (REA)

REA experiment was run at each time point itself (of incubations being studied) with fresh Micrococcus *lysodeikticus* solution. However, when same final dilution was saved at 4°C and ran after 24-hours it gave nearly identical values. So in future it may be possible to aliquot samples at each time point and run them all together at a go in the end. It will save considerable time and efforts.

Since enzyme activity critically depends on ionic strength, so question arises as to what happens when osmolytes are present. Answer lies in the fact that we dilute initial reaction mixture from 120 µM to 700 nM, i.e. 171 times, so even if initial assay has had 500 mM osmolyte, in the final reaction, it is no more than 3 mM which is negligible as far as its effect on ionic strength is concerned [as we have seen, Lysozyme activity depends a lot on ionic strength, though mechanism of that is not clear, nor seems to have been studied adequately in literature]. Moreover, what we are measuring is ratio-metric, with concentration of osmolyte remaining identical, only amount of folded HEWL varying, so it does not matter what the osmolyte concentration was.

3.10 Thioflavin-T Fluorescence for measuring extent of Amyloid

Clinical literature of detection of *ex vivo* amyloid came up with two gold standards for amyloid detection. One was birefringence with dye Congo-Red (CR) and another was fluorescence with Thioflavin-T [9]. Later on they were adapted for *in vitro* studies [10-12]. In a paper in 2001, it was argued that Congo-Red is not an amyloid specific dye and that it binds to native and partially folded proteins and in some cases also leads to oligomerization though they themselves have doubt about it [13]. Because of this paper, now Congo-Red is not used much.

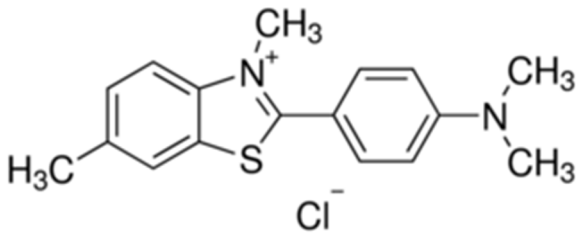


Figure-3.8 Structure of Thioflavin-T

Harry Levine in 1993 [11] reported that Thioflavin-T associates with amyloid fibrils rapidly (within a matter of seconds) and changes the UV absorption and fluorescence emission properties of Th-T. While free Th-T has absorption maxima at 385 nm and emission maxima at 445 nm, that of dye bound to amyloid is 450 nm and 482 nm respectively. This change is only upon binding to amyloid state.

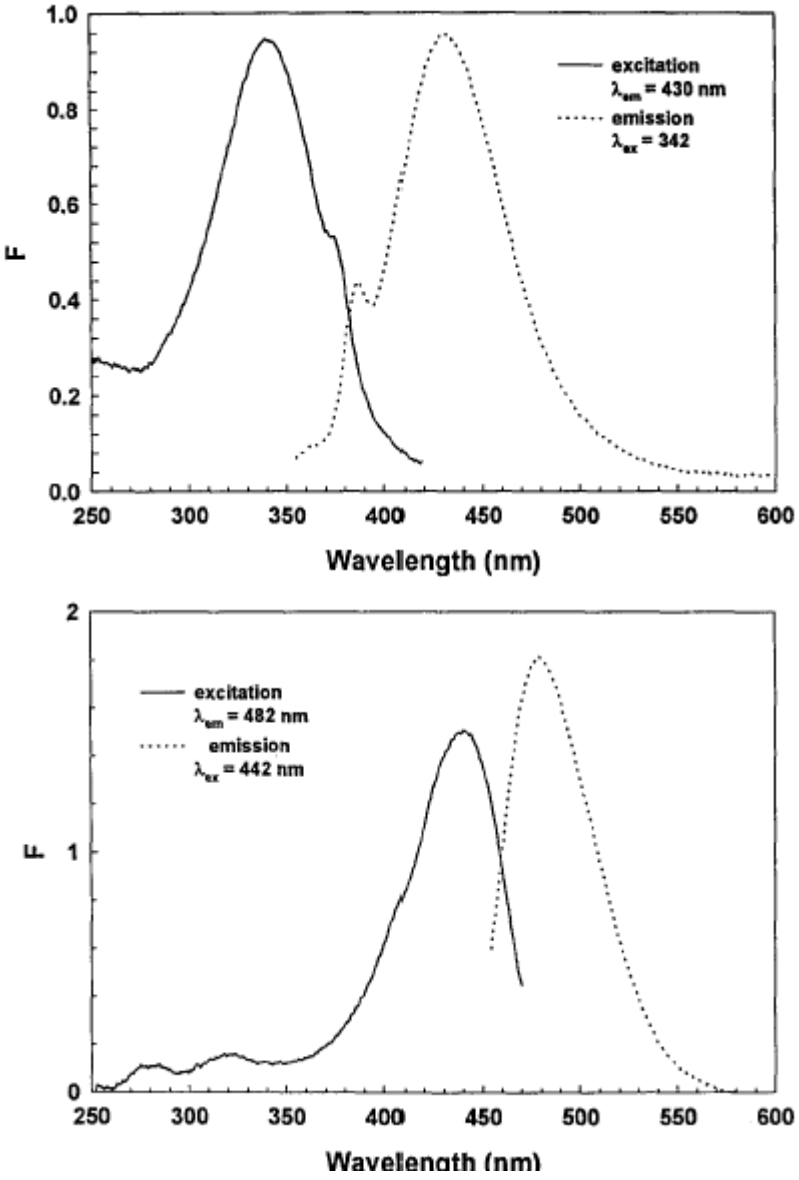


Figure-3.9 Fluorescence spectra of free and Aβ(1-42)-bound Thioflavin-T (ThT): 5 µM ThT in 50 mM glycine-NaOH, pH 8.5; 2 µM peptide when present. Excitation/emission fluorimeter slits = 3/10 nm. Solid line, excitation; dotted line, emission spectra. *(Top) 5 µM* ThT alone. *(Bottom) 5 µM* ThT + 2 µM Aβ(1-42) amyloid fibrils.

Of course Thioflavin-T has advantage that to some extent, it can be used to quantify amount of amyloid [12] .There is a flurry of papers in recent years to investigate mechanism of Thioflavin-T [14-18]. While not going into the merits of these papers [yet to study them in great detail], we do however still feel that after fixating 3-4 well studied *in vitro* amyloid forms, simple pH titration of UV-absorbance and fluorescence across wide range of pH, a similar temperature titration, changing stoichiometric ratios of protein to Thioflavin-T over a wide range, changing polarity of solvents, addition of commonly used co-solvents in protein biophysical chemistry and use of some synthetic modifications of the dye could be an alternate way to redress the mechanism issue. Though we did not study this line of investigation, we did touch upon them to some extent. In our work with ethanol, in final Th-T assay there was ~8-9% ethanol [at pH 8.5]. Likewise, though recommended pH of Thioflavin-T assay was 8.5, when we added pH 12.2 sample, the final pH became 10.9 which is way too above of the recommended and yet exhibited fluorescence enhancement at 482 nm. To counter this abrupt change in pH, we changed buffer from 20 mM Gly-Gly pH 8.5 to 100 mM Gly-Gly pH 8.5 but that lowered fluorescence intensity at 482 nm markedly. This line was investigation was therefore too preliminary as it was not the focus of the thesis. Nevertheless, we feel that this approach can be a fulfilling exercise and shed valuable inputs into the mechanism of Thioflavin-T. Settlement of mechanism of Thioflavin-T requires urgency because it is so widely used. And use of Congo-Red needs to be revived

One point which we would like to emphasize is that, Th-T data is not alone in this thesis. It is further complemented with steady state anisotropy data which measures extent of oligomerization. Both have been done together in this thesis. We thus have a truth table like case

S. No.	**r_{ss}**	**Th-T**
1	Low	Low
2	Low	High
3	High	Low
4	High	High

In the above hypothetical table, barring scenario (2), all other cases are of self-complementary nature. Indeed, in this thesis we have never encountered scenario (2) and encountered only cases 1,3 &4.

Continuous wave steady-state fluorescence spectra were recorded using a HORIBA Fluoromax-3 spectrofluorimeter. Fluorescence spectra were obtained from samples in quartz cuvettes with 10 mm length The protein was diluted to 12 µM in 20mM pH 8.5 [Gly-Gly] buffer. To this was added Thioflavin-T stock (in Milli-Q) to final concentration of 20 µM. The pH for final assay was 10.9 [when pH 12.2 sample was used].

The Th-T stock (in Milli-Q) was 1.0 mM, with concentration measured by absorbance at 380 nm in Ethanol with extinction coefficient 26620 cm^{-1} [19].
For recording Thioflavin-T fluorescence above sample (mix of protein and Thioflavin-T dye in pH 8.5 Gly-Gly buffer) was gently vortexed for 30-seconds and immediately placed in cuvette. Sample was excited at 450 nm. Excitation slit was 1 nm and emission slit was 10 nm. Emission was measured from 470-550 nm. Emission intensity at 482 nm [which was maxima of this graph] was noted down [after deducting blank with buffer alone].

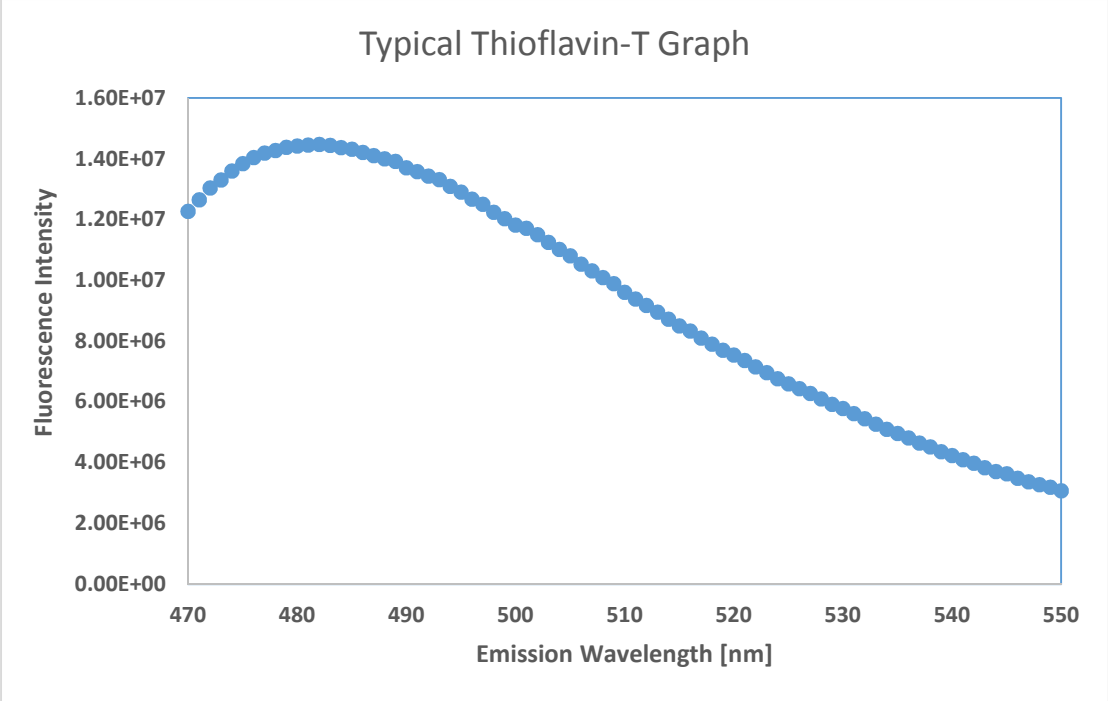


Figure-3.10 Typical Thioflavin-T Spectra

3.11 Theory of Steady State Anisotropy (ss-Anisotropy) Experiments

Our group has shown that steady state anisotropy experiments can be successfully used to have a rough idea of oligomer/amyloid size [20]. While there are other techniques for the same such as Scattering Methods. But they require specialized and costly instruments and very careful experimentation and careful keeping of instruments. Whereas ss-anisotropy experiments can be performed with steady state fluorimeter which are found in almost all institutions. All that is required is an accessory to be fitted with this instrument and this accessory is not very expensive. This instrumentation does not require any special housekeeping. And running ss-anisotropy experiments is perhaps one of the simplest amongst all of experimental sciences. Just put entire reaction mixture in quartz cuvette [or let reaction mixture incubate in cuvette itself, no dilution, no mixing, nothing] and duration of experiment is barely few seconds.

Despite this simplicity it is surprising that this technique has barely been used by protein aggregation community. One reason could be that in many cases, amyloids are insoluble

and thus not amenable to steady state fluorescence anisotropy measurements as they cause scattering. This is one more reason as to why; we have been interested in soluble amyloids as far as possible. However though not done in this thesis, perhaps the most accurate experiment would have been to use Osmometer.

Anisotropy measurements are based on the principle of photoselective excitation of fluorophores by polarized light. This selective excitation results in a partially oriented population of fluorophores (photoselection), and in partially polarized fluorescence emission. Emission also occurs with the light polarized along a fixed axis in the fluorophore. The fluorescence anisotropy (r) is defined

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$$

Where $I_{\parallel}$ and $I_{\perp}$ are the fluorescence intensities of the vertically and horizontally polarized emission, when the sample is excited with vertically polarized light.

Experimentally we use an accessory fitted with Horiba Fluoromax-3 which is a L-Format with dual Glan-Thompson polarizers placed at the entrance and the exit of T-box and a PMT (Photo-Multiplier Tube) operating at a voltage of 950 V in photon counting mode. Excitation was at 380 nm with slit width 1 nm and emission was at 444 nm with slit width of 4 nm. In L-format we measure four components; I_{VV}, I_{VH}, I_{HV} & I_{HH} where subscripts refer to the orientation of excitation and emission polarizers; e.g., I_{HV} corresponds to horizontally polarized excitation and vertically polarized emission.

I_{HV}/I_{HH} gives G factor (which is the ratio of the sensitivities of the detection system for vertically and horizontally polarized light) which is calculated during course of measurements. Once this G value becomes available, the anisotropy is calculated from the following equation:

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}}$$

Rotational diffusion decreases values of measured anisotropy if rotational correlation time is comparable to excited-state lifetime of fluorophore.

Steady state fluorescence anisotropy (r_{ss}) gives a time integrated average value of the rotational motion of molecules in the excited state:

$$r = \frac{\int_0^\infty I(t)r(t)dt}{\int_o^\infty I(t)dt}$$

Anisotropy measurements provide information on the size and shape (not molecular weight) of proteins from following equation (also called Perrin's equation)

$$r = \frac{r_0}{1 + \left(\frac{\tau}{\theta}\right)}$$

where r_0 is the fundamental anisotropy of fluorophore & τ is the life time of fluorophore (which for Dansyl is ~11 ns at pH 7.0 and pH 12.2) [20] . Value of θ for monomeric HEWL is 22 ns [21] . Thus life time of dansyl and rotational correlation time of monomeric HEWL (and that of small oligomers of HEWL) is comparable. However, for higher order oligomers, they are no more comparable and that could explain why values of measured r_{ss} is not changing as much as expected. Another reason for lower r_{ss} values is the fast rotation around dansyl-HEWL bond. Ideally we should do (separate) labelling with 2-3 fluorophores with widely different life times (e.g. one could be Dansyl, another could be pyrene sulfonamides with longer life-time up to ~30 nanoseconds and yet higher values [perhaps by use of transition metal-ligand complexes but they should pass the criteria of being able to withstand alkaline pH (12.2)]. We did not do all this, so what we have been looking is highly qualitative. Using above equation (anisotropy is related to rotational correlation time θ). This rotational correlation time is related to molecular volume V (assuming spherical shape) which in turn is related to molecular weight M by

$$\theta = \frac{\eta V}{RT} = \frac{\eta M}{RT}(\bar{v} + h)$$

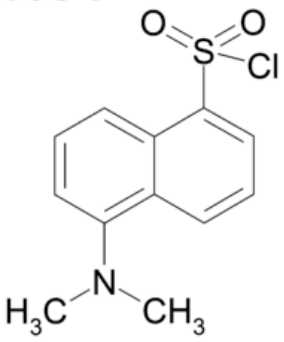
Here η is viscosity, $\bar{v}$ is specific volume of the of protein (volume per gram of protein) which is typically 0,.73 ml/g and h is hydration which is typically 0.23 g H_2O per gram of protein.

The r_{ss} is a dimensionless quantity and does not depend on concentration of fluorophore.

3.11.1 Protein Labelling with Dansyl dye

For fluorescence experiments we need a fluorescent group. Tryptophan does offer such possibility. But since there are four tryptophan

residues in HEWL so it can't be used. Therefore, external labelling is essential. One such group is Dansyl which survives the alkaline pH used and shown to work in previous studies from our group [20] . Dansyl-Cl group reacts with amino group. The structure of Dansyl-chloride is



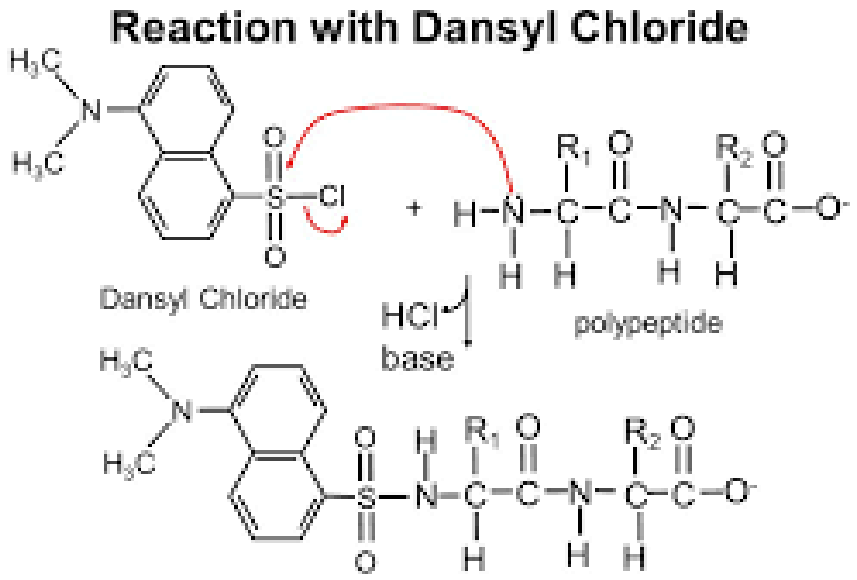


Figure 3.12: Mechanism of Dansyl Chloride Reaction with Proteins

For Dansyl-labelling, fresh $NaHCO_3$ buffer (100 mM, pH 9.0) was made. In 1.0 ml buffer in Round Bottom flask was stirred 10 mg HEWL at 4°C in cold room. In a separate vial 6 mg Dansyl Chloride was mixed in 200 µL DMF. This did not dissolve completely and remained suspended. This mixture of Dansyl Chloride was added dropwise slowly (over 20 minutes) to HEWL mixture. This mixture was stirred for 3:00 hours. The solution remained cloudy and yellowish till end. A PD-10 [Protein Desalting] column of Amersham which was pre-equilibrated with pH 7.0 buffer was used to isolate protein from unbound dye. UV-Vis absorbance values were recorded at 280 nm (ε=38,400 M^{-1} $cm^{-}1$) and 380 nm (ε=16,000 M^{-1} $cm^{-}1$) to determine extent of labelling. The same turned out to be typically 20% [i.e., one in each five HEWL got labelled with Dansyl].

Though it is believed that Dansyl-Chloride reacts with Lysine side chains (there are six Lysine residues in HEWL). However, if it is so, any of six Lysine residues will react (and sometimes more than one Lysine will react) thus leading to a very heterogeneous mixture). However, we believe (keeping in mind that the pH used in the reaction mixture was 9.0). At this pH, Lysine residues (pKa ~10.5) will still be protonated whereas N-terminal amino group (with pKa ~7.5) will be un-protonated. This un-protonated amino group will be more nucleophilic than protonated Lysine residues (in which case lone pair of nitrogen will not be available). So it is expected that Dansyl will be conjugated to only N-terminal. This will thus make mixture homogeneous which is advantageous.

The assay for steady state fluorescence anisotropy was as follows

---

pH 12.2 [with or without osmolyte] + 118 µM un-labelled HEWL + 2 µM of above labelled HEWL [Dansyl conjugated HEWL].

3.13 The abbreviations used in next chapters to denote results from various techniques is as under

REA refers to Residual Enzymatic Activity, i.e., the remaining enzymatic activity.

r_{ss} OR ss-anisotropy refers to steady state fluorescence anisotropy of partially (2 µM) dansyl-conjugated and remaining being unlabelled HEWL. Excitation wavelength was 380 nm. Emission was recorded at 444 nm. Excitation band pass was 1 nm and that for emission was 4 nm.

ThT refers to Thioflavin-T fluorescence [value of steady state fluorescence at 482 nm] when excited at 450 nm with excitation slit 1 nm and emission slit 10 nm.

References:

1 Goda, S. *et al.* Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein science* 9, 369-375, doi:10.1110/ps.9.2.369 (**2000**).

2 Yonezawa, Y. *et al.* An Insight into the pathway of the amyloid fibril formation of hen egg white lysozyme obtained from a small-angle X-ray and neutron scattering study. *Journal of molecular biology* 323, 237-251 (**2002**).

3 Szymanska, A., Hornowski, T. & Slosarek, G. Denaturation and aggregation of lysozyme in water-ethanol solution. *Acta biochimica Polonica* 59, 317-321 (**2012**).

4 Blake, C. C. *et al.* Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature* 206, 757-761 (**1965**).

5 Phillips, D. C. The three-dimensional structure of an enzyme molecule. *Scientific American* 215, 78-90 (**1966**).

6 Davies, R. C., Neuberger, A. & Wilson, B. M. The dependence of lysozyme activity on pH and ionic strength. *Biochimica et biophysica acta* 178, 294-305 (**1969**).

7 Sophianopoulos, A. J. & Van Holde, K. E. Evidence for dimerization of lysozyme in alkaline solution. *The Journal of biological chemistry* 236, PC82-PC83 (**1961**).

8 Sophianopoulos, A. J. & Vanholde, K. E. Physical Studies Of Muramidase (Lysozyme). Ii. Ph-Dependent Dimerization. *The Journal of biological chemistry* 239, 2516-2524 (**1964**).

9 Kelenyi, G. On the histochemistry of azo group-free thiazole dyes. *The journal of histochemistry and cytochemistry* 15, 172-180, doi:10.1177/15.3.172 (**1967**).

10 Naiki, H., Higuchi, K., Hosokawa, M. & Takeda, T. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavin T1. *Analytical biochemistry* 177, 244-249 (**1989**).

11 LeVine, H., 3rd. Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. *Protein science* 2, 404-410, doi:10.1002/pro.5560020312 (**1993**).

12 LeVine, H., 3rd. Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods in enzymology* 309, 274-284 (**1999**).

13 Khurana, R., Uversky, V. N., Nielsen, L. & Fink, A. L. Is Congo red an amyloid-specific dye? *The Journal of biological chemistry* 276, 22715-22721, doi:10.1074/jbc.M011499200 (**2001**).

14 Khurana, R. *et al.* Mechanism of thioflavin T binding to amyloid fibrils. *Journal of structural biology* 151, 229-238, doi:10.1016/j.jsb.2005.06.006 (**2005**).

15 Krebs, M. R., Bromley, E. H. & Donald, A. M. The binding of thioflavin-T to amyloid fibrils: localisation and implications. *Journal of structural biology* 149, 30-37, doi:10.1016/j.jsb.2004.08.002 (**2005**).

16 Sabate, R., Rodriguez-Santiago, L., Sodupe, M., Saupe, S. J. & Ventura, S. Thioflavin-T excimer formation upon interaction with amyloid fibers. *Chemical communications* 49, 5745-5747, doi:10.1039/c3cc42040j (**2013**).

17 Rybicka, A. *et al.* Thioflavin T: Electronic Circular Dichroism and Circularly Polarized Luminescence Induced by Amyloid Fibrils. *Chemphyschem* 17, 2931-2937, doi:10.1002/cphc.201600235 (**2016**).

18 Qin, Z. *et al.* Kinetic Mechanism of Thioflavin T Binding onto the Amyloid Fibril of Hen Egg White Lysozyme. *Langmuir* 33, 5398-5405, doi:10.1021/acs.langmuir.7b00221 (**2017**).

19 Nilsson, M. R. Techniques to study amyloid fibril formation in vitro. *Methods* 34, 151-160, doi:10.1016/j.ymeth.2004.03.012 (**2004**).

20 Homchaudhuri, L., Kumar, S. & Swaminathan, R. Slow aggregation of lysozyme in alkaline pH monitored in real time employing the fluorescence anisotropy of covalently labelled dansyl probe. *FEBS letters* 580, 2097-2101, doi:10.1016/j.febslet.2006.03.012 (**2006**).

21 Creighton, T. E. *Proteins: Structures and Molecular Properties*. Second edn, (W. H. Freeman and Company, **1993**).

Chapter-4

RESULTS: Perturbation of Amyloidogenic Conditions of Lysozyme

Since the original idea was to make results free from condition employed, hence plan was made to use all the four conditions, viz., alkaline pH, acidic pH/elevated temperature, concentrated ethanol and Guandinium Hydrochloride/elevated temperature. However, before that is done, it was decided to probe these conditions in detail. As it happens, barring acidic pH/elevated temperature, other conditions are barely studied beyond their discovery. We decided to therefore study these conditions in detail before introducing osmolytes. We decided to do this by systematically varying each of the physico-chemical parameters. It was expected that this will provide mechanistic insight into the aggregation process.

While there are many different conditions for making *in vitro* amyloids of Lysozyme, the following are focus of this chapter:

a. Acidic pH / elevated temperature [1,2]
b. Alkaline pH at room temperature [3]
c. Ethanol/water mixture with or without salt [4-6]
d. In presence of Guandinium Hydrochloride at elevated pH and with mechanical agitation [7]

It can be noticed that for each of the four amyloidogenic conditions, following biophysical parameters were used

1. Temperature. We call it 'Temperature Force'
2. High or low pH OR concentrations of denaturant [ethanol or Guandinium Hydrochloride. We call it 'Physico-chemical Force'
3. Mechanical Force

In this thesis, we focussed only on first two, viz., Temperature Force and Physico-chemical force. We did not study mechanical force which is akin to an extremely developed and rigorously quantitative field of Fluid Dynamics but has not been studied that way (as we feel).

Once we fixated at four conditions and two 'forces'; next choice was the approach. We took clue from physical organic chemistry. In here, we decided to change each of these two forces by small increments and measure the results on three crucial stages of protein unfolding [measured by residual enzymatic activity], oligomerization [by steady state fluorescence anisotropy] and quantity of amyloid [by Thioflavin-T fluorescence]. It is

important to note that both of these forces are to be adjusted [mere adjustment of one will not do]. We feel this approach to be precursor to simulation. Though due to paucity of time we focussed only on four conditions of lysozyme, same needs to be done for other amyloidogenic conditions of other proteins.

4.1 Alkaline pH

4.1.1 HEWL and Human Lysozyme behave similarly

Since amyloidogenic behaviour was observed with point mutants of Human Lysozyme; therefore, most obvious choice for the studies would have been the use of same very mutants. However, since mutants are not commercially available and difficult to express and purify in large scale, the next obvious choice would have been to use wild type Human Lysozyme. But human Lysozyme was expensive and its supply from Sigma-Aldrich was not regular. Thus choice fell to Hen Egg White Lysozyme (HEWL). Another big advantage of using HEWL was that in literature almost all biophysical studies with Lysozyme have been carried out with HEWL, making a plethora of data available for integrating the data generated from this thesis for interpretation and analysis. Further, HEWL is available cheap commercially.

However, to make sure that results of this thesis obtained with HEWL indeed 'do hold' for wild type human lysozyme was demonstrated by comparative studies of Thioflavin-T fluorescence of HEWL and Human Lysozyme at different pH (in alkaline range). Since Thioflavin-T fluorescence measures the final species in the multimodal pathway, so it can reasonably be argued that any similarity at this stage signifies similarity at all intermediate stages. While designing this experiment the gap between pH was kept as low as 0.2. The Result of this experiment is plotted in Figure-4.1.

For a series of buffers between pH 11.0 to pH 12.8 in the increment of buffers were made as per the following table taken from the website http://delloyd.50megs.com/moreinfo/buffers2.html [8].

100 mL 50 mM K_2HPO_4 + mls of 0.1 N NaOH	
pH	mls of 0.1-N NaOH added
11.0	8.2
11.2	12.6
11.4	18.2
11.6	27.0
11.8	38.8
12.0	53.8

50 mL 200 mM KCl + mls of 0.2 N NaOH	
pH	mls of 0.2-N NaOH added
12.20	20.4
12.40	32.4
12.60	51.2
12.80	82.4

This table was found compatible with calculations using Henderson–Hasselbalch equation. Subsequently they were cross-checked with pH meter and found in good accordance.

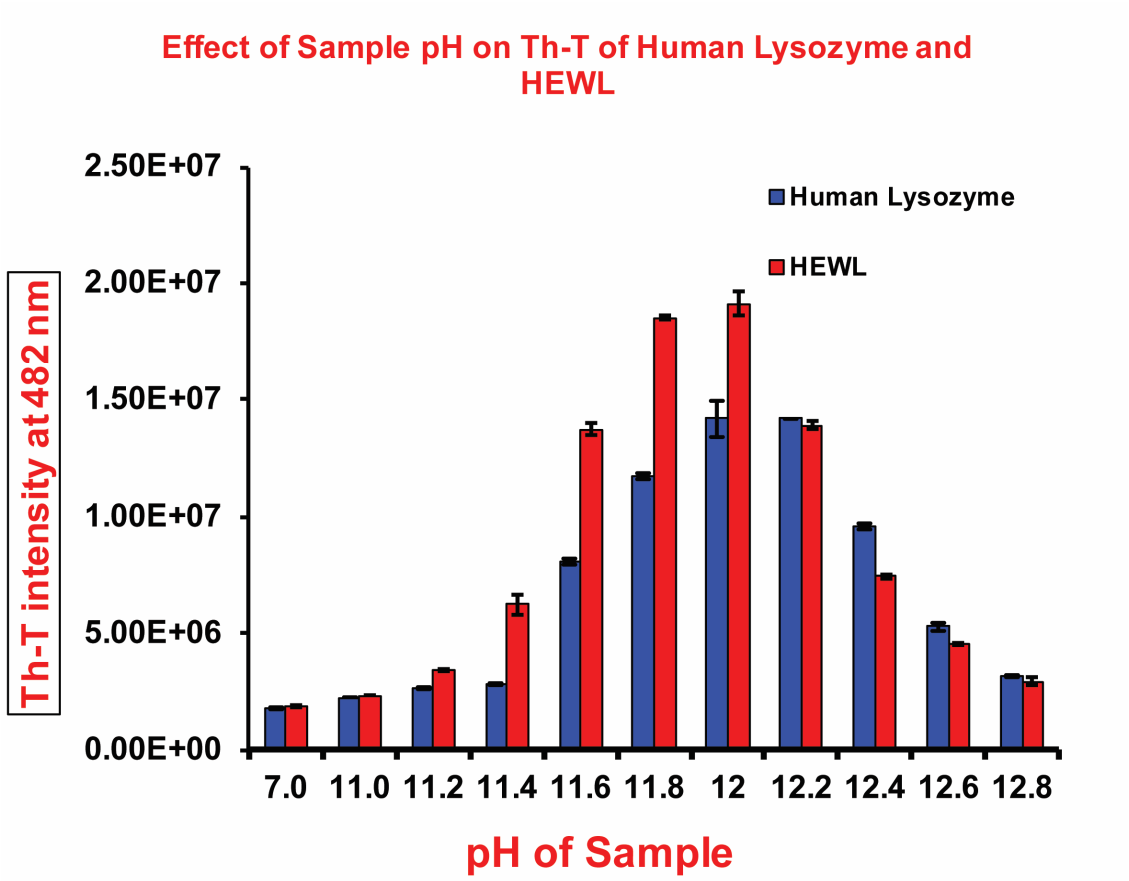


Figure-4.1: Thioflavin T fluorescence at different pH at 25 C incubated for 96 hours of HEWL and human Lysozyme 120 µM each. EX 450 nm. Slits for excitation was 1 nm and that for emission was 10 nm.

Close similarity between HEWL and Human Lysozyme is in marked contrast to the fact that there is only 60% sequence homology and there is significant difference between their behaviour at acidic pH [HEWL forms amyloid at pH 2.0/57°C whereas Human Lysozyme does so at significantly higher pH of pH 3.0/60°C] [2,9].

4.1.2 Perturbation of pH

Chemically speaking, the reaction mixture is 120 μM HEWL in 50 mM Potassium phosphate buffer, pH 12.2 at 25°C in a 2 mL cylindrical micro centrifuge tube, kept without any mechanical agitation. Under alkaline condition, the solution remained fully transparent without any trace of gel or particulate material (to naked eye). Because of this apparent homogeneity and also not to interfere with inherent processes undergoing at mesoscopic levels, aliquots were withdrawn directly without any remixing.

4.1.3 Effect of pH on HEWL REA

Figure 4.2 depicts how the REA of HEWL changes with time of incubation under different pH values of the medium from 11.0 to 12.2. It is noticed that entire activity is retained after 400 hours at pH 11.0, a dip in activity is palpable for all pH values higher than pH 11.0 in comparison. For pH 12.2, nearly all enzymatic activity is lost near 140 hours unlike at pH 11.4 and 11.8, highlighting the harsh nature of the condition. At pH 11.0 there is virtually no loss in activity even after 3-weeks. But beyond pH 11.0, the activity drops rapidly. Previous work in literature (Sophianopoulos JBC1964) have not reported activity beyond pH 11.0 at which point they also report oligomers. Then there is a very old study where they have managed to get crystals at pH 11.4 [10].

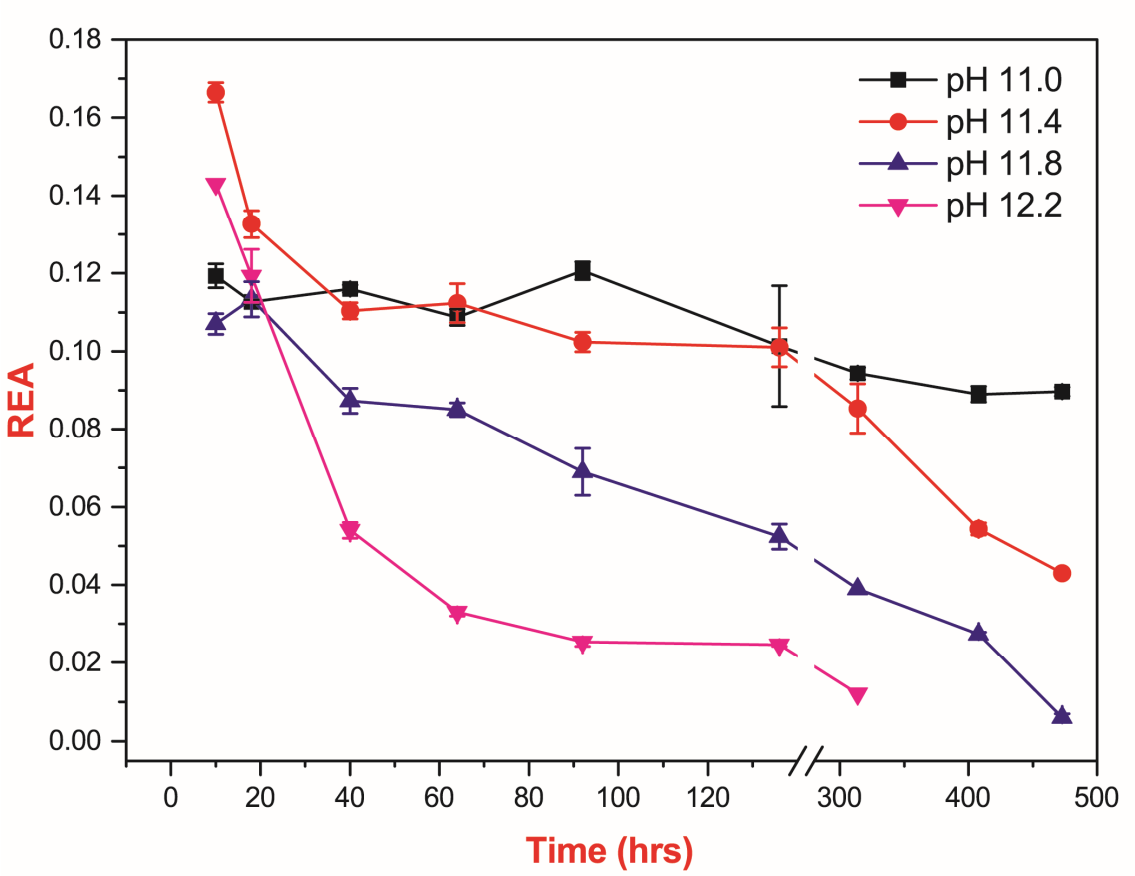


Figure-4.2: REA (in terms of slopes for first 30 s) of HEWL (120 µM) incubated at diff pH at 25°C

4.1.4 Effect of pH on r_{ss} of dansyl-labelled HEWL

The effect of pH on HEWL oligomer size and growth kinetics (as revealed by r_{ss} vs. time plots) is shown in Figure 4.3. Starting from pH 12.2 where bountiful production of oligomers is accompanied by fast growth kinetics, a significant decline in both the growth rate and size of oligomers is evident as one proceeds to lower pH values. At pH 11.4 and 11.0 the oligomer growth rates appear not much different compared to pH 7. It is clear from the plot that significant oligomer formation occurs only if pH is 11.8 or higher.

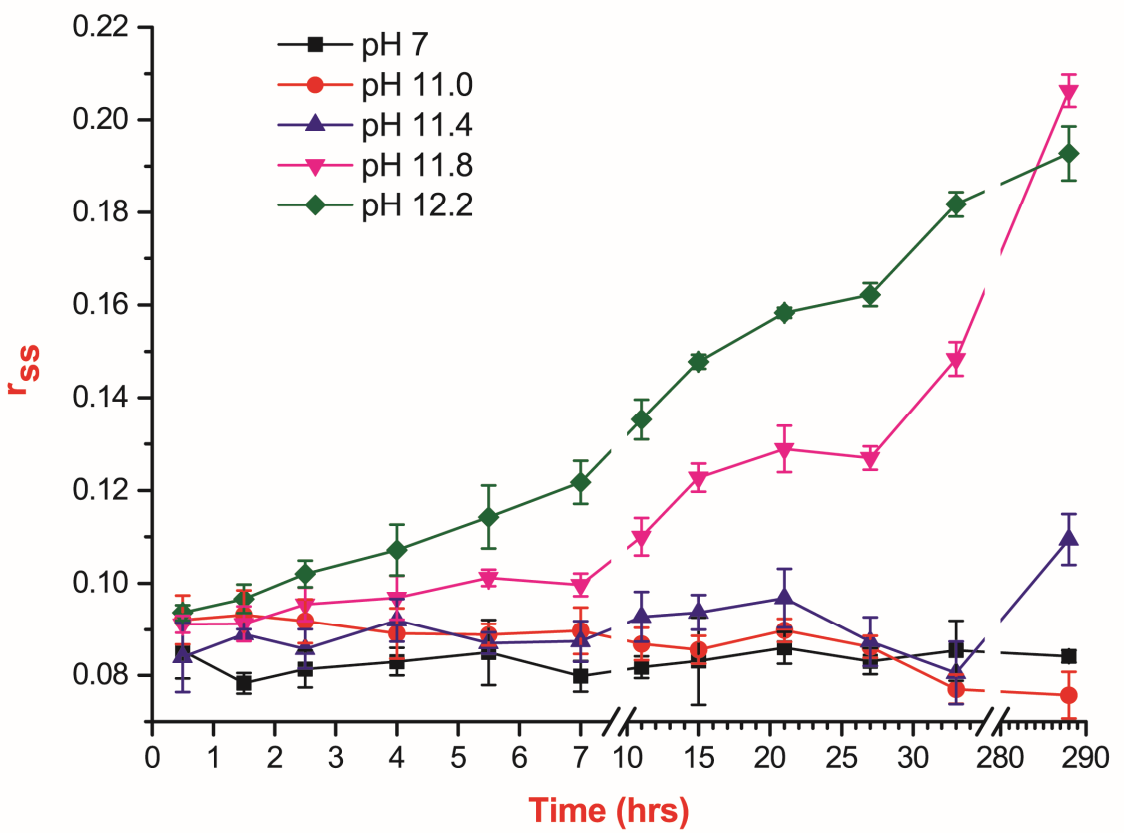


Figure-4.3: Steady state anisotropy of dansyl probe conjugated HEWL incubated at diff pH at 25 C for 120 µM HEWL (118 µM unlabeled HEWL + 2 µM Dansyl labelled HEWL). Ex 380 nm, Em 444 nm. Slits were 1 nm (excitation) and 4 nm (emission).

4.1.5 Effect of pH on Thioflavin T fluorescence (measured after 10 days)

The amyloid population in incubated samples of HEWL and human lysozyme was measured for different pH populations of incubation medium. Figure 4.1 shows the gradual rise and subsequent decline of Thioflavin-T fluorescence as one moves from pH 7.0 to 12.8. Interestingly the trend observed for HEWL nearly mirrors the trend seen with human lysozyme. This suggests that both these proteins probably follow a similar mechanism leading to amyloid formation. Since Thioflavin-T measures that the last product of kinetic

pathway (amyloid), so the similarity between human lysozyme and HEWL at this level justifies the use of HEWL instead of human Lysozyme in this thesis.

4.1.6 Effect of sample temperature on HEWL REA

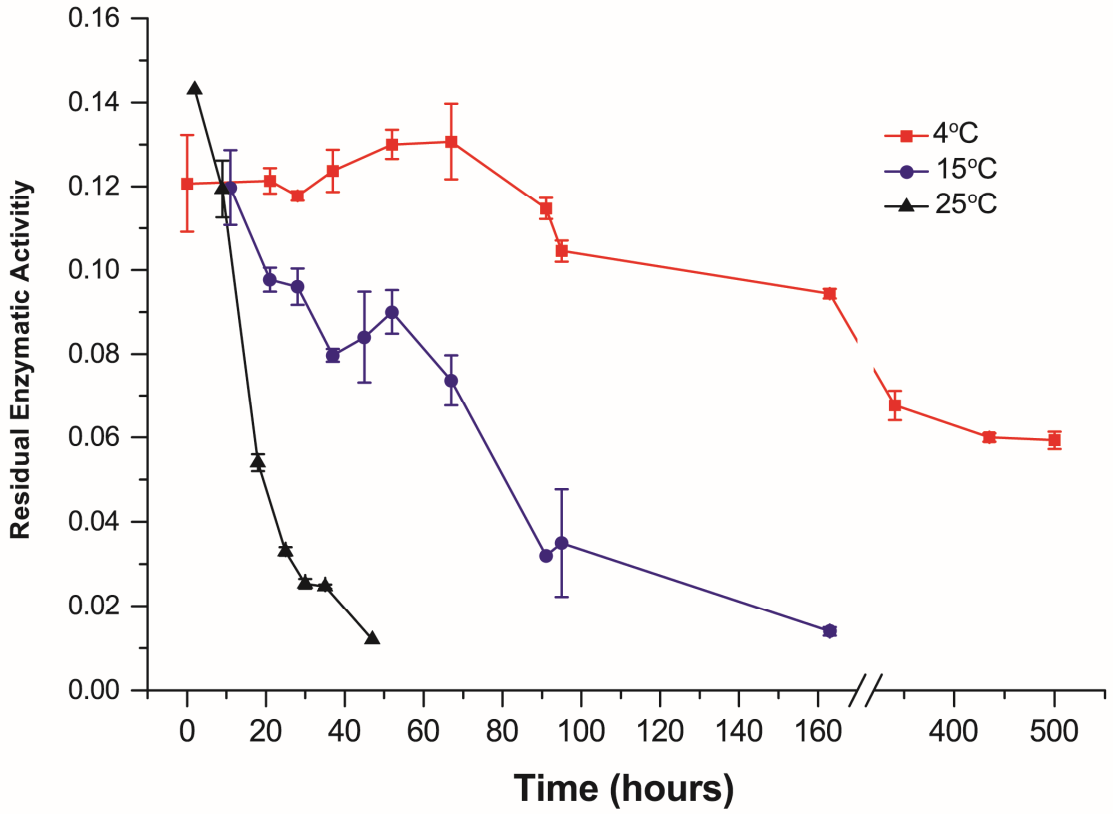


Figure-4.4: REA (slope of first 30 s) of HEWL 120 µM incubated in pH 12.2 at different temperatures at different time intervals.

Figure 4.4 shows the effect of sample incubation temperature on the activity of HEWL exposed to pH 12.2. As expected, a higher temperature like 25°C has a quick and dramatic effect in abolishing activity. The rate of decrease in activity is slowed considerably as the temperature is lowered to 15°C and even more slowed at 4°C. This effect is most likely due to reduced rate of unfolding at lower temperatures (see later). It

can be inferred from the figure that T_m (melting temperature of protein) at pH 12.2 is between 4-15 degree Celsius.

4.1.7 Effect of sample temperature on r_{ss} of dansyl-labelled HEWL

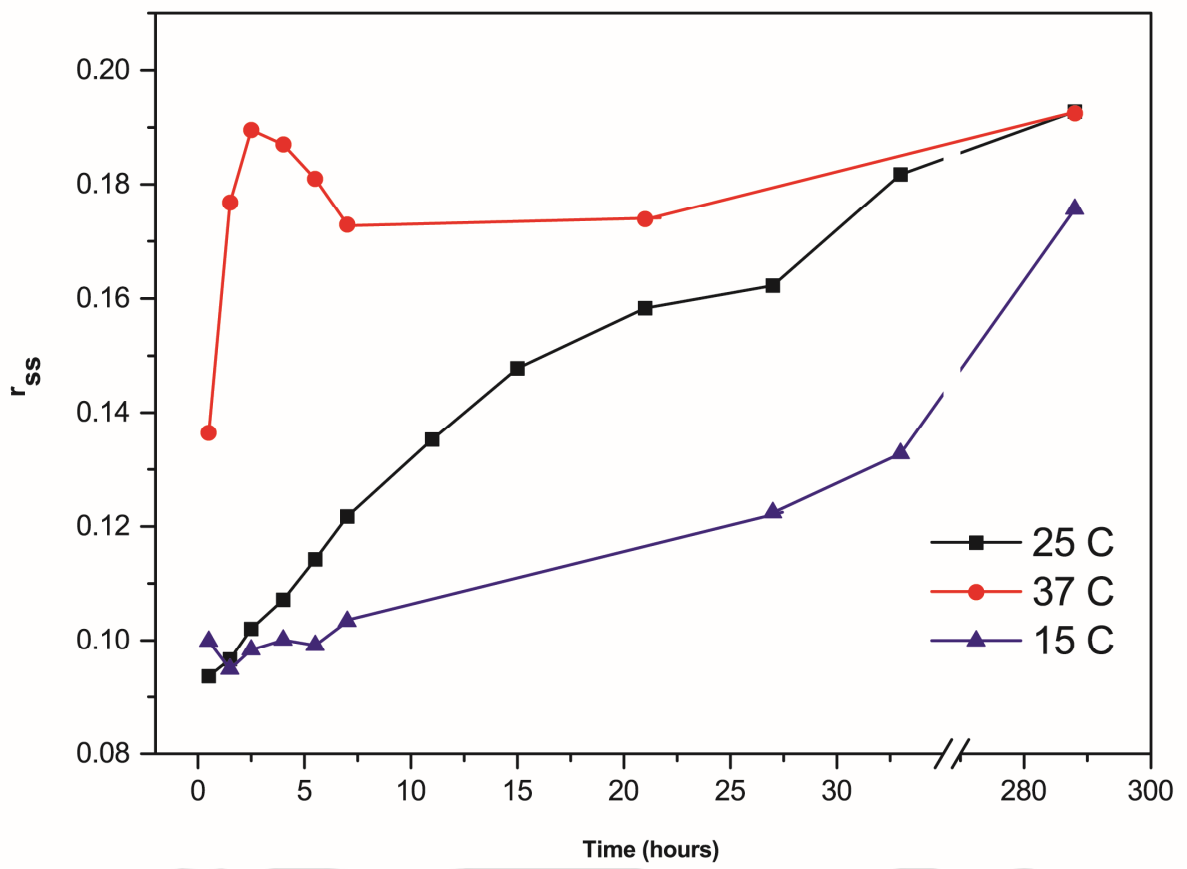


Figure-4.5: Steady state anisotropy of Dansyl conjugated HEWL (118 + 2) incubated at pH 12.2 at different temps for diff time intervals. Ex 380 nm, Em 444 nm Slits for excitation was 1 nm, while that for emission was 4 nm.

Figure 4.5 shows the influence of sample incubation temperature on the growth kinetics of HEWL oligomers. The growth rate is quite rapid at 37°C. At room temperature (25°C) the rate appears moderate, while at lower temperature (15°C) it is indeed slow to begin with but appears to accelerate at later times (well beyond 33 hours). Interestingly at long duration, HEWL incubated at all three temperatures appear to converge on the

oligomer size. These experiments clearly demonstrate that HEWL oligomer formation at pH 12.2 is strongly temperature dependent.

4.1.8 Effect of sample temperature on Thioflavin-T fluorescence

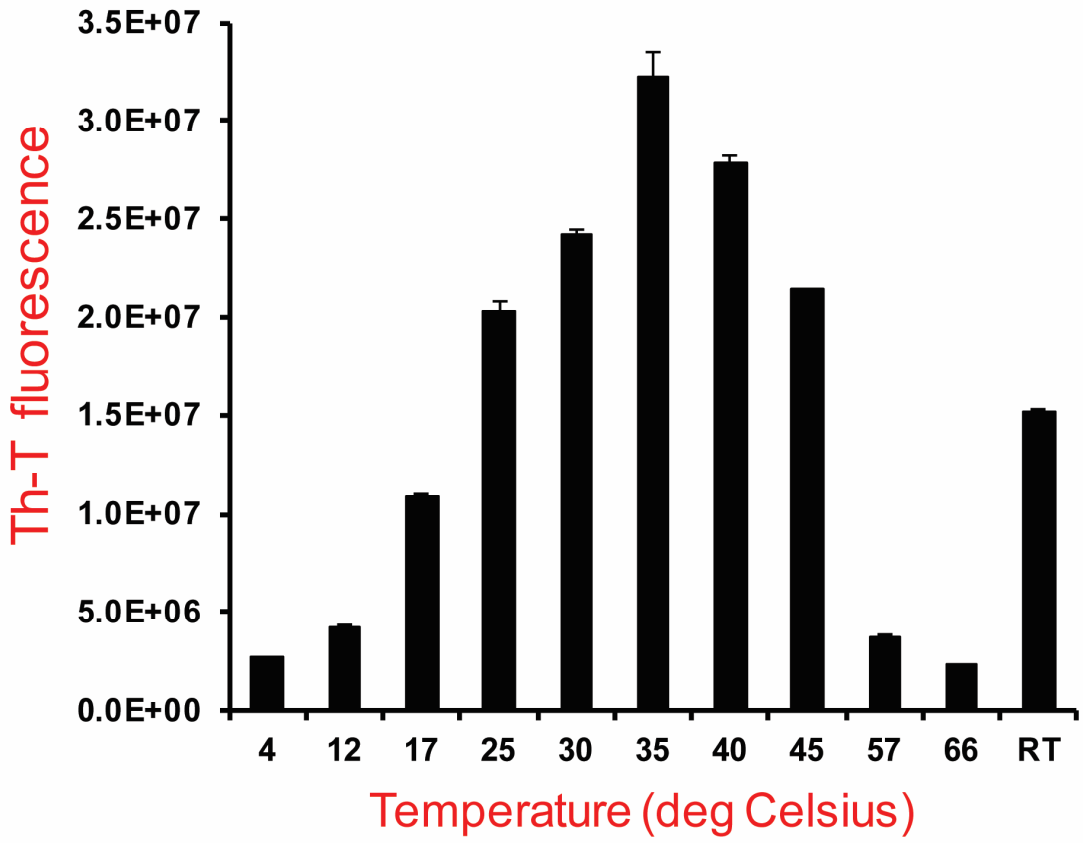


Figure-4.6: ThT fluorescence of HEWL (120 µM) incubated at pH 12.2 for 96 hours at different indicated temperatures. RT refers to room temperature.

In Figure 4.6, effect of HEWL sample incubation temperature on Thioflavin-T fluorescence measured after 4 days is shown. A bell shaped profile is noticeable. Maximum Thioflavin-T fluorescence was observed at 35°C. At higher temperatures, a decrease is observed as we move from 40°C to 66°C. This implies that while mild increase in

temperature favours aggregation and amyloid formation, a further increase probably destabilizes the oligomers and subsequently diminishes aggregation and amyloid formation. This experiment needs to be repeated with more incubation times rather than singular 96-hours. That way we can find whether there are two processes of aggregation and dis-aggregation are working at higher pH or not. That way we can also find the condition which gives rise to maximum amyloid and that can be projected as Standatd Condition to make Lysozyme amyloid.

4.1.9 Effect of HEWL concentration on REA

In Figure 4.7, the decrease in enzymatic activity of HEWL is plotted for different concentrations of the protein in the alkaline medium (pH 12.2), along with control (120 μM, pH 7.0). In comparison to control, all samples reveal a steady decrease in catalytic rate with time. Interestingly, the trend exhibited by different HEWL concentrations are nearly similar or superimposable with each other. Such a concentration independent phenomenon which diminishes the activity is likely to be denaturation of the protein. This experiment is hitherto the only published data arising out of this thesis [11].

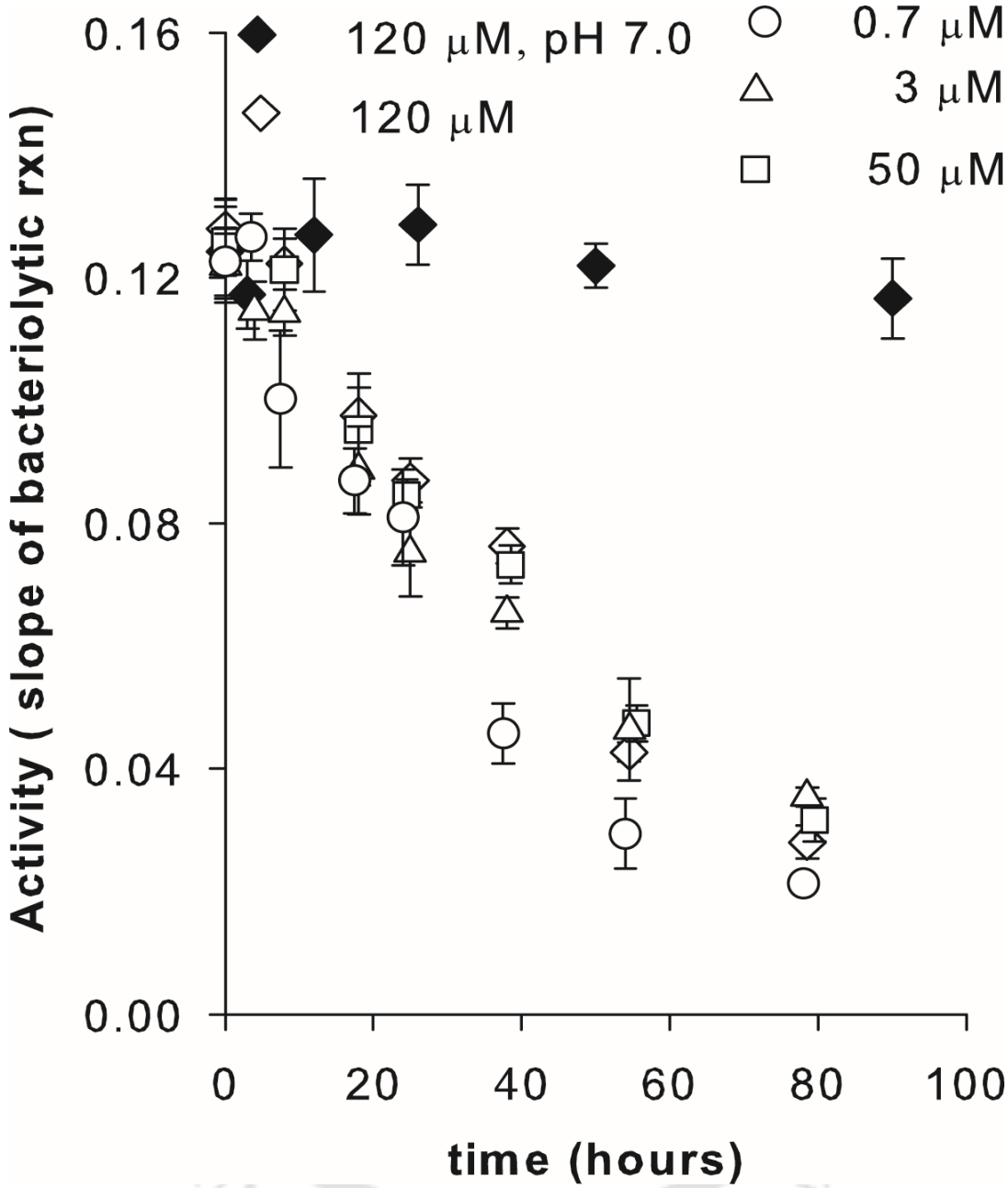


Figure-4.7: REA of HEWL incubated at different pH (7.0 and 12.2) and different concentrations as indicated in figure.

4.1.10 Effect of seeding

4.1.10 A. Effect of seeding observed using r_{ss} of dansyl-labelled HEWL

In order to observe the effect of seeding on HEWL oligomer growth rate, wherein preformed aggregates of HEWL are added to freshly incubated HEWL samples in pH 12.2, the r_{ss} of dansyl-labelled HEWL was measured. Figure 4.8 shows the r_{ss} profile for both fresh sample and fresh sample with seeding. No significant difference in growth kinetics is observed, suggesting that seeding does not influence the growth kinetics of HEWL oligomers.

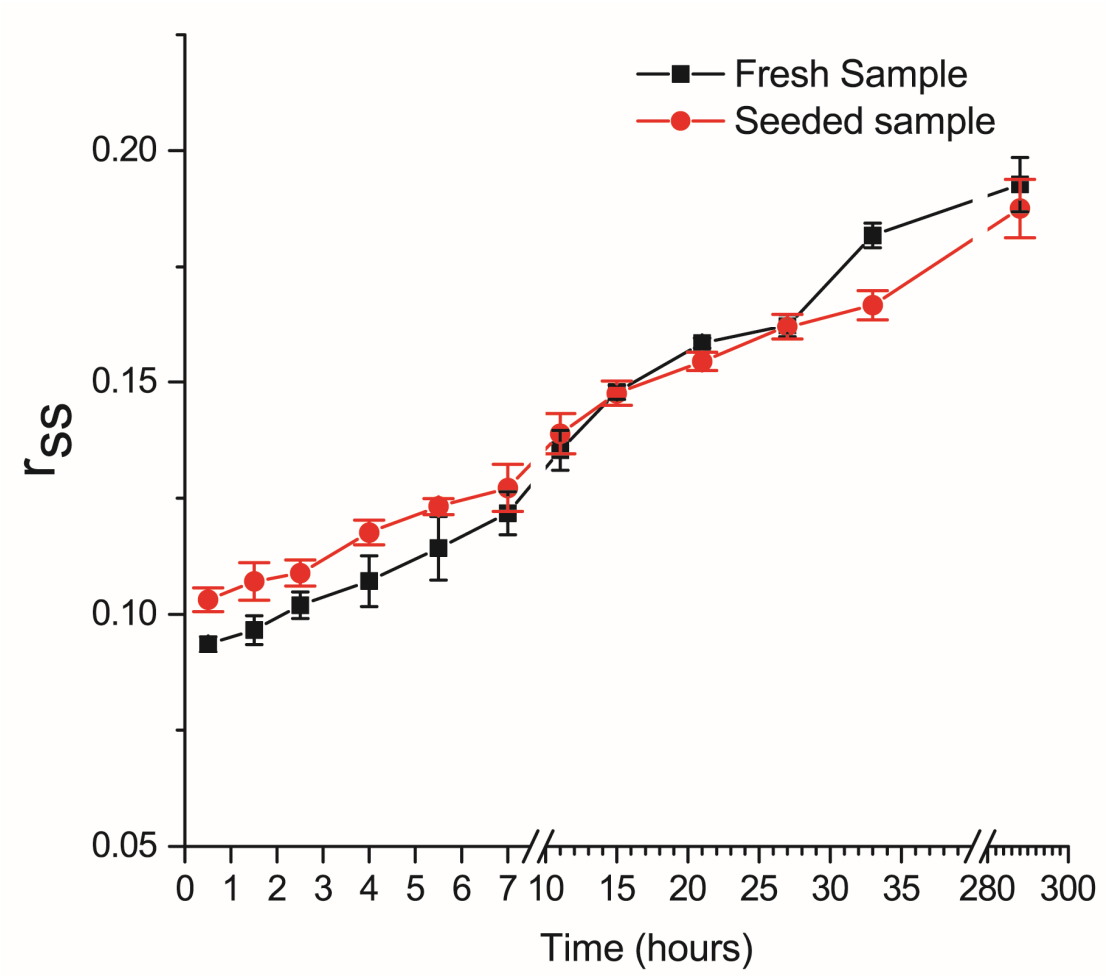


Figure-4.8: Steady state anisotropy of dansyl-conjugated HEWL incubated at pH 12.2 in 25°C for different times as indicated. Data is shown for freshly prepared sample (118 µM unlabeled HEWL + 2 µM dansyl-conjugated HEWL) and sample seeded with 12 µM (remaining 106 µM unlabeled with 2 µM dansyl-labelled of fresh sample) of 30-days old unlabeled sample of 120 µM HEWL in pH 12.2 at 25°C.

4.1.10 B. Effect of seeding on Thioflavin T fluorescence

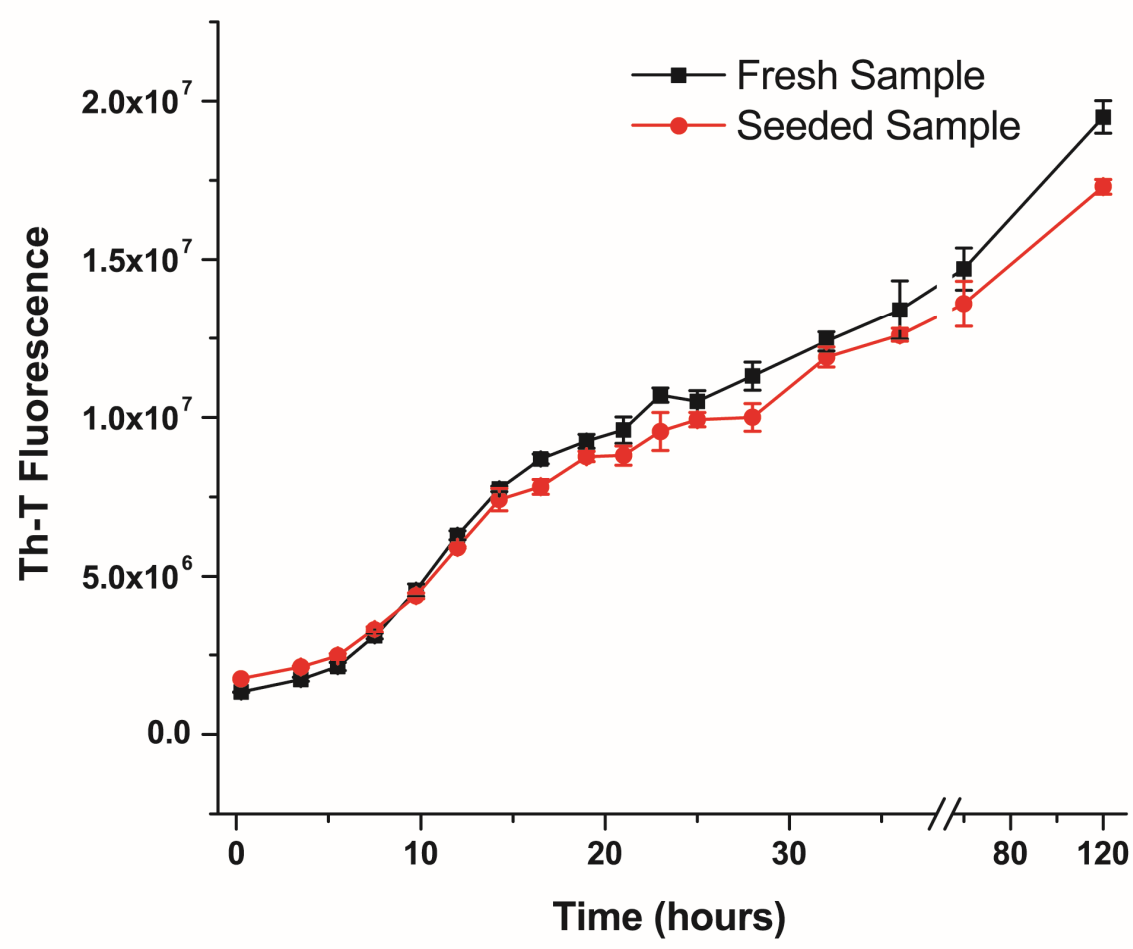


Figure-4.9: ThT fluorescence of fresh HEWL (120 µM) and seeded HEWL (12 µM seed of 30 days old sample + 108 µM fresh sample) at pH 12.2 at 25°C.

In a similar experiment as detailed above, effect of seeding was investigated by measuring the Thioflavin-T fluorescence in HEWL samples. As evident from Figure 4.9, no appreciable change in the Thioflavin-T fluorescence is seen between normal and seeded sample. This again suggests that seeding has no role in influencing the growth of amyloid aggregates in HEWL at pH 12.2.

4.1.10 C. Conclusions & Prognosis of Seeding Experiments

Current paradigm is that seeding promotes amyloidogenesis in general. However, as the following set of experiments [4.1.10 & 4.1.11] suggests; same is not the case at alkaline pH. In the following experiments 1-month old seeds were used. However as lateral addition of Dansyl [Please refer to section 4.1.13] suggests that these one month seeds have been so much mutilated that they have lost any adhesive tendency. Therefore, it is desirable to repeat these seeding experiments with fresh seeds no more than a week old. We favour a model as per which oligomerization depends on concentration of monomer simply because of 'Order' of the process; more monomer concentration means that for a given monomer chances of finding complementary surface is more but seeding has no role to play in oligomerization [and in that sense it is isodesmic]. However structural transition to amyloid is a slower process akin to beginning of formation of helix (where a helical twist has to begin). But once such a process has begun it will proceed quickly. Seeds therefore provide species with initial twist (nucleation) upon which further growth is accelerated. That is why amyloid growth is often seeded and actually even cross-seeded because eventually same cross beta-spine twist is found in all amyloids. In present study we can see this. Whereas oligomerization has shot up right from the beginning, and there is significant oligomerization within 6-hours; on the contrary there is very little amyloid in first 6-hours but once it begins, it goes on to build up rapidly.

4.1.11 Dimerization of HEWL at different pH observed using r_{ss} of dansyl-labelled HEWL

Figure 4.10 reveals the steady state fluorescence anisotropy of dansyl-labelled HEWL incubated at different pH conditions. It is noticeable that proceeding from pH 2, r_{ss}

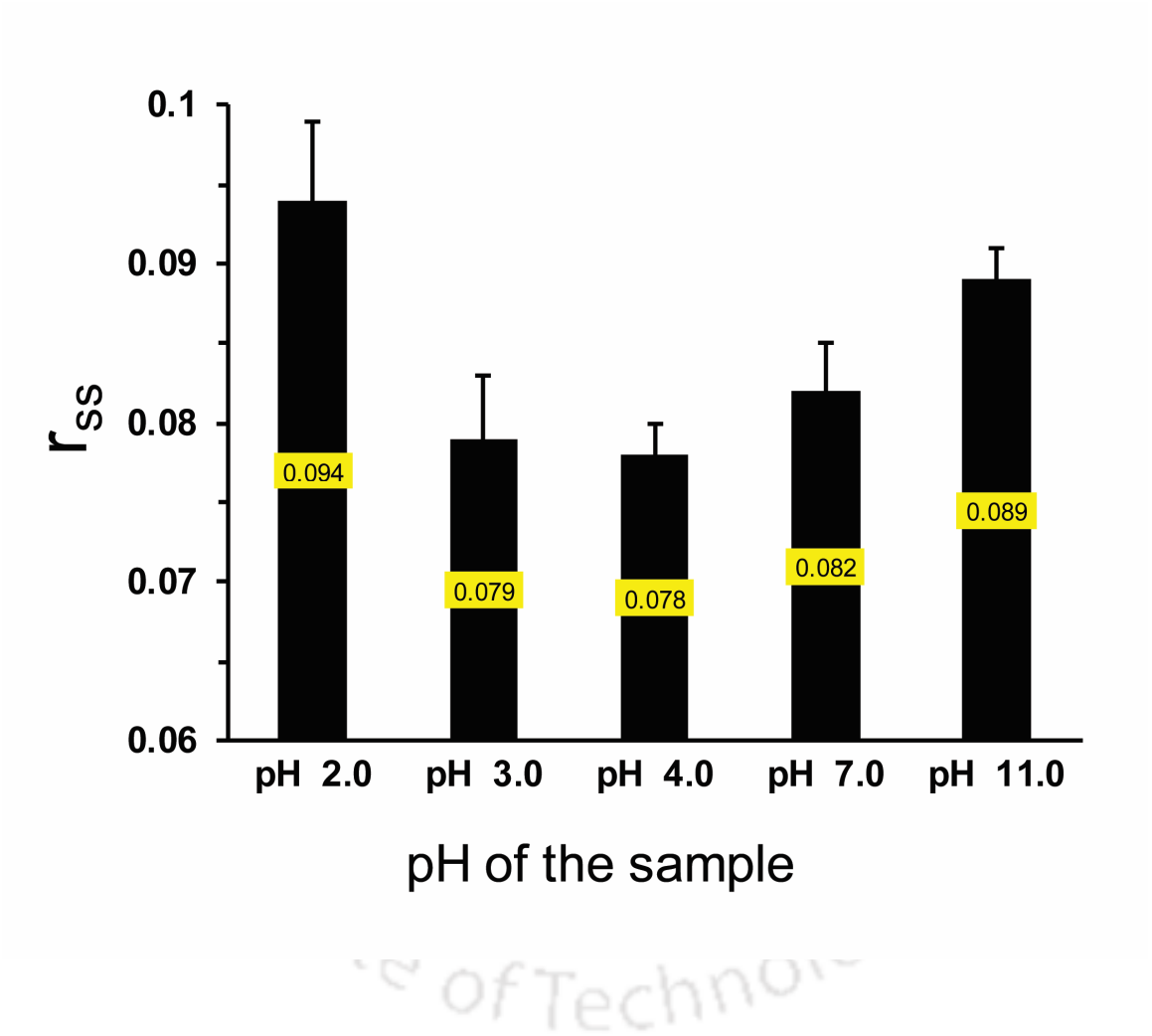


Figure-4.10: Steady state anisotropy of dansyl probe conjugated HEWL (118 µM unlabeled HEWL and 2 µM dansyl-conjugated HEWL) incubated at indicated pH for 15 minutes' duration. r_{ss} values are indicated in the bars.

reaches a minimum at pH 4.0, where it is likely to be monomeric. Thereafter, it shows a tendency to rise owing to formation of oligomers. This shows that HEWL is most monomeric between pH 3.0 and 4.0 and partially dimeric at pH 7.0 which is an important

result. Presence of HEWL dimers at alkaline pH has been previously reported by Sophianopoulos and co-workers [12,13].

4.1.12 Effect of temperature jump on Thioflavin T fluorescence

Temperature jump experiments are commonly done in Physical Organic Chemistry. For this experiment, initially 5 samples were made. 4°C in duplicate, 15°C in duplicate and 25°C single, each at pH 12.2 and incubated for 48 hours. Subsequently, temperature of one 4°C and one 15°C were enhanced to 25°C and further incubated for 48 hours. The remaining three samples remained at indicated temperatures, namely, 4°C, 15°C and 25°C for further 48 hours. ThT of all the samples was recorded after 96 hours.

Figure 4.11 shows the influence of temperature on Thioflavin-T fluorescence arising from presence of HEWL oligomers at pH 12.2. It was observed that after 48 hours in 4°C, a jump to 25°C elicits a strong fluorescence nearly equivalent to that observed with incubation at 25°C all along. A similar jump from 15°C to 25°C triggers only a moderate increase in Thioflavin-T fluorescence. Perhaps this suggests that while samples at 4°C behaves as un-aggregated fresh HEWL samples, those at 15°C may have undergone irreversible structural/conformational alterations that prevent them from acquiring fibrillar structure that is similar to samples at 25°C. This experiment suggests that at 4°C sample is nearly frozen but retained its ability to fully aggregate when allowed to reach 25°C deg. Whereas, 15°C sample, was essentially dead and did not pick up when the opportunity was presented by enhancing the temperature.

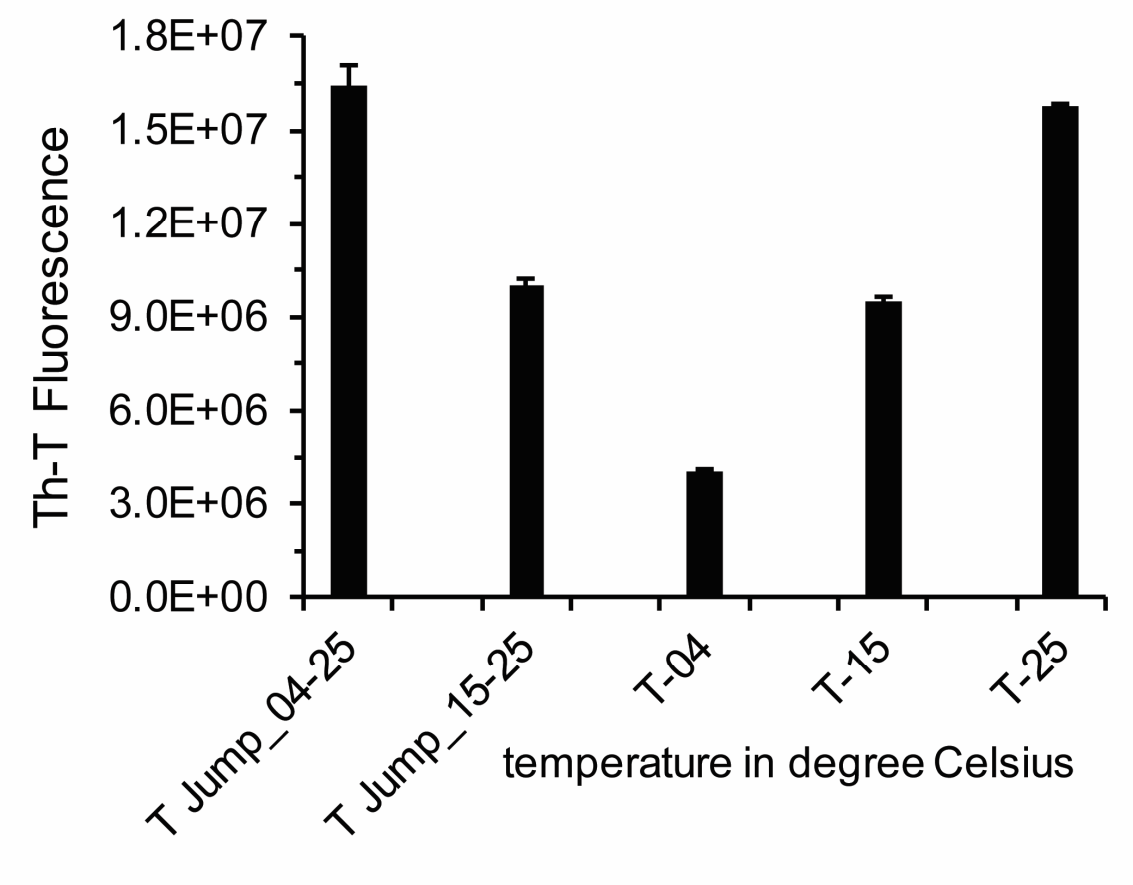


Figure-4.11: HEWL 120 µM incubated at 4°C or 15°C for 48 hours. Subsequently temperature was changed to 25°C for 48 hours. Then ThT fluorescence was recorded. For T-04, T-15 and T-25, the temperature remained constant over 96 hours' period.

4.1.13 Lateral addition of dansyl-labelled HEWL

This experiment aims to determine if labelled fresh HEWL added at later times during aggregation gets incorporated in and if they do get in, with what kinetics.

4.1.13.1 How this experiment was carried out?

First a reaction with unlabelled HEWL [all 120 µM unlabelled HEWL at pH 12.2] was done for indicated [as indicated in figure legend] number of hours. Now an aliquot of

it was taken out and to that was added 2 µM [final concentration] Dansyl-labelled HEWL. Then r_{ss} kinetics of this new reaction mixture was carried out. It is this kinetics [with differently aged unlabelled HEWL] which is plotted in the following figure.

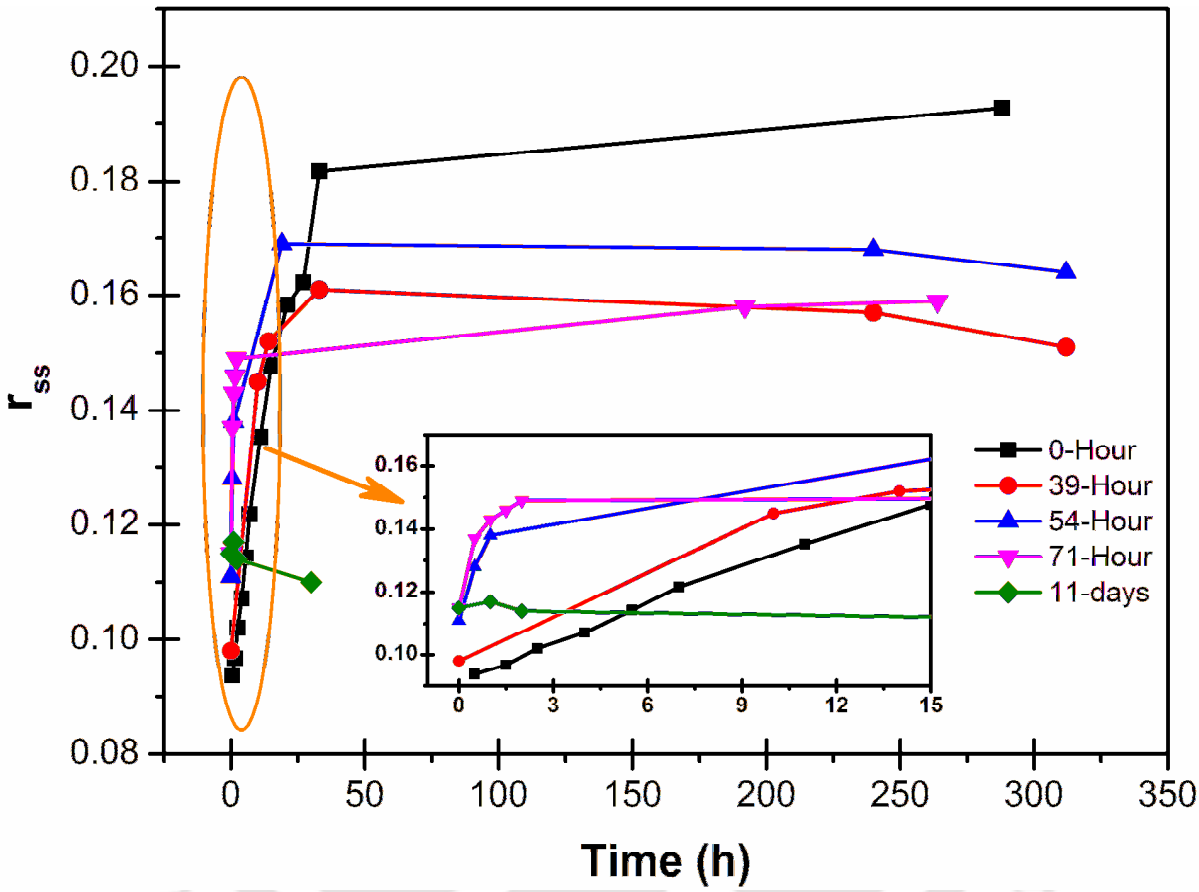


Figure-4.12: **Lateral addition of dansyl-labelled HEWL**

4.1.13.2 Conclusions from this experiment:

Zero point r_{ss} with 71-hour sample is close to 0.12 so here we are talking of dansyl-HEWL incorporation rapidly with a substantial size oligomer. Here two points are inherent.

a. Average oligomer size is substantial. Actually this could be alternate way [and more precise] to tell how big is average oligomer at different time points as compared to its cumulative r_{ss}.

b. Incorporation of Dansyl-HEWL monomer is rapid. In a way what we are talking is 'order of reaction' between [polymer strand] and [monomer] (brackets refer to their concentrations).

The r_{ss} builds up rapidly with 71-hour sample reaching 0.15 in mere 2:00 hours something which a zero-hour sample takes 15-hours to reach.

The 11-days old sample [oligomer and/or amyloid] is dead. It does not incorporate dansyl conjugated HEWL anymore. Could it be that 11-days old is mostly amyloid and less of unstructured oligomer and it could be that amyloid is kind of an irreversible zipping process as a result of which:

a. All/Most chemical groups get masked
b. There is lateral self-association of individual threads into a fibre after which dissociation and re-association process of threads [from either terminal of thread or from even between part of the thread] is fully stopped. It is expected that initiation of beta zipper is slow process but once it starts, it proceeds rapidly and that explains seeding experiments in literature. It is reasoned that oligomerization is necessarily an isodesmic process whereas conformational reorganization to cross beta spine conformation could be catalytic

Other 'hour' old samples show intermediate behaviour between '0' and '71'.

4.2 Ethanol Condition

4.2.1 REA of 120 µM HEWL incubated at indicated times and indicated concentrations of ETHANOL in water (v/v).

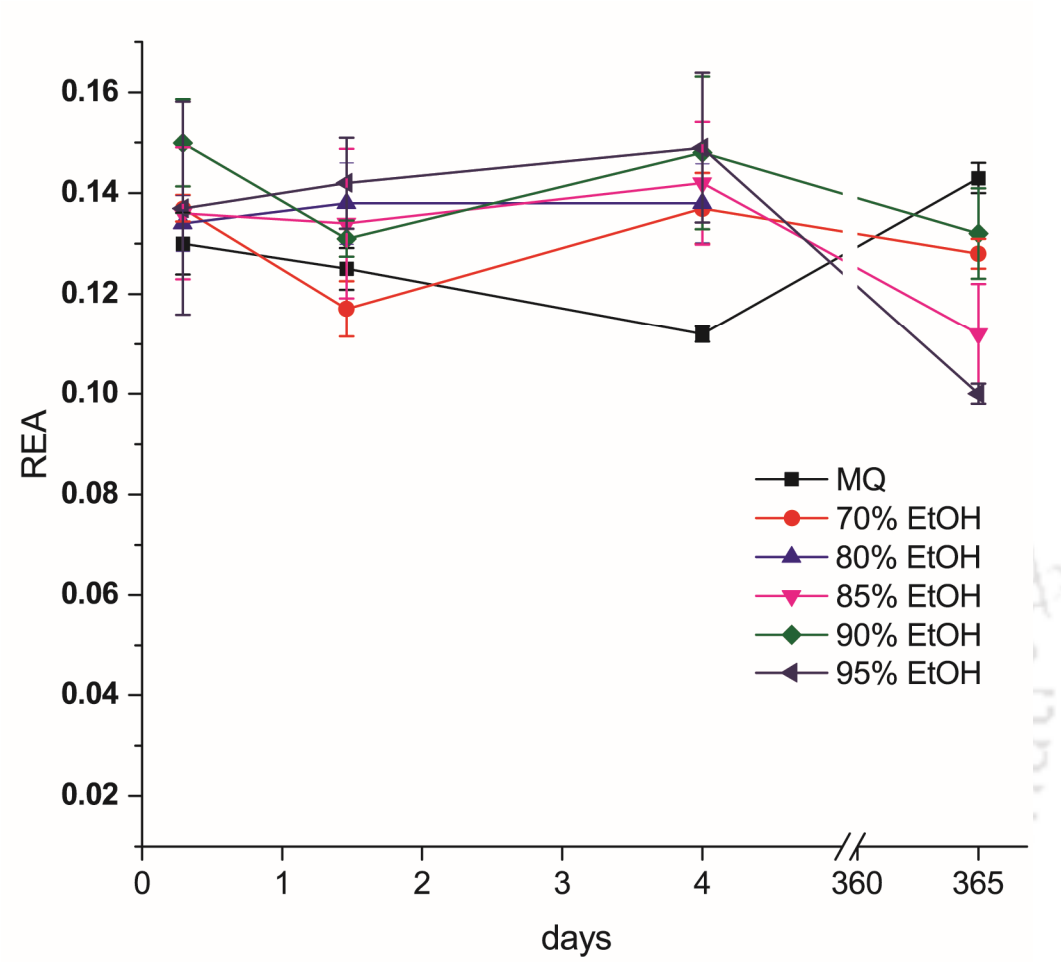


Figure-4.13: REA of 120 µM HEWL incubated at indicated concentration of ethanol (v/v) in water at indicated number of days. MQ refers to deionized water.

Figure-4.13 shows the influence of different concentrations of ethanol on the enzymatic activity of HEWL (120 µM). By and large the activity shows no major decline at any condition even after a period of one year. The 80% sample after one year was lost and so could not be recorded. This implies that structure and active site conformation of HEWL is not significantly altered in prolonged presence of ethanol.

4.2.2 Steady State Anisotropy of dansyl-labelled HEWL in presence of Ethanol

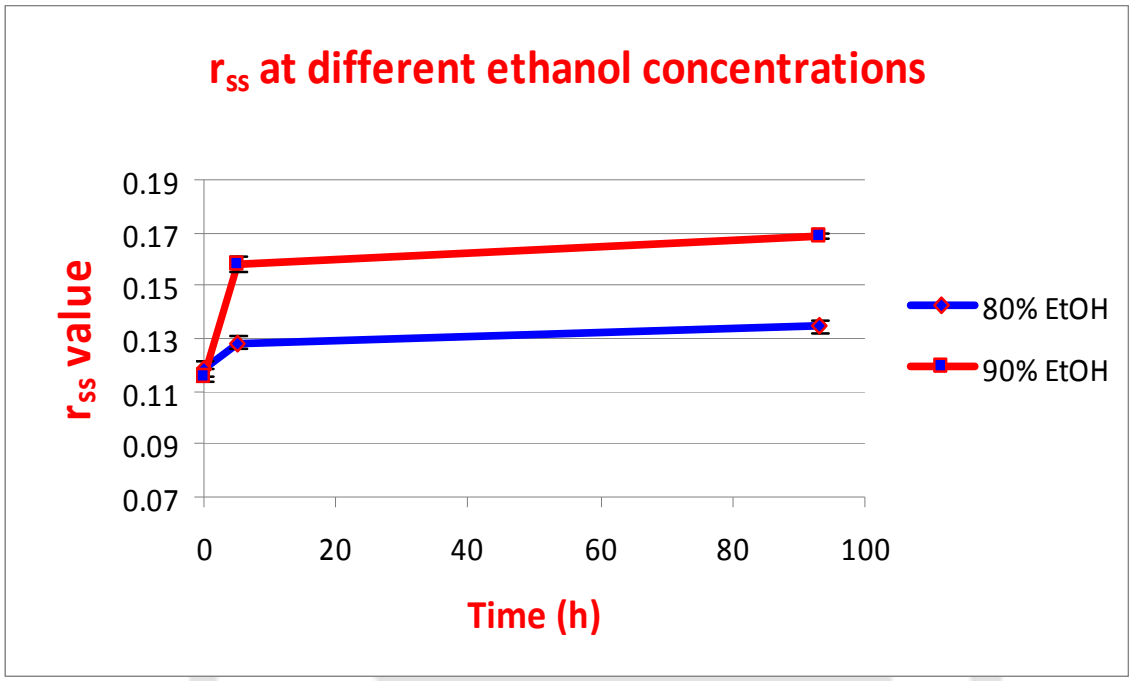


Figure-4.14: Steady state anisotropy of ethanol condition

Figure-4.14 shows that r_{ss} of dansyl-labelled HEWL in presence of 90% ethanol /water (V/V) shows a significant rise initially followed by a plateau-like profile. The effect is less pronounced at 80% ethanol/water (V/V) suggesting a moderate influence on oligomer formation at this condition. This is consistent with the nearly complete activity observed under this condition (see Figure-4.13). Life time of Dansyl at 80% & 90% Ethanol is yet to be measured.

4.2.3 Light scattering from HEWL samples in presence of Ethanol

Figure-4.15 reveals the extent of static light scattering [3] from 6 days and one-year-old samples of HEWL incubated in different concentrations of ethanol. With exception of 95% ethanol sample, all other samples show low scattering intensity implying minimal presence of large oligomers. One year old sample reveals maximum OD at 350 nm for 95% sample and a slightly less OD for 70% sample. Other data reveal negligible population of large oligomers. Clearly presence of 95% ethanol has a significant influence in promoting large oligomer formation. A typical static light scattering experiment is done as described in the following figure.

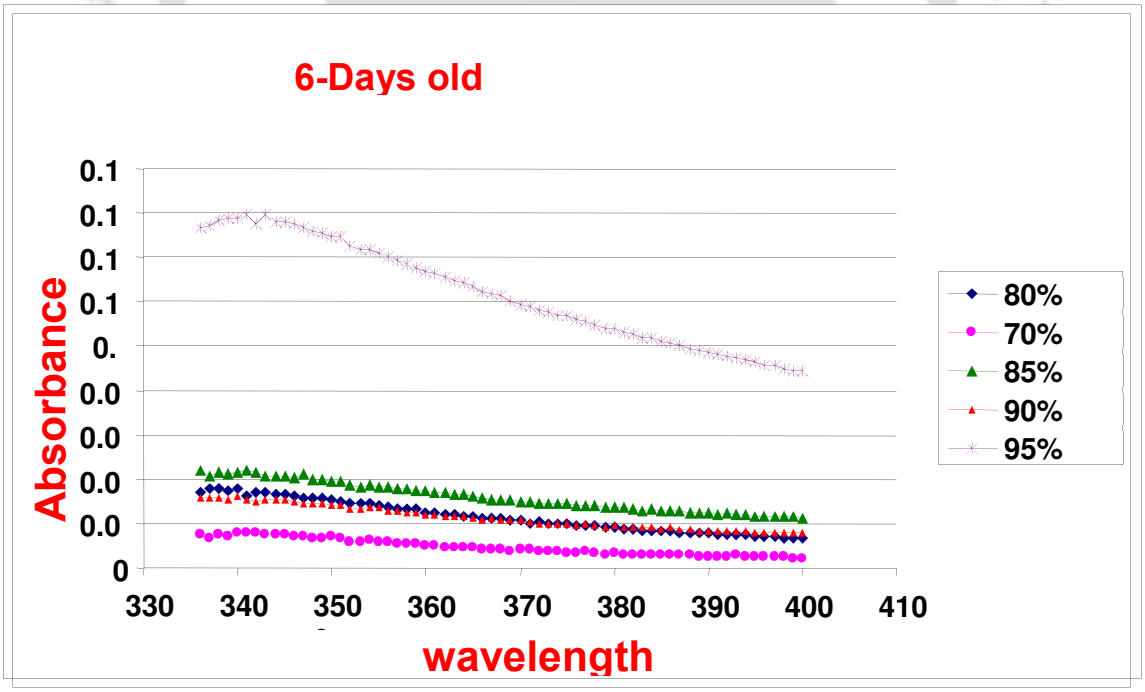


Figure-4.15 Static Light Scattering with different ethanol concentrations

A. Light scattering (measured at 350 nm) with HEWL incubated in ETHANOL for One Year

%	Abs (a.u.)
MQ	0.04
70%	0.13
80%	0.04
85%	0.02
90%	0.02
95%	0.15

B. Light scattering (measured at 350 nm) with HEWL incubated in ETHANOL for 6 days

%	Abs (a.u.)
MQ	0.04
70%	0.01
80%	0.03
85%	0.04
90%	0.03
95%	0.15

Table-4.1: Static light scattering in UV-Vis spectrophotometer measured at 350 nm wavelength for 120 µM HEWL at indicated Ethanol concentrations in water (v/v). Table A shows data for 6 days and Table B shows data for 1-year-old sample incubated at room temperature.

4.2.4 Thioflavin-T of HEWL samples incubated at Ethanol condition

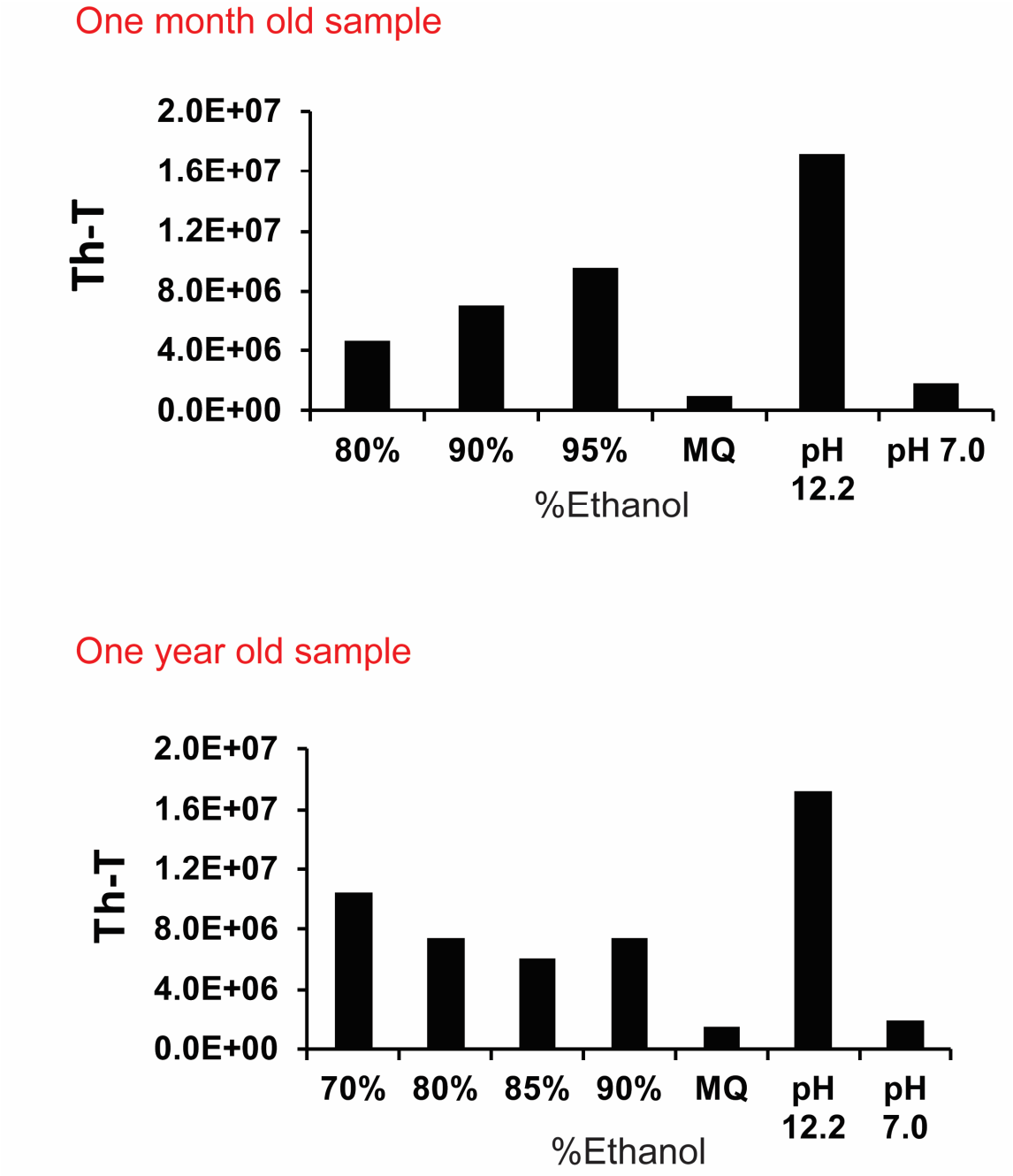


Figure-4.16: **120 µM HEWL incubated at indicated concentrations of ethanol in water (V/V) for 1-month old sample (top) and 1-year-old sample (bottom). 12 µM HEWL with 10 µM ThT in pH 7 buffer. pH 7 and 12.2 samples were not from this experiment but from a separate experiment, but shown here for comparison as to where ethanol condition stands vis-à-vis pH 12.2 in amyloid formation.**

Figure 4.16 top panel shows only a moderate fluorescence intensity from Thioflavin-T samples of HEWL incubated at 80, 90 and 95% ethanol in comparison to pH 12.2. Interestingly after a duration of one year (Figure-4.16 bottom panel), 70% ethanol sample along with 80 and 85% begin to show more Thioflavin-T fluorescence which is comparable to that seen with 90% ethanol. This suggests that ethanol can exert its effects at a much later stage also. This in a way is directly relevant to physiological situations where there is mild perturbation to protein milieu and symptoms appear late in life [for patients of lysozyme amyloid, they appear after fifth decade of life]. Given a time of one year, even 70% goes on to form amyloid. This has significance that apparently even minutely perturbing condition can form amyloid over physiological time scale. Goda et al. (Goda,2000) worked with precipitates but in our case all are soluble except 95% case. Thus we have succeeded in producing soluble amyloid working with which has its own advantages. Even for ethanol condition (like pH 12.2 condition), oligomerization is very rapid, thus it appears that oligomerization is isodesmic even in this case and could actually be the paradigm (across all amyloidogenic conditions) that oligomerization is isodesmic and seeding independent whereas seeding plays a role only (seeds) in the cross beta-spine conformational transition. Moreover, Amyloid formation takes place despite no apparent loss in protein activity even after one year. Thus we have a case of amyloid with fully active protein. This poses difficult questions regarding conformational transformations. Goda et al. (Goda,2000) mentions 50% activity loss in a week [they had salt in it], but here for a year no loss in protein activity. So our salt free condition is less harsh, gives rise to soluble amyloid so perhaps a better version for future replications of ethanol condition. Why not use methanol [that is another variation we need to attempt]?

4.3 Acidic pH:

Only Residual Enzymatic Activity (REA) and steady state anisotropy (r_{ss}) experiments were done for this condition because they have not been done in literature. Thioflavin-T experiments have already been done ample number of times in literature [1,2,9,14] for this condition which has been extensively studied in literature and has nearly swamped all other amyloidogenic conditions. We don't know how stable is sulphonamide bond of Dansyl-conjugated Lysozyme at acidic pH of 2.0 especially when the sample is heated to 57^{o}C. Unfortunately, steady state fluorescence anisotropy experiment is very preliminary and there are not enough number of data points. The reason for that is we did not have heating accessory for fluorimeter. Since r_{ss} value depends a lot on temperature. Neither did we measure life time of Dansyl at pH 2.0. We uniformly (across all data points) took out sample from hot air oven (57^{o}C), let it stay at room temperature for 30-min and then measured r_{ss}. Thus there is uniformity across all data points as far as temperature of the sample is concerned.

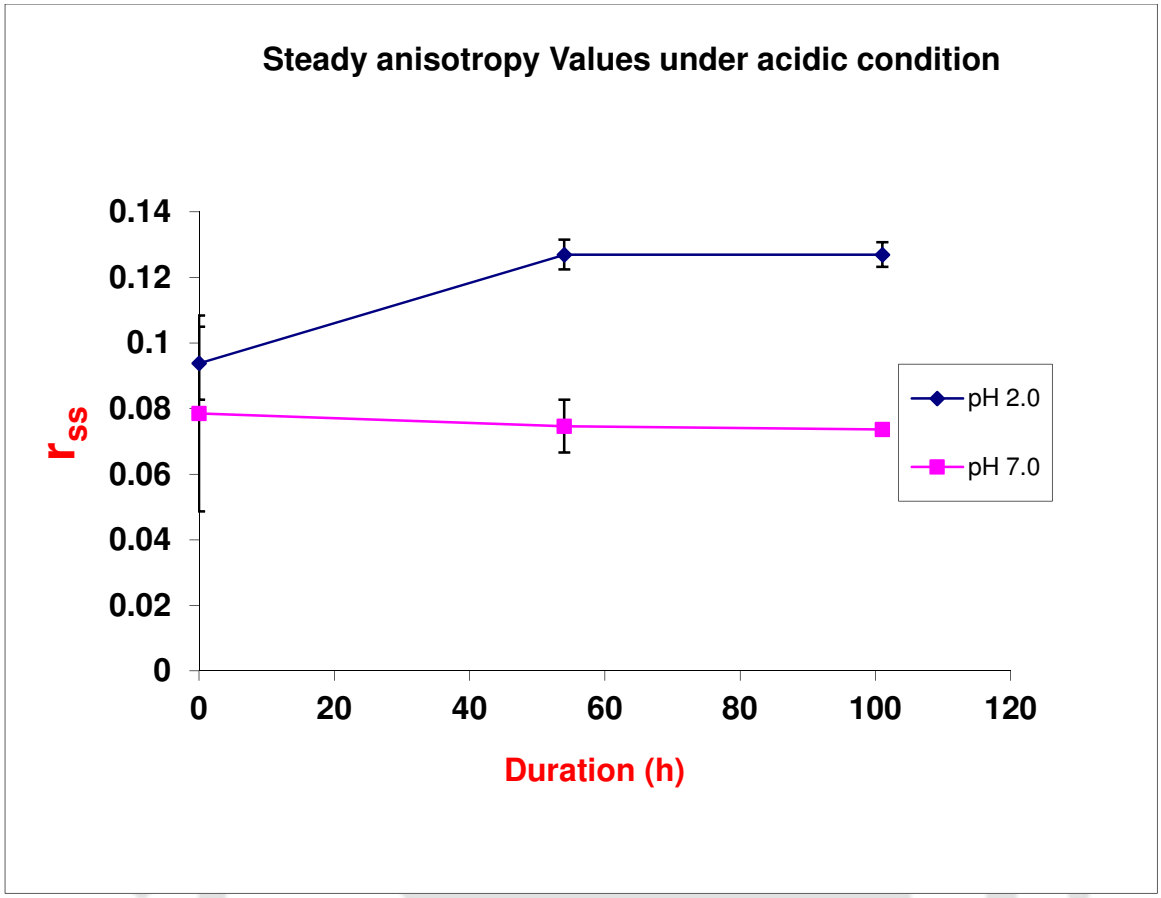


Figure-4.17: **Plot of steady state anisotropy value at different times (HEWL concentration 600 µM, pH 2.0 in Glycine buffer, incubated at 57°C but cooled at room temperature for 30-min before recording ss-anisotropy because ss-anisotropy is strongly temperature dependent.**

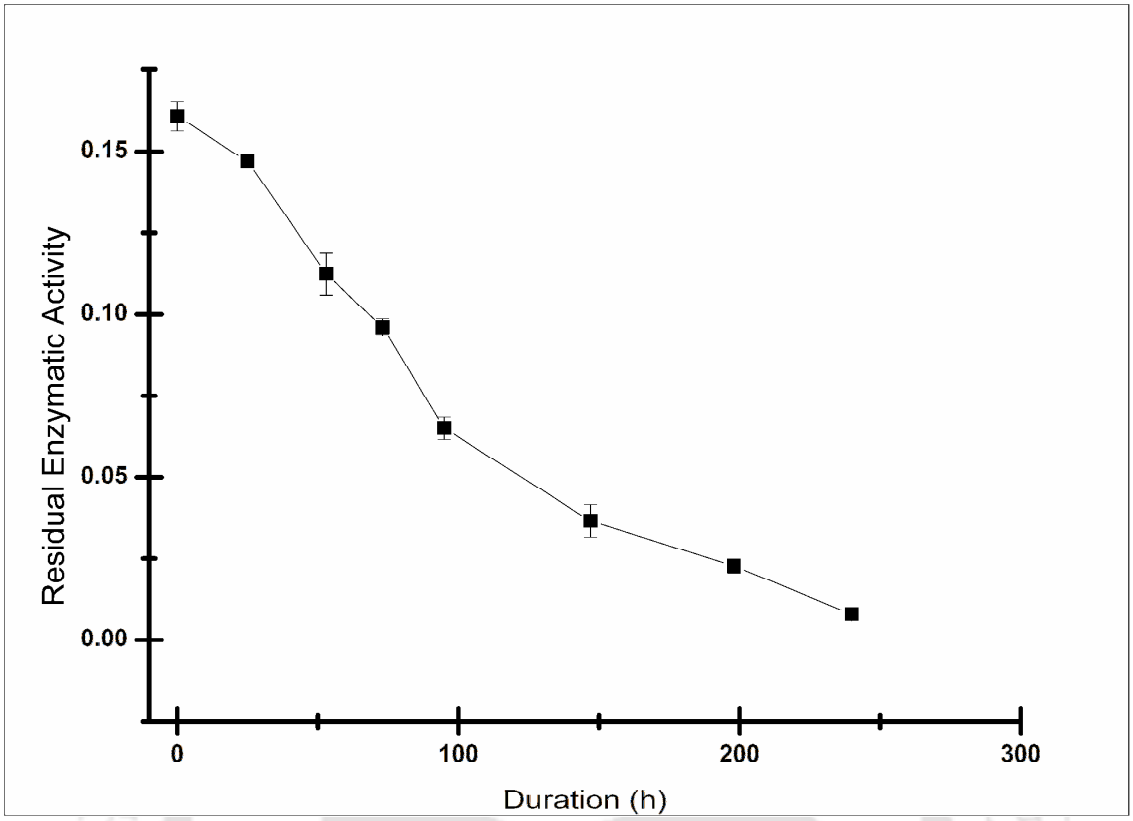


Figure 4.18: REA of 600 µM HEWL incubated at pH 2.0 at 57°C

Clearly enzyme activity falls off much slowly at pH 2.0/57°C as compared to pH 12.2 at 25°C, so we can possibly say that acidic condition is less detrimental as far as conformation of folded protein is concerned. Corollary to this observation could be that melting temperature (T_m) of HEWL at alkaline pH is less than 25°C but more than 4°C where protein is fully stable even after weeks.

4.4 Guanidinium Condition:

Guanidinium Hydrochloride 6-M at 25°C is known to fully denature the protein to the extent of causing fully random coil conformation and is used as such before measurements such as absorbance at 280 nm. Moreover, Guanidinium Hydrochloride is often used as

reversible unfolding of proteins and measurement of thermodynamic parameters of protein unfolding. Therefore, this condition is biophysically appealing. So the fact that at 3-M Guanidinium Hydrochloride at 45°C, the HEWL at 120 µM concentration forms amyloid, prompted us to study this condition. Use of residual enzymatic activity for 5-M Guanidinium Hydrochloride led to following table

Time (h)	REA Value
0	0.235± 0.011
0.25 h	0.230 ± 0.021
1.33 h	0.222 ± 0.023
10 h	0.237 ± 0.018
31 h	0.242 ± 0.017
28-days	0

Table 4.2: Residual Enzymatic Assay Kinetics at 5-M Guanidinium Hydrochloride at 45°C for 120 µM HEWL

So clearly the HEWL was still fully folded even after 24-hours at 5-M Guanidinium Hydrochloride at 45°C which was surprising since 6-M Guanidinium Hydrochloride at 25°C is thought to completely unfold the proteins. It needs mention that unlike previous

cases here the final concentration of HEWL in the assay was 120 nM and not 70 nM and hence the high value of zero point REA.

Same experiment with 3-M Guanidinium Hydrochloride led to the following Table.

Time (h)	REA Value
0 h	0.230 ± 0.012
0.33 h	0.219 ± 0.012
0.66 h	0.216 ± 0.010
1 h	0.233 ± 0.019
2 h	0.227 ± 0.008
4 h	0.223 ± 0.005
12 h	0.224 ± 0.010
33 h	0.229 ± 0.005
22-days	0.206 ± 0.015

Table 4.3: Residual Enzymatic Assay Kinetics at 3-M Guanidinium Hydrochloride at 45°C for 120 µM HEWL

So at 3-M Guanidinium Hydrochloride at 45°C, HEWL is still nearly fully folded even after 3-weeks.

Next we did static light scattering at 350 nm and Thioflavin-T measurements and results are mentioned in the following Table

Concentration of Guanidinium	**Scattering Value**	**Thioflavin-T Value**
0-M	0	1.0
3-M	0.01	1.32
5-M	0	1.83

Table 4.4: Static Light Scattering and Thioflavin-T Values at different Guanidinium Hydrochloride Concentrations (for 120 µM HEWL at 45°C)

Clearly there is no scattering whereas in the original paper [7], there were precipitates. Likewise, Th-T values are not great. Then we looked at the original paper. We found that in addition to Guanidinium Hydrochloride and high temperature, they also did stirring [a fact not mentioned prominently]. When we did shaking, we also found precipitates. So mere Guanidinium Hydrochloride and temperature are not enough to make amyloid and simultaneous mechanical agitation is also necessary.

References

1 Morozova-Roche, L. A. *et al.* Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *Journal of structural biology* 130, 339-351, doi:10.1006/jsbi.2000.4264 (**2000**).

2 Krebs, M. R. *et al.* Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the beta-domain. *Journal of molecular biology* 300, 541-549, doi:10.1006/jmbi.2000.3862 (**2000**).

3 Homchaudhuri, L., Kumar, S. & Swaminathan, R. Slow aggregation of lysozyme in alkaline pH monitored in real time employing the fluorescence anisotropy of covalently labelled dansyl probe. *FEBS letters* 580, 2097-2101, doi:10.1016/j.febslet.2006.03.012 (**2006**).

4 Goda, S. *et al.* Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein science* 9, 369-375, doi:10.1110/ps.9.2.369 (**2000**).

5 Yonezawa, Y. *et al.* An Insight into the pathway of the amyloid fibril formation of hen egg white lysozyme obtained from a small-angle X-ray and neutron scattering study. *Journal of molecular biology* 323, 237-251 (**2002**).

6 Fujiwara, S., Matsumoto, F. & Yonezawa, Y. Effects of Salt Concentration on Association of the Amyloid Protofilaments of Hen Egg White Lysozyme Studied by Time-resolved Neutron Scattering. *Journal of molecular biology* 331, 21-28, doi:http://dx.doi.org/10.1016/S0022-2836(03)00722-8 (**2003**).

7 Vernaglia, B. A., Huang, J. & Clark, E. D. Guanidine hydrochloride can induce amyloid fibril formation from hen egg-white lysozyme. *Biomacromolecules* 5, 1362-1370, doi:10.1021/bm0498979 (**2004**).

8 Delloyd. *Preparation of pH buffer solutions*, <http://delloyd.50megs.com/moreinfo/buffers2.html>

9 Frare, E. *et al.* Characterization of oligomeric species on the aggregation pathway of human lysozyme. *Journal of molecular biology* 387, 17-27, doi:10.1016/j.jmb.2009.01.049 (**2009**).

10 Alderton, G. & Fevold, H. L. Direct crystallization of lysozyme from egg white and some crystalline salts of lysozyme. *The Journal of biological chemistry* 164, 1-5 (**1946**).

11 Ravi, V. K., Swain, T., Chandra, N. & Swaminathan, R. On the Characterization of Intermediates in the Isodesmic Aggregation Pathway of Hen Lysozyme at Alkaline pH. *PloS one* 9, doi:10.1371/journal.pone.0087256 (**2014**).

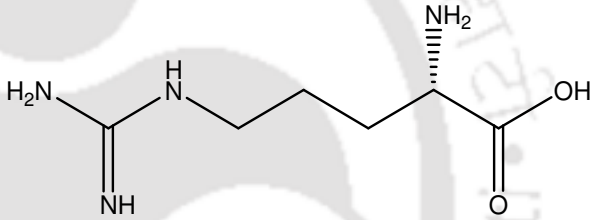
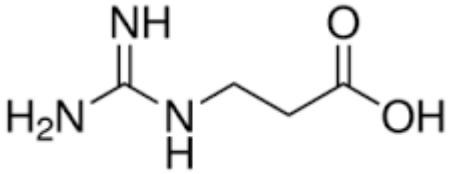
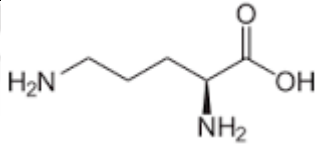
12 Sophianopoulos, A. J. & Van Holde, K. E. Evidence for dimerization of lysozyme in alkaline solution. *The Journal of biological chemistry* 236, PC82-PC83 (**1961**).

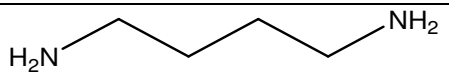
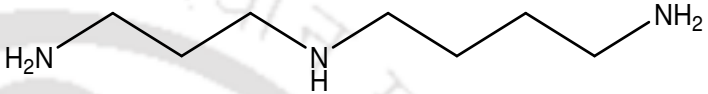
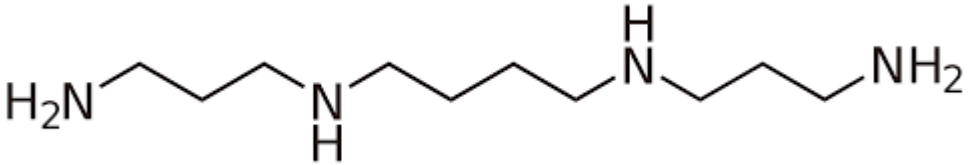
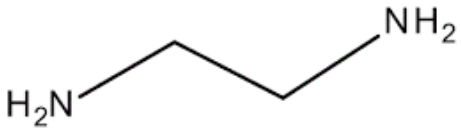
13 Sophianopoulos, A. J. & Vanholde, K. E. Physical Studies Of Muramidase (Lysozyme). Ii. Ph-Dependent Dimerization. *The Journal of biological chemistry* 239, 2516-2524 (**1964**).

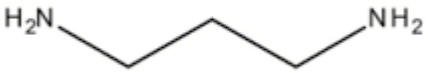
14 Frare, E. *et al.* Identification of the core structure of lysozyme amyloid fibrils by proteolysis. *Journal of molecular biology* 361, 551-561, doi:10.1016/j.jmb.2006.06.055 (**2006**).

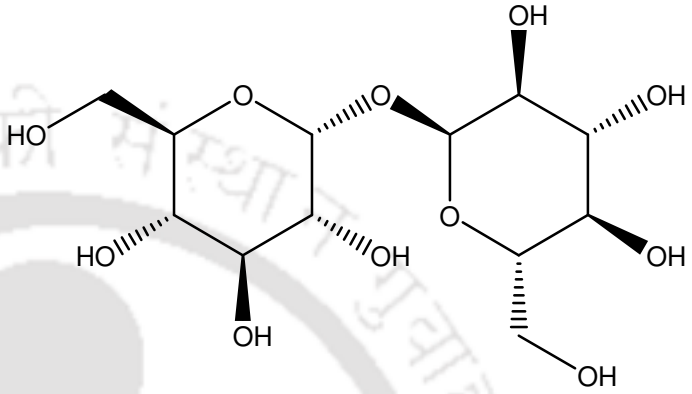
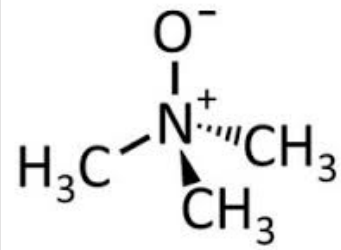
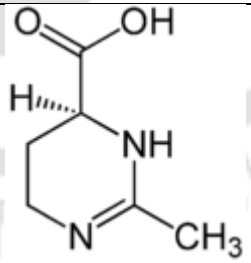
Chapter-5

Results: Concomitant use of Osmolytes with Amyloidogenic conditions of Lysozyme

Osmolyte Generic Name	Name of Individual molecules	Techniques used with indicated Molecule			Structure of Molecule
		r_{ss}	REA	ThT	
Arginine and analogues	Arginine	Y	Y	Y	 Arginine
	Guanidino Propionic acid		Y	Y	
	Ornithine		Y	Y	

Polyamines	Putrescine		Y	Y	 Putrescine
	Spermidine	Y	Y	Y	 Spermidine
	Spermine		Y	Y	
Putrescine Homologs	1,2-diaminoethane		Y	Y	

	1,3-diaminopropane		Y	Y	
	1,5-diaminopentane (Cadaverine)		Y	Y	

Trehalose	Trehalose	Y	Y	Y	 Trehalose
TMAO	TMAO	Y	Y	Y	
Ectoine	Ectoine	Y	Y	Y	

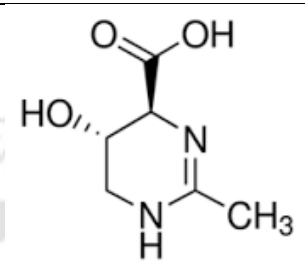
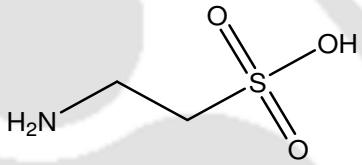
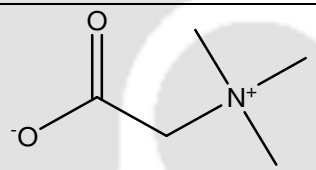
	Hydroxy- Ectoine		Y		
Taurine	Taurine	Y		Y	 Taurine
Betaine	Betaine		Y	Y	 Betaine

Table 5.1: Osmolytes [and their Analogs] Used in Present Study, their Chemical Structure and Techniques Used with Them

While these are all natural molecules. Some of them are available as pharmaceutical products. Which means if their efficacy can be established, they can be used in clinical studies very quickly

Figure 5.1: **This Figure shows Arginine, Ornithine, Guanidino Propionoic Acid, Trehalose, Taurine and Ectoine as available in Pharmaceutical sector**

5.1 Single Osmolyte at Alkaline pH

5.1.1 L-Arginine

While cell is full of amino acids and as such they have been projected as one entire class of compatible osmolytes [1,2] (by compatible osmolytes, it is meant that they don't interfere with metabolic processes); most studies have focussed only on Arginine (and to some extent Proline).

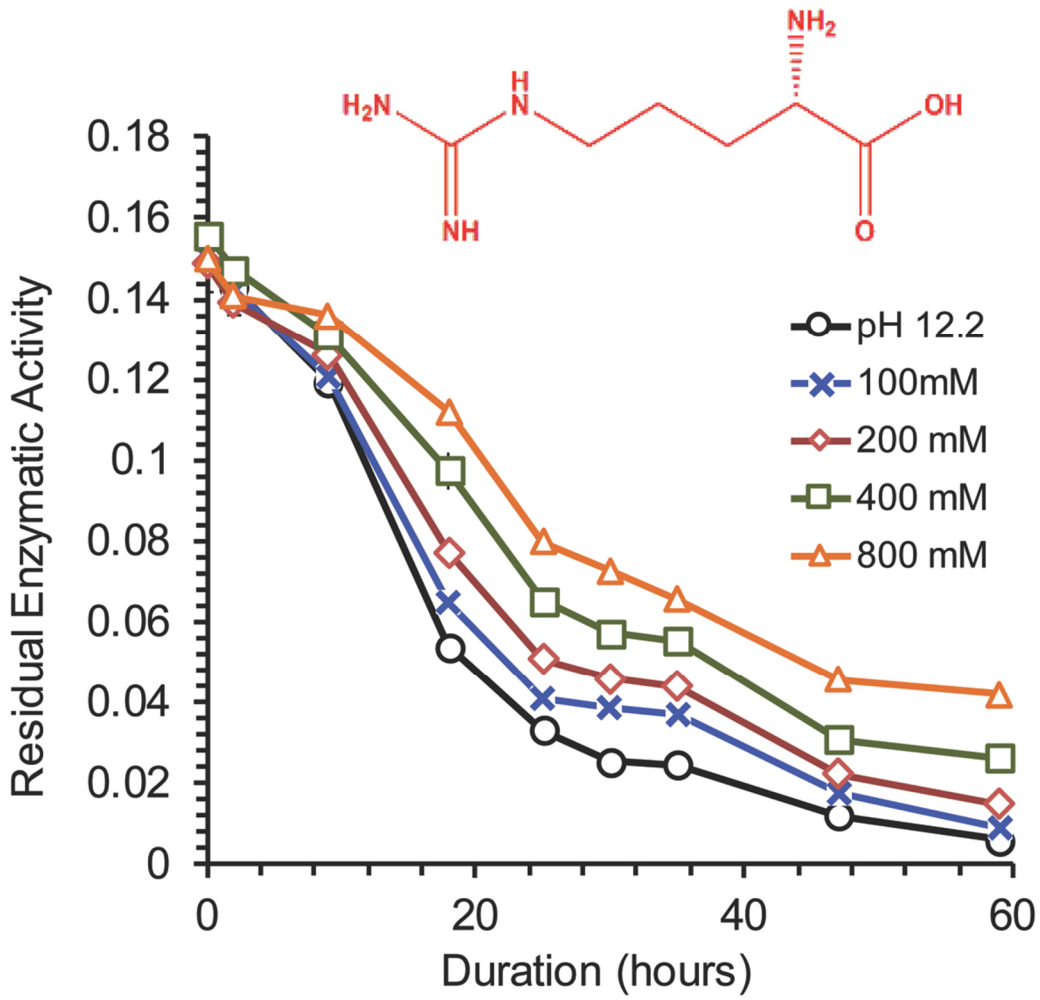


Figure: 5.2: REA of HEWL 120 µM incubated at pH 12.2 at 25 C (for all samples) with different indicated concentrations of ARGININE

Figure 5.2 shows the influence of increasing arginine concentration on the residual enzymatic activity (REA) of HEWL measured at different times since incubation in alkaline pH (12.2). It is noticeable that significant HEWL activity is retained with increasing arginine concentration in the aggregation prone condition. This implies that presence of arginine offers protection to the native state of HEWL even at pH 12.2. One striking feature of this graph is that there is no crossover of lines. Due to this well behaved nature of the plot, arginine has been used as additional control apart from pH 12.2 in case of other osmolytes subsequently.

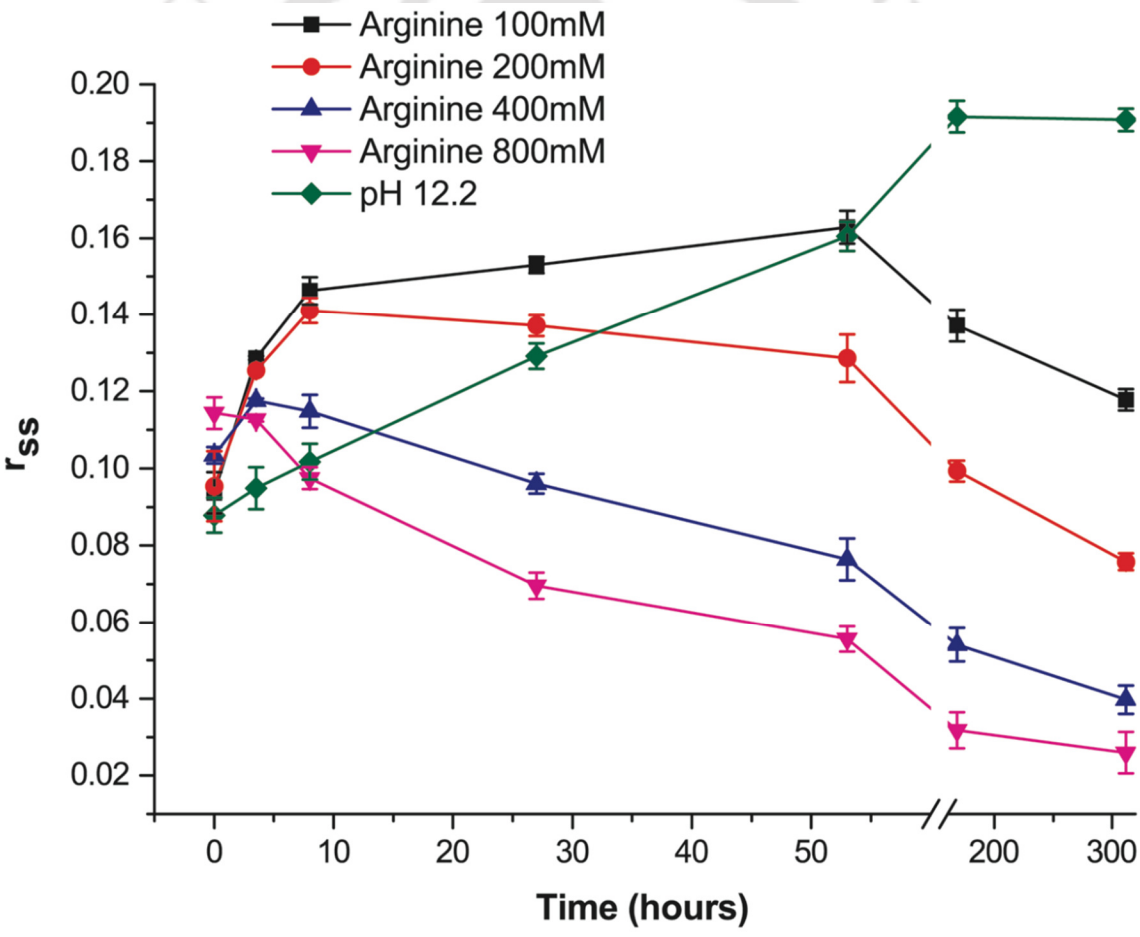


Figure 5.3: r_{ss} **of dansyl-conjugated HEWL (2 µM) + 118 µM Unlabelled HEWL at pH 12.2 and 25°C incubation shown for indicated concentrations of arginine.**

Figure 5.3 reveals the change in r_{ss} as a function of time in presence of arginine. In contrast to control (pH 12.2 with no arginine) r_{ss} shows an increase followed by a steady decline with increasing arginine concentration. This suggests that HEWL oligomer formation is inhibited by arginine. For lower concentrations (100—200 mM), initial growth is followed by saturation and later possible breakup of oligomers beyond 100 hours. For higher arginine concentrations (400 and 800 mM), this initial growth is subdued or nearly gone. Very low values of r_{ss} (for high concentrations, at long durations) is perplexing because it is even less than monomer (0.08), but can possibly be explained if we assume that globularity has given way to heavily extended random coil (complete loss of conformation is supported by near zero value of enzyme activity at these high time points). Similar observations have been recorded earlier [3].

Figure 5.4 shows a steady decline in Thioflavin-T fluorescence (measured after 2-4 days) with increasing presence of arginine in the HEWL samples aggregating at pH 12.2. This implies a significant decline in population of both amyloid and amorphous aggregates of HEWL in the medium. In all the above cases, the trends clearly indicate that higher the concentration of arginine, better is the inhibition of aggregation.

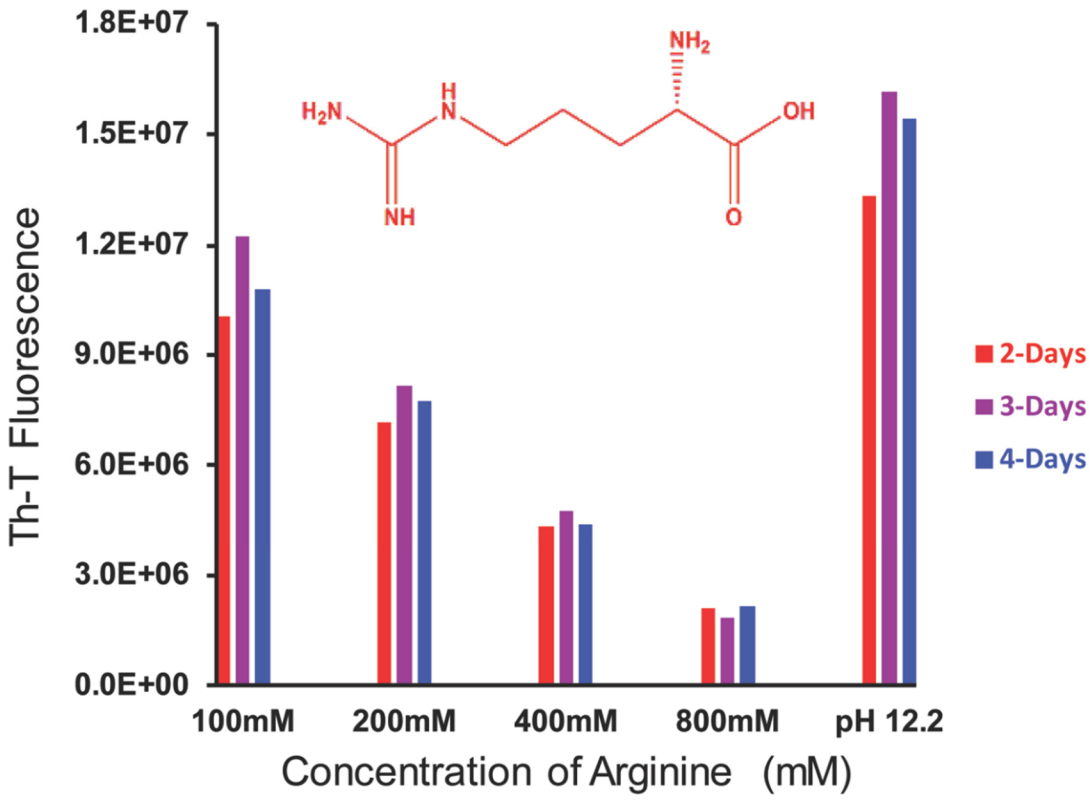


Figure 5.4: **ThT of HEWL (120 µM) at pH 12.2 and 25°C incubation for indicated concentration of Arginine and indicated number of days.**

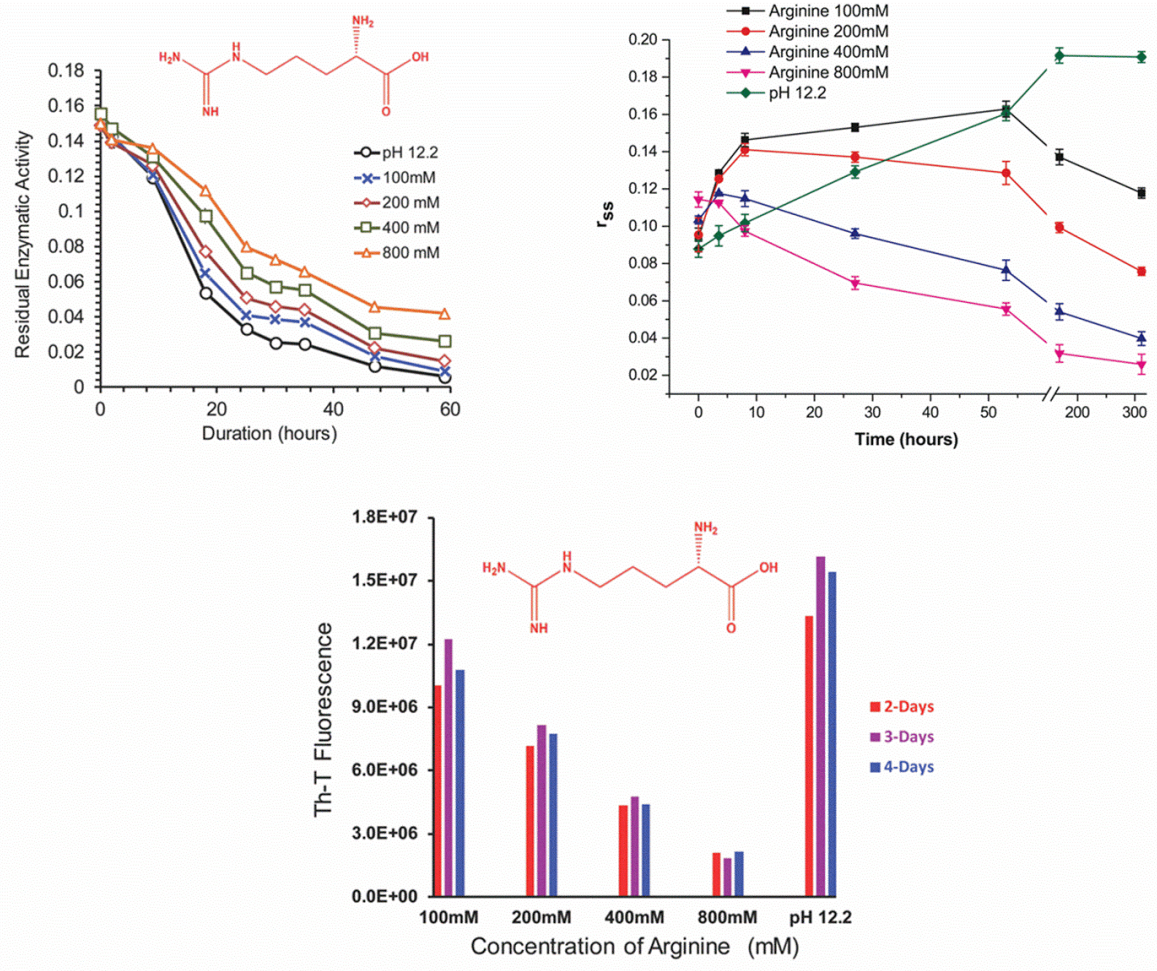


Figure-5.5 Combined Figure with REA, rss & Th-T of Arginine

Figure 5.5 summarizes the 'approach' of this chapter. First there is a osmolyte with a chemical structure followed by three plots, viz., REA, rss, Th-T. Roughly what we are seeking is a kind of QSAR [quantitative structure activity relationships] as they seek in pharmaceutical industry. But as of now it is 'Qualitative'. In order for it to become 'Quantitative' combinations of more proteins; more osmolytes and their structural analogues need to be studied

5.1.2 Arginine analogues

Figure 5.6 depicts the decrease in enzymatic activity of HEWL in presence of arginine and two of its analogues, Guanidino propionic acid and ornithine.

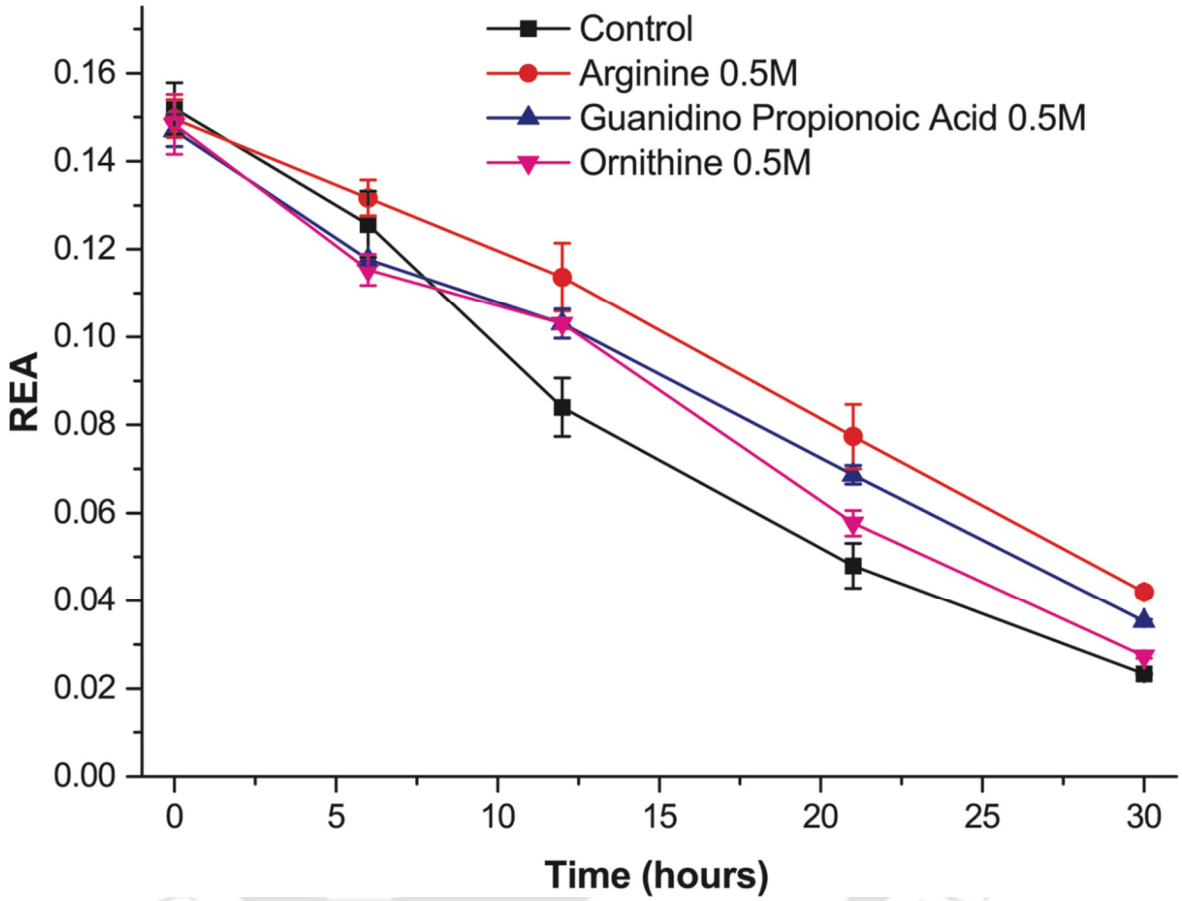


Figure 5.6: REA of HEWL (120 µM) at pH 12.2 and 25°C incubation for indicated times and concentrations of arginine analogues (500 mM) mentioned. Control refers to pH 12.2 sample alone.

Exact analogue of arginine, namely Guanidino butyric acid was not soluble in pH 12.2 buffer [when heated it did dissolve but precipitated out upon cooling], hence Guanidino propionic acid was used. It is observed that in comparison to arginine, Guanidino propionic acid is nearly as potent. Ornithine is however weaker in comparison.

This suggests that Guanidino group plays a significant role in arresting amyloidogenesis, while the amino group of arginine appears not so important.

Figure 5.7 & Figure 5.8 shows the influence of arginine analogues on Thioflavin-T fluorescence. In Figure 5.7, it is observed that while Guanidino propionic acid is nearly as potent as arginine suppressing amyloid formation after six days; Ornithine is not so effective in arresting amyloid formation.

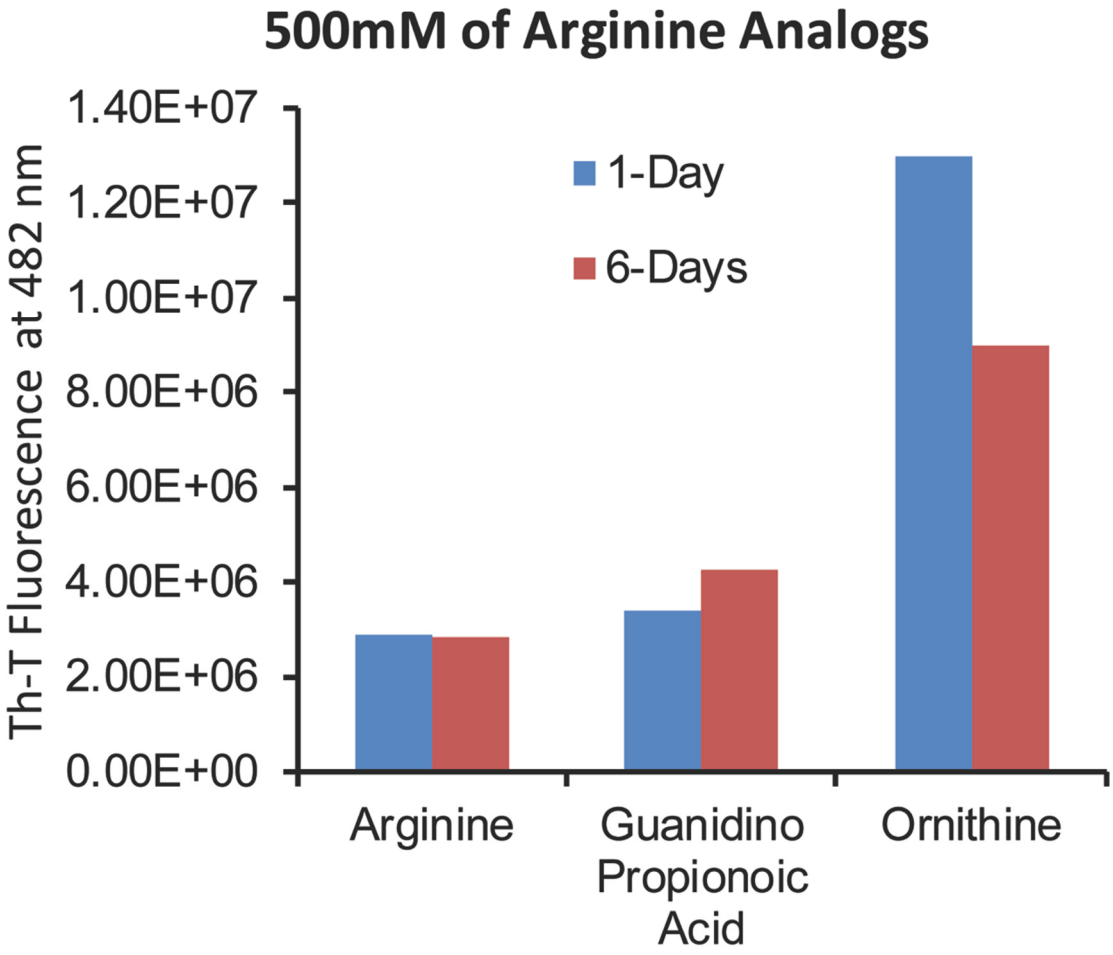
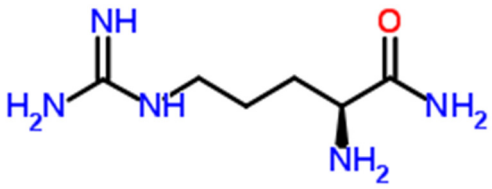


Figure- 5.7: ThT of HEWL (120 µM) at pH 12.2 and 25°C incubation for indicated days and indicated for Arginine Analogues [concentration of Arginine and its analogues is 500 mM]

Figure 5.8 reveals that Argininamide is more potent than arginine in suppressing amyloid formation in HEWL. Structure of Argininamide is as under



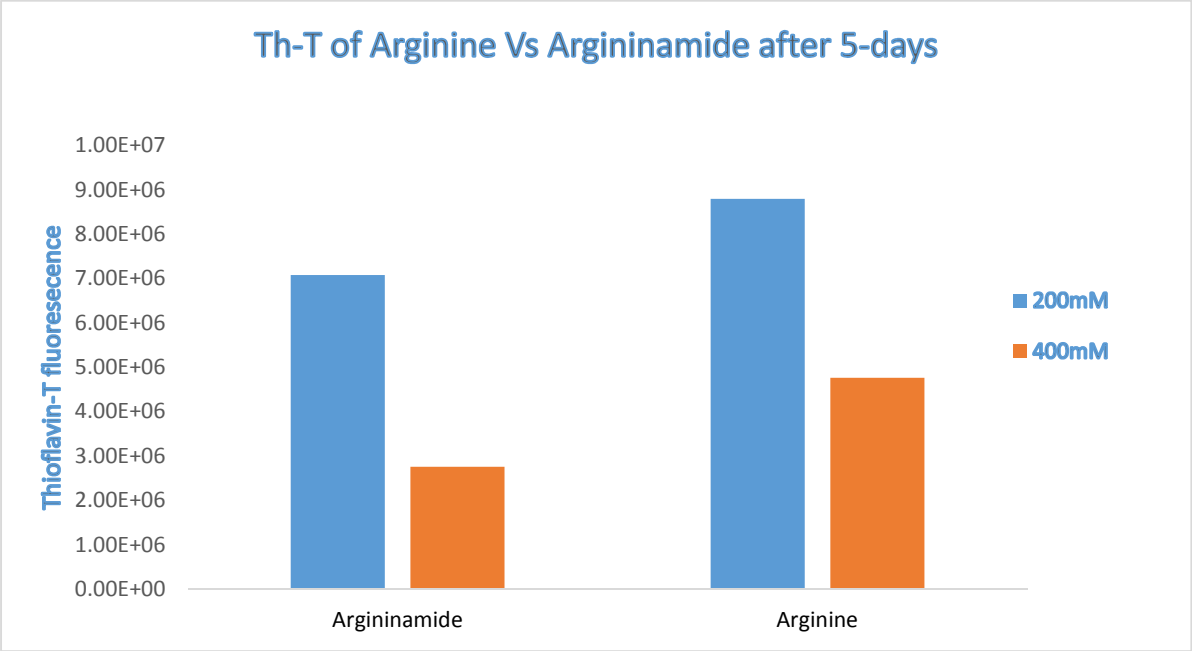


Figure-5.8: Th-T of HEWL (120 µM) at pH 12.2 and 25°C incubation for indicated concentrations with Arginine & Argininamide [recorded after 5-days]

5.1.3 Polyamines

Figure 5.9 A & 5.9 B show the effect of polyamines Putrescine, Spermidine and Spermine (each at concentration of 500 mM) on the REA of HEWL incubated at pH 12.2. Figure 5.9 A data is directly from the experiment. Here we can see the classic Hofmeister series in action. Putrescine with 2-NH_2 groups makes effective (at concentration of 500 mM) NH_2 concentration 1.0 M and yields (at zero-time point) same REA value as pH 12.2 control and with Arginine. However, in case of Spermidine (with three amino groups) the effective amino concentration is 1.5 M and in case of Spermine (with four amino groups) the effective amino concentration is 2.0 M. At such high concentrations there is precipitation of protein thus reducing zero-point values with Spermidine and Spermine.

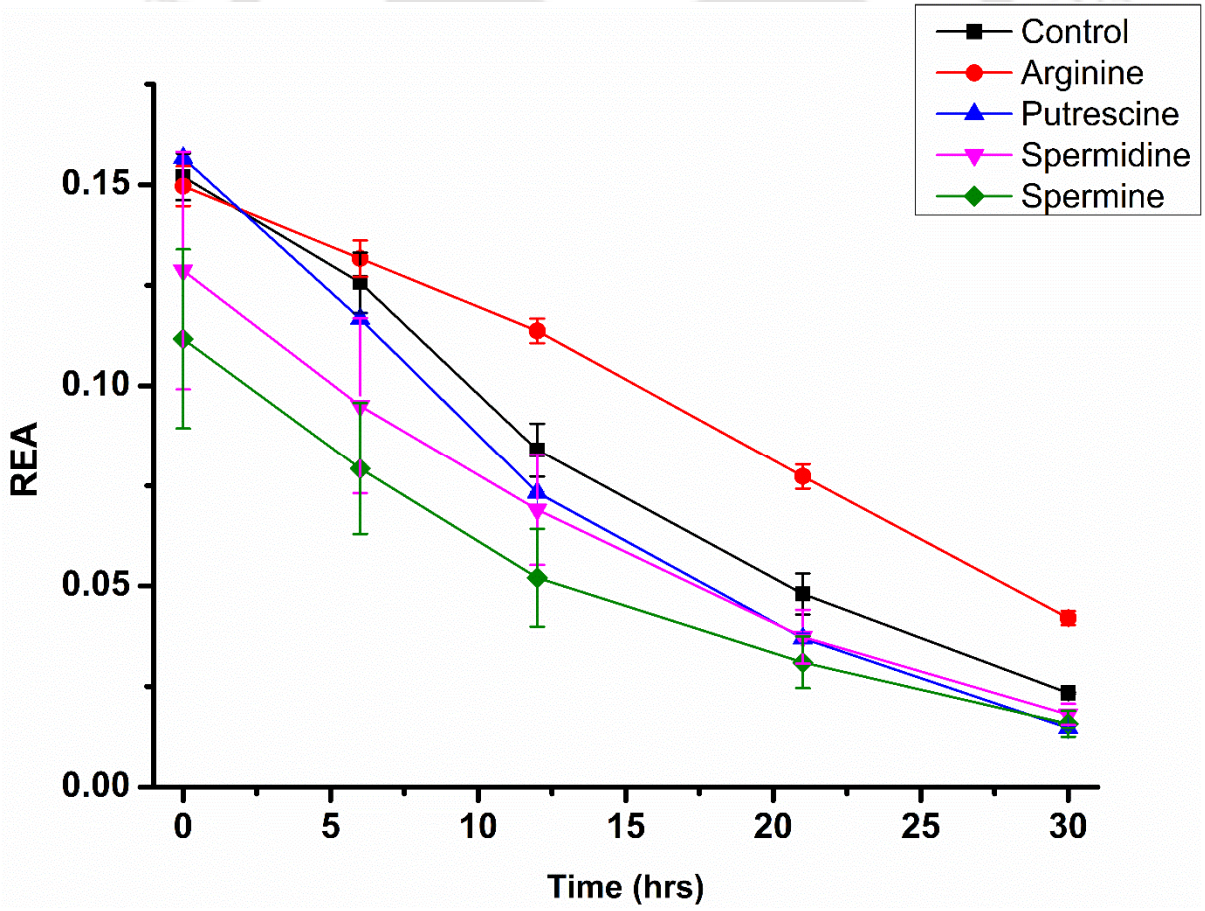


Figure-5.9 A: REA of HEWL at pH 12.2 in presence of 500 mM of Polyamines

One more noticeable fact is that whereas such precipitation is known with ammonium ions, in this case even un-protonated amine [which is the case at pH 12.2] does the same. So it is not necessarily a salt effect and also suggests that we do need to move beyond ions and need a Hofmeister series for uncharged species. Because of this precipitation, we re-normalized zero-point value of Spermidine and Spermine to that of Putrescine by way of multiplication [and multiplied all REA values at other time-points with same multiplicative factor]. The resulting plot is Figure-5.9 B. From this figure we can see that all polyamines offer nearly no protection for enzymatic activity of HEWL (as compared to pH 12.2 control).

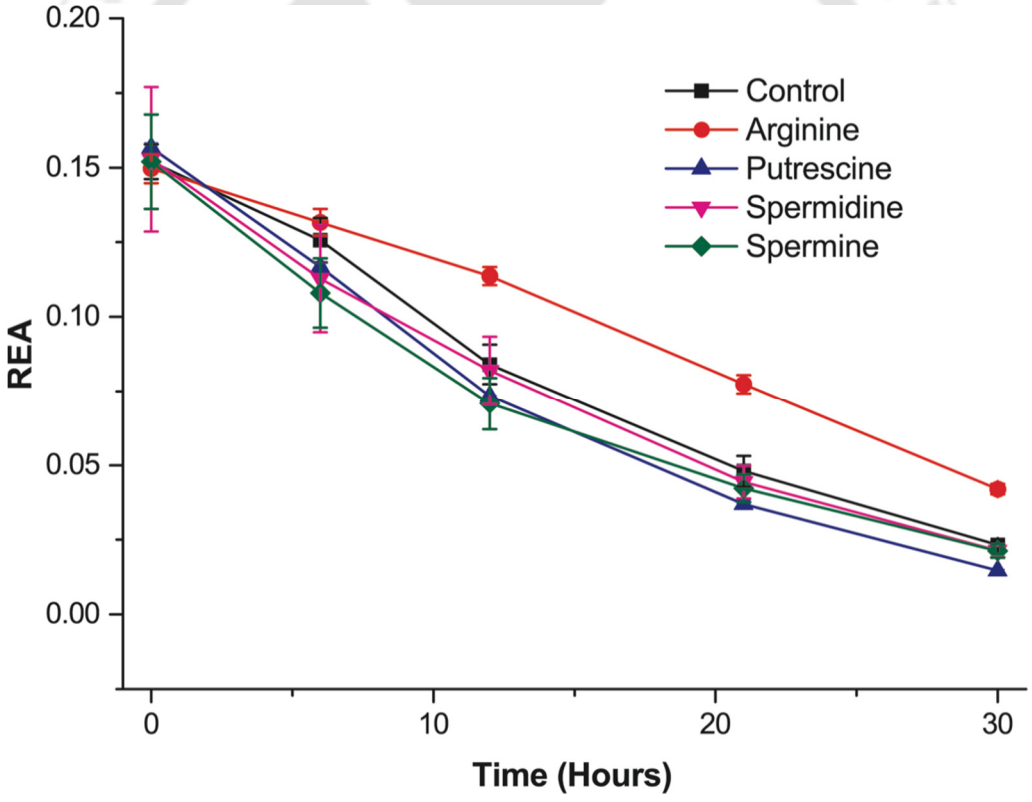


Figure 5.9 B: REA of HEWL at pH 12.2 in presence of 500 mM of Polyamine after re-normalization of Spermidine and Spermine to match the same of Putrescine at zero-time point

The effect of spermidine on the fluorescence anisotropy of dansyl-labelled HEWL is shown in Figure 5.10. In comparison to control (pH 12.2), the growth kinetics in presence of spermidine is markedly reduced. Interestingly after 26 hours, r_{ss} for 200 and 400 mM concentrations slows down and begins to fall drastically. For lower concentrations, this happens much later at 150 hours. This clearly demonstrates the dramatic influence of spermidine in halting the growth of HEWL oligomers.

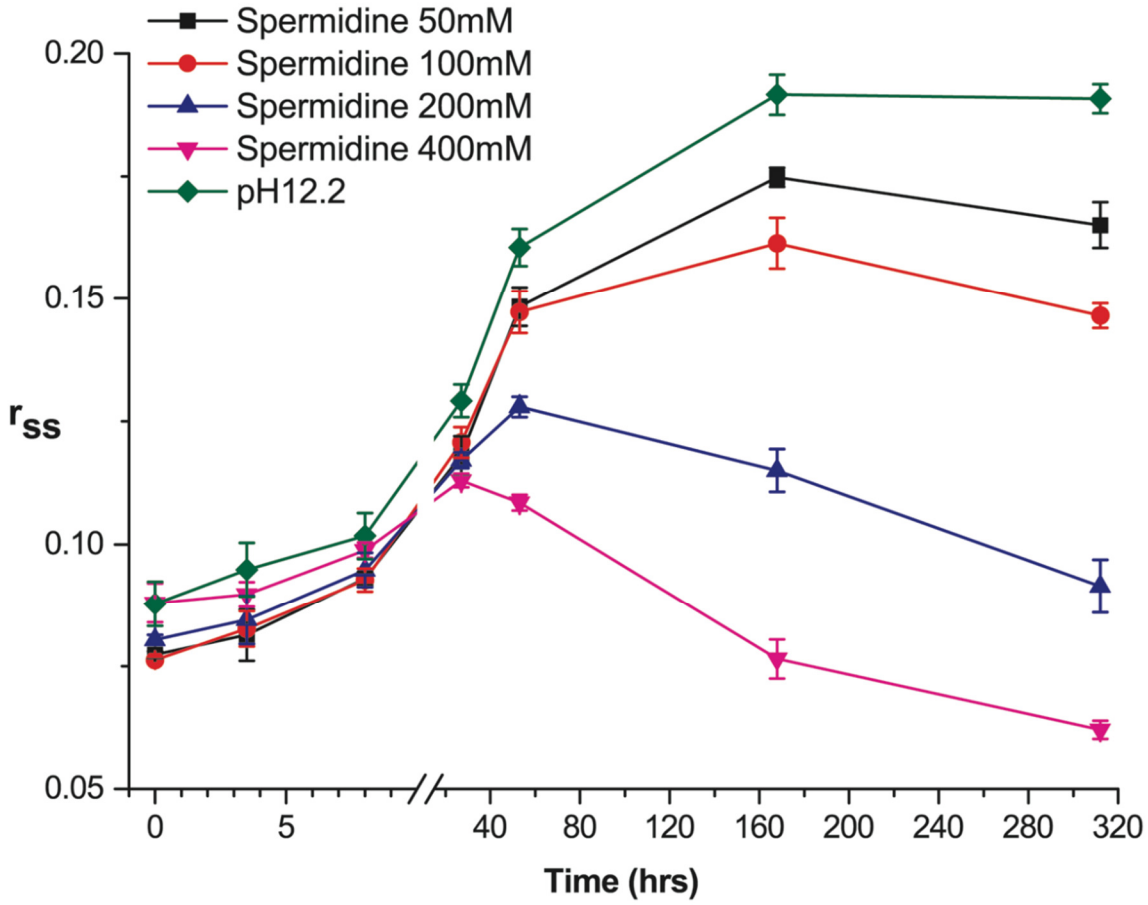


Figure 5.10: **Kinetics of ss-Anisotropy with indicated concentrations of Spermidine**

5.1.4 Putrescine Analogues

Since structure of Putrescine is deceptively simple, this could be ideal to make use of many analogues. In this thesis we have used only available analogues; many more can

be synthesized by taking advantage of simple chemical structure. While 1-Amino, 4-Butanol [a Putrescine derivative] could be used for this study, other analogues Adipic acid and 5-Amino Valeric Acid could not be used due to insolubility problem.

Figure 5.11 shows the influence of putrescine analogues 1,2 & 1,3 diamines on the Residual Enzymatic Activity of HEWL. In comparison to arginine, the analogues offer poor protection to HEWL. There is virtually no difference as compared to Putrescine

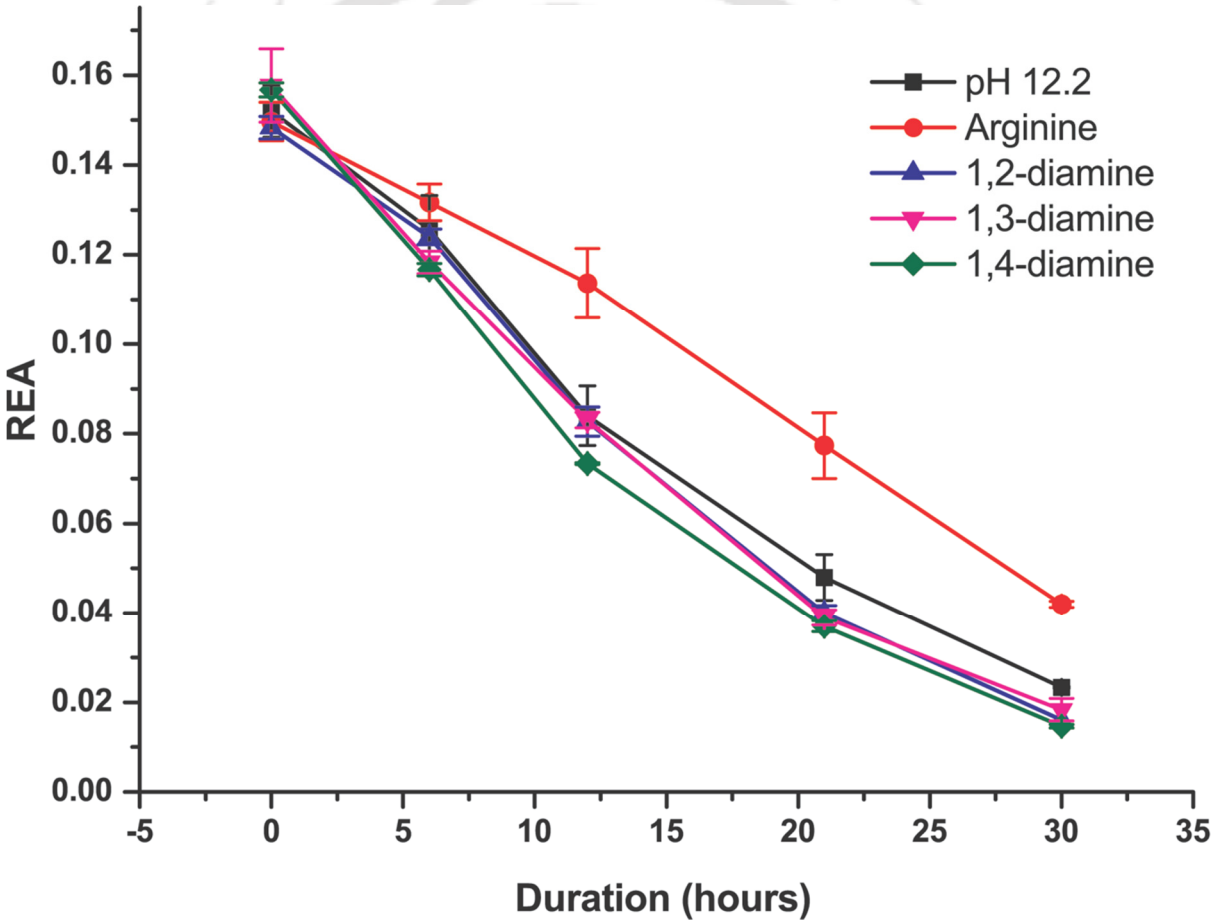


Figure-5.11: Kinetics of REA in presence of indicated concentrations of lower homologs of Putrescine

Figure 5.12 shows that all polyamines are potent suppressors of amyloid formation in comparison to arginine. Figure 5.12 reveals Thioflavin-T fluorescence in presence of

additives at 500 mM concentration after 1 and 6 days of incubation. It is observed that analogue 1,4-Diamine is the most potent in suppressing amyloid formation in comparison to analogues 1,3 and 1,2-Diamine which appear better than arginine.

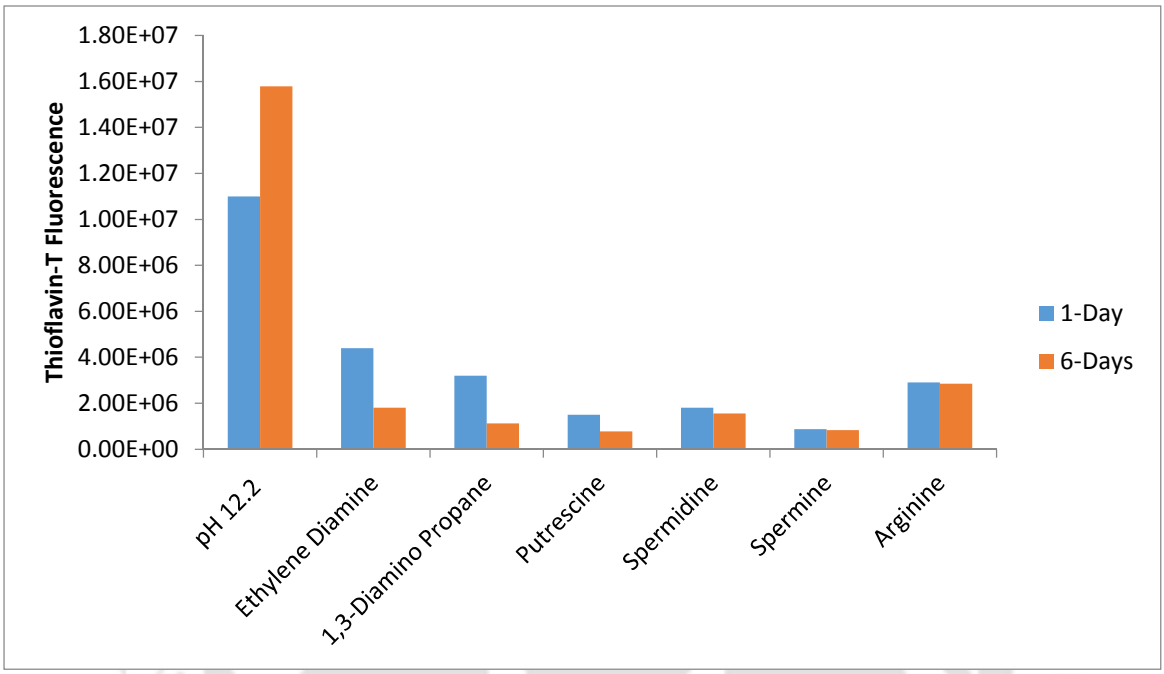


Figure 5.12: Th-T with 500mM of various Polyamines and their lower homologs

Figure 5.13 shows that analogues 1,4; 1,5; 1,6 and 1,7 diamines are potent suppressors of amyloid formation at 200 and 400 mM concentrations after an incubation period of 5 days. analogues 1,2 and 1,3 appear not so potent while 1-Amino 4-butanol is moderate in suppressing amyloid formation in comparison to analogues above.

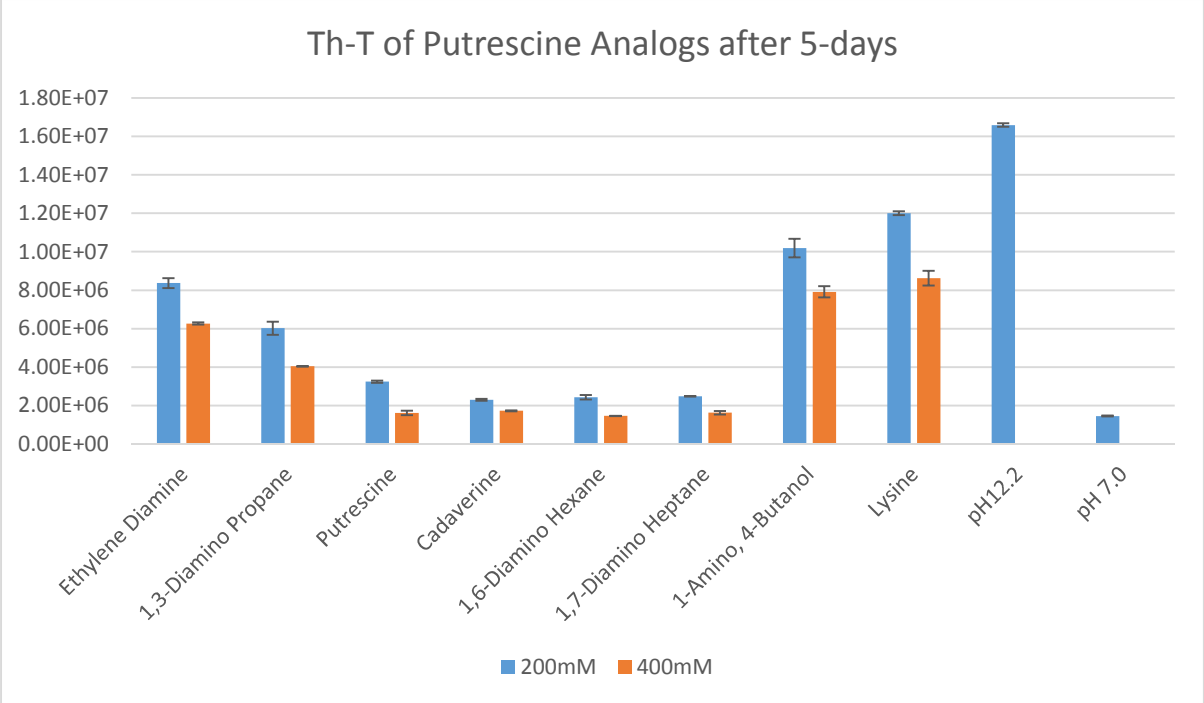


Figure 5.13: Th-T with 200 & 400mM concentrations of Putrescine and its structural and functional analogues

Figure 5.14 shows the effect low concentrations of the additive (50 and 100 mM after an incubation period of 21 and 2 days, respectively. Here too 1,4-Diamine appears more potent in suppressing amyloid formation in comparison to 1,2 and 1,3-Diamine which are a shade better than arginine at same low concentration.

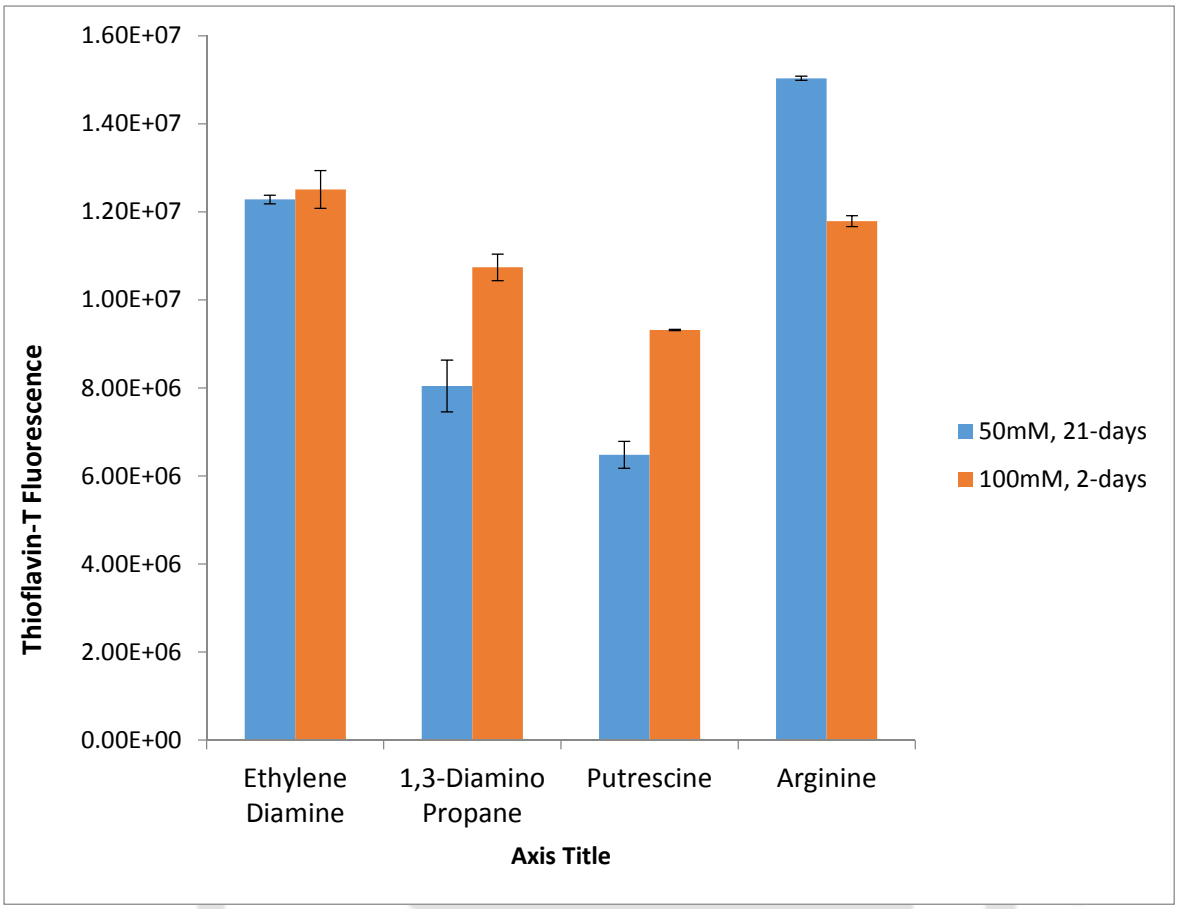


Figure 5.14: Th-T with very low concentrations of Putrescine and its lower homologs

5.1.5 Testing of fluorescence quenching by polyamines: Because Amines are quenchers of fluorescence, quenching of Thioflavin-T fluorescence by amines was investigated. As already mentioned before, Thioflavin-T with polyamines has been used at concentrations from 50 mM to 500 mM. Experiments were conducted in the following way: Fibrils made in pH 12.2 were diluted by half either with pH 12.2 buffer or with 1 M of amines. Thus the extent to which polyamines quench Thioflavin-T fluorescence appears to be moderate (Figure-5.15). Hence significantly reduced

Thioflavin-T fluorescence in Figure 5.12 (and likewise Figure 5.13 & Figure 5.14) indicates decreased amyloid population rather than quenching of Th-T fluorescence.

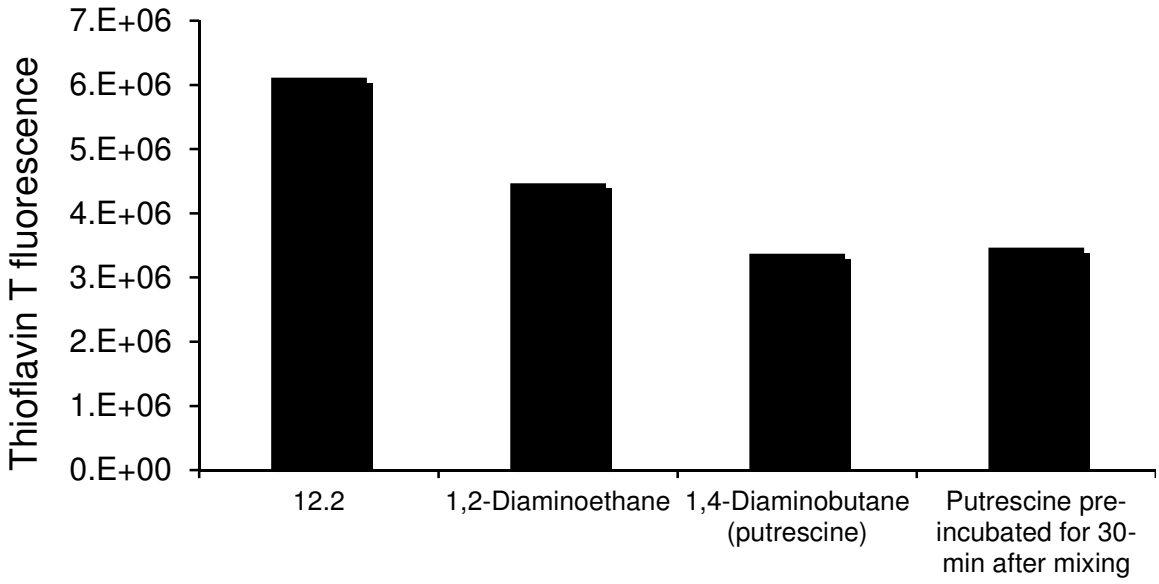


Figure 5.15: Testing of quenching by polyamines. Quenching of ThT fluorescence by externally added amines (500 mM) is shown. The plot also shows effect observed after putrescine incubated with fibrils for 30 minutes

5.1.6 Trehalose

Trehalose has mostly been studied in plants undergoing stressed conditions such a drought. Trehalose is one molecule which has been used at very high concentrations in rat model of Huntington disease [4] so from toxicology point of view it appears safe. Figure 5.16 shows that trehalose at 500 mM is not significantly better than control in retaining HEWL activity. However, at 900 mM it offers better protection, nearly similar to 500 mM arginine. This suggests that unlike arginine where the concentration dependence profile on HEWL protection at pH 12.2 was kind of linear, a more non-linear and criticality based trend is observed with trehalose.

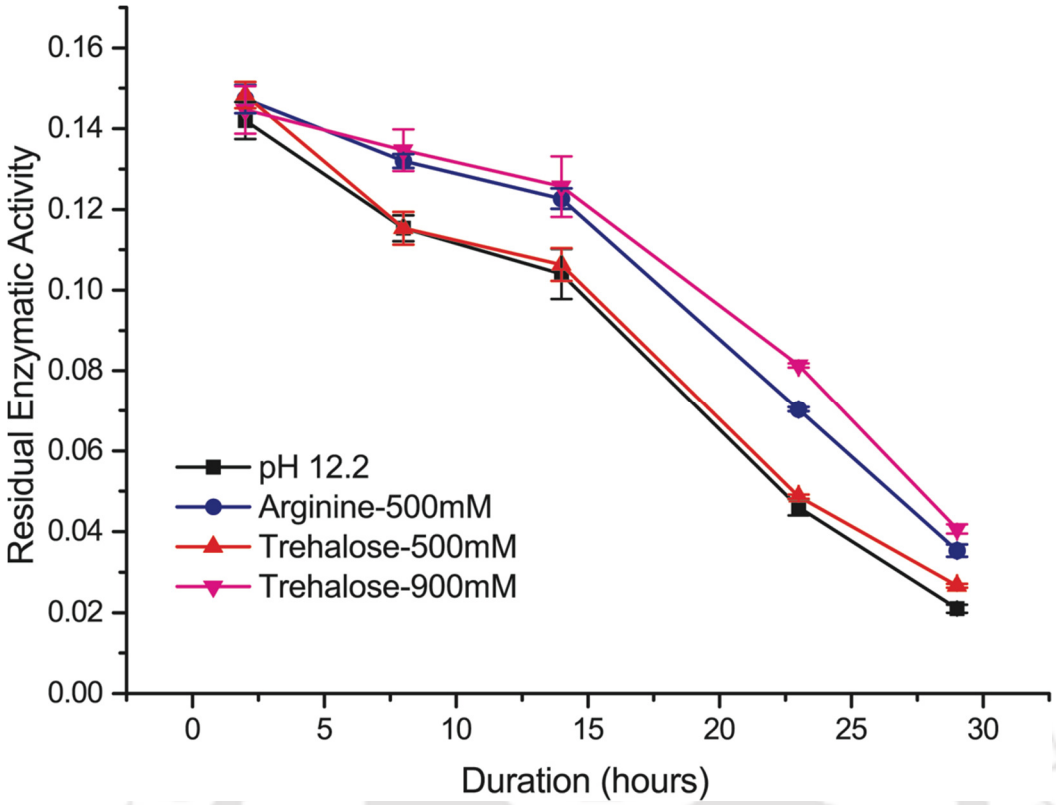


Figure 5.16: The residual enzymatic activity of HEWL is shown for different concentrations of trehalose in pH 12.2 for different incubation times as indicated. pH 12.2 and arginine samples are shown as control.

The r_{ss} plot (Figure 5.17) shows that trehalose has only a marginal effect in arresting the growth of oligomers, which occurs very late at 300 hours. At initial times (0-50 hours), no significant effect of trehalose on oligomer growth kinetics is evident.

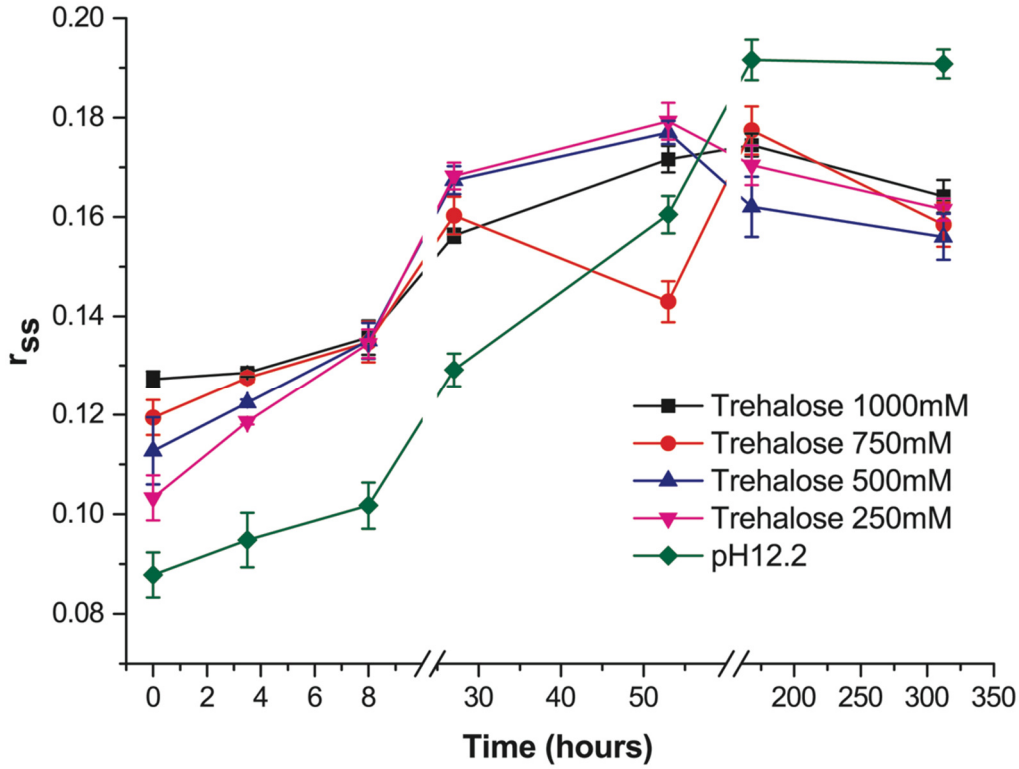


Figure 5.17: The ss-anisotropy of dansyl-conjugated HEWL is shown for different concentrations of trehalose in pH 12.2 for different incubation times as indicated.

Trehalose favours oligomer formation in a concentration dependent way. A rapid increase is observed even at early minutes of incubation so at higher concentrations, Trehalose favours oligomerization. Amyloid formation appears significantly less with increasing concentration trehalose as revealed by Figure 5.18. However, structural reorganisation to final amyloid is being prevented almost independent of concentrations in the range 500-1000 mM (figure 5.18). Concentration dependency profile of REA and Th-T is very different. In REA, there is nearly no difference (from control) at 500 mM but shoots up rapidly thereafter. In Th-T already significant difference at 500 mM which does not improve much thereafter. It can be noted that despite high r_{ss}, the Th-T value is low which means that lot of oligomers but little amyloid [and such differential studies are rarely done in literature].

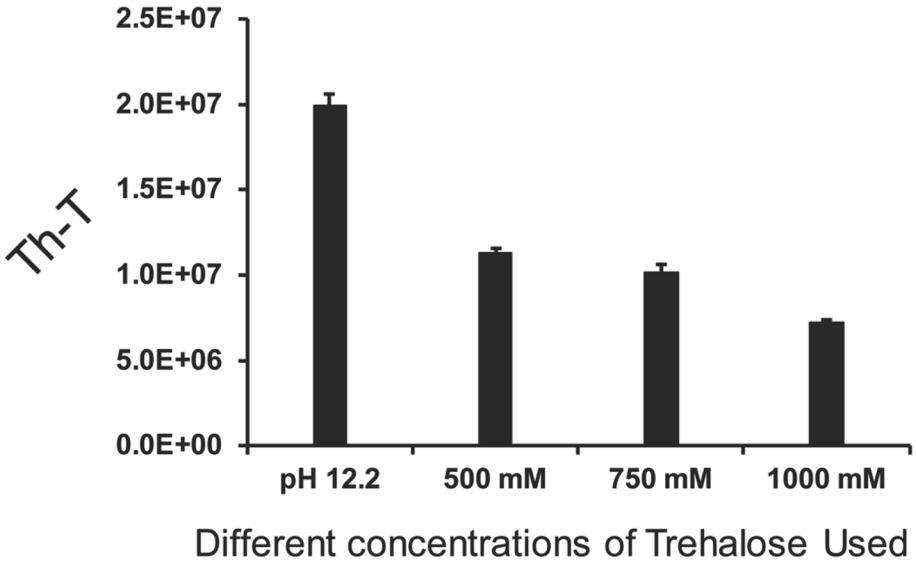


Figure 5.18: The ThT fluorescence observed in HEWL aggregates is shown for different concentrations of Trehalose in pH 12.2 after 3- days

5.1.7 Tri Methyl Amine Oxide (TMAO)

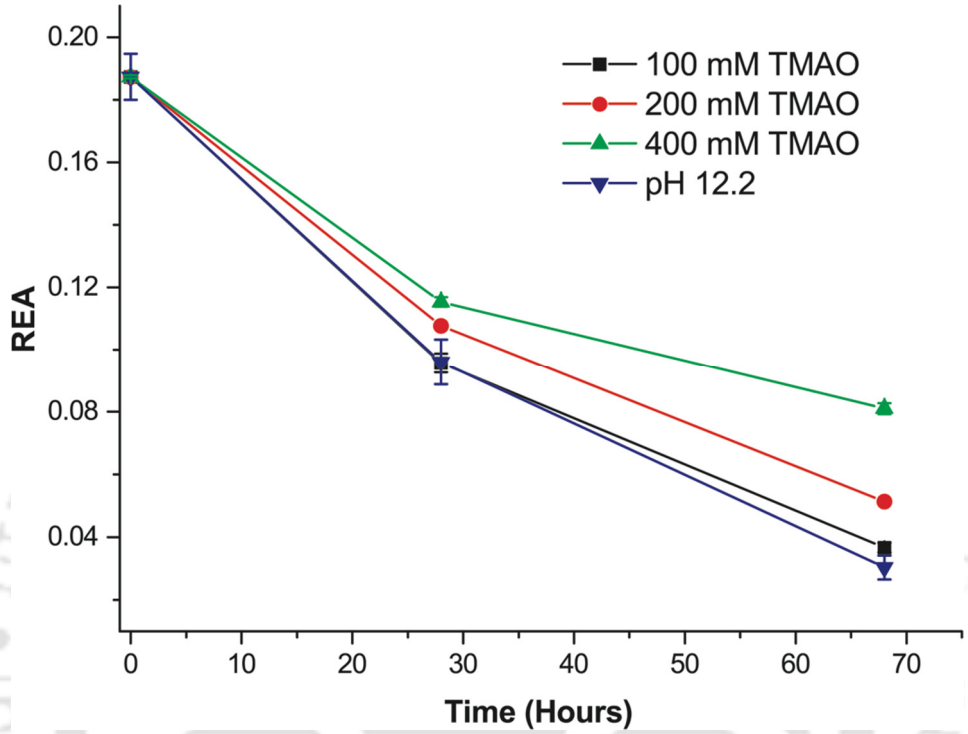


Figure 5.19: The residual enzymatic activity of HEWL is shown for different concentrations of TMAO in pH 12.2 for different incubation times as indicated.

TMAO is the most popular osmolyte in literature till date. Historically the reason for that is because it is found in high quantity in kidney where it helps to overcome effect of urea which is very damaging for proteins. In a recent paper [5], it has been argued that concentration of TMAO in deep sea fish correlates with depth and determinant to what depth fishes can be found. Despite such large number of studies; Not even mono and dimethyl analogues of TMAO seem to have been thoroughly compared with TMAO side-by-side which is surprising [unless they are unstable]. Figure 5.19 reveals that increasing concentrations of TMAO preserves the enzymatic activity of HEWL in comparison to

control (pH 12.2). The r_{ss} profile (Figure 5.20) profile shows that oligomerization is favoured at concentrations 200mM and above especially in early hours. However, as the Th-T profile shows, Amyloid formation appears significantly less in presence of high concentrations of TMAO (Figure 5.21). Thus we are dealing with a scenario where oligomerization occurs but with less amyloid content. So we are dealing with amorphous aggregate. TMAO appears to prefer formation of oligomers with increasing concentrations, like Trehalose. However rapid formation on exposure at early minutes is not observed like Trehalose, whereas reorganisation to amyloid is being prevented at concentrations 400 mM and above.

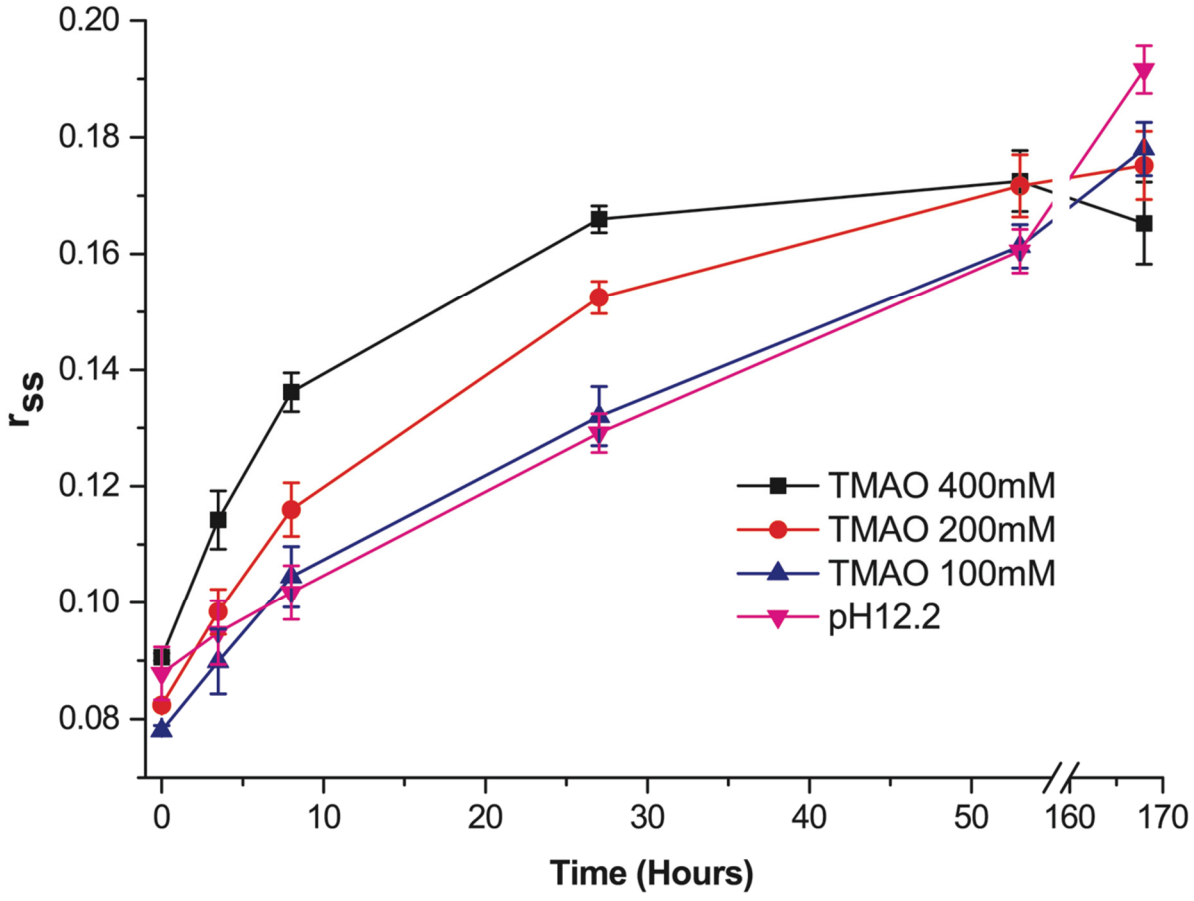


Figure 5.20: **The r_{ss} of dansyl-conjugated HEWL is shown for different concentrations of TMAO in pH 12.2 for different incubation times as indicated.**

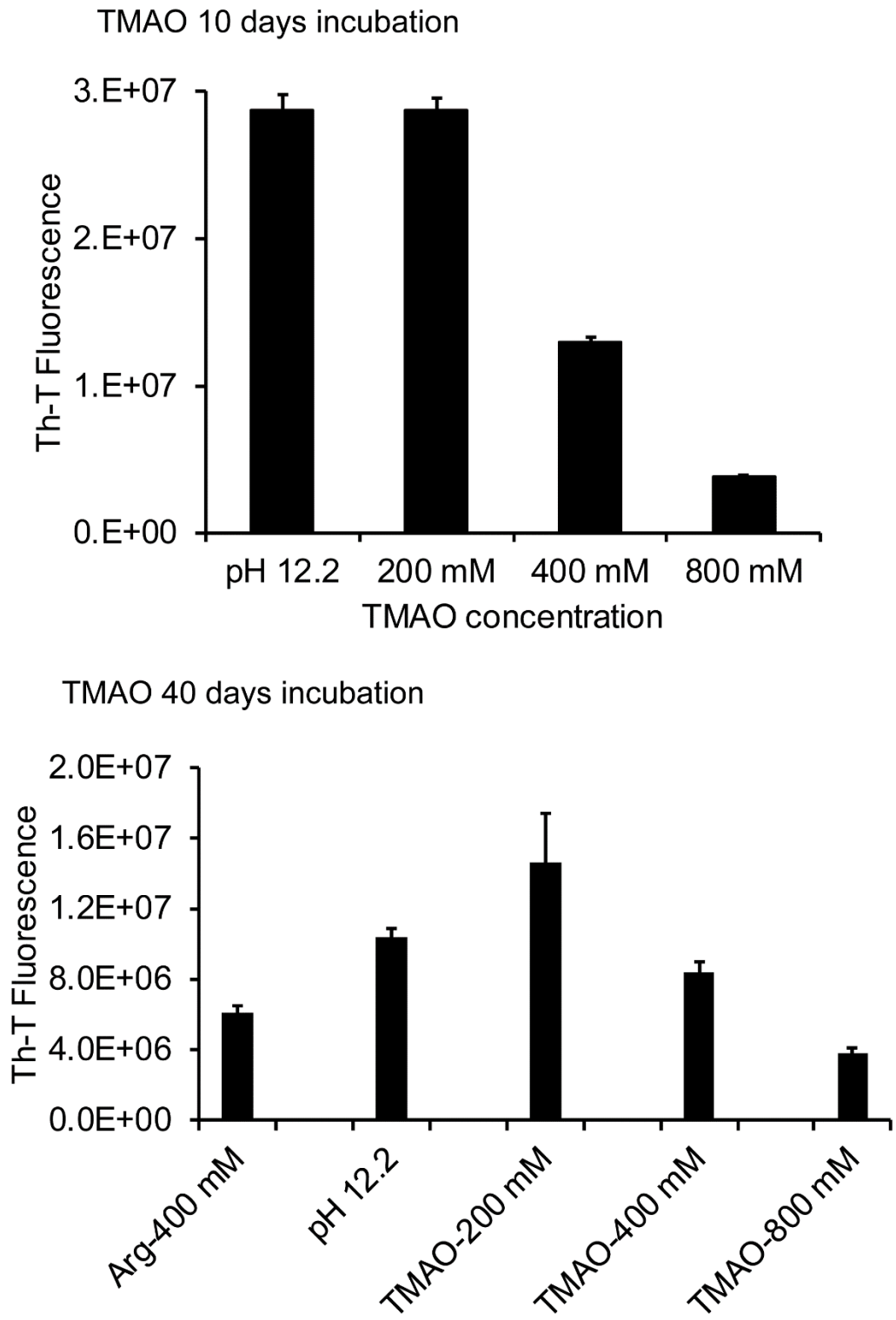


Figure 5.21: Th-T in presence of different concentrations of TMAO after 10 & 40 days [Both experiments are from different batches and hence not plotted together]

5.1.8 Ectoine and Hydroxy-Ectoine

Ectoine was discovered from halophilic bacteria in 1985 [6] and that makes it a molecule of most recent origin as far as this thesis is concerned. There are only few biological studies with it. This molecule has previously been used with Aβ [7] and Insulin [8] to counteract amyloidogenicity. Figure 5.22 shows the influence of additives Ectoine and Hydroxy-Ectoine on the REA of HEWL incubated at pH 12.2. In comparison to control, concentrations as low as 100 mM or 200 mM of Ectoine are effective in preserving the enzymatic activity in HEWL. This protection is lot more as compared to any other osmolyte used in this study, so much so that at incubation times beyond 440 hours [about 19-days] there is still about half of the activity present for 100 mM Ectoine. For concentration even higher it is as if there is near zero loss in activity, Hydroxy-Ectoine is even more dramatic at 100 mM. Thus both Ectoine and Hydroxy-Ectoine appear extremely potent in protecting the HEWL catalytic activity for a long duration under harsh pH condition of pH 12.2. No other osmolyte scores anywhere near to them as far as preserving HEWL activity at pH 12.2.

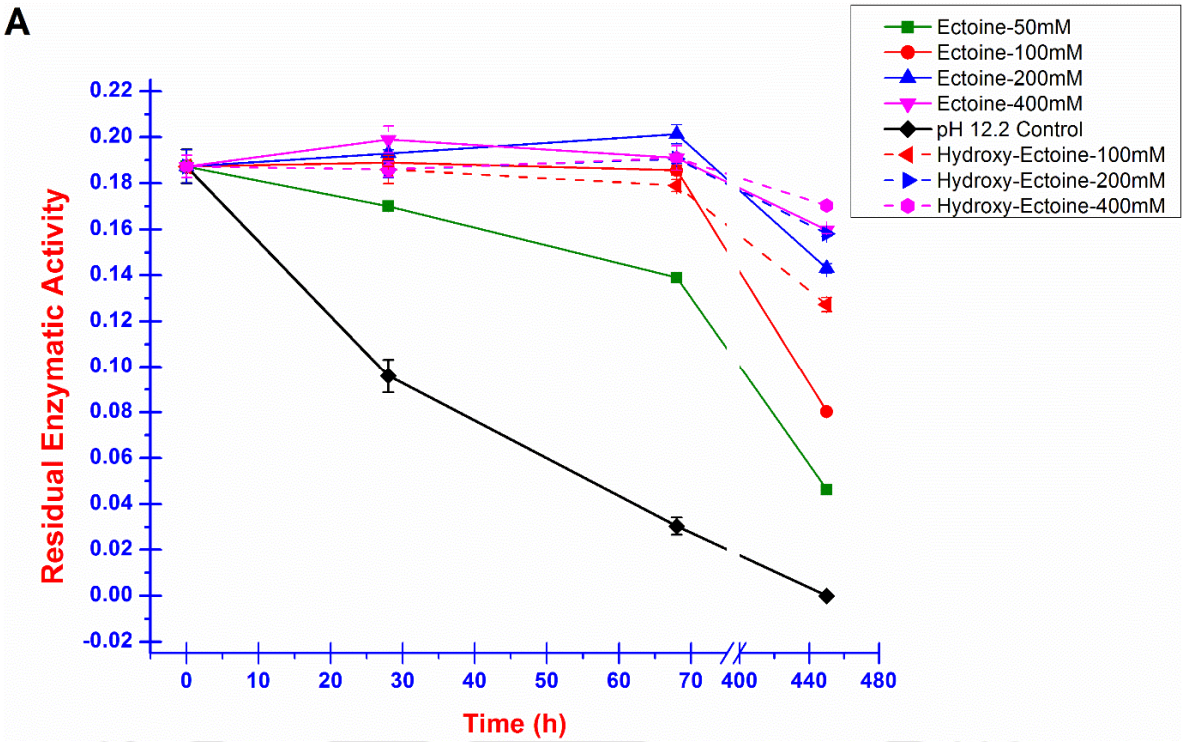


Figure 5.22: Kinetics of REA in presence of indicated concentrations of Ectoine & Hydroxy-Ectoine

Figure 5.23 reveals that increasing concentration of Ectoine has an arresting influence on the growth of HEWL oligomers. At 100 mM, the growth is completely abolished, which is indeed remarkable. As shown in Figure 5.24, the amyloid content in HEWL samples at pH 12.2 appears more and more diminished as the concentration of Ectoine is increased. Concentration as low as 25 mM which is something easy to attain physiologically also slows down oligomerization significantly.

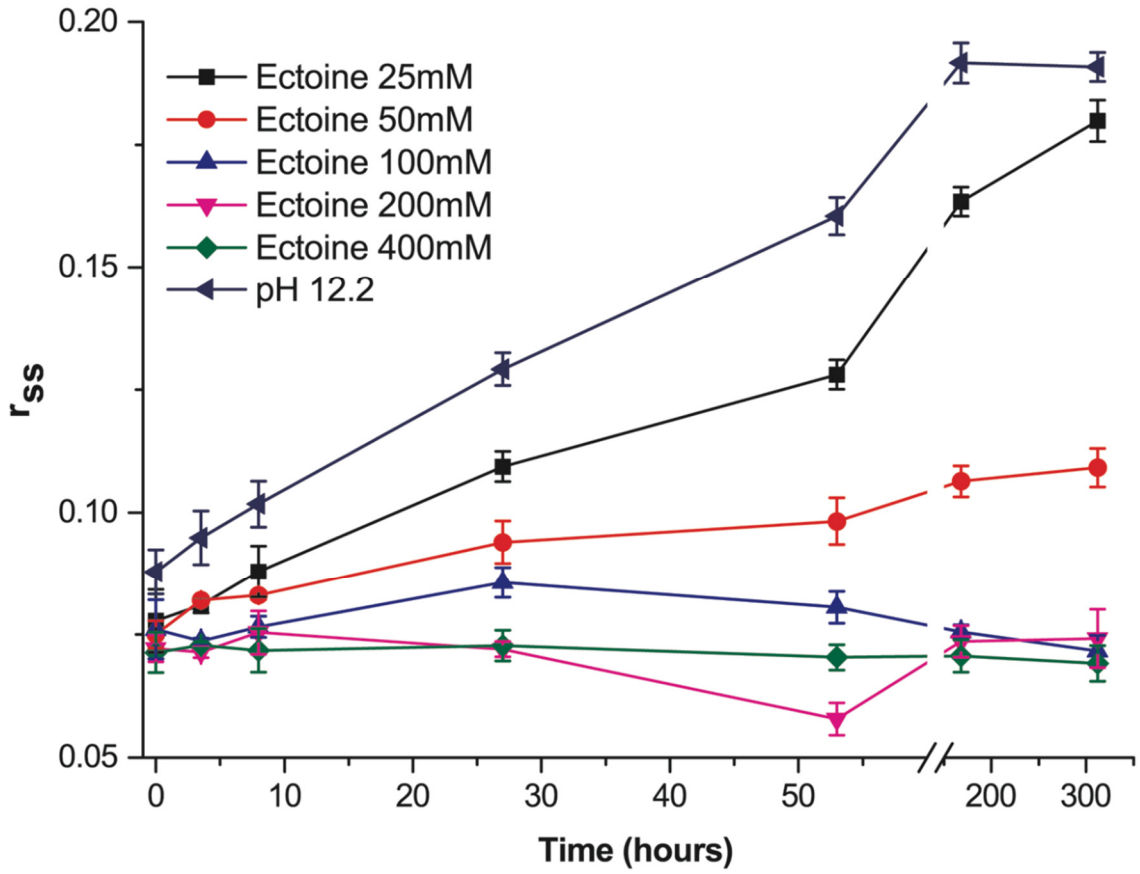


Figure 5.23: Kinetics of ss-Anisotropy in presence of indicated concentrations of Ectoine

Figure 5.24 shows that while 25 mM & 50 mM Ectoine have no effect on Th-T, 100 mM is quite effective and 200 mM brings it to near baseline.

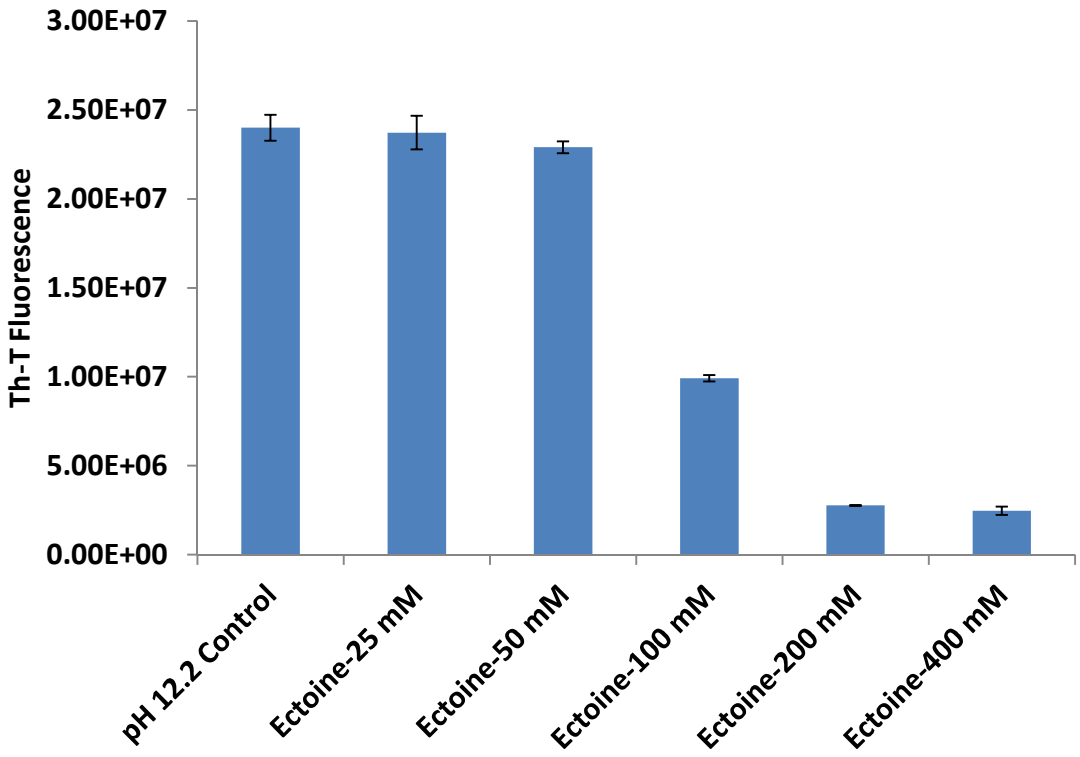


Figure 5.24: Th-T in presence of indicated concentrations of Ectoine

The REA, r_{ss} and Thioflavin-T data clearly highlight the potency of Ectoine in arresting the aggregation of HEWL at pH 12.2.

5.1.9 Taurine

Taurine shows an anomalous effect on growth kinetics of HEWL oligomers at initial times (Figure 5.25). At 50 mM, its effect is most pronounced in halting oligomer growth, while at 400 mM no influence is seen on the growth kinetics until 25 hours. However, at 150 and later at 300 hours the trends are reversed.

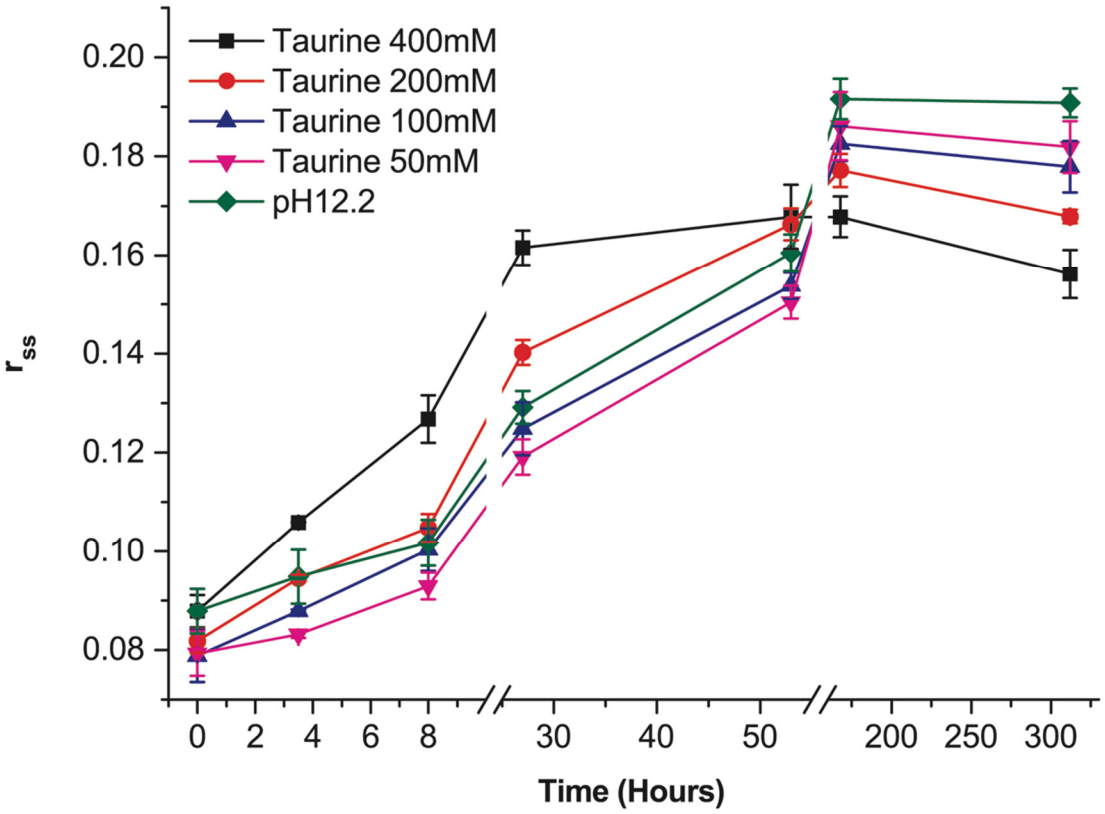


Figure 5.25: Kinetics of ss-Anisotropy in presence of indicated concentrations of Taurine

Figure 5.26 shows a significant dip in Thioflavin-T fluorescence in presence of 200 mM taurine. Further decrease is observed with 400 mM in comparison to Arginine at same concentration. Beyond 400 mM no significant change is noticeable.

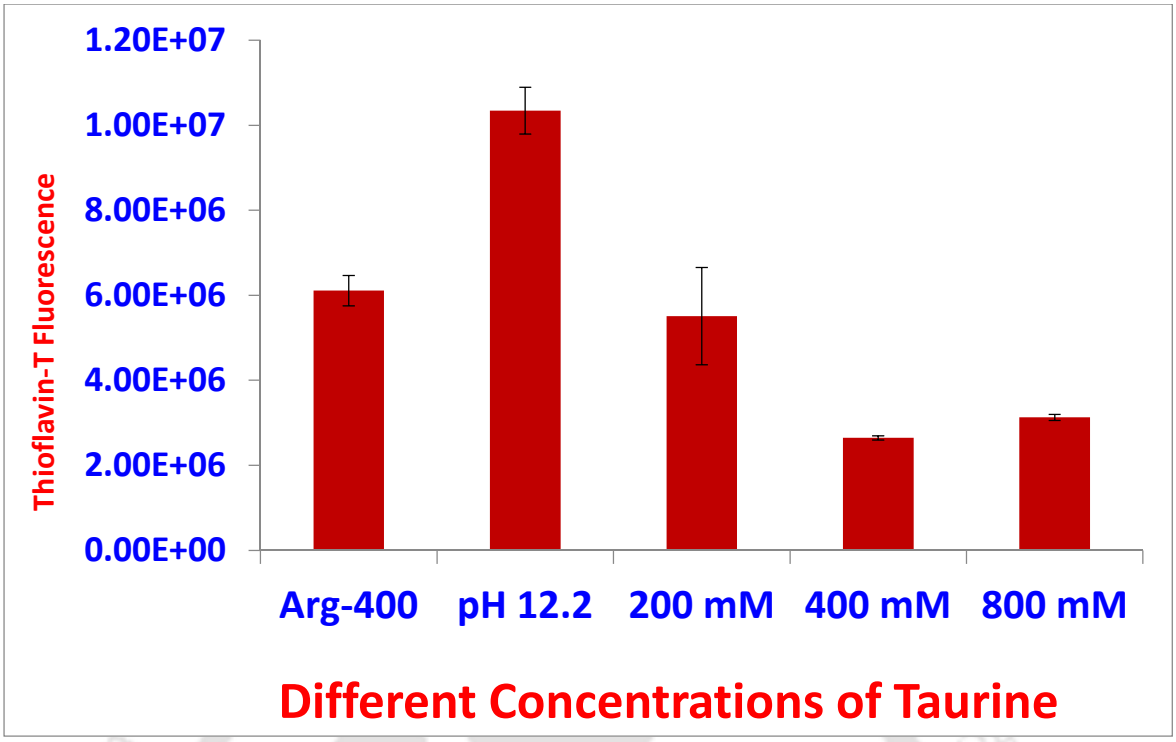


Figure 5.26: Th-T in presence of indicated concentrations of Taurine after 40 days

5.1.10 Betaine

Betaine occurs in plants under stressed conditions such as Drought or high salinity. REA with it lead to a concentration dependent protection (Figure 5.27).

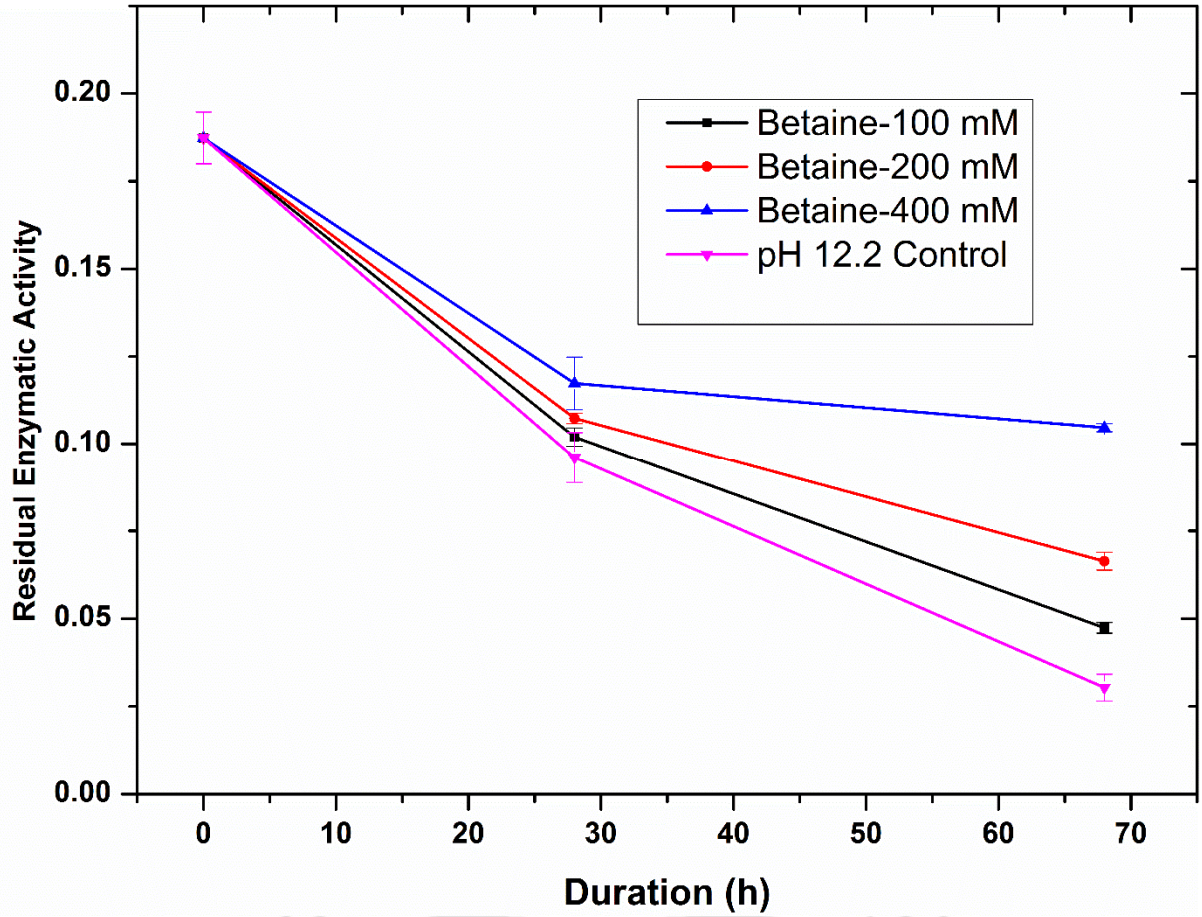


Figure 5.27: Kinetics of REA in presence of indicated concentrations of Betaine

Figure 5.28 shows that betaine has no significant influence in changing the amyloid population in HEWL incubated at pH 12.2 for 10 days. However, after 40 days of incubation effects are significant at 400 and 800 mM concentrations of Betaine.

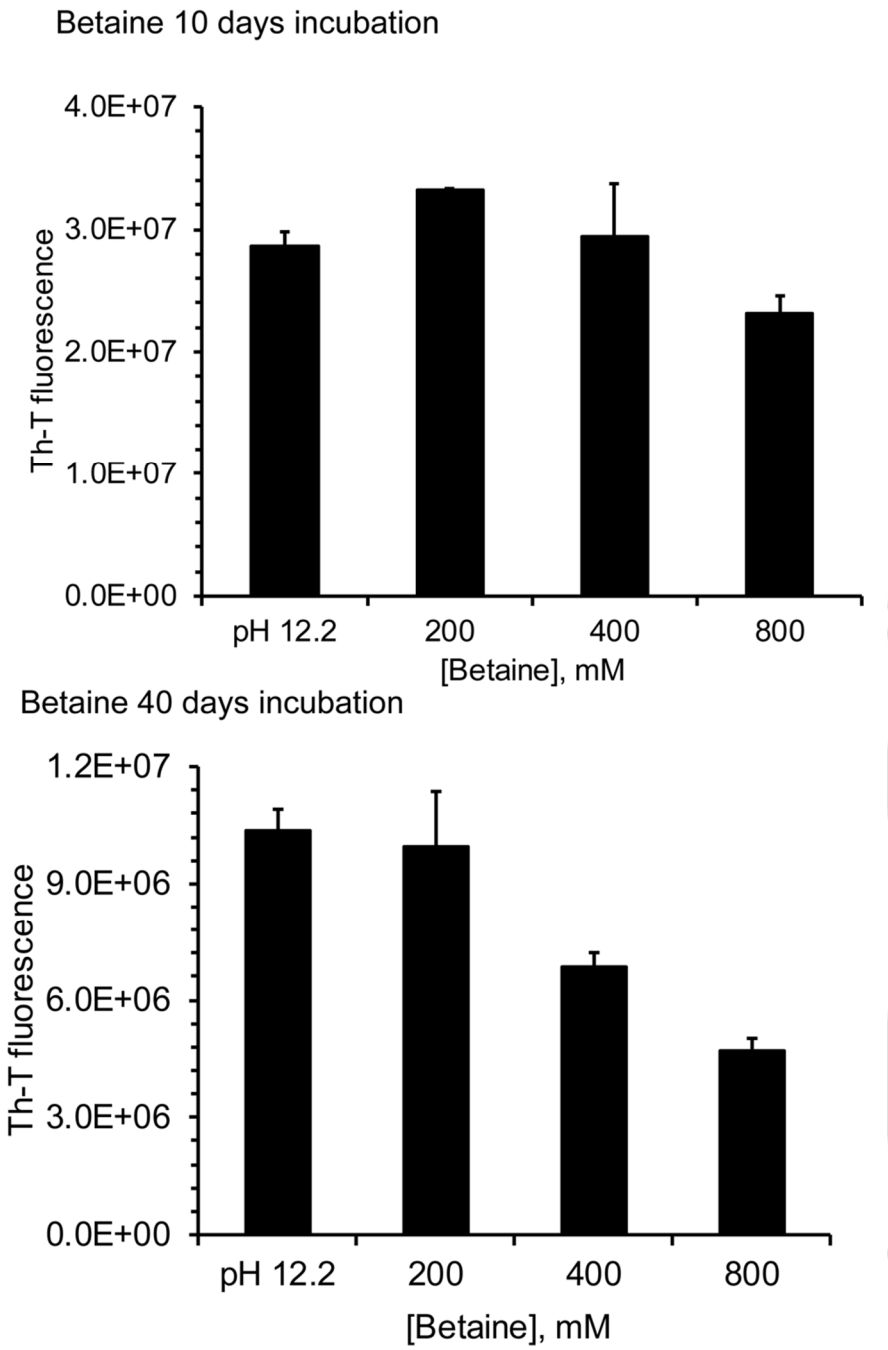


Figure 5.28: **Th-T in presence of indicated concentrations of Betaine after 10 days & 40 days [both experiments are from different batches and hence not plotted together]**

5.2 Mixture of Osmolytes (under Alkaline Condition)

We now try to investigate if a mixture of different osmolytes used in this thesis, taken in suitable proportion can exhibit synergistic affects. Table 5.2 shows a comparison of different compositions among **A**rginine, **S**permidine, **E**ctoine and **T**rehalose.

No. of Components	Arginine (mM)	Spermidine (mM)	Ectoine (mM)	Trehalose (mM)	Relative Fluorescence Intensity of Th-T
	pH 12.2 control				16.4
	pH 7.0 control				1.0
1	200	x	x	x	8.21
1	x	100	x	x	7.03
1	x	x	100	x	6.17
1	x	x	x	500	8.40
2	100	50	x	x	8.41
2	100	x	x	250	0.97
2	100	x	50	x	3.93
2	x	50	x	250	1.13
2	x	50	50	x	4.48
2	x	x	50	250	0.84
3	66	33	x	166	1.07
3	66	33	33	x	6.33
3	x	33	33	166	0.92
4	50	25	25	125	0.92

Table 5.2: Relative Thioflavin-T intensity of various permutations of the mixtures of four osmolytes,viz, Arginine, Spermidine, Ectoine & Trehalose [with 120 µM HEWL at pH 12.2, 25°C for 14-days]

Taken alone [mono component], to achieve near 50% reduction in Th-T fluorescence, approx. 200 mM Arginine OR 100 mM Spermidine OR 100 mM Ectoine OR 500 mM Trehalose is required. When we increase number of components, two effects are noticeable [apart from one significant finding already described that nearly every osmolyte has inhibitory effect especially at higher concentrations]. Effect of more than one components is more than additive if not less. This has potentially interesting ramifications if we see that each biological cell or extracellular matrix has hundreds of osmolytes even if at small concentrations. As an example, following table [given on next page] details composition of small molecules in E. *coli*. Cytoplasm [9].

Their additive effect is thus enough to offset amyloidogenic process and may even break pre-formed amyloids. The increasing potency of the mixture in comparison to single component, clearly indicates a synergistic effect in action of osmolytes when combined together.

Going by this hypothesis we propose a tonic, nothing but a cocktail of naturally occurring osmolytes as a therapy against amyloid diseases. Moreover, since all components of this putative tonic are naturally occurring as well as in small amounts, thus this can already be given to patients whose amyloid disease is in advanced state. Going one step further, we even feel that one reason of aging and age related diseases is that many of our body protein loose structure and function and as our work shows that osmolytes work even at first step of preserving the protein conformation so it could be plausible that such osmolyte cocktails can be administered to even healthy aged individuals to delay their age related diseases. However, for that residual enzymatic activity experiment with mixed osmolytes in needed to be done.

metabolite	mM	metabolite	mM
glutamate	96	S-adenosyl-L-methionine	0.18
glutathione	17	phosphoenolpyruvate	0.18
fructose-1,6-bisphosphate	15	threonine	0.18
ATP	9.6	FAD	0.17
UDP-N-acetyl-glucosamine	9.2	methionine	0.14
hexose-P	8.8	2,3-dihydroxybenzoicacid	0.14
UTP	8.3	NADPH	0.12
GTP	4.9	fumarate	0.11
dTTP	4.6	phenylpyruvate	0.090
aspartate	4.2	NADH	0.083
valine	4.0	N-acetyl-glucosamine-1P	0.082
glutamine	3.8	serine	0.068
6-phospho-D-gluconate	3.8	histidine	0.068
CTP	2.7	flavinmononucleotide	0.054
NAD	2.6	4-hydroxybenzoate	0.052
alanine	2.5	dGMP	0.051
UDP-glucose	2.5	glycerolphosphate	0.049
glutathionedisulfide	2.4	N-acetyl-ornithine	0.043
uridine	2.1	gluconate	0.042
citrate	2.0	malonyl-CoA	0.035
UDP	1.8	cyclic-AMP	0.035
malate	1.7	dCTP	0.034
3-phosphoglycerate	1.5	tyrosine	0.029
glycerate	1.4	inosine-diphosphate	0.024
coenzyme-A	1.4	GMP	0.024
citrulline	1.4	acetoacetyl-CoA	0.022
pentose-P	1.3	riboflavin	0.019
glucosamine-6_phosphate	1.2	phenylalanine	0.018
acetylphosphate	1.1	aconitate	0.016
gluconolactone	1.0	dATP	0.016
GDP	0.68	cytosine	0.014
acetyl-CoA	0.61	shikimate	0.014
carbamyl-aspartate	0.59	histidinol	0.013
succinate	0.57	tryptophan	0.012
arginine	0.57	dihydroorotate	0.012
UDP-glucaronate	0.57	quinolinate	0.012
ADP	0.55	ornithine	0.010
asparagine	0.51	dAMP	0.0088
2-ketoglutarate	0.44	adenosine-phosphosulfate	0.0066
lysine	0.40	myo-inositol	0.0057
proline	0.38	propionyl-CoA	0.0053
dTDP	0.38	ADP-glucose	0.0043
dihydroxyacetone-phosphate	0.37	anthranilate	0.0035
homocysteine	0.37	deoxyadenosine	0.0028
CMP	0.36	cytidine	0.0026
isoleucine+leuicine	0.30	NADP+	0.0021
deoxyribose-5-P	0.30	guanosine	0.0016
AMP	0.28	adenine	0.0015
inosine-monophosphate	0.27	deoxyguanosine	0.00052
PRPP	0.26	adenosine	0.00013
succinyl-CoA	0.23		
inosine-triphosphate	0.20	Sum	231
guanine	0.19		

Figure 5.29: Composition of a Typical Metabolome [9]

Another striking effect of this table is that of Trehalose. In whichever multi-component Trehalose is there as one component there is 100% reduction of Th-T fluorescence. Why this profound effect will be the focus of further investigations.

Recently a paper has appeared which describes synergistic effect of two osmolytes (Proline & Sorbitol) as inhibitor of aggregation of Lysozyme & Insulin [10]. But they have used only two osmolytes and only single amyloidogenic condition of Lysozyme (viz. Guandinium Hydrochloride condition).

Apart from the application point of view, this mixed osmolyte experiment tests mechanistic aspects. There is a paper which suggests that arginine exerts its effect by making supramolecular assemblies e. However, such supramolecular assemblies are unlikely is mixed osmolyte case.

In present case we have studied combined effect of only four osmolytes. In future we would like to study combinations of more than four osmolytes but using more osmolytes together will make so many combinations that they will be difficult to study in a single experiment especially time consuming experiments like residual enzymatic activity.

One experiment is proposed wherein cytoplasm (especially from extremophiles) is used instead of mixture of osmolytes.

5.3 Osmolytes at acidic pH condition

Since the idea behind the thesis was to generalize the findings and keep them free from amyloidogenic condition used. Therefore, we just did not stick to alkaline condition and were open for other conditions especially with acidic condition which has got

maximum attention in literature. However, as the following table shows, the reaction mixtures were not as well behaved as in alkaline condition. In alkaline condition for all osmolytes [even at very high concentrations] the liquid was homogeneous, free from any particulate material and free flowing [not viscous] and that too from beginning till forever. The same was not found to be the case with acidic condition except for all concentrations of Taurine & Trehalose and 50 mM of Ectoine as the flowing table shows.

S. No.	Osmolyte	Physical appearance of the sample reaction mixture
1	**pH 7.0**	**heavy amorphous ppt**
2	**pH 2.0**	**very weak gel**
3	**Arginine-50mM**	weak gel with little ppt (long needle like crystals)
4	**Arginine-100mM**	crystalline gel
5	**Arginine-200 mM**	crystalline (more) gel
6	**Arginine-400 mM**	ppt
7	**Spermidine-50mM**	very strong gel (almost transparent)
8	**Spermidine-100mM**	crystalline gel
9	**Spermidine-200mM**	very strong crystalline gel
10	**Spermidine-400mM**	heavy amorphous ppt
11	**Ectoine-50mM**	**clear solution**
12	**Ectoine-100mM**	transparent solution with few long needle like crystals
13	**Ectoine-200mM**	solution with crystalline ppt

14	**Ectoine-400 mM**	solution with crystalline ppt
15	**TMAO-50mM**	solution with few needle like crystals
16	**TMAO-100mM**	solution with more crystals
17	**TMAO-200mM**	Solution with more crystals
18	**TMAO-400mM**	solution with more crystals
19	**Taurine-50mM**	**clear solution**
20	**Taurine-100mM**	**clear solution**
21	**Taurine-200mM**	**clear solution**
22	**Taurine-400mM**	**clear solution**
23	**Trehalose-500mM**	**clear solution**
24	**Trehalose-1000mM**	**clear solution**

Table 5.3: Physical appearance of the reaction mixtures containing different Osmolytes after 7- days under acidic pH (pH 2.0, 57°C, 600 µM HEWL]. The concentrations of Osmolytes (in mM) are indicated after the name (separated by a Dash)

For this reason, only Trehalose was studied in detail. Studies with Taurine is future plan. Trehalose has emerged as a promising molecule for pharmaceutical interventions and has already been used in *in vivo* work [4] that too at high concentrations. Prognosis however is that if we do indeed wish to use osmolytes under acidic condition, we have to make a departure from STANDARD CONDITION [viz., 600 µM HEWL, 57°C, pH 2.0] which is not amenable for such studies. In that case, we can either lower HEWL concentration OR lower the pH so that temperature is brought down. Perhaps this will yield a case with fewer precipitates though incubation period may go up.

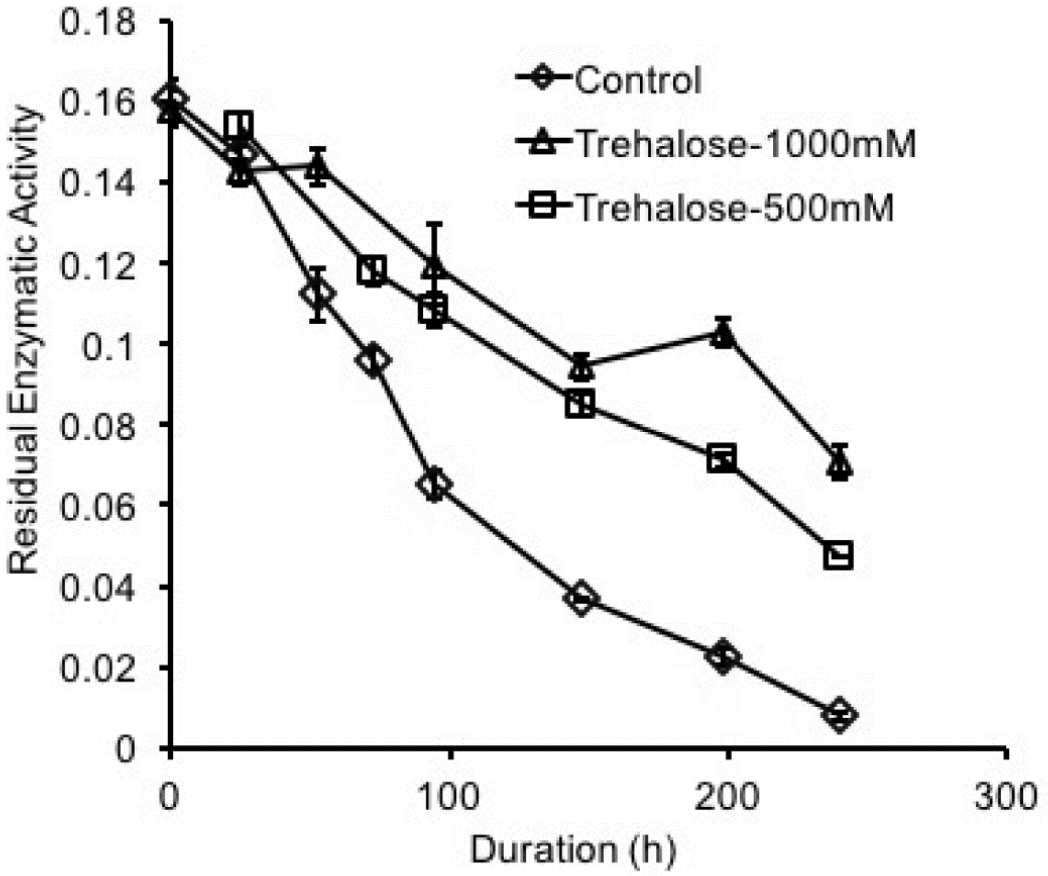


Figure 5.30: Kinetics of REA under acidic condition for indicated concentrations of Trehalose

Figure 5.30 shows that Trehalose at 500 mM and at also 1000 mM are effective in preserving HEWL activity in acidic pH in comparison to control. At 500 mM it is only marginally better but at 1000 mM it is dramatically superior. What is however most striking being the similarity of Trehalose behaviour between Alkaline and Acidic condition(s).

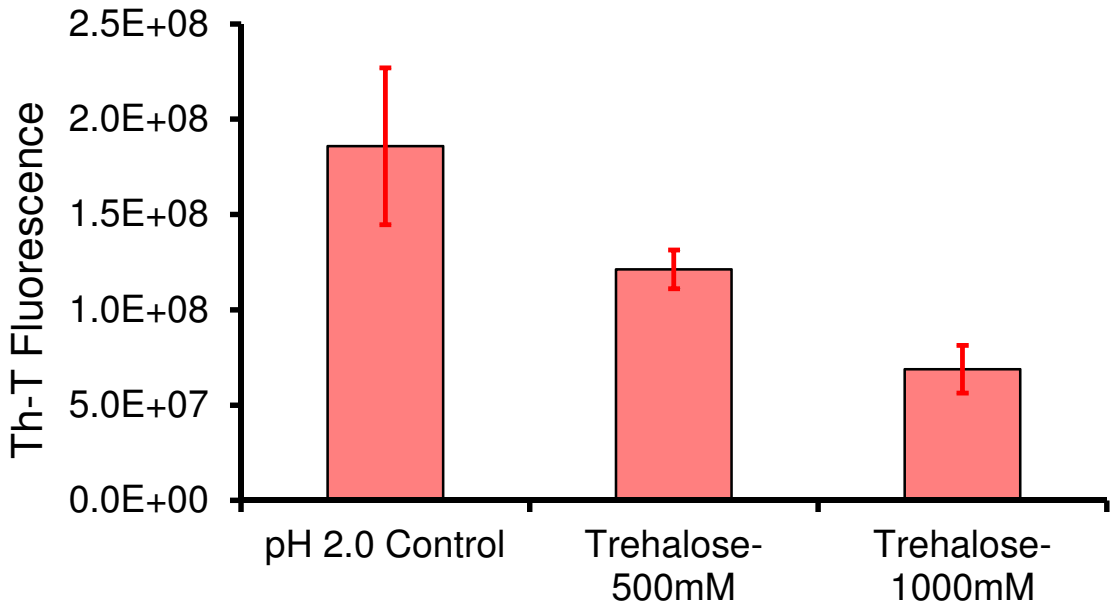


Figure-5.31: Th-T under acidic condition for indicated concentrations of Trehalose after incubation for 21-days

Figure-5.31 shows that Trehalose both at 500 and 1000 mM concentrations (after 21 days of incubation) brings about a significant reduction in Thioflavin-T fluorescence intensity in HEWL samples incubated at pH 2.0. This hints at significant reduction in amyloid population in presence of Trehalose. Similarity of Trehalose data with alkaline as well as acidic condition point towards generality of osmolyte behaviour. This generality/non-generality was the reason to work with many different orthogonal conditions so that focus can be kept on osmolytes and not on some mechanism of one chosen amyloidogenic condition.

Conclusions

Results with osmolyte clearly indicate that osmolytes have much varied and complicated relationship with proteins than previously assumed. There is no all-encompassing 'Osmophobic effect'. Result with mixed osmolyte might have significant potential about medical applications.

References:

1 Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. & Somero, G. N. Living With Water-Stress - Evolution Of Osmolyte Systems. *Science (New York, N.Y.)* 217, 1214-1222, doi:10.1126/science.7112124 (**1982**).

2 Yancey, P. H. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *The Journal of experimental biology* 208, 2819-2830, doi:10.1242/jeb.01730 (**2005**).

3 Homchaudhuri, L., Kumar, S. & Swaminathan, R. Slow aggregation of lysozyme in alkaline pH monitored in real time employing the fluorescence anisotropy of covalently labelled dansyl probe. *FEBS letters* 580, 2097-2101, doi:10.1016/j.febslet.2006.03.012 (**2006**).

4 Tanaka, M. *et al.* Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease. *Nature medicine* 10, 148-154, doi:10.1038/nm985 (**2004**).

5 Yancey, P. H., Gerringer, M. E., Drazen, J. C., Rowden, A. A. & Jamieson, A. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proceedings of the National Academy of Sciences of the United States of America* 111, 4461-4465, doi:10.1073/pnas.1322003111 (**2014**).

6 Schuh, W., Puff, H., Galinski, E. A. & Truper, H. G. THE CRYSTAL-STRUCTURE OF ECTOINE, A NOVEL AMINO-ACID OF POTENTIAL OSMOREGULATORY FUNCTION. *Z.Naturforsch.(C)* 40, 780-784 (**1985**).

7 Kanapathipillai, M., Lentzen, G., Sierks, M. & Park, C. B. Ectoine and hydroxyectoine inhibit aggregation and neurotoxicity of Alzheimer's beta-amyloid. *FEBS letters* 579, 4775-4780, doi:10.1016/j.febslet.2005.07.057 (**2005**).

8 Arora, A., Ha, C. & Park, C. B. Inhibition of insulin amyloid formation by small stress molecules. *FEBS letters* 564, 121-125, doi:10.1016/s0014-5793(04)00326-6 (**2004**).

9 Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature chemical biology* 5, 593-599, doi:10.1038/nchembio.186 (**2009**).

10 Choudhary, S., Save, S. N., Kishore, N. & Hosur, R. V. Synergistic Inhibition of Protein Fibrillation by Proline and Sorbitol: Biophysical Investigations. *PloS one* 11, e0166487, doi:10.1371/journal.pone.0166487 (**2016**).

Chapter-6

Numeric Code for Amino Acids and its Applications

As mentioned in "Thesis Objectives" [quoted below] an easy way to memorize and comprehend sequence of HEWL was sought and resulted into a Numeric Code which is described below. Later on it was found that it has applications with regard to various 'Amino Acid Scales'. As a proof of concept, this has been applied to transmembrane tendency of Human Rhodopsin.

"It became obvious in the beginning of my experimental work with Lysozyme that I would be working exclusively on this protein during entire tenure of my PhD. As a curiosity it occurred to me that If I am going to spend so many years on a single protein which is not very long, why not memorize its primary sequence. The idea was that it could help in some way to analyse the results. This has been shown that Lysozyme at acidic pH/elevated temperature starts breaking up into pieces (Frare, Mossuto et al. 2006). This was done at a single time point but it can be expected that multiple time points would yield several overlapping pieces of increasingly smaller sizes. The same can be expected at alkaline pH. In the presence of different osmolytes, this breaking down patterns can be different. This sort of breakdown analysis was a research plan which we wanted to but could not be pursue [and now it is a future plan]. But had it been done, it would have been facilitated with memorization of sequence. Off course one can take recourse to bioinformatics software which everyone does. But we reasoned that having imprinted sequence in our mind will help us do the thing which a computer can't. I started the efforts by beginning with alphabetic code. But it turned out to be extremely difficult. Then came the idea of numeric code. Once the idea sparked, we found out that it is not merely useful in memorizing the primary sequence but also has applications in other domain of 'Amino Acid Scale'."

A) Facile memorization of primary and secondary structure of proteins

A novel way of protein sequence coding is proposed. It has two advantages over conventional coding (one letter alphabets). On the outset this seems to be a useless exercise but it has its own advantages. This is based on numbers (**0** - **9**) and their underline counterparts (**0** - **9**). It is expected that this can lead to divide long protein sequences into small sequences of numbers (and their underlined counterparts) which will then help to memorize primary and secondary structures. If we can somehow find a way to connect this to tertiary structure (by contact map information or something else) then essentially we can have photographic memory of protein structure imprinted on brain. This memorization is extremely difficult with present alphabetic one letter code because we are naturally good at remembering numbers. Moreover, with mathematics being language of science, numbers are more scientific than alphabets. Numbers are easier to comprehend than alphabets. Even a common man used to remember many 10-digit phone numbers before mobile revolution. There are many professions in which professionals remember large set of numbers. Those in professions such as PCO's owners used to remember large number of STD codes or railway reservation or enquiry counter people

remember train numbers of large number of trains. In a recent movie "Life of Pi' a small boy memorizes a very large sequence of numbers.

Questions might be raised as to why memorize sequences. But for those working on Biophysics of a small protein it will probably be better to do so since it will take not more than a couple of weeks to do so (simple program can be written which will transform 1-letter code to numeric code, or in absence of it a simple 'Find and Replace' option of MS-WORD will do). Strings of numbers are easier to memorize than string of letters because making words of short sequences often makes no sense more so because out of 5-vowls of A, E, I, O, U; O & U don't correspond to any amino acid nor does vowel 'I' occur frequently enough thus leaving only A & E as only vowels used frequently. With such shortage of vowels; making meaningful words is very difficult. It is anticipated that entire primary and secondary structure (especially for shorter peptide) can be memorized quickly and with excellent recall after long times.

Scanning large sequences with speed and perfection is not easy task but such scanning with numeric code is very easy. An easy psychological experiment is proposed to be performed wherein 10+2 students (in 2-groups of 20 or so with similar academic standing) are asked to memorize numeric or alphabetic code each and see which group memorizes faster and which has better recall say, after one month.

Only problem is that whereas there are 20-AA's but with single digits we have only 10 of them (0-9). Nevertheless, we are at the threshold of luck that we have 20-AA's so remaining 10 can be taken as underline of 0-9. This way we can express all AA's. Had there been 21-AA's the approach would have failed.

Strings can be memorized in short sequences which often happens to be case with coil and beta-sheets. We can split memorization process into two parts.

1. Strings of integers
2. Then putting 'underline' underneath some of them (and make them red coloured)

Here it needs to be mentioned that occurrence of underlined numbers can be reduced significantly if going by 'frequency of occurrence criteria in each doublet'. We can allot underlined number to lesser occurring AA's (though this has not been done in this thesis), e.g., frequently occurring G as 0 and less occurring W as **0**.

Incredibly large number of sequences are being deposited in PDB. It is envisioned that if we have huge computer display covering nearly entire wall of a large room and displaying

nothing but numbers corresponding to numeric code, we can decipher patterns which even a computer can't. In a way this exercise is analogous to code-breaking.

The way pairing has been done in this thesis is fairly rationale but can be made more rationale if do pairing of AA's in such a way that they are nearest mutational partners.

This approach does not wish to replace present alphabetic approach but merely to 'transform' sequences of proteins of interest using a simple program [even a program is not mandatory, a simple 'Find and Replace option in MS-WORD will do] just prior to any memorization, cross-comparison, looking for sequence homology, looking for particular AA (this approach is likely to be more beneficial when 2 or more AA's are being looked simultaneously.

Now after putting 'underlines' and secondary structure element, these numbers have been grouped as single digit, double digit or triple digit. There are quite a few books which illustrate how strings of numbers can be memorized quickly and effectively using various types of associations which are not possible with alphabets. In a way numbers are by far the best and readymade 'mnemonics' whereas alphabets are not.

Table 6.1: The proposed scheme is as follows

1-Letter Alphabetic Code	Numeric Code	Criteria of assigning Numeric Code
G	0	Smallest and largest AA with frequently occurring G as 0
W	0	
A	1	Sequential aliphatic side-chains
V	1	
L	2	Isomers
I	2	
F	3	Sequential aromatic AA's
Y	3	
D	4	Carboxylic acid side-chains
E	4	
N	5	Amide side-chains
Q	5	
R	6	Residues with positive charges
K	6	
S	7	Residues with –OH groups
T	7	
C	8	Residues containing S
M	8	
H	9	Residues with 5-membered rings
P	9	

Using above coding scheme (which is arbitrary), Numeric Code for Human Lysozyme is as under:

Table-6.2: Numeric Code of Human Lysozyme

Positions of Residues	Number of residues	Secondary structure	Residues [One letter alphabetic code]	Residues [Numeric Code]
1-4	4	Random Coil	KVFE	6 1 3 4
5-14	10	Helix	RCELARTLKR	6 8 4 2 1 6 7 2 6 6
15-24	10	Helix	LGMDGYRGIS	2 0 8 4 0 3 6 0 2 7
25-36	12	Helix	LANWMCLAKWES	2 1 5 0 8 8 2 1 6 0 4 7
37-42	6	Random Coil	GYNTRA	0 3 5 7 6 1
43-45	3	Beta Strand	TNY	7 5 3
46	1	Random Coil	N	5
47-49	3	Turn	AGD	1 0 4
50-51	2	Random Coil	RS	6 7
52-54	3	Beta Strand	TDY	7 4 3
55-58	4	Turn	GIFQ	0 2 3 5
59-60	2	Random Coil	IN	2 5
61-63	3	Turn	SRY	7 6 3
64-66	3	Beta Strand	WCN	0 8 5
67-69	3	Random Coil	DGK	4 0 6

70-72	3	Beta Strand	TPG	7 9 0
73-80	8	Random Coil	AVNACHLS	1 1 5 1 8 9 2 7
81-85	5	Helix	CSALL	8 7 1 2 2
86-88	3	Beta Strand	QDN	5 4 5
89	1	Random Coil	I	2
90-101	12	Helix	ADAVACAKRVVR	1 4 1 1 1 8 1 6 6 1 1 6
102-104	3	Random Coil	DPQ	4 9 5
105-108	4	Helix	GIRA	0 2 6 1
109	1	Random Coil	W	0
110-115	6	Helix	VAWRNR	1 1 0 6 5 6
116-118	3	Turn	CQN	8 5 5
119-121	3	Random Coil	RDV	6 4 1
122-125	4	Helix	RQYV	6 5 3 1
126-130	5	Random Coil	QGCGV	5 0 8 0 1

So far numeric code was applied to Lysozyme from Homo sapiens only. In next page we have listed first 63 residues of Lysozyme from seven different species. Limitation of A-4 papers restricts us to 63 residues. On a large display such as a 1-metre long poster entire sequence of even a 500 amino acid protein can be put and on Y-axis same from many species can be put. Our argument is that such display can help us to find patterns which a computer can't find (much like code breaking).

A,B,C etc. in the following refers to Species as given below

A) Sus scrofa (wild boar, Eurasian wild pig)
B) Rattus norvegicus (brown rat, common rat, street rat, sewer rat)
C) Mus musculus (house mouse, laboratory mouse)
D) Pan troglodytes (common chimpanzee)
E) Homo sapiens (Human)
F) Danio rerio (Zebrafish)
G) Gallus gallus (Chicken)

A) 6 1 3 4 6 8 4 3 1 6 2 2 6 6 7 0 8 4 0 3 6 0 1 7 2 1 5 0 1 8 2 1 6 0 4 7 4 3 5 7 6 1 2 5 9 5 1 0 7 7 4 3 0 2 3 5 2 5 7 6 3 0

B) 6 2 3 4 6 8 4 3 1 6 7 2 6 6 5 0 8 7 0 3 3 0 1 7 2 1 4 0 1 8 2 1 5 9 4 7 5 3 5 7 5 1 6 5 3 5 9 0 4 5 7 7 4 3 0 2 3 5 2 5 7 6

C) 6 1 3 5 6 8 4 2 1 6 2 2 6 6 5 0 8 4 0 3 6 0 1 6 2 1 4 0 1 8 2 1 5 9 4 7 5 3 5 7 6 1 7 5 3 5 6 0 4 6 7 7 4 3 0 2 3 5 2 5 7 6

D) 6 1 3 4 6 8 4 2 1 6 7 2 6 6 2 0 8 4 0 3 6 0 2 7 2 1 5 0 8 8 2 1 6 0 4 7 0 3 5 7 6 1 7 5 3 5 1 0 4 6 7 7 4 3 0 2 3 5 2 5 7 6

E) 6 1 3 4 6 8 4 2 1 6 7 2 6 6 2 0 8 4 0 3 6 0 2 7 2 1 5 0 8 8 2 1 6 0 4 7 0 3 5 7 6 1 7 5 3 5 1 0 4 6 7 7 4 3 0 2 3 5 2 5 7 6

F) 6 7 2 0 6 8 4 1 3 6 2 3 6 5 4 0 2 4 0 3 4 0 3 7 2 0 5 3 1 8 7 1 3 0 4 7 6 3 6 7 9 6 1 6 7 1 4 7 0 6 4 3 0 2 3 5 2 5 7 3 6 0

G) 6 1 3 0 6 8 4 2 1 1 1 8 6 6 9 0 2 4 5 3 6 0 3 7 2 0 5 0 1 8 1 1 6 3 4 7 5 3 5 7 5 1 7 5 6 5 7 4 0 7 7 4 7 0 2 2 5 2 5 7 6 0

Figure-6.1: Numeric Code as applied to first 63-residues of LYSOZYME from seven species as defined on top of the figure

B) Visual Bioinformatics Analysis

A new numeric amino acid coding scheme is proposed for visual bioinformatics analysis. This scheme has advantage, that it can allow bioinformatics analysis to be done without computers. Since brain is by far the best computer in doing abstract thinking so it is expected that we can find patterns which a computer can't [until we incorporate it in computer's logic]. Thus such analysis can be forerunner to bioinformatics programming. The idea is same as described above, i.e. according a number instead of one letter alphabetic code. As a proof of concept, this time we will apply it to transmembrane portions of a protein Human Rhodopsin [Human Rhodopsin [NCBI Reference Sequence: NP_000530.1, UniProt Code: P08100] which is a 348 amino acids protein.

There are many Hydrophobicity scales which are used in predicting transmembrane portions of transmembrane proteins (Creighton 1993). More recently a scale has even come about "Amino acid scale: Transmembrane tendency" (Zhao and London 2006). Since this scale is of most recent origin and directly used for transmembrane tendency so it has been adopted in this chapter.

We propose that numeric amino acid coding scheme is better than Hydrophobicity/ Hydropathy plots/transmembrane tendency widely used in literature. While hydrophobicity plot predicts membrane portions graphically but it masks the identity of residues; whereas in our approach even identity is displayed, allowing us to use our mental faculties and make more conclusions than what hydropathy plot allows,

Table 6.3: Scheme of Numeric Code for Transmembrane Tendency

[Reference for this Table (except Numeric Code) is from (Zhao and London 2006)]

3-letter Alphabetic Code	1-Letter Alphabetic Code	Genomic Hydrophobicity (GH) Scale	Transmembrane (TM) Tendency Scale	Numeric Code for Transmembrane (TM) Tendency
Lys	K	-2.53	-3.46	0
Asp	D	-3.11	-3.27	0
Glu	E	-3.66	-2.90	1
Arg	R	-1.91	-2.57	1
Gln	Q	-1.68	-1.84	2
Asn	N	-0.98	-1.62	2
Pro	P	-0.25	-1.44	3
His	H	-2.22	-1.44	3
Ser	S	0.73	-0.53	4
Thr	T	-0.27	-0.32	4
Cys	C	-0.08	-0.30	5
Gly	G	0.00	-0.19	5
Ala	A	0.15	0.38	6
Tyr	Y	-1.02	0.49	6
Met	M	1.11	1.40	7
Val	V	1.11	1.46	7
Trp	W	1.00	1.53	8
Leu	L	1.68	1.82	8
Ile	I	1.72	1.97	9
Phe	F	2.00	1.98	9

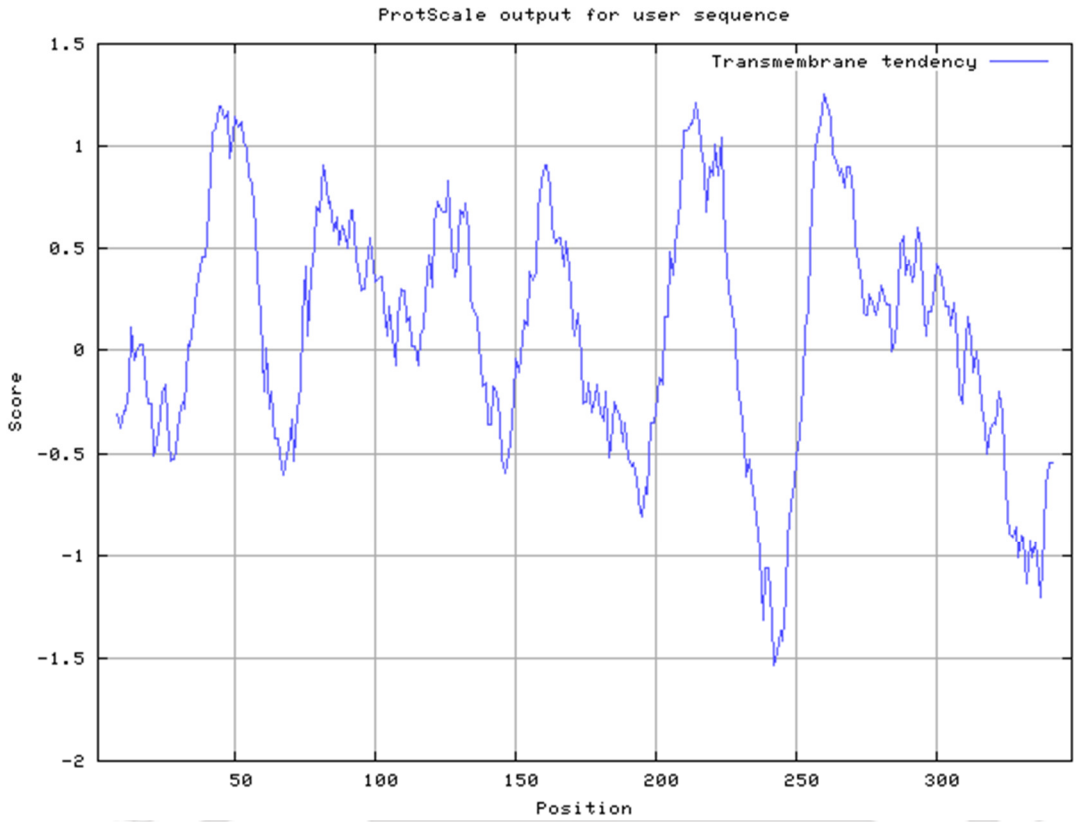


Figure 6.2 Transmembrane tendency of Human Rhodopsin as generated from http://web.expasy.org/protscale/ [with window size of 15]

While above figure protrudes out transmembrane regions [out of seven such regions, six are clearly visible]; but it completely masks identity of amino acids. However, when attempted with numeric code [as done on next pages], the same is revealed.

Table 6.4: Topology along with Numeric Code for Human Rhodopsin

Location Of Segments	Position (s) of Residues	No. of Resi-dues	Alphabetic Code of Segment	Numeric Code of Segment
Extracellular (EC-1)	**1 – 36**	36	MNGTEGPNFYVPFSNATGVVR SPFEYPQYYLAEPWQ	7 2 5 4 1 5 3 2 9 6 7 3 9 4 2 6 4 5 7 7 1 4 3 9 1 6 3 2 6 6 8 6 1 3 8 2
Transmembrane (TM-1)	**37 – 61**	25	FSMLAAYMFLLIVLGFPINFL TLYV	9 4 7 8 6 6 6 7 9 8 8 9 7 8 5 9 3 9 2 9 8 4 8 6 7
Cytoplasmic (CP-1)	**62 – 73**	12	TVQHKKLRTPLN	4 7 2 3 0 0 8 1 4 3 8 2
Transmembrane (TM-2)	**74 – 98**	25	YILLNLAVADLFMVLGGFTST LYTS	6 9 8 8 2 8 6 7 6 0 8 9 7 7 8 5 5 9 4 4 4 8 6 4 4
Extracellular (EC-2)	**99 – 113**	15	LHGYFVFGPTGCNLE	8 3 5 6 9 7 9 5 3 4 5 5 2 8 1
Transmembrane (TM-3)	**114 – 133**	20	GFFATLGGEIALWSLVVLAI	5 9 9 6 4 8 5 5 1 9 6 8 8 4 8 7 7 8 6 9
Cytoplasmic (CP-2)	**134 – 152**	19	ERYVVVCKPMSNFRFGENH	1 1 6 7 7 7 5 0 3 7 4 2 9 1 9 5 1 2 3
Transmembrane (TM-4)	**153 – 176**	24	AIMGVAFTWVMALACAAPPLA GWS	6 9 7 5 7 6 9 4 8 7 7 6 8 6 5 6 6 3 3 8 6 5 8 4
Extracellular (EC-3)	**177 – 202**	26	RYIPEGLQCSCGIDYYTLKPE VNNES	1 6 9 3 1 5 8 2 5 4 5 5 9 0 6 6 4 8 0 3 1 7 2 2 1 4

Transmembrane (TM-5)	**203 – 230**	28	FVIYMFVVHFTIPMIIIFFCY GQLVFTV	9 7 9 6 7 9 7 7 3 9 4 9 3 7 9 9 9 9 9 5 6 5 2 8 7 9 4 7
Cytoplasmic (CP-3)	**231 – 252**	22	KEAAAQQQESATTQKAEKEVT R	0 1 6 6 6 2 2 2 1 4 6 4 4 2 0 6 1 0 1 7 4 1
Transmembrane (TM-6)	**253 – 276**	24	MVIIMVIAFLICWVPYASVAF YIF	7 7 9 9 7 7 9 6 9 8 9 5 8 7 3 6 6 4 7 6 9 6 9 9
Extracellular (EC-4)	**277 – 284**	8	THQGSNFG	4 3 2 5 4 2 9 5
Transmembrane (TM-7)	**285 – 309**	25	PIFMTIPAFFAKSAAIYNPVI YIMM	3 9 9 7 4 9 3 6 9 9 6 0 4 6 6 9 6 2 3 7 9 6 9 7 7
Cytoplasmic (CP-4)	**310 – 348**	39	NKQFRNCMLTTICCGKNPLGD DEASATVSKTETSQVAPA	2 0 2 9 1 2 5 7 8 4 4 9 5 5 5 0 2 3 8 5 0 0 1 6 4 6 4 7 4 0 4 1 4 4 2 7 6 3 6

In the above Numeric Code, in the Transmembrane segments, residues Gly (5), Ala (6), Tyr (6), Met (7), Val (7), Trp (8), Leu (8), Ile (9) & Phe (9) are considered REGULAR as far as Transmembrane (TM) Tendency is considered [their values on TM Scale varies from -0.19 to 1.98]. Ser (4), Thr (4) and Cys (5) with TM Tendency values [-0.53 to -0.30] are considered MARGINALLY IRREGULAR and marked in Red. All other residues are considered HIGHLY IRREGULAR and have been marked in Red as well as Highlighted in Yellow. This is not really arbitrary. MARGINALLY IRREULAR residues are only slightly more unsuitable compared to Glycine [for which Global Hydrophobicity value in this paper is 0.00, and actually zero in many other Hydrophobicity Scales] as far as TM Tendency Scale is considered. Whereas the next ones have considerably lower values

Done this way we observe the following:

Table-6.5: Analysis of Transmembrane Segments

S. No.	Transmembrane Segments (TM)	Total No. of Residues in the Segment	No. of Marginally Irregular Residues (marked in Red in the above Table) & (Their Percentage with respect to Total number of Residues)	No. of Highly Irregular Residues (marked in Yellow as well as Highlighted in Green in the above Table) & (Their Percentage with respect to Total number of Residues)
1	Transmembrane-1 (TM-1)	25	2 (8%)	1 (4%)
2	Transmembrane-2 (TM-2)	25	5 (20%)	2 (8%)
3	Transmembrane-3 (TM-3)	20	2 (10%)	0 (0%)
4	Transmembrane-4 (TM-4)	24	2 (8.3%)	2 (8.3%)
5	Transmembrane-5 (TM-5)	28	2 (7.1%)	3 (10.7%)
6	Transmembrane-6 (TM-6)	24	2 (8.3%)	1 (4.2%)
7	Transmembrane-7 (TM-7)	25	2 (8%)	4 (16%)

One out of context observation is that length of all transmembrane segments is between 20-28 [indeed out of 7-segments, length of 5-segments is 24-25]. Such sharp number indicates that they have similar secondary and tertiary structure while they traverse through lipid bilayer.

From the above table we can see that barring one case, Number of Marginally Irregular residues is below 10% and same for Highly Irregular Residues is even lesser. This sort of analysis may help us design proteins if we wish to further stabilize interaction of transmembrane segment to lipid bilayer or vice-versa. This concept can even be applied to proteins which are not transmembrane but whose segments are partly embedded into lipid bilayers.

With this approach, we can make reasonable guess about segmental or overall hydrophobicity merely by looking at sequence.

What we have shown is only one area of Bioinformatics, viz., hydrophobicity / hydropathy/ transmembrane tendency plots; the same can be applied in the area of secondary structure prediction. Over there we need the scale in increasing/decreasing order of 'Relative Helical Tendencies

Eventually we can generalize this concept and apply it to any Amino Acid Scale. An **amino acid scale** is defined by a numerical value assigned to each type of amino acid. The most frequently used scales are the hydrophobicity or hydrophilicity scales and the secondary structure conformational parameters scales, but many other scales exist which are based on different chemical and physical properties of the amino acids. A program [hosted at http://web.expasy.org/protscale/] provides 57 predefined scales entered from the literature.

References:

Creighton, T. E. (**1993**). Proteins: Structures and Molecular Properties, W. H. Freeman and Company.

Frare, E., M. F. Mossuto, P. Polverino de Laureto, M. Dumoulin, C. M. Dobson and A. Fontana (**2006**). "Identification of the core structure of lysozyme amyloid fibrils by proteolysis." J Mol Biol 361(3): 551-561.

Zhao, G. and E. London (**2006**). "An amino acid "transmembrane tendency" scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: relationship to biological hydrophobicity." Protein Sci 15(8): 1987-2001

Chapter-7

Discussion & Future Directions

7.1 Matrix Formalism of Kinetics Data

Rather than an end in itself, this thesis is more like a proof of concept. Concept here is the requirement of a curated and mathematically intense data. This is pre-requisite if any theoretical formalism is attempted. It was felt that the research field of "inhibition of protein aggregation with small molecules" is fragmented. People have been using combination of different protein(s) [including those not known to form amyloid under *in vivo* conditions] and different sets of small molecule(s). Moreover, most of the studies have centred around few molecules such as TMAO, Betaine and Proline etc. This makes cross-comparison difficult. It was felt that if we fix the protein [that too a protein which is innately amyloidogenic as well as for which a wealth of biophysical data is available] and pit not just a couple but as many as possible osmolytes against it. Another drawback of papers is to use just one *in vitro* amyloidogenic condition even though many such conditions are known. This thesis also touches the cornerstone of many diverse fields, one being as old a field as Hofmeister series and another a new field of metabolomics. Yet another advantage with Lysozyme is that it is an enzyme so even the step of protein unfolding can be studied with simple enzyme assay.

So for I have used merely three techniques. However, I am satisfied to some extent owing to following reasons.

1. Use of techniques in 'kinetics' manner as rigorously as we could anticipate and execute with limitations such as instrument's availability in a shared scenario and as much as I could do with solo efforts (this 'Kinetics' is a major pillar of Physical Chemistry).
2. Demonstrate the power of residual enzymatic activity vis-à-vis first step of conformational fluctuations/unfolding [something not available with any other naturally amyloid forming protein] with significant reduction in error bars.
3. Simultaneous use of steady state fluorescence anisotropy and Thioflavin-T to decipher the message regarding oligomerization to structural reorganization.

Moreover, in these studies a wide range of concentrations were used for each Osmolyte as this will help to determine 'order' of the steps. Plans are underway to use repertoire of more biophysical techniques (but continuing in the form of Kinetics rather than discreet

and keeping Δt as short as possible, avoid big jump between two time points, and keeping duration of overall experiments to be fairly long). Since most of the data is 'Kinetics' in nature, hence an attempt is being made to tabulate all this data in the form of a mathematical matrix. This is expected to be useful in many ways. First of all, this will facilitate a very concise presentation. These will therefore facilitate easy comparison across each osmolyte which appears daunting if attempted merely with diagrams. Basically a [m x n] matrix where 'm' is number of kinetic data points and 'n' is the number of methods employed, and present guideline of our research is to make both 'm' & 'n' large.

However, we must make it clear that mainly due to being sole experimental hand and partially due to late realization of aforementioned idea we don't have as many time points as we wish we should have had. We have then tried to be able to reduce error bars to the extent that such numerization makes sense. So this idea is still in infancy.

The grand aim of this matrix (when made sufficiently elaborate to satisfy an experimentalist) would be to be able to superimpose snap shots of computer simulations onto this 'experimental space' to judge accuracy of simulation and also to guide simulation in right direction. It goes without saying that unless this desired synthesis of rigorous experimentation and simulation takes place, 'understanding resolution' would be low.

The real potential for this matrix approach is when techniques are automated. Techniques like REA, Th-T and r_{ss} are easily amenable to robotics. With such automation, huge amount of data can be generated for same reaction mixture. We have already seen in thesis that trends are not monotone but quite complex [that too at different time scales]. Once we do such robotics with simple instruments like UV Spectrophotometer and Steady State Fluorimeter, we can study more number of osmolytes and more number of proteins [and various amyloidogenic conditions for each protein] in short time and in such a scenario we can't envision any other mode of analysis other than matrix formalism. Any theoretical formalism will make sense only when such a plethora of experimental data is available and our contention is that it is not the case presently.

Table-7.1: Matrix Formalism of Kinetic Time Points with Multiple Osmolytes at different concentrations using Multiple Techniques

Time	O_1 C_1 T_1	O_1 C_2 T_1	O_1 C_3 T_1	O_1 C_4 T_1	O_1 C_1 T_2	O_1 C_2 T_2	- - -	- - -	O_1 C_1 T_n	O_1 C_2 T_n	- - -	O_2 C_1 T_1	- - -	- - -	- - -		O_z C_1 T_1	O_z C_2 T_1	- - -	- - -	O_z C_4 T_n
t_1																					
t_2																					
t_3																					
-																					
-																					
-																					
t_{m-1}																					
t_m																					

O: Osmolyte (1,2, _ _ _ _ _ z) [O_1, O_2, O_3, _ _ _ _ _ _, O_z]

C: Concentration of Osmolyte (~ 4 different concentrations) [C_1, C_2, C_3, C_4]

T: Data from Technique number 1,2 ,3 _ _ _ _, n [T_1, T_2 T_3, _ _ _ _, T_n]

t_1, t_2, t_3, _ _ _ _ t_{m-1}, t_m: Time points since beginning of incubation

There has been only some conceptual development from my side with even smaller set of data in its support. But quest for larger of such matrices can be realized mainly due to large reservoir of techniques now at dispense and enormous literature data currently available and thousands of osmolytes available AND partly with easy Robotics solution (or in its absence by simply doubling manpower). What we would like to emphasize is that if Robotics is uneconomical then simply doubling man power will do, one working in daytime and another in night. However, since I was sole hand, that being the main reason as to why idea of matrix formalism' remained on 'drawing board' and could not be put to execution.

Science is not just visualization but numbers. Even in this thesis we have so many figures and despite having glanced over them for long periods, it is difficult to recall finer features in them. Actually experiments never produced figures, they produced numbers which were plotted due to automation of this process. In very old journal articles we find extensive tables and very few figures but with availability of plotting software's, we invariably go about plotting these numbers without a second thought. But in the process, these numbers are lost. This lacuna is addressed with above proposal.

****

7.2 Prognosis with Ethanol, Acidic & Guanidinium Conditions

Since the original idea was to make results free from condition employed, hence plan was made to use all the four conditions, viz., alkaline pH, acidic pH/elevated temperature, concentrated ethanol and Guandinium Hydrochloride/elevated temperature in the osmolyte work. However, before that is done, it was decided to probe these conditions in detail. As is happens barring acidic pH/elevated temperature, other conditions are barely studied beyond their discovery. We decided to therefore study these conditions in detail before introducing osmolytes. We decided to do this by systematically varying each of the physico-chemical parameters. It was expected that this will provide mechanistic insight into the aggregation process.

While the plan was to use osmolytes not only with alkaline condition but with other conditions as well. However, it turned out that only alkaline condition was well behaved: all osmolytes could be used with it even at high concentrations. There was no problem of precipitation or gelation (from beginning till end). Thus a claim can perhaps be made [we have already made a claim that Lysozyme is the model protein to study amyloidogenesis [1]]

that alkaline pH is the best condition to study osmolytes. One more reason for this claim is the fact that for any theoretical formalism we need a solution, not a precipitate or gel. Theoretical chemistry for solution is lot more developed than that of gel or bi-phasic. Add these two together, we reach conclusion "HEWL at alkaline pH is one of the best conditions to study effect of osmolytes". This is one major message of this thesis. As is clear from Table 5.3, acidic pH could not be used for many osmolytes [only Taurine & Trehalose were well behaved from beginning till end]. Prognosis however is that if we do indeed wish to use osmolytes under acidic condition, we have to make a departure from STANDARD CONDITION [viz., 600 μM HEWL, 57°C, pH 2.0] which is not amenable for such studies. In that case, we can either lower HEWL concentration OR lower the pH so that temperature is brought down. Perhaps this will yield a case with fewer precipitates though incubation period may go up.

Ethanol condition is known to be very susceptible to salts, so it was not expected that osmolytes [most of which are salts with exception of Trehalose, Taurine and Ectoine] will go well with and instead cause precipitation of HEWL. Actually it was planned to even change [under ethanol condition] the concentration of HEWL [as part of perturbation study] but could not be done.

For making ethanol samples [ethanol/water mixture of different v/v], it was noticed that if we try to dissolve HEWL powder in pre-mixed ethanol/water mixture, it fails to dissolve despite sonication or mild heating. Therefore, firstly HEWL powder was dissolved in water. Then this was added to ethanol [ethanol used was absolute ethanol from Fluka] and topped up with remaining water to achieve desired v/v. The way we conducted our ethanol experiments was different from its discoverers. They also used salt [NaCl at such low concentrations as 10 mM] in addition to ethanol and obtained amyloid precipitates [2]. However, their claim of using salt-free HEWL is wrong since they [like us] have used HEWL from Sigma which has 5% buffer salts in it. But we reasoned that precipitates are very difficult to study. Moreover, perturbation of salts is something we would like to do later and in this thesis focussed on using no exogenous salt [though 5% buffer salts are there]. Reason for not using salt [exogenous] was that we were interested in soluble amyloid [manifested as clear solution to naked eye though exhibiting Tyndall effect when viewed with 650 nm, 5 mW Laser pointer in orthogonal direction]. Certainly soluble amyloid is far from ideal solution wherein particle size is supposed to be tending to zero. By soluble amyloid we mean that particle size is very small [not visible to naked eye. It is to be noted

that vision acuity for adults is ~30 µm] and remains suspended with Brownian motion thus appearing homogeneous and mono-phasic. Though far from ideal solution, it is nevertheless a solution and hence extremely well developed principles of solution state physical chemistry can be applied [and many techniques working for solution state can be made use of]. On the contrary our knowledge for biphasic solution-solid equilibria is limited and so are techniques at hand. A very good and relevant example is applicability of steady state fluorescence anisotropy (r_{ss}). We could use r_{ss} only because it was a solution. Had it been solid or biphasic suspension, r_{ss} was ruled out due to problems of scattering. Detailed study with osmolytes under ethanol condition was not carried out and this remains future plan. Such studies are likely to be marred by the fact that many of the molecules used in this thesis are salts which when used with ethanol/water mixture are known to give precipitates (which could be amyloid but difficult to study and as of now we are interested in soluble amyloids). An elementary solubility experiment did however show that TMAO and Spermine [in free base form, not in salt form] did dissolve easily up to ~500 mM concentration in absolute ethanol; whereas Ectoine and Trehalose did not dissolve much. Of course here we are talking about solubility in absolute ethanol. Chances are they and other osmolytes could dissolve more in 80-90% Ethanol/Water [v/v] mixture OR [just as we circumvented solubility problem with HEWL as described above] first we dissolve them in pure water and dilute in ethanol OR even if that fails, we might opt to work with low concentration solutions. In this thesis we have only changed ethanol concentration and perturbation of temperature still remains on drawing board. Once that is done and result in amyloidogenesis at higher temperatures, we can make use of the same and it is likely that many osmolytes dissolve at such high temperatures. One perturbation in the form of HEWL concentration is also desired and it is very much possible that lower HEWL concentration might well tolerate osmolytes without causing precipitates.

So in future studies, we have at least two osmolytes which are likely to work under ethanol condition.

That leaves behind only Guanidinium Hydrochloride condition. However, investigation into this condition led to so many surprises, that idea to pursue osmolyte with it could not be pursued. For example, 5-M Guanidinium Hydrochloride at 45^{o}C, still had REA even after several hours. Mere dipping into 6-M Guandinium Hydrochloride is used as a gold standard for denaturation of proteins but here 5-M Guanidinium Hydrochloride

at 45ºC defies this. Obvious next step would be to see if 6-M Guanidinium Hydrochloride at 25ºC defies this as well [though that seems highly plausible]. If so, it would be against one of the fundamental postulates of protein chemistry. Nevertheless, engaging Guanidinium Hydrochloride in reverse direction is clearly way forward. This line has already been pursued by Nand Kishore et al. [3]. In doing so, let Guanidinium Hydrochloride (with shaking) make turbid aggregates and see if (and to what extent) same can be prevented by adding different concentrations of different osmolytes. Alternatively, if we want shaking-free condition, we may explore increasing the temperature (vast scope to do so, as it is aqueous) or increase concentration of HEWL.

****

Philosophically we feel; we have had too much of reductionist approach with proteins being in buffers. It is time we put them back to where they belong to, i.e., in the company of many small molecules which is what is now fast emerging field of metabolomics. However, what is being proposed is actually one of the oldest experiment of biophysics viz. Hofmeister series [4,5]. It is surprising that there is no satisfactory theory yet for Hofmeister effect even after 130 years. Partial blame goes to the fact that this series deals only with anions and cations. What we know there are un-ionized molecules as well, e.g. Trehalose and Ectoine in present study. Ions are far too hard to study theoretically as compared to un-ionized molecules as more number of forces are at interplay [columbic attraction/repulsion, hydrogen bonds, salt bridge etc.] along with those applicable for un-ionized molecules; whereas for un-ionized molecules lesser number of forces [6 7] are at play. So it is time that we move beyond 'ions' to neutral molecules' and evolve a new Hofmeister category for un-ionized molecules. It is contemplated that to understand Hofmeister effect we should focus on this 'un-ionized' category.

****

While bacterial assays are very simple and very cheap, it is impossible to rule out re-folding of protein as we dilute it to pH 7.0 buffer [though we have ruled it out for alkaline condition] where bacterial assay works especially when we work with Guandinium Hydrochloride condition [which is known to be reversible in nature]. To rule it out we have to move to a synthetic substrate of lysozyme. But they are very expensive substrates. These are 4-Methylumbelliferyl β-D-N',N''-triacetyl-chitotrioside (a fluorogenic substrate) [costing INR 20,000 for 5 mg] and 4-Nitrophenyl β-D-N,N',N'' triacetylchitotriose

[costing INR 47,000 for 5 mg]. On the contrary Micrococcus *lysodeikticus* is very cheap. Nevertheless, in the future we plan to use these substrates. But since they are not natural substrates how loyally they convey information about folded protein becomes somewhat controversial.

****

One surprising outcome of literature search was that for such a popular protein where more than 10,00 papers have been published in area of biophysics alone, the pI in mentioned across papers is 11.35 and this is what Sigma-Aldrich fact sheet quotes. This pI was determined in 1950's with a very old technique of Free Zone Electrophoresis (with moving boundary) way back in 1951 at a time when even protein sequence was not known with precision. The theoretical pI using ExPasy Protparam software is 9.4. this is massive difference. There is thus vast difference between theoretical pI and experimental pI. We pondered over it and decided to determine pI using 2D-gel electrophoresis but postponed it as future plan. Very recently however we came across a paper wherein they have determined pI using Isoelectric Focussing taking advantage of the synthesis of a new, strongly basic Immobiline (pKa 10.3 at 10^{o}C). thus helping to formulate a new pH 10-11 recipe for focusing very alkaline proteins, not amenable to fractionation with conventional isoelectric focusing in carrier ampholyte buffers [8]. It is time the Lysozyme community takes this paper into account, corrects the widely propagating error and does away with hitherto assumed pI of 11.35. Perhaps this is time, a database be started in protein chemistry along the lines of journal 'Organic Syntheses' [which describes experimentally verified syntheses of important organic molecules] which should describe experimentally verified data and procedures of very important proteins. Repetition (using novel methodology if available) by peers is likely to be more useful than mere self-repetition. This is kind of 'Peer Review' of 'Experiments' rather than mere 'peer review' of manuscripts. To some extent I did follow this. It was reported from our group in a previous paper [9] that Th-T at alkaline pH could not be recorded up to 48-hours due to problem of precipitation and could be recorded only after 48-hours. When I repeated, this was not the case and rigorous Th-T kinetics could be done even before 48-hours. This has been published in [1].

****

One major emphasis of this thesis is respect for all amyloidogenic conditions. It is sad that conditions like Guandinium Hydrochloride and ethanol are nearly died down

whereas alkaline condition is being studied only by our group. It is as if the condition has been patented with no other group willing to use them in their studies. Is not it surprising that such vital experiment such as Thioflavin-T was not done previously for ethanol condition. Again Goda condition emphasizes ionic strength but they have overlooked the fact that the HEWL which they have used is from Sigma and is not salt free as they have assumed but as per Sigma Fact Sheet 5% of it is buffer salts [not mere simple NaCl [which they have used to alter ionic strength] but many different salts. It is because of this approach; conditions fail to reach their detailed investigations. Mechanistic pathway deciphered/speculated using just one aggregation condition can't and should not be generalized since actual mechanism could drastically be skewed by the role of physiochemical forces at helm.

****

Amyloid is typically associated with precipitates. But then precipitates are very difficult to study, for example Goda et al [2] could not even do such a basic experiment as Thioflavin-T. What we have focussed in this thesis are soluble amyloid [amyloid which make clear solution to naked eye though at more microscopic levels, they do manifest Tyndall effect]. We did not use any added salt with a result that we managed to get soluble amyloid. Actually use of different amount of salt is 'another perturbation' which we will study in future.

****

At a time when fluid dynamics is such vastly researched subject, merely a phrase 'mechanical agitation' [as in Guandinium Hydrochloride condition] is used and no effort is made to even document RPM [rounds per minute] in case a shaker is used as mode of mechanical agitation. Entirely different types of forces are at fray when mechanical stirring is carried out [again at very different speeds]. We feel that application of experimental and theoretical fluid mechanics can go a long way to understand not only the kinetics of amyloidogenic pathways but protein dynamics in general. It can be envisioned that at low force it may favour oligomers/amyloid but if forces are very high they may break the same. However, to begin with application of Fluid Mechanics can be attempted on protein unfolding problem. May be it will fulfil the void in protein folding-unfolding research by opening an entirely different line of investigation.

****

While sticking to Lysozyme alone as protein, my only desire has been to use many different Osmolytes (majority of published literature deals with only two such Osmolytes one being TMAO (Tri-methyl Amine Oxide) and the other Betaine). Not even mono and dimethyl analogues of TMAO seem to have been thoroughly compared with TMAO side-by-side. Our other desire has been to use many different biophysical methods but then apart from 'as many as possible', our insistence has been on methods which would be more appropriate for chosen research problem. Now we can see that some such simple Properties introduced at school level books (some even in experimental curricula albeit with primitive experimental setups) like osmotic pressure, viscosity, surface tension and refractive index could be more useful than many sophisticated ones currently sought after.

****

Now it clearly seems that the abundantly naturally occurring (and thus extremely cheap) molecules of my studies could be tried to prevent/sufficiently delay Cataract in large population and resulting blindness. Recently a molecule (Lanosterol) has been discovered which clears cataract [10]. Either osmolyte tonic alone or as an adjunct to this molecule could heal cataract without surgery. Then there is talk about chaperone healing to overcome protein aggregation. At a time when there is so much talk about Patent etc., there is a naturally a hope for a natural and thus extremely cheap solution (more so because much lesser quantities would be required due to topical nature of application. Moreover, due to topical nature, toxicity due to ingestion can be avoided. In a way cataract could be first real test for trying out anti-amyloid nature of these molecules before attempting them for other amyloid diseases. As far as tonic concept with mixture of cocktails is concerned, the best disease to target is cataract [which affects a large population of World and is a leading cause of blindness in India] because of it being a localized disease [that too in an organ which is directly accessible without a need of a surgical opening]. It is surprising however that despite being an amyloid disease, cataract is barely studied as such despite the fact that it can be a truly LIVE amyloid model [without any need of transgenics, normal Laboratory mice can be used and cataract can be induced by exposure to UV light] which can be studied in real time and over long periods of time without any surgery, and efficacy of molecules be studied.

****

It must be emphasized that so far, only 3% of whole sea has been explored and not all invertebrates (again we need not limit to animal species but can look at plants or microbes under stressed conditions) have been thoroughly screened (especially at different embryonic stages). Recently a fish has been found at astonishingly low depth of ~8000 metres under Sea [11]. So it can be assumed that many more osmolytes are yet to be discovered, if not as a chemical entity, definitely in abundance and in this aspect the budding field of 'Metabolomics' [12] [13] appears to be very promising. 'Osmolytes' concept came into prominence in scientific literature after publication of an insight review by Yancey et al. in 1982. This late arrival is surprising because osmosis happens to be the first example of second law of thermodynamic and thoroughly studied in Plant Sciences throughout nineteenth century. Something as simple as a Plant Seed stores cells for years under a very dry state (and yet proteins and cells are 'happy' in such a dry state; this dryness definitely is a stressed condition for a milieu used to large amount of water, therefore seeds need to be studied using metabolomics)

****

Therefore, protein aggregation requires application of lot more techniques [experimental, theoretical modelling or simulation preferably all hand-in-hand way] before any claim for any significant study be made. However, it is not merely use of technique but to allow each technique to talk with as much verbosity as possible and at the same time reducing vagueness. It is where we believe the above 'matrix formalism' scores over pictorial representation with fewer data points.

As of now we have used merely three techniques. However, we are satisfied to some extent owing to following reasons:

a. Use of techniques in 'Kinetics' manner as rigorously as we could anticipate and execute with limitations such as instrument's availability (this 'Kinetics aspect' is major pillar of Physical Chemistry) and limitation of being single-handed [for incubations running over several days].
b. Demonstrate the power of residual enzymatic activity vis-à-vis first step of conformational fluctuations/unfolding [something not available with any other naturally amyloid forming protein] with significant reduction in error bars.
c. Simultaneous use of anisotropy kinetics and Th-T to decipher the message regarding oligomerization to structural reorganization.

d. It should be noted that apart from their usual interpretation [viz., oligomerization for anisotropy and extent of amyloid for Th-T] they also serve in supporting roles, e.g., Th-T does confirm that large aggregates have been formed [thus supporting findings of anisotropy]. This was reason we did not invest time in another supporting and qualitative technique of AFM/TEM. It should however be noted that excellent AFM mapping of aggregates by our colleague [from a similar research problem of Lysozyme amyloid under alkaline condition] Dr Vijay Ravi has confirmed formation/non-formation [and even 'extent' of formation for Lysozyme at amyloidogenic condition as well as when inhibited by Chitotriose] [14]. While kinetics of aggregation is being looked into by fluorescence anisotropy and Thioflavin-T but Th-T actually tells about final reorganized fibrils manifesting fluorescence properties which is like one of the 'postulates' for being classified as 'amyloid'. Therefore, to augment present work with 'another postulate meeting Congo-Red criteria' should be tried out and we are looking for ways to do that (and in a quantifiable way). However, extent of oligomerization (regardless of structural transition) is best deciphered by fluorescence anisotropy but in a qualitative way. We have spent considerable time pondering over use of light scattering methods. But now it is apparent that scattering methods involve too many approximations [apart from difficulty in sample handling such as absolutely dust free sample, procurement of expensive and more worryingly specialized instrumentation least likely to please funding bodies and then there is not just one but many different types of scattering such as DLS, SAXS, SANS etc. and all of them need to be applied in tandem. A very good paper in this regard is [15]. While this is an excellent paper, herein lies the travesty as well. They could use such sophisticated tools but neither did they use gold standard of Th-T nor salt free condition]. On the contrary Osmometer seems to be straightforward, highly sensitive, directly informative and most importantly hardly troubled by ANY APPROXIMATION because it depends on nothing but simply number of particles (regardless of size/shape of molecules). To the best of our knowledge, Osmometer has never been used in the field of protein aggregation though it is widely used in the field of essentially sibling field of polymers. Therefore, it would be worth trying out.

****

Thermodynamics & Kinetics of Protein Aggregation

Thermodynamics and Kinetics are two pillars on which classical physical chemistry rests. Yet we were awed to see that neither has been done in majority of papers dealing with protein aggregation.

However, if we are allowed to use one more concept/technique; it would be Thermodynamics. Thermodynamics is perhaps the only subject which is covered at length in both physical (including engineering) as well as Chemical Sciences in the first year of undergraduate curricula. Therefore, we are too keen get into it. Classical Physical Chemistry typically rests on Kinetics & Thermodynamics. If we have to reach to the core of mechanistic aspects, Thermodynamics is indispensable yet only a handful of papers in the field of protein aggregation deal with Thermodynamics in a serious way [3]. With a case already made for Kinetics in previous paragraphs, focus therefore shifts for Thermodynamics. If pursued, these measurements will connect us to the Thermodynamics of mechanistic investigation. Most suitable method for that seems to be micro-calorimeter. In its absence we can try non-direct methods such as melting of tryptophan/dansyl by fluorescence [16]; Guanidinium hydrochloride or Urea induced perturbations [17] [18] or a very simple method of 'second derivative method of absorption coefficient' [16]. We have tried this 'second derivative of absorption coefficient' experiment but results are too primitive to be included in this thesis.

Though not explicitly mentioned [19] as to why 57°C was used in the paper, it is known from an earlier reference [20] that it is T_m of HEWL at pH 2.0. The same paper [with rigorous use of micro-calorimeter] finds that T_m values are 68°C at pH 2.5 and 78°C at pH 4.5 which kind of suggests a way out of solubility problems encountered with most osmolytes at acidic pH. If we lower pH even lesser than pH 2.0, lesser temperature will be required and that may circumvent solubility problems. This paper by Privalov is classic and with availability of micro-calorimetry [though we could not get hold of one, a DSC is available at institute but requires large amount of sample], this should be the first experiment to be carried for any new amyloidogenic condition. Therefore, one of the first experiments in our pipeline is to determine T_m at pH 12.2 [and actually in the entire spectrum ranging from pH 11.0 to 12.8], at various ethanol/water mixtures and in presence of various Guandinium Hydrochloride concentrations and in the presence of osmolytes. We did however attempt a double differential method using absorbance spectroscopy but the results are too

elementary to be included in this thesis. However, T_m is not the only attraction we have, but entire plethora of physico-chemical parameters. It does make sense to hypothesize that once we introduce osmolytes, many of these parameters will vary. It is really surprising that thermodynamics and kinetics are two pillars of classical physical chemistry but detailed thermodynamics work is rarely carried out. Second pillar has not to do with availability of instrument but with approach. It requires the need to require as many timed data points as possible and to speak mathematically nearly 'continuously'. We did try to address it to some extent but were limited by instrument's availability in a shared scenario and bodily limits. However simple robotics could be a solution at a time when major advances are being made in big data analytics or in its absence doubling man power, one working in day shift and another in night shift.

Entry into Thermodynamics does not begin with first law of thermodynamics (and measurement of enthalpy, heat capacity etc.) but with zeroth law (which gives concept of Temperature). Accordingly, we can begin thermodynamics by doing measurements at different temperatures. If this be so, we can say that we did venture into Thermodynamics when we did temperature perturbations. But more such studies are needed, e.g., doing temperature perturbations in the presence of osmolytes.

****

Use of Osmometer & Viscometer

In this endeavour of pondering over applicability of different techniques, we have come to realize that for a process similar to polymerization, the simple technique of Osmometry is directly applicable since it makes little approximations unlike other techniques and is also highly sensitive without much instrumental sophistication. But we are yet to see any paper utilizing osmometry for protein aggregation. Another conceptually simple technique of Viscosity has also been used sparingly.

It was figured out that since essentially we are studying polymerization process so we need to heavily borrow techniques from the field of polymer science. Two such techniques are

a. intrinsic viscosity as an indicator of mass
b. surprisingly a very simple and a very direct technique for molecule mass determination without any of the approximations as used in viscosity or light

scatting measurements. This technique is thus deceptively simple. A small calculation has shown that pressure difference upon protein aggregation would be within reach of sensitivity of the osmometer by a huge margin.

Among colligative effects (osmotic pressure, freezing point depression & boiling point elevation); changes in osmotic pressure are the most sensitive (almost 1000 times over cryoscopic & ebullioscopic methods as per Flory's book written in 1953 [21]; But we are yet to see any application of osmometry in protein aggregation.

Pioneer of Macromolecules [Flory] gave an advice nearly 60-years ago, "the measurement of viscosity is so much simpler than execution of any of the absolute methods that the establishment of the intrinsic viscosity-molecular weight relationship is almost always a primary objective in a molecular weight investigation on a polymer series" [21]. But there are very few studies using this technique for protein aggregation (at least to my chosen sub-fields of 'Lysozyme' and 'Inhibition of Aggregation'). Recently Rheological measurements of Lysozyme in ethanol/water mixture have been carried out [22] but they have used very high concentrations of HEWL [8.2-21.0 mg/ml] whereas initial Goda's work as well us have used it at a much lower concentration of 2.0 mg/ml.. Such viscosity measurements are rare but should be done more commonly. We do have plan to do such measurement especially with osmolytes under acidic conditions where problem of gelation is common. Likewise, we do have plan to study effect of osmolytes on surface tension, a kind of study that is also rare but should be common.

****

Approach with regard to osmolyte use is similar to QSAR [Quantitative Structure Activity Relationship] widely used in pharmaceutical research. Though it is qualitative in approach as of now as far as this thesis or more generally the literature is concerned but it is hoped that if we introduce robotics and employ matrix approach to study large number of protein and large number of osmolytes, it will start turning 'QUANTITATIVE'.

****

Ammonium ion is perhaps the most familiar ion when it comes to Hofmeister series. It is perhaps first time that we see that un-ionized amine group performs in a similar way [as Polyamines exist at pH 12.2 in case of both primary and secondary amine groups].

****

While for mentioned reasons we could use HEWL and not Point mutants of human lysozyme. But eventually we will try to transfer knowledge gained on HEWL to Human Lysozyme and eventually like to work with amyloidogenic point mutants of Lysozyme. But we need not stop there. till now it is *in vitro*. More knowledge can be gained if we go *in vivo*. In order to go *in vivo*, we need transgenic animals. Currently transgenic models are mostly rodents centred around neurodegenerative disorders. But this approach has limitations. These are difficult to study unless animals are sacrificed [no live surgery is possible]. Another problem is complicated anatomy of brain further confounded by little insight into its understanding. What we propose it to use *in vivo* carrying mutations in Lysozyme gene. If this be the goal, what better than using chicken as a model. After all, one of the most studied protein in Biophysics comes from Chicken; so findings of thousands of papers with HEWL can be DIRECTLY corroborated. A back of envelope calculation has shown that the chicken egg has roughly 100-200 µM HEWL. Also a big mammal can be made to carry mutant Lysozyme gene using transgenics or CRISPR-CAS technology. If really they develop amyloid in viscera like humans, they can be ideal system to study amyloid formation [LIVE without sacrificing] *in vivo* as well as studying effect of inhibitors *in vivo*.

****

One major reason for choosing to work with HEWL was the fact, that it is favourite among biophysics community so lot of biophysical data is available. It is known to be easily crystallisable and small enough [only 129 AA's] to be workable with NMR. In fact, HEWL was one of the first proteins whose structures were known from both crystallography & NMR. While these traits of HEWL could not be made use of in this thesis but then it can be a future plan. It has been advised by peers to use ESMS but exactly how such an experiment is to be carried out is yet to be finalized.

****

A recent paper [23] suggests that pH 2.0 at 57ºC supplemented with 5% ethanol; out of 38 studied proteins 25 formed amyloid. In addition, 3 more did so if it was supplemented with 5% trifluoro ethanol instead of ethanol thus leading to 28 out of 38. It thus makes the condition kind of generalized so perhaps condition itself is amyloidogenic rather than the protein being amyloidogenic. But mechanistic pathway deciphered/speculated using just

one condition can't and should not be generalized since actual mechanism could be drastically skewed by the role of physicochemical forces at helm.

****

The thesis can barely be completed without mention on 1922 discovery by Fleming "on a remarkable bacterio-lytic element found in tissues and secretions" [24]. It reports both Lysozyme and Micrococcus *lysodeikticus* which is still the most widely used method for its enzymatic assay. Though discovered in nasal secretions, they found its high concentrations in egg-white of Hen, leading to its common acronym HEWL. In this paper, Fleming also reported his experiments with heating as well as effect of acid and alkali so in a way he was already on his path to make amyloid.

****

Use of osmolytes as has another advantage [apart from desired use as 'Inhibitors']. In a way they are acting as 'PROBES' towards discerning mechanistic aspects of the multi-step cascade and indeed they have done as shown in this thesis.

****

A Unified Model of Amyloidogenesis

The results with r_{ss} at different pH ranging from pH 2.0 to pH 11.0 indicate dimer at both pH 2.0 and pH 11.0 [Figure-4.10]. Sophianopoulos reported dimerization at pH 11.0 [25,26]. Then there is protein expression work where it is often found that protein behaves as dimer or even multimer on SDS-PAGE. Even under ethanol condition there is a report of dimer [27] . All this leads to an interesting hypothesis. Perhaps first step of all aggregations is the formation of dimer which under harsher conditions makes higher oligomers. Under slightly harsh [or in case of concentrated solutions, whereby proteins molecules are close and even exert intermolecular forces on each other], a reversible dimerization takes place. We know that proteins in solution exist as ensemble of many dynamic conformations. As soon as a fully folded protein undergoes such conformational transition, it creates two complementary open surfaces [which can be hydrophobic OR hydrogen bond donor and acceptor or their salt bridge counterparts]. Now these surfaces may patch up within same molecule leading back to fully folded protein. Such conformational fluctuations will be small at physiological pH but as we depart from physiological, such occurrences of 'opening up' will increase. At this juncture they may close with intramolecular association

but if concentrations of proteins in solution is high, then sticky patches of one protein will stick to complementary part of another and this will lead to dimer. We therefore speculate that dimerization is necessary first step of all protein aggregations. Situation at this step is like bi-functional monomer leading to a dimer. If we take above concept one step further, and move to harsher condition, conformational opening can occur at two or more places simultaneously. Situation is akin to a poly-functional monomer leading to a polymer. This hypothesis of dimer always preceding oligomer is also appealing from molecular modelling point of view. Modelling of entire oligomerization cascade at microscopic level is next to impossible but studying monomer-dimer equilibria quantitatively and at microscopic level is achievable. This monomer-dimer equilibrium can be a theoretical model system to study effect of osmolytes. Best starting point for such studies could be to crystallize HEWL at a pH where it is known to form a dimer, viz., pH 11.0. We did attempt crystallization in a rudimentary way. We started with 20 mg/ml solution at pH 11.0. After 10-days it gave needle like crystals. This raises hope [and here lysozyme's nature of readily forming crystals scores again] that we may be able to make single crystals of lysozyme at pH 11.0, the threshold from where it moves from dimer to oligomer. Later on we figured out a very old paper where they were successful in making a crystal at pH 11.0 [28]. So this indeed is realistic plan. In any case structure of a dimer would be interesting, as it will reveal the site with maximum propensity to 'open up'.

****

One major controversy in the literature is as to what is more toxic, an oligomer or an amyloid? We believe that both are toxic but oligomer is more toxic. Our reasoning is as follows: amyloid fibril is toxic because of its interaction and rupturing of lipid membrane leading to cell death but that predominantly occurs with neuronal plaques because there is no way to flush out amyloid fibrils. So they stay there and eventually fuse the neuronal cells. But as far as lysozyme is concerned, since deposition occurs in viscera so a mechanism is available to them to flush out of body using hepatic, renal or colon mode.

So there is lesser damage with amyloid fibril in body as compared to brain. However, oligomer on the other hand [with a set of chemically very diverse groups and loads of such chemical groups are present in them] can mess up with anything they come in contact with. Just like body makes over 10^{12} types of antibodies, similar scale of conformational environments [chemically and stereo-chemically] are created by oligomers

thus the power to annihilate proteins and cells and even tissues. That is why we believe oligomers are more toxic than amyloid fibrils. Our further hypothesis is that evolutionarily amyloid is body's safety mechanism. Since oligomers are more lethal so living beings have found to escape their lethal effects by zipping them into an amyloid so as to bury most of their functional groups and more so the vast space of conformational microenvironments by restricting stereo-chemistries. So to escape lethal effects of oligomers, either cells degrade them [just as they degrade unfolded proteins] perhaps we should not only have been talking about unfolded protein degradation machinery but an integrated 'unfolded protein and oligomer degradation machinery']. Perhaps studies of this oligomer degradation machinery is stillbirth as of now. Thus we have two scenarios to get rid of oligomer. Preferably to degrade oligomer and if that fails initiate plan B to zip them into a less lethal amyloid form. A recent paper by Eisenberg [29] talks about 'AMYLOME' hypothesis according to which most of the proteins are capable of forming amyloid

****

Studying at alkaline pH brought its own share of occupational benefits inadvertently. Proteins are rarely studied at such high pH so surprises were anything but unexpected. One is the fact, that now we know that Polyamines (or for that matter the ammonium ions in Hofmeister series) don't work because of being cation but by unionized amine moiety as well [which is what remains at 12.2 for a pKa of ~10.0]. Likewise, we stuck down a theory which suggests that mode of arginine action is interaction of positive charge of Guanidino group with tryptophan of proteins [30]. At such high pH, even Guanidino group of Arginine [pKa ~12.0] tends to deprotonate and hence it rules out any mechanism of lysozyme action owing to charge at Guanidino group. This high pH also poses challenge to existing theories of Thioflavin-T. As we tested that pH in final assay with Thioflavin-T was 10.9 and not 8.5, yet it exhibited significant Thioflavin-T fluorescence. Future theories of Th-T must take it into account.

****

Phosphate has three pKa's viz., 2.12, 7.21, & 12.22. So it can buffer pH 2.0 & pH 12.2 [both amyloidogenic] and pH 7.0 [physiological & non-amyloidogenic]. So we have a major advantage that phosphate buffer provides opportunity to oscillate pH between pH 12.2, pH 2.0 [both known to be amyloidogenic] and pH 7.0 [physiological and non-amyloidogenic. Thus we are armed with a powerful pH switch by which we can alter between two amyloidogenic state and non-amyloidogenic state. Thus we can carry out

experiments wherein we can proceed to make oligomer/amyloid at one amyloidogenic state, change pH *in situ* to pH 7.0 [and see whether such a process is halted or keeps on proceeding] or change pH to other extreme and monitor what happens now when entire ionic atmosphere is reversed. Essentially this experiment is similar to the temperature jump experiment carried out in the thesis but far more powerful because we can try to tame ionic environment. Though these switch experiments were not carried out, initial experiments suggested that in the presence of HEWL, the pH could be switched back and forth without any hysteresis.

****

Two experiments which were planned but could not be executed are:

a. Pre-soak osmolytes overnight with HEWL before changing pH to 12.2 [or instead, simply store them at 4^{o}C where we have shown that the process is frozen and then simply allow it to come to 25^{o}C] and see if the pattern of unfolding, oligomerization and amyloidogenesis is any different as compared to present case when everything is mixed all at a time.
b. Yet another experiment which could not be carried out; was to add osmolytes not from beginning but at different time points in between. Again this has bearing to the 'switch' concept introduced in the previous section. But in particular we can find out whether osmolytes can slow down/halt/reverse the aggregation process [as data from Spermidine shows] and if so in which time frame. Moreover, this will resemble physiological scenario where medicines are administered long after aggregation process has gone on [to such an extent that they can be diagnosed].

****

Fragmentation of protein was one issue which we did not study. It is known that at pH 2.0/57^{o}C, fragmentation does occur [and in the process they identified a peptide from central part; which forms amyloid in matter of 24-hours as compared to 2-weeks for full length HEWL]. In future we would like to make this peptide and use it for alkaline, ethanol and Guandinium Hydrochloride conditions. While we can't rule out fragmentation at pH 12.2, we can hope that much shorter incubation [~48-hours] and at being room temperature, it will not be profound. In this aspect, it is worth noting that HEWL even at pH 12.2 [but at 4^{o}C], retained full activity even after weeks. The pH rupturing ionic environments and hence affecting unfolding may be frozen in this case if T_m is really more than 4^{o}C. However,

it is likely to have much less effect on slowing of chemical fragmentation. So the fact that protein retains near full activity even after several weeks at 4°C (pH 12.2) or at pH 11.0 (25°C) points to a good extent that there is no significant chemical fragmentation at alkaline pH for 2-3 days. However, there could be significant fragmentation leading after several days and that could explain as to why 11-days old sample [in Lateral Addition experiment] lost all the associative tendency. However, if we wish to get rid of chemical fragmentation completely we can use ethanol or Guandinium Hydrochloride condition but then they have been nearly abandoned.

****

One perturbation which we would like to do is to move beyond and use other homologous alcohols [Methanol, Ethanol, n-Propanol, iso-Propanol, n-Butanol, tert-Butanol] and mixture of them [with or without water]. Idea is not merely to change polarity but chemical environments at micro level.

So far most people have focussed on single amyloidogenic condition. We can move beyond and use mixed amyloidogenic condition. In essence, it is similar to mixed osmolyte case because both are essentially co-solvents [here old debate of solute Vs solvent merges, at such lower concentrations of 100-200 mM dissolved in 55 M water, it is to be called solute or solvent; perhaps if we are talking in Molar range, it can be called solvent such as Guandinium Hydrochloride or as people use Molar range of osmolytes such as Proline, it can be called co-solvent but at 100-500 mM concentrations used in this thesis better to call it solute. Though it is nomenclature but makes big difference when we attempt theoretical formalism]. So far only one group has tried to combine more than one amyloidogenic condition [31,32]. They found that "lysozyme at pH 7.0 and 11.0 are nearly stable to the addition of up to 45% t-butanol, but treatment of the alkali-induced molten globule form of HEWL [AMGL] with 20% t-butanol caused the formation of amyloid-like fibrils as evidenced by enhanced Thioflavin T binding and DLS measurements". They used only t-Butanol at 45% in pH 12.7 [31]. In yet another paper they tested effect of ionic strengths (2-M KCl) at pH 12.6 [32]. Since they are the only workers to be working at such high pH apart from our group; so their work is directly correlated to ours. Work of Hameed et al & Ansari et al is thus a very good example of mixed amyloidogenic condition for Lysozyme. We tried to do it in a more systematic way; using pH 12.2 and adding to it 10%, 20%, 30% & 40% of Methanol, Ethanol, iso-Propanol & tert-Butanol but results are too elementary to be included in this thesis. Similar in spirit (though at a much lower concentration of 5%

ethanol) is the use of 5% ethanol at pH 2.0 which markedly increased amyloidogenic potential at pH 2.0 [to the tune that 25 out of 36 proteins formed amyloid]. Then there are many solvents typically used in biophysics of proteins and with whom there are lots of data [including that with Lysozyme] are hitherto untouched as far as amyloid of Lysozyme is concerned. Take for example, DMSO [33], Acetonitrile and so forth. Their isolation has to be ended and they have to be 'touched' if we want to reach to the core of mechanism. But here again 'soluble amyloid' scores.

****

Amongst all amino acids two have drawn maximum attention. One is Arginine and another is Proline. We feel that other amino acids (especially those with charged side-chains) need not be ignored based on few previous studies and should be studied more often (all in parallel,) because they are bread and butter for cells. we did study Lysine & Ornithine in this study whilst focussing on Arginine. Mere presence of amino acid is not important; what matters more is the concentrations at which they occur. Here we have to watch out for what Metabolomics of various normal and diseased cells come up with. Just why amino acids act as protectant of proteins? We feel it is because amino acids make proteins so they have common features. All charged side chain amino acids have a positively charged amino group, a negatively charged carboxyl group, another charged side chain and an intervening hydrophobic part of variable length. It is plausible that these charges engage opposite charges on a growing oligomer. It should be noticed that engagement by minimum two groups is essential to provide this sort of 'coating'. This active coating hypothesis can easily explain the protection they offer to proteins against adverse environment such as acidic or alkaline pH. The coating hypothesis also explains protection against elongation of oligomers by subduing opportunities available. While at very high or very low pH typically used in *in vitro* models one of them becomes uncharged, same does not happen at physiological pH and hence this coating hypothesis is likely to be more beneficial *in vivo* as compared to *in vitro* models. This active coating hypothesis is in contrast to Osmophobic effect theory/preferential hydration theory. This coating hypothesis gains traction in the light of results we obtained with Polyamines. The 1,2-diamino ethane & 1,3-diamino propane had lesser inhibition on amyloid whereas all higher homologues had nearly constant and very high inhibition. This result can be explained by arguing that lower homologues could not engage two negatively charged groups simultaneously whereas higher homologues could do that (either by full intervening coating or with puckering). Of

course we note that at pH 12.2, amino groups of polyamines are not expected to carry full-fledged negative charge. So some alternative mechanism could be operative. This alternative mechanism could be the hydrogen bonding.

****

Amino group is the best Hofmeister cation and Sulphate is the best Hofmeister anion. Why? One reason could be that the charge is evenly spread over [due to resonance] over a large surface area, thus helping them interact with proteins/oligomer more. Mechanism of Hofmeister effect [e.g., salting out in presence of ammonium ions] is not known. One simple reason could be that ammonium ions mask all exposed negatively charged residues especially with its charge spread out over large surface area. And due to decrease in effective charge on protein, it becomes less soluble.

Hofmeister series for cations is as under [6]:

$NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+} >$ Guandinium

The hypothesis of charge spreading out over large surface area can explain trend observed with Potassium, Sodium and Lithium ions (and very markedly with ammonium ions).

****

It is reprehensible that protein Lysozyme has been confined to one paragraph of standard Immunology textbooks. Neither mechanism is mentioned nor the secretion sites all over body. Such huge amount of Lysozyme found in egg-white must have a bearing to safety of human embryos (or perhaps even in paediatric subjects]) where rigorous immunological mechanisms have not yet been kicked off. Perhaps Immunology textbooks should provide more space to Lysozyme (along with the field of 'antimicrobial peptides' which has recently caught attention of scientific community). This is all the more important at a time when world is at the crux of shortage and saturation of antibiotics. It can be envisioned that perhaps Lysozyme can pave way to a class of antibiotics. One could be a synthetic oligomer of Lysozyme. Lysozyme is not just confined to nasal secretions as is presented in textbooks but even in many visceral Organs [34] . There is certainly vast scope of anatomical, physiological and biochemical research into Lysozyme (off course not restricting to Humans but into a wide variety of species because it is found even in bacteriophage). Yet another surprising feature of Lysozyme is that Human and Chicken Lysozyme have merely 60% sequence homology; yet their tertiary structures can be nearly superimposed on each

other. Just why Nature allowed sequence divergence but not the structure needs to be looked into.

****

Deep Sea research is difficult [35]. Going as deep as up to 8 kilometres requires sophisticated ships and catching equipments. But once deep sea animals are caught, can't we make a breeding lab for them on surface (much like Fisheries). It should be possible to create such high pressure, darkness, availability of organic materials as available in their natural habitat. This could be a costly venture but will accelerate understanding into deep sea ecology.

****

The results with mixture of osmolytes could have wide ranging ramifications. For one; using a great many of them in minuscule will help them keep below their individual toxicological level. For example, the most popular osmolyte TMAO (which occurs in two vastly different stressed conditions of urea-laden kidney and high pressure deep sea) and which has been focus on many theoretical and experimental biophysical investigations. Yet high serum concentrations of TMAO (and also choline & Betaine) in the form of dietary supplement in mice triggered cardiovascular diseases [36]. So TMAO, Betaine and Choline at very high concentrations are unlikely to work. But if they are part of a cocktail with hundreds of osmolytes (secondary or tertiary metabolites), then their concentrations can be lowered to much below toxicological level. Same generalizes for all other molecules. Perhaps it is these osmolyte cocktails which keeps proteins happy in body and in old age it is likely that osmolytes are pruned as each one requires a cascade of biosynthetic pathways and therefore their syntheses are energetically and metabolically demanding. So exogenous supply of mixture of osmolytes might not only help with protein aggregation and misfolding diseases but also counter effects of ageing.

****

References

1 Swaminathan, R., Ravi, V. K., Kumar, S., Kumar, M. V. S. & Chandra, N. in *Advances in Protein Chemistry and Structural Biolog* (ed R. Donev) 84, 63-111 (Elsevier Academic Press Inc, **2011**).

2 Goda, S. *et al.* Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein science* 9, 369-375, doi:10.1110/ps.9.2.369 (**2000**).

3 Choudhary, S. & Kishore, N. Addressing mechanism of fibrillization/aggregation and its prevention in presence of osmolytes: spectroscopic and calorimetric approach. *PloS one* 9, e104600, doi:10.1371/journal.pone.0104600 (**2014**).

4 Hofmeister, F. 11. Zur Lehrevon der Wirkung der Salze. *Archiv fuer experimentelle Pathologic and Pharmakologie* XXIV, 247-260 (**1887**).

5 Kunz, W., Henle, J. & Ninham, B. W. 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. *Current Opinion in Colloid & Interface Science* 9, 19-37, doi:10.1016/j.cocis.2004.05.005 (**2004**).

6 Creighton, T. E. *Proteins: Structures and Molecular Properties*. Second edn, (W. H. Freeman and Company, **1993**).

7 Karshikoff, A. *Non-Covalent Interactions in Proteins*. (Imperial College Press, **2006**).

8 Gelfi, C., Bossi, M. L., Bjellqvist, B. & Righetti, P. G. Isoelectric focusing in immobilized pH gradients in the pH 10-11 range. *Journal of biochemical and biophysical methods* 15, 41-48 (**1987**).

9 Kumar, S., Ravi, V. K. & Swaminathan, R. How do surfactants and DTT affect the size, dynamics, activity and growth of soluble lysozyme aggregates? *The Biochemical journal* 415, 275-288, doi:10.1042/bj20071499 (**2008**).

10 Zhao, L. *et al.* Lanosterol reverses protein aggregation in cataracts. *Nature* 523, 607-611, doi:10.1038/nature14650 (**2015**).

11 Yancey, P. H. *Deep-Sea Biology*, <http://people.whitman.edu/~yancey/deepsea.html> (**2017**).

12 Kopka, J. *et al.* GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics (Oxford, England)* 21, 1635-1638, doi:10.1093/bioinformatics/bti236 (**2005**).

13 Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature chemical biology* 5, 593-599, doi:10.1038/nchembio.186 (**2009**).

14 Ravi, V. K. *Biophysical Investigations of Hen Lysozyme Aggregation and its inhibition at Alkaline pH* PhD thesis, Indian Institute of Technology Guwahati, (**2011**).

15 Fujiwara, S., Matsumoto, F. & Yonezawa, Y. Effects of Salt Concentration on Association of the Amyloid Protofilaments of Hen Egg White Lysozyme Studied by Time-resolved Neutron Scattering. *Journal of molecular biology* 331, 21-28, doi:http://dx.doi.org/10.1016/S0022-2836(03)00722-8 (**2003**).

16 Creighton, T. E. *Protein Structure: A Practical Approach*. Second Edition edn, (Oxford University Press, **1997**).

17 Greene, R. F., Jr. & Pace, C. N. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *The Journal of biological chemistry* 249, 5388-5393 (**1974**).

18 Thomson, J. A., Shirley, B. A., Grimsley, G. R. & Pace, C. N. Conformational stability and mechanism of folding of ribonuclease T1. *The Journal of biological chemistry* 264, 11614-11620 (**1989**).

19 Morozova-Roche, L. A. *et al.* Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *Journal of structural biology* 130, 339-351, doi:10.1006/jsbi.2000.4264 (**2000**).

20 Privalov, P. L. & Khechinashvili, N. N. A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *Journal of molecular biology* 86, 665-684 (**1974**).

21 Flory, P. J. *Principles of Polymer Chemistry*. (Cornell University Press, 1953).

22 Szymanska, A., Hornowski, T. & Slosarek, G. Denaturation and aggregation of lysozyme in water-ethanol solution. *Acta biochimica Polonica* 59, 317-321 (**2012**).

23 Aso, Y., Shiraki, K. & Takagi, M. Systematic analysis of aggregates from 38 kinds of non disease-related proteins: Identifying the intrinsic propensity of polypeptides to form amyloid fibrils. *Bioscience Biotechnology and Biochemistry* 71, 1313-1321, doi:10.1271/bbb.60718 (**2007**).

24 Fleming, A. On A Remarkable Bacteriolytic Element Found In Tissues And Secretions. *Proceedings of the Royal Society of London. Series B* 93, 306-317 (**1922**).

25 Sophianopoulos, A. J., Rhodes, C. K., Holcomb, D. N. & Van Holde, K. E. Physical studies of lysozyme. I. Characterization. *The Journal of biological chemistry* 237, 1107-1112 (**1962**).

26 Sophianopoulos, A. J. & Vanholde, K. E. Physical Studies Of Muramidase (Lysozyme). Ii. Ph-Dependent Dimerization. *The Journal of biological chemistry* 239, 2516-2524 (**1964**).

27 Yonezawa, Y. *et al.* An Insight into the pathway of the amyloid fibril formation of hen egg white lysozyme obtained from a small-angle X-ray and neutron scattering study. *Journal of molecular biology* 323, 237-251 (**2002**).

28 ALDERTON, G., WARD, W. H. & FEVOLD, H. L. ISOLATION OF LYSOZYME FROM EGG WHITE. *Journal of Biological Chemistry* 157, 43-58 (**1945**).

29 Goldschmidt, L., Teng, P. K., Riek, R. & Eisenberg, D. Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3487-3492, doi:10.1073/pnas.0915166107 (**2010**).

30 Tsumoto, K. *et al.* Role of arginine in protein refolding, solubilization, and purification. *Biotechnology progress* 20, 1301-1308, doi:10.1021/bp0498793 (**2004**).

31 Hameed, M., Ahmad, B., Khan, R. H., Andrabi, K. I. & Fazili, K. M. Tertiary butanol induced amyloidogenesis of hen egg white lysozyme (HEWL) is facilitated by aggregation-prone alkali-induced molten globule like conformational state. *Protein and peptide letters* 16, 56-60 (**2009**).

32 Ansari, M. A. *et al.* Identification and characterization of molten globule-like state of hen egg-white lysozyme in presence of salts under alkaline conditions. *Protein and peptide letters* 17, 11-17 (**2010**).

33 Bhattacharjya, S. & Balaram, P. Effects of organic solvents on protein structures: observation of a structured helical core in hen egg-white lysozyme in aqueous dimethylsulfoxide. *Proteins* 29, 492-507 (**1997**).

34 Trexler, A. J. & Nilsson, M. R. The formation of amyloid fibrils from proteins in the lysozyme family. *Current protein & peptide science* 8, 537-557 (**2007**).

35 Yancey,P.H. DEEP-SEA BIOLOGY. ***http://people.whitman.edu/~yancey/deepsea.html*.**

36 Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472, 57-63, doi:10.1038/nature09922 (**2011**).

Publications from this Thesis:

1. Swaminathan, R., Ravi, V. K., Kumar, S., Kumar, M. V. S. and **Chandra, N.** (2011) Lysozyme: A Model protein for amyloid research. *Adv. Prot. Chem. Struct. Biol.,* Donev, R.M. (editor), Vol. 84, Academic Press, 2011, pp. 63-111. ISBN: 978-0-12-386483-3

2. Ravi, V. K., Swain, T., **Chandra, N.** and Swaminathan, R. (2014) On the characterization of intermediates in the isodesmic aggregation pathway of hen lysozyme at alkaline pH. *PLoS ONE* 9(1): e87256 doi 10.1371/journal.pone.0087256.

3. Chhabra, G., **Chandra, N.** and Swaminathan, R. (2017) Osmolytes: Key Players in Regulating Protein Aggregation. Cellular Osmolytes: From Chaperoning Protein Folding to Clinical Perspectives. Pages 97-119. Singh, R.K., Dar, L. and Ali, T. (Editors)

Manuscripts to be submitted:

1. **Chandra, N.** and Swaminathan, R., (xxxx) Inhibition of Hen Lysozyme Amyloidogenesis in the presence of multiple Osmolytes

2. **Chandra, N.** and Swaminathan R. (xxxx) Perturbation of amyloidogenic conditions of Hen Lysozyme

3. **Chandra, N.** and Swaminathan, R., (xxxx) Numeric Code of Amino Acids and Applications thereof.

Conferences:

1. **Chandra, N** and Swaminathan, R. Inhibition of Lysozyme Amyloidogenesis by Osmolytes. Biophysical Society 56th Annual Meeting, February 25-29, 2012 in San Diego, California (USA)

2. **Chandra, N** and Swaminathan, R. Matrix formalism to emphasize 'Kinetics' aspect in multinomial cascade leading to Protein aggregation. Theoretical Chemistry Symposium 2012. December 19-22, 2012 Department of Chemistry, IIT Guwahati (INDIA)